

## Isolation and Characterization of the 2S Albumin Gene and Promoter from Grapevine

Z.T. Li, D.J. Gray, S.A. Dhekney, M. Dutt, M. Van Aman, J. Tattersall and K.T. Kelley  
Mid-Florida Research and Education Center  
University of Florida/IFAS  
2725 Binion Road  
Apopka, FL 32703-8504  
USA

**Keywords:** *Vitis vinifera*, seed-specific promoter, gene activation, *cis*-acting elements, transformation

### Abstract

High-throughput sequencing of cDNA libraries has resulted in millions of expressed sequence tags (ESTs) from plants. To exploit such a valuable molecular resource for functional analysis of genes and genetic elements, we developed an improved thermal asymmetric interlaced (TAIL) PCR technique. We demonstrated its usefulness by recovering the complete seed-specific 2S albumin gene and promoter using a partial EST and genomic DNA of *Vitis vinifera* grapevine. The 2S albumin *VvAlb1* (*V. vinifera* 2S albumin 1) gene obtained from different cultivars encompasses a coding region of 504 to 540 nucleotides corresponding to a deduced amino acid sequence of 167 to 179 residues. This deduced protein contains up to 30% glutamine residues and 8 cysteine residues arranged in a pattern highly conserved among 2S albumins for disulphide bond formation. DNA sequence alignment revealed that the same *VvAlb1* gene among different grape cultivars varied greatly, including an insertion of up to 36 bp near the 3' end of the gene sequence isolated from *V. vinifera* 'Thompson Seedless'. DNA sequence analysis indicated that a number of highly conserved seed-specific regulatory motifs were present throughout a 2.2 kb region 5' upstream of the transcription start site of the *VvAlb1* gene. Comparative analysis of promoter activity using both EGFP and GUS genes directed by various artificially truncated promoter fragments established the function of several promoter regions in regulating seed-specific gene expression in both transiently and stably transformed grape SE. In particular, a 0.4 kbp promoter fragment (-1 to -404) supported the highest level of transient expression but failed to produce any detectable level of expression in stably transformed SE. However, an opposite situation was found when a 2.2 kbp promoter fragment was used. In addition, all transgenic SE lines harboring this long promoter fragment, but not other shorter fragments, showed arrested embryo development. Plants were regenerated successfully from all transgenic SE lines and are being evaluated for temporal and spatial regulation of promoter activity. These results exemplify TAIL-PCR as a cost-effective and target-specific method to utilize the vast amount of molecular information now available.

### INTRODUCTION

Technological improvements for the genetic transformation of grape have significantly enhanced our ability to transfer foreign genes into the *Vitis* genome. This has facilitated efforts to genetically engineer grape in part by supporting genetic/genomic studies to determine gene functions and mechanisms of gene expression and regulation. International efforts to sequence various cDNA libraries in grapevine have produced a wealth of molecular information, e.g., over 183,000 expressed sequence tags (ESTs) of *Vitis vinifera* ([http://www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList\\_sz.html#EST](http://www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList_sz.html#EST)) have been deposited in the GenBank database of National Center for Biotechnology Information (NCBI). These sequence data are useful for the analysis of gene expression profiles, development of molecular markers, and the isolation and functional analysis of unknown genes and genetic elements.

In order to isolate seed-specific promoters, we modified the thermal asymmetric interlaced PCR (TAIL-PCR) procedure originally described by Liu and Whittier (1995) and amplified several genomic DNA sequences utilizing the sequence of GenBank accession No. AY267256 - an EST corresponding to a 2S albumin gene. Sequence analysis of the amplification products revealed the complete *VvAlb1* gene encoding a 2S albumin precursor polypeptide and a 0.6 kb 5' flanking sequence conferring seed-specific promoter activity (Li and Gray, 2005).

In this study, we analyzed a 2.2 kb 5' sequence flanking the *VvAlb1* gene for promoter activity via progressive 5' deletions. By using both transiently and stably transformed grape somatic embryos (SE) and plants, the importance of several regions of the extended *VvAlb1* promoter in regulating temporal and spatial gene expression was revealed. The implication of genetic elements within promoter regions on efficacious DNA-protein interactions and tissue-specific gene activation is also discussed.

## MATERIALS AND METHODS

### Plant Materials and *Agrobacterium*-Mediated Transformation

Somatic embryos were induced from in vitro-grown leaves of *V. vinifera* 'Thompson Seedless' and maintained as described previously (Gray, 1995; Li et al., 2001). Somatic embryos at the mid-cotyledonary stage of development were used in all transformation experiments.

A modified *Agrobacterium*-mediated transformation procedure was used to introduce test constructs into grape (Li et al., 2005). After co-cultivation with *Agrobacterium*, SE were placed on selection medium to induce transgenic SE that were subsequently used for regeneration of transgenic plants (Li et al., 2001). Somatic embryos, 3 d post-co-cultivation, were collected, processed, and utilized in transient gene expression studies.

### Source of Promoter Sequences and Construction of Plant Transformation Vectors

A 2.2 kb DNA fragment 5' upstream of the *VvAlb1* gene was previously amplified from *V. vinifera* 'Merlot' (Li and Gray, 2005) and separated into 5 fragments by progressive deletion at the 5' end. These fragments were designated as -2145, -1711, -1071, -551, and -357 based on their 5' end positions relative to the transcription initiation site. These promoter fragments were end-modified to incorporate a unique *HindIII* site at the 5' end, a unique *KpnI* site at the 3' end, and then cloned into a pBIN19-based binary vector. This resulted in a series of binary constructs referred to as A1 (-2145), A2 (-1711), A3 (-1071), SAG (-551) and A4 (-357) (Fig. 1). In each construct, a GUS expression unit under the control of the respective promoter fragment and a bi-functional EGFP/NPTII expression unit controlled by a doubly-enhanced CaMV 35S promoter (Mitsuhara et al., 1996) were arranged in tandem within the T-DNA region. A positive control vector (pd35G) containing a GUS gene controlled by a doubly-enhanced CaMV 35S promoter (Kay et al., 1987), and a negative control vector (pD35S) lacking a GUS expression unit (Li et al., 2001, 2004) were included for comparison. All binary vectors were introduced into *A. tumefaciens* strain EHA105 via the freeze-thaw method (Burrow et al., 1990).

### Monitoring GFP Expression and GUS Assays

GFP expression was monitored as described by Li et al. (2001). Multiple images of GFP expression in transformed tissues were recorded and representative images analyzed. Somatic embryos and leaf explants were collected and assayed for GUS activity according to Jefferson (1987). For abscisic acid (ABA) induction experiments, stable transgenic SE at the mid-cotyledonary stage of development were divided into two groups. One group was placed on medium containing 10  $\mu$ M ABA and the other on identical medium lacking ABA for 5 d. These SE were then assayed for GUS activity. For quantitative expression analyses, three replicates of at least 30 SE were used for each treatment. Experiments were repeated two to three times depending on the availability of explant materials.

## RESULTS AND DISCUSSION

### Sequence Structure of an Extended 2S Albumin Gene Promoter from Grape

We previously determined that a 0.6-kb 5' flanking sequence of the *VvAlb1* gene contained a number of *cis*-acting elements associated with seed-specific promoters in other species (Li and Gray, 2005). However, this promoter sequence lacked the AT-rich regions commonly associated with DNA curvature and gene activation (Wu and Crothers, 1984; Koo et al., 1986).

To determine whether sequences upstream of the previously tested promoter fragment contained additional regulatory elements, a larger amplified fragment (2182 bp in length) was cloned and sequenced. This fragment contained the 2145-bp 5' flanking sequence (-1 to -2145) and a 37-bp 5' untranslated leader (UTR) sequence (+1 to +37) of the *VvAlb1* gene.

Sequence analysis revealed a region rich in A- and T-tracts (at least A<sub>4</sub> or T<sub>4</sub>) located between nucleotide positions -1070 and -1710. In addition, a relatively small number of sporadic A- and T-tracts were present in the region spanning nucleotide positions -350 to -550.

### Transient GUS Expression Analysis

Results of the fluorogenic analysis of transient GUS activity revealed a relatively high level of expression in the control vector pd35G (data not shown). However, much lower levels of GUS activity were observed in all *VvAlb1* gene promoter-containing vectors. Similarly low levels of promoter activity were found in  $\beta$ -*phaseolin* promoter-controlling gene expression for major seed storage protein in common bean, *Phaseolus vulgaris* (Chandrasekharan et al., 2003; Grace et al., 2004). Among the tested vectors, A4 produced a highest level of activity, followed by SAG and A3. Further reductions in GUS activity were observed in vectors A2 and A1 (Fig. 2). Noticeably, two promoter fragments -551 (SAG) and -357 (A4) resulted in a significant difference in GUS activity. Inspection of the DNA sequence from -551 (SAG) to -357 (A4) revealed that this region harbored a copy of H-box crucial for stress-responsive and tissue-specific gene expression (Sugimoto et al., 2000). However, the removal of the H-box from the -551 fragment in SAG appeared to enhance the level of transient GUS expression (Fig. 2, SAG vs. A4).

### Stable GUS Expression Analysis

A total of 14 stable transgenic SE lines containing various *VvAlb1* gene promoters were selected for GUS expression analysis based on their uniformity for GFP expression. These lines included 4 each of A1 and SAG, and 3, 2 and 1 from A2, A4 and A3, respectively. All transgenic SE of A1 lines showed arrested embryo development with reduced cotyledonary structure, whereas SE containing other *VvAlb1* promoter fragments exhibited normal embryo morphology (data not shown).

As illustrated in Fig. 3, only the A1 lines yielded a significantly higher level of stable GUS activity among the tested vectors. These results suggest that the DNA region from -2145 to -1710 (435 bp 5' region) contained *cis*-acting or structural elements vital for promoter activity in a stable expression system.

GUS assays using leaf tissues revealed that transgenic plants of A1-, A2- and A3-lines did not exhibit GUS expression. Only traces of GUS activity were detected from plants of SAG- and A4-lines, suggesting a leaky expression in these lines (data not shown).

Examination of the 435 bp 5' region revealed one copy of the F2 motif previously shown to be involved in seed-specific gene activation (Vincentz et al., 1997). Two additional copies of the F2 motif were also present downstream. Whether all three F2 motifs are required to induce DNA-protein interactions and the specific role of each element in gene activation await further examination. Additionally, a copy of the GT2 motif, a binding site for transcription factor GT-2, was found within this 435 bp 5' region. GT-2 is a well-characterized plant transcriptional activator with trihelix DNA-binding

capability and is responsible for light down-regulation and tissue-specific gene expression (Nagano et al., 2001; Perisic and Lam, 1992). However, the involvement of GT-2-mediated transcriptional activation in seed-specific promoters has not been fully substantiated.

Global structural changes may also have contributed to the various degrees of promoter activity associated with promoter deletions. An A/T-rich region was identified between -1710 (5' end of A2) and -1070 (5' end of A3). This region contained numerous A/T homopolymeric tracts that are well known for inducing DNA curvature (Wu and Crothers, 1984; Koo et al., 1986). It is likely that the formation of productive promoter geometry involves the 435 bp 5' region that provides binding sites for regulatory proteins while the downstream A/T tract-rich region facilitates the bending of the DNA molecule necessary for gene transcriptional activation. The deletion of the 435 bp 5' region removed anchoring points and the ability of the promoter to develop the functional structure required for stable temporal and spatial expression. Comparative analysis of DNA sequences with those of other seed-specific promoters such as *Napin B* and  $\beta$ -*Phaseolin* indicated that the pattern of A/T-tract distribution was unique in the *VvAlb1* promoter (Ericson et al., 1991; Slightom et al., 1983). Thus, the *VvAlb1* promoter may be useful as a model to provide insights into mechanisms governing seed-specific gene activation in grape.

#### **Effect of ABA on Expression Activation**

Abscisic acid plays an important role in regulating seed-specific promoters (Ezcurra et al., 2000; Li et al., 1999). A region (-46 to -915) of the *VvAlb1* promoter contained several sequence motifs homologous to ABA-responsive elements (ABREs) reported previously (Shen and Ho, 1995). As illustrated in Fig. 4, no change in GUS activity was observed in SE from A2, A3, SAG and A4 lines after ABA treatment. In contrast, the GUS activity of SE from A1 lines increased by more than 50% after ABA treatment. These results indicated that the activity of the *VvAlb1* promoter can be modulated by ABA. However, ABA-responsiveness necessitates the fully functional promoter structure contained in the -2145 fragment.

#### **ACKNOWLEDGEMENTS**

This research was supported by the Florida Agricultural Experiment Station and the Florida Department of Agriculture and Consumer Services' Viticulture Trust Fund.

#### **Literature Cited**

- Burrow, M.D., Chlan, C.A., Sen, P. and Murai, N. 1990. High frequency generation of transgenic tobacco plants after modified leaf disk cocultivation with *Agrobacterium tumefaciens*. *Plant Mol. Biol. Rep.* 8:124-139.
- Chandrasekharan, M.B., Li, G., Bishop, K.J. and Hall, T.C. 2003. S phase progression is required for transcriptional activation of the  $\beta$ -phaseolin promoter. *J. Biol. Chem.* 278:45397-45405.
- Ericson, M.L., Muren, E., Gustavsson, H.O., Josefsson, G. and Rask, L. 1991. Analysis of the promoter region of napin genes from *Brassica napus* demonstrates binding of nuclear protein in vitro to a conserved sequence motif. *Eur. J. Biochem.* 197:741-746.
- Ezcurra, I., Wycliffe, P., Nehlin, L., Ellerstrom, M. and Rask, L. 2000. Transactivation of the *Brassica napus* napin promoter by ABI3 requires interaction of the conserved B2 and B3 domains of ABI3 with different cis-elements: B2 mediates activation through an ABRE, whereas B3 interacts with an RY/G-box. *Plant J.* 24:57-66.
- Grace, M.L., Chandrasekharan, M.B., Hall, T.C. and Crowe, A.J. 2004. Sequence and spacing of TATA box elements are critical for accurate initiation from the  $\beta$ -phaseolin promoter. *J. Biol. Chem.* 279:8102-8110.
- Gray, D.J. 1995. Somatic embryogenesis in grape. p.191-217. In: S.M. Jain, P.K. Gupta and R.J. Newton (eds.), *Somatic Embryogenesis in Woody Plants*, Vol. 1, Kluwer Academic Publisher, Dordrecht, The Netherlands.

- Jefferson, R.A. 1987. Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol. Biol. Rep.* 5:387-405.
- Kay, R., Chan, A. and McPherson, J. 1987. Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Sci.* 236:1299-1302.
- Koo, H., Wu, H. and Crothers, D.M. 1986. DNA bending at adenine-thymine tracts. *Nature* 320:501-506.
- Li, G., Bishop, K.J., Chandrasekharan, M.B. and Hall, T.C. 1999.  $\beta$ -Phaseolin gene activation is a two-step process: PvALF-facilitated chromatin modification followed by abscisic acid-mediated gene activation. *PNAS* 96:7104-7109.
- Li, Z., Jayasankar, S. and Gray, D.J. 2001. Expression of a bifunctional green fluorescent protein (GFP) fusion marker under the control of three constitutive promoters and enhanced derivatives in transgenic grape (*Vitis vinifera*). *Plant Sci.* 160:877-887.
- Li, Z., Jayasankar, S. and Gray, D.J. 2004. Bi-directional duplex promoters with duplicated enhancers significantly increase transgene expression in grape and tobacco. *Transgenic Res.* 13:143-154.
- Li, Z.T. and Gray, D.J. 2005. Isolation by improved thermal asymmetric interlaced PCR and characterization of a seed-specific 2S albumin gene and its promoter from grape (*Vitis vinifera* L.). *Genome* 48:312-320.
- Li, Z.T., Dhekney, S., Dutt, M., Van Aman, M., Tattersall, J., Kelley, K.T. and Gray, D.J. 2005. Optimizing *Agrobacterium*-mediated transformation of grapevine. *In Vitro Cell. Dev. Biol. - Plant* (Accepted).
- Liu, Y.G. and Whittier, R.F. 1995. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* 25:674-681.
- Mitsuhara, I., Ugaki, M., Hirochika, H., Ohshima, M., Murakami, T., Gotoh, Y., Katayose, Y., Nakamura, S., Honkura, R., Nishimiya, S., Ueno, K., Mochizuki, A., Tanimoto, H., Tsugawa, H., Otsuki, Y. and Ohashi, Y. 1996. Efficient promoter cassettes for enhanced expression of foreign genes in dicotyledonous and monocotyledonous plants. *Plant Cell Physiol.* 37:49-59.
- Nagano, Y., Inaba, T., Furuhashi, H. and Sasaki, Y. 2001. Trihelix DNA-binding protein with specificities for two distinct *cis*-elements. *J. Biol. Chem.* 276:22238-22242.
- Perisic, O. and Lam, E. 1992. A tobacco DNA binding protein that interacts with a light-responsive box II element. *Plant Cell* 4:831-838.
- Shen, Q. and Ho, T.D. 1995. Functional dissection of an abscisic acid (ABA)-inducible gene reveals two independent ABA-responsive complexes each containing a G-box and a novel *cis*-acting element. *Plant Cell* 7:295-307.
- Slightom, J.L., Sun, S.M. and Hall, T.C. 1983. Complete nucleotide sequence of a French bean storage protein gene: Phaseolin. *PNAS* 80:1897-1901.
- Sugimoto, K., Takeda, S. and Hirochika, H. 2000. MYB-related transcription factor NtMYB2 induced by wounding and elicitors is a regulator of the tobacco retro-transposon Tto1 and defense-related genes. *Plant Cell* 12:2511-2527.
- Vincenz, M., Leite, A., Neshich, G., Vriend, G., Mattar, C., Barros, L., Weinberg, D., De Almeida, E.R., Pae de Carvalho, M., Aragao, F. and Gander, E.S. 1997. ACGT and vicilin core sequences in a promoter domain required for seed-specific expression of a 2S storage protein gene are recognized by the opaque-2 regulatory protein. *Plant Mol. Biol.* 34:879-889.
- Wu, H. and Crothers, D.M. 1984. The locus of sequence-directed and protein-induced DNA bending. *Nature* 308:509-513.

## Figures

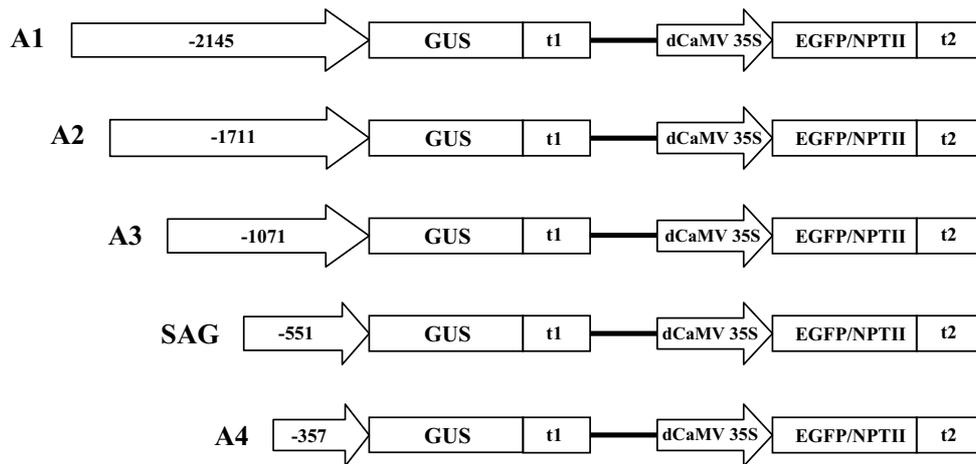


Fig. 1. Schematic illustration of transformation vectors containing the *VvAlb1* promoter fragments. GUS,  $\beta$ -glucuronidase gene; t1, terminator and polyadenylation signal sequences from the *NOS* gene of *Agrobacterium*; EGFP-NPTII, a bifunctional EGFP-NPTII fusion gene (Li et al., 2001); t2, terminator and polyadenylation signal sequences from the CaMV 35S transcript. All vectors were constructed based on pBIN19 binary backbone.

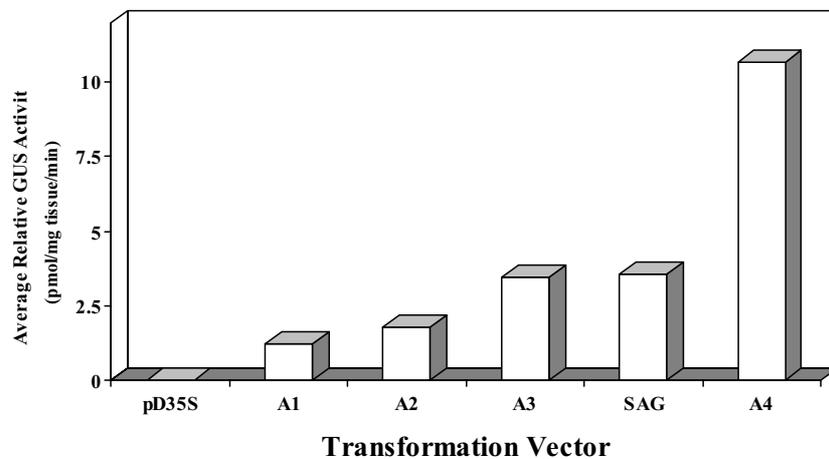


Fig. 2. Transient GUS expression analysis of SE after transformation with *VvAlb1* promoter-containing vectors. Relative GUS activity was derived from 4 independent experiments. Bar values represent the average activity level of each transformation vector.

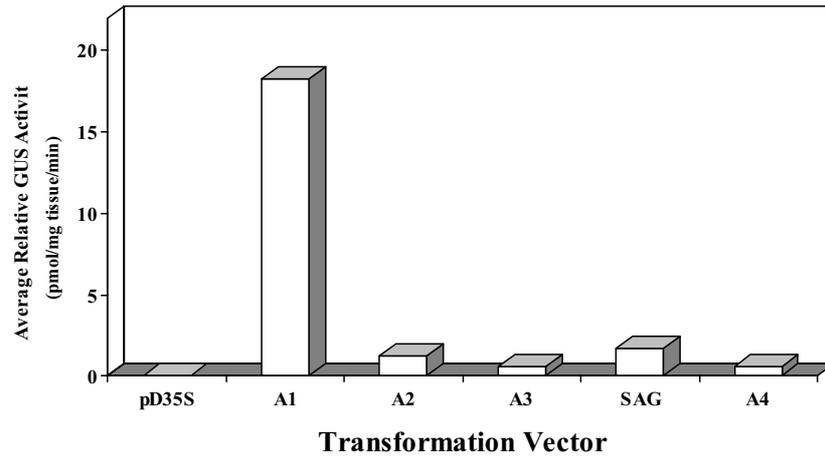


Fig. 3. GUS expression analysis of stable transgenic SE lines recovered after transformation with *VvAlb1* promoter-containing constructs. Relative GUS activity from each line was assayed twice. Bar values represent the average activity level of each individual transgenic line.

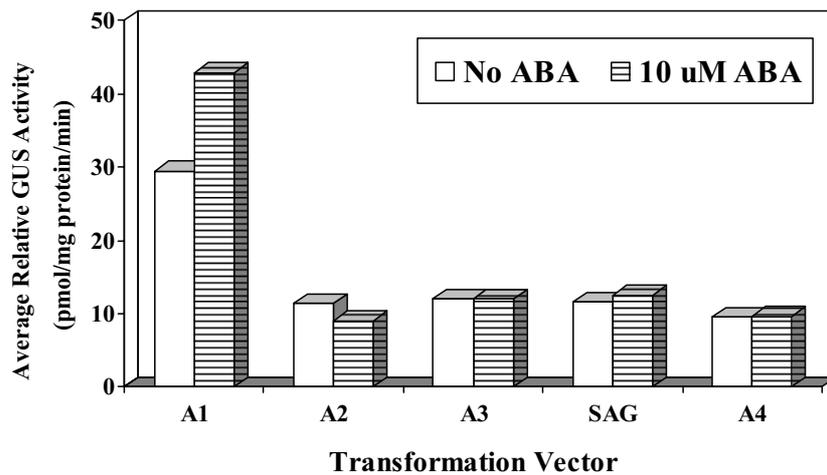


Fig. 4. GUS expression analysis of stable transgenic SE lines with and without ABA treatment. SE at the mid-cotyledonary stage of development were selected from each individual transgenic line and divided into two groups. One group was moved to 10  $\mu$ M ABA-containing medium and the other cultured on an identical medium but lacking ABA for 5 d. SE were then subject to GUS activity assay. Bar values represent the average activity level of each individual transgenic line from 2 repeated assays.

