

A novel promoter from soybean that is active in a complex developmental pattern with and without its proximal 650 base pairs

Martina V. Strömvik¹, Vijaya P. Sundararaman^{1,2} and Lila O. Vodkin^{1,*}

¹Department of Crop Sciences, Edward R. Madigan Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA (*author for correspondence); ²current address: Global Clinical Communications, Clinical Reasearch & Development, Wyeth-Ayerst Research, Radnor, PA 19087, USA

Received 26 May 1999; accepted in revised form 28 July 1999

Key words: fruit-specific gene expression, *Glycine max*, major latex protein homologue, pod development, promoter analysis, transgenic plants

Abstract

We report the isolation of a novel soybean gene, *Msg*, which is highly expressed in developing soybean pods. The gene shows significant homology to a family of fruit- and flower-specific genes, designated the major latex protein (MLP) homologues, so far reported in only a few species and whose functions are unknown. The MLPs are more distantly related to a group of pathogenesis-related proteins (IPR or PR-10) whose functions are likewise unknown. This is the first report of a MLP homologue in a plant for which there is already an IPR-protein reported. We performed an analysis of the *Msg* promoter with 14 different promoter fragments ranging from 0.65 kb to 2.26 kb, fused to the *uidA* (GUS) gene. High transient expression was obtained with all the constructs upon particle bombardment in soybean and green bean pods. Stable *Arabidopsis* transformants were obtained with the *Agrobacterium* vacuum infiltration method. The promoter is fully active in *Arabidopsis* only in plants transformed with the 2.26 kb fragment promoter, expressing GUS in nectaries, nodes, short style and in guard cells of the silique, pedicel and stem but not in mature leaves. Surprisingly, the proximal 650 bp TATA-containing region cannot function on its own in *Arabidopsis* and can be deleted without a change in expression pattern in both *Arabidopsis* and soybean. Thus, tissue-specific regions of the complex *Msg* promoter reside in the distal 5' regions upstream of a dispensable TATA box in contrast to many examples of tissue-specific elements that reside much closer to the TATA box.

Introduction

Combinatorial regulation is the result of *trans*-acting factors interacting with the specific set of *cis* regions (boxes or elements) in a promoter, leading to tissue-specific gene expression (Singh, 1998). The *trans*-acting factors themselves may be expressed tissue-specifically or constitutively. Many previously reported promoters are shown to have elements that control the levels of gene expression (quantitative regions, enhancers) at distal 5' sites and tissue-specific regions proximal to the TATA box (Stougaard *et al.*,

1987; Bustos *et al.*, 1991; Zhao *et al.*, 1994; Hamilton *et al.*, 1998; Ruíz-Rivero and Prat, 1998). The TATA box is the binding site for the TBP (TATA-box-binding protein) which is part of the transcription-initiation complex. In a recent report it was shown that in a mammalian system, naturally occurring transcription-initiation complexes that lack TBP or TBP-like proteins could initiate transcription from both TATA-containing and TATA-lacking promoters (Wieczorek *et al.*, 1998). We report the discovery of a novel soybean gene, the *Msg* gene, whose 5'-flanking promoter has multiple tissue-specific *cis* regions that extend to over 2 kb upstream of the start site of the gene. Even more unusual is that the TATA box and 650 bases of proximal 5' sequence of this promoter can be

The nucleotide sequence data reported will appear in the GenBank, EMBL and DDBJ Nucleotide Sequence Databases under the accession number AJ239127.

eliminated without affecting the basic tissue-specific expression patterns of the reporter gene.

This unusual gene was found in a search for pod-specific promoters that can potentially be used to direct expression of anti-fungal genes in the pod. The pod is responsible for protecting the developing seeds from pathogens, and especially fungal pathogens can attack soybean pods, leading to decreased seed yield and quality. Many seed-borne fungi will also kill the next generation of seedlings. Pod and stem blight (*Diaporthe/Phomopsis*) (Kulik, 1984), pod and seed decay (*Alternaria*) and anthracnose (*Glomerella*) (Schneider *et al.*, 1974) are some examples of fungal pathogens on soybean. By expressing anti-fungal proteins in the soybean pod but not in the seed itself, the fungus may be inhibited and the risk is small for carry-over of foreign substances (i.e. anti-fungal proteins) in the seed product.

There is no guarantee that a non-constitutive promoter from a heterologous species will regulate gene expression as anticipated when introduced into soybean. To date, apart from the seed-specific (Chen *et al.*, 1989; Cho *et al.*, 1995) and nodule-specific promoters (Stougaard *et al.*, 1987; Lauridsen *et al.*, 1993), not very many soybean promoters have been characterized. Thus, there is a need for a larger variety of promoters with defined patterns of tissue-specific expression. Using conventional RNA blotting, we screened 160 random inserts from a pod cDNA library for tissue expression patterns. One cDNA had a high level of expression in the pod and was used to isolate and identify a novel soybean gene, *Msg*, that is related to the major latex proteins (MLPs) of poppy seed and intracellular pathogenesis-related proteins (IPRs) from a number of plants. Deletion analysis of the *Msg* promoter in transgenic *Arabidopsis* revealed the complex pattern of its tissue-specific expression and that the native TATA box is dispensable.

Materials and methods

DNA and RNA methods

Standard laboratory procedures were used for DNA and RNA extractions, blotting, DNA probe labeling, PCR and restriction enzyme digestion of DNA as described (Sambrook *et al.*, 1989; Lindstrom *et al.*, 1990; Cho *et al.*, 1995; Todd and Vodkin, 1996). Sequencing and primer syntheses were performed by the Genetic Engineer-

ing Facility at the Biotechnology Center, University of Illinois. The DNA sequence was analyzed using the functions at the BCM Search Launcher on the Internet (<http://gc.bcm.tmc.edu:8088/search-launcher.html>). The putative amino acid sequence was analyzed by using Proteomics tools at the ExPASy web site (<http://expasy.hcuge.ch/>).

Construction and screening of the cDNA library

A cDNA library was constructed using the Superscript Lambda system kit obtained from BRL (Life Technologies, Gaithersburg, MD). About 1.8 μ g of mRNA from young soybean pods (UC75, an isolate of cv. Clark), isolated by the PolyATtract mRNA isolation kit (Promega), was used to construct the cDNA library according to the manufacturer's instructions. The size-fractionated cDNAs were ligated to λ ZipLox, *NotI*-*Sall* arms, packaged *in vitro* with the λ -packaging system (BRL) and used to infect *Escherichia coli* Y1090(ZL) cells.

Bacteriophage clones containing cDNA inserts were identified by PCR. Primary phage plaques were picked from the library and resuspended in 500 μ l of SM buffer (100 mM NaCl, 8 mM MgSO₄, 50 mM Tris-HCl pH 7.5, 0.01% gelatin) and 20 μ l of chloroform. Aliquots of individual phage stocks (5 μ l volume) were subjected to quick-freeze-thaw in liquid nitrogen for a total of three times. These DNA samples were then amplified by PCR using M13 forward and reverse primers. After amplification, the PCR products were electrophoresed through a 0.7% agarose gel in 1 $\times$ TA (40 mM Tris, pH 7.8) (Sambrook *et al.*, 1989). After staining the gels with ethidium bromide the fragments of interest were cut out and purified by passing through a 0.2 μ m syringe filter (Gelman Sciences #4192, Ann Arbor, MI) (Lu *et al.*, 1994).

The syringe-purified PCR products were used as probes for northern blots. Probes were prepared by random primer reaction and labeled with α -³²P-dATP as described (Feinberg and Vogelstein, 1982). Total RNA from soybean seed coats, immature cotyledons, leaves, stems, roots, young and old pods along with appropriate markers were electrophoresed on gels and blotted onto supported nitrocellulose membranes (Optitrans; Schleicher & Schuell, Keene, NH; cat. 68390). The RNA on the membranes was hybridized to the labeled probe at 68 °C and detected on X-ray film (Hyperfilm MP; Amersham LifeScience, Buckinghamshire, UK; product RPN1678H).

The cDNA library was also screened using the radioactively labeled PCR product of the VS-107 cDNA clone as a probe. Three plates with an average of 40 000 pfu per plate were screened using the Y1090 *E. coli* strain as a host. The total number of positive plaques was counted and 12 positives were isolated by secondary and tertiary screening. The vectors were *in vivo* excised into plasmids pZLVS107:1-pZLVS107:12 by mixing 25 μ l of phage solution with 100 μ l of *E. coli* DH10B(ZIP) cells, incubated at room temperature and plated on LB plates containing 10 mM MgCl₂ and 100 μ g/ml ampicillin. Plasmid DNA was extracted using the QIAprep Spin Miniprep kit (Qiagen, Santa Clarita, CA, cat. 27104).

Screening a genomic library

A soybean genomic library in the Lambda Fix II Vector (Stratagene, La Jolla, CA, cat. 946103) was screened using the PCR product of the pod-specific VS-107 cDNA clone as a probe and the *E. coli* host strain XL1-Blue MRA(P2). The library was made from DNA of cv. Williams 82 partially digested with *Sau*3A I. The host bacteria were grown, infected and plated as described in the library instruction manual. An average of 45 000 pfu were plated on each of 12 Petri plates (size 150 mm $\times$ 15 mm, Fisherbrand cat. 08-757-14) making the total number of phage screened ca. 540 000. Plaque lifts were performed using 0.45 μ m, 137 mm NitroPure, supported nitrocellulose membranes (Micron Separations, Westboro, MA; cat. WP4HY13750) and following the procedure in the library instruction manual. The probe was radio-labeled and hybridized to the DNA on the membranes at 68 °C. Positive clones were detected on X-ray film. After the tertiary screening, the isolated phage plaques were diluted and stored as described for the cDNA clones.

Mapping, subcloning and sequencing of genomic lambda clones

The genomic lambda clones were cultured (500 ml) as described (Sambrook *et al.*, 1989), and DNA was extracted using the Qiagen Lambda Maxi Kit (cat. 12562). The lambda DNA was restriction-digested and maps were combined of results from X-ray-RFLP data and from the FLASH fluorescent kit data (Stratagene, cat. 200381). The plasmid DNA (pGEM3zf+, Promega, Madison, WI) and the 16 kb lambda MSGb DNA were double-digested with *Sa*II and *Eco*RI. Two fragments resulted from the lambda

MSGb, a 3.2 kb fragment containing the coding region (pMSG250) and a 6.3 kb fragment containing the 5'-flanking (promoter) region (pMSG251). The digested DNAs were purified by QIAquick PCR Purification kit (Qiagen, No. 28104) and then ligated at 12 °C overnight. The ligation products were transformed into DH5 α using standard procedures (Sambrook *et al.*, 1989) and selected on LB plates containing ampicillin (50 μ g/ml). Positive clones were identified following plasmid mini-prep (Zhou *et al.*, 1990) and restriction enzyme digest.

Constructs for promoter analysis

Primers were designed to anneal at -2256 (MSG1F), -1684 (MSG4F), -1220 (MSG5F), -651 (MSG8F), -634 (MSG2Ratg and MSG3R), 0 (MSG6Ratg) and +3 (MSG7R) relative to the start codon (Figure 3). A *Hind*III restriction enzyme site was included as an overhang in the 5'-end primers (MSG1F, MSG4F, MSG5F and MSG8F) and a *Bam*HI site was included in the 3'-end primers (MSG2Ratg, MSG3R, MSG6Ratg and MSG7R). Two sets of deletion fragments were made using either MSG2Ratg/MSG3R (A92atg-F219 constructs) primers or MSG6Ratg/MSG7R (G305atg-N353 constructs) as the reverse primer. The forward primers used to obtain 5' deletions were the MSG1F for the A92atg, B103, G305atg, H358 constructs, MSG4F for the C211atg, D162, I372atg, J312 constructs, MSG5F for the E178atg, F219, K340atg, L318 constructs and finally the MSG8F for the M347atg and N353 constructs. The resulting constructs, name and size can be seen in Figure 6. For the A92atg-F219 constructs the pMSG251 subclone was used as a PCR template and for the G305atg-N353 constructs the lambda MSGb DNA was used. The PCR product was purified using the QIAquick PCR Purification kit and the fragments and the recipient plasmid pBI101 (Bevan, 1984; Jefferson *et al.*, 1987) were double-digested with *Bam*HI and *Hind*III. The fragments and linear plasmid were again purified with the QIAquick PCR kit and then ligated at 12 °C overnight. The ligation products were transformed into *E. coli* DH5 α . Plasmid DNA was extracted using Qiagen Plasmid Maxi kit (Qiagen, No. 12162). Restriction enzyme digests with *Bam*HI and *Hind*III, and PCR with the cloning primers, confirmed the transformants. The plasmid DNAs containing the inserts were used for particle bombardment of soybean and green bean pods and for freeze-thaw transforma-

tion of *Agrobacterium tumefaciens* (Holsters *et al.*, 1978).

Agrobacterium infiltration/floral dipping/in planta transformation

The *Agrobacterium* infiltration procedure was adapted from Bechtold *et al.* (1993) and Clough and Bent (1998). *Arabidopsis thaliana* ecotype Columbia was sown in pots covered with window nylon mesh and cultured in 18–24 h light until bolting. The main shoot was cut down to promote side shoots forming. The transformation procedure was performed when unopened flower buds were numerous (about 2–5 days after the main shoot cutting). *Agrobacterium tumefaciens*, strain GV3101 (pMP90) (Koncz and Schell, 1986) transformed with the *Msg* promoter:GUS constructs was grown to a high density ($OD_{600} > 1.2$) in YT medium containing kanamycin (0.05 mg/ml). The cells were harvested by centrifugation at $5500 \times g$ and resuspended in infiltration medium containing $1/2 \times$ MS salts and Gamborg's vitamins, 44 nM benzyl-amino purine, 5% sugar and 0.005% Silwet L-77, pH 6.3. A beaker containing the infiltration media and the suspended bacteria was placed in a dessicator and the potted, bolting plants were placed inverted in it. Vacuum was applied briefly. The pots were returned to the 18–24 h light culture condition and maintained until siliques were dried and opened and the seeds ready to harvest. The transformed seeds were selected on $1/2 \times$ MS plates containing kanamycin (50 μ g/ml). Histochemical GUS assays were performed on flowering plants and as described (Jefferson, 1987) and modified as in Stomp (1992). The plant tissue was immersed in GUS assay buffer containing 5-bromo-4-chloro-3-indoyl- β -D-glucuronic acid (X-gluc) (Jersey Lab Supply, Livingston, NJ) and incubated at 37 °C overnight.

Particle bombardment

Pods and leaves of field-grown soybeans, *Glycine max* (L.) Merr., and pods of green bean, *Phaseolus vulgaris* (L.), from the grocery store were collected and surface-sterilized in a 50% bleach, 0.1% Tween 20 solution. After the seeds and seed coats were removed, the inside of the pods were shot using gold particles (1 μ m diameter with 2.5 μ g of plasmid DNA/shot) with a BioRad Biolistic PDS-1000/He gun (DuPont) (Klein *et al.*, 1987) or a particle inflow gun (Finer *et al.*, 1992). As a positive control for

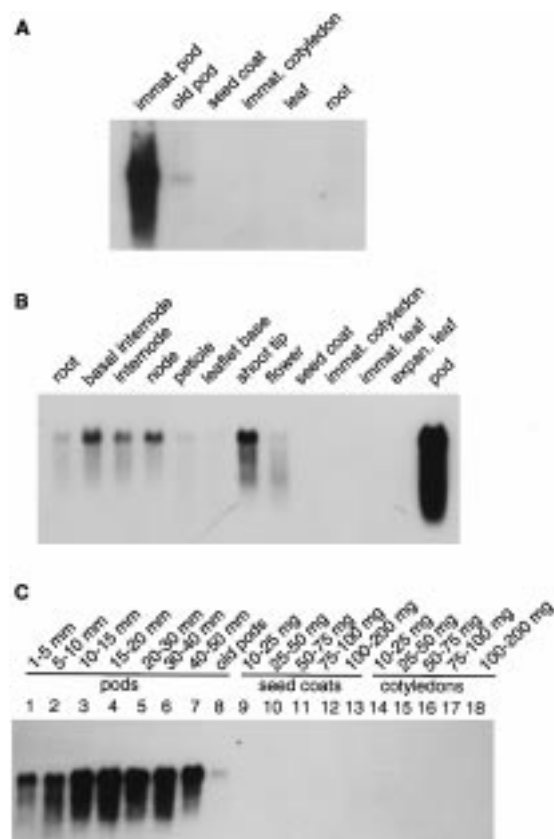


Figure 1. Preferential expression of cDNA clone VS-107 in developing soybean pods. A, B. Total RNA, 10 μ g per lane, from the indicated tissues and stages of development were hybridized with the labeled insert of cDNA clone VS-107 from the pod cDNA library. Panel B is overexposed to illustrate the weak expression that is present in roots, nodes, stems, shoot tips and flowers as compared to the high expression in pods. C. Total RNA from the different stages of development of pods (lanes 1–8, pod length indicated), seed coats (lanes 9–13, from seed of the indicated fresh weight range) and cotyledons (lanes 14–18, from seed of the indicated fresh weight) was hybridized to cDNA clone VS-107.

transient GUS expression, the plasmid pZA300 encoding a CaMV35S-*uidA* construct was used. The pZA300 plasmid was a gift from J.H. Zhou and A.G. Atherly, Iowa State University. Soybean cotyledons, seed coats and leaves were also bombarded with the same constructs. Two days after the bombardment, a histochemical GUS assay was performed as described above.

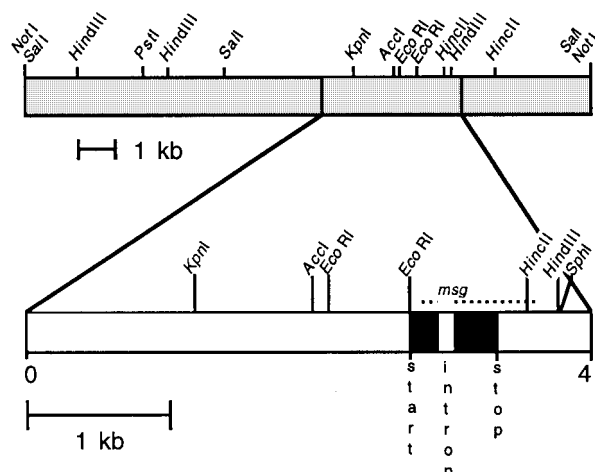


Figure 2. Map of the 16 kb *Msg* genomic fragment representing the pod-specific cDNA clone VS-107. The enlarged segment represents the 4 kb of the MSGb clone that have been sequenced and the dark boxes represents the two exons of the *Msg* gene, interrupted by a small intron. The sequence of the original cDNA clone (VS-107) is marked with a dashed line.

Results

Identification of an abundant mRNA preferentially expressed in soybean pods

We constructed a lambda cDNA library of soybean pod mRNA and isolated 160 random lambda clones. The insert from each clone was labeled and used to examine its expression pattern by hybridization to total RNA extracted from soybean roots, stems, leaves, seed coats, cotyledons and pods. Figure 1A shows a northern blot probed with the insert from the random cDNA clone VS-107. This clone is highly expressed in the developing pods while virtually no expression can be detected in the other tissues assayed. Sometimes a very weak expression is observed in roots, nodes and stem as seen in the overexposed northern blot in Figure 1B. In Figure 1C, the pod specificity is demonstrated by hybridization of VS-107 to RNA extracted from pods (lanes 1–8), cotyledons (lanes 14–18) and seed coats (lanes 9–13) at different stages of development. No expression of VS-107 can be detected at any time or stage of development in the seed coats or cotyledons. In the pods, however, the expression is high at all stages of pod expansion up to full length of 4 to 5 cm. Expression of VS-107 mRNA declines in older pods as the seeds reach full maturity before dehydration.

Isolation and characterization of *Msg* from a genomic library

Using the VS-107 cDNA clone as a probe, a soybean genomic lambda library (Stratagene) was screened. Out of about 500 000 genomic lambda clones screened, two positive clones were isolated and designated MSGa and MSGb. Restriction mapping showed that the 16 kb MSGb insert contained more flanking DNA than the 13 kb MSGa clone and was therefore chosen for further investigation. Figure 2 shows a schematic view of the lambda MSGb insert and the 4 kb that have been sequenced, containing the coding region of *Msg* gene that corresponds to the VS-107 cDNA.

Figure 3 shows the sequence of *Msg* and about 2.8 kb of the 5'-flanking region. The coding region of the gene is 459 bp long and it is interrupted by an intron of 105 bp at position +190 to +295 in the figure. About –2 kb (–1871 to –2021 bp) from the start of translation, there are five conserved direct repeats of 30 bp each in a row with a consensus sequence of 5'-TAAGACCCTTTAAGGGCGGTGAGTTAATTA-3'. These repeats have the potential to form hairpin structures and/or act as binding sites interacting with the regulation of the gene. The putative TATA box (TATATATA) is positioned from –89 to –96 from the translational start and the poly(A) signal (AATAAA) from +593 to +598, 25 bp downstream of the stop codon (TAA). When aligning the cDNA sequence to the genomic sequence, the 300 bp long 3'-untranslated region is evident. The original cDNA (VS-107) was not full-length, it begins at position +85, and the intron was spliced following the GT-AG rule. We have since isolated full-length cDNAs from the original cDNA library with the VS-107 as a probe. The *Msg* gene, when translated, would yield a 17.7 kDa protein (153 aa) with a predicted pI of 6.25.

Msg has homology to the major latex proteins, a family of fruit- and flower-specific genes

Database analysis of the coding sequence indicates that the gene belongs to a family of fruit- and flower-specific genes, the major latex protein homologues. These have so far been reported in opium poppy (*Papaver somniferum*; Nessler and Von der Haar, 1990), tobacco (*Nicotiana tabacum*; Meeks-Wagner *et al.*, 1989; Neale *et al.*, 1990), bell pepper (*Capsicum annum*; Pozueta-Romero *et al.*, 1995), musk melon (*Cucumis melo*; Aggelis *et al.*, 1997), and wild strawberry (*Fragaria vesca*; Nam *et al.*, 1999). There

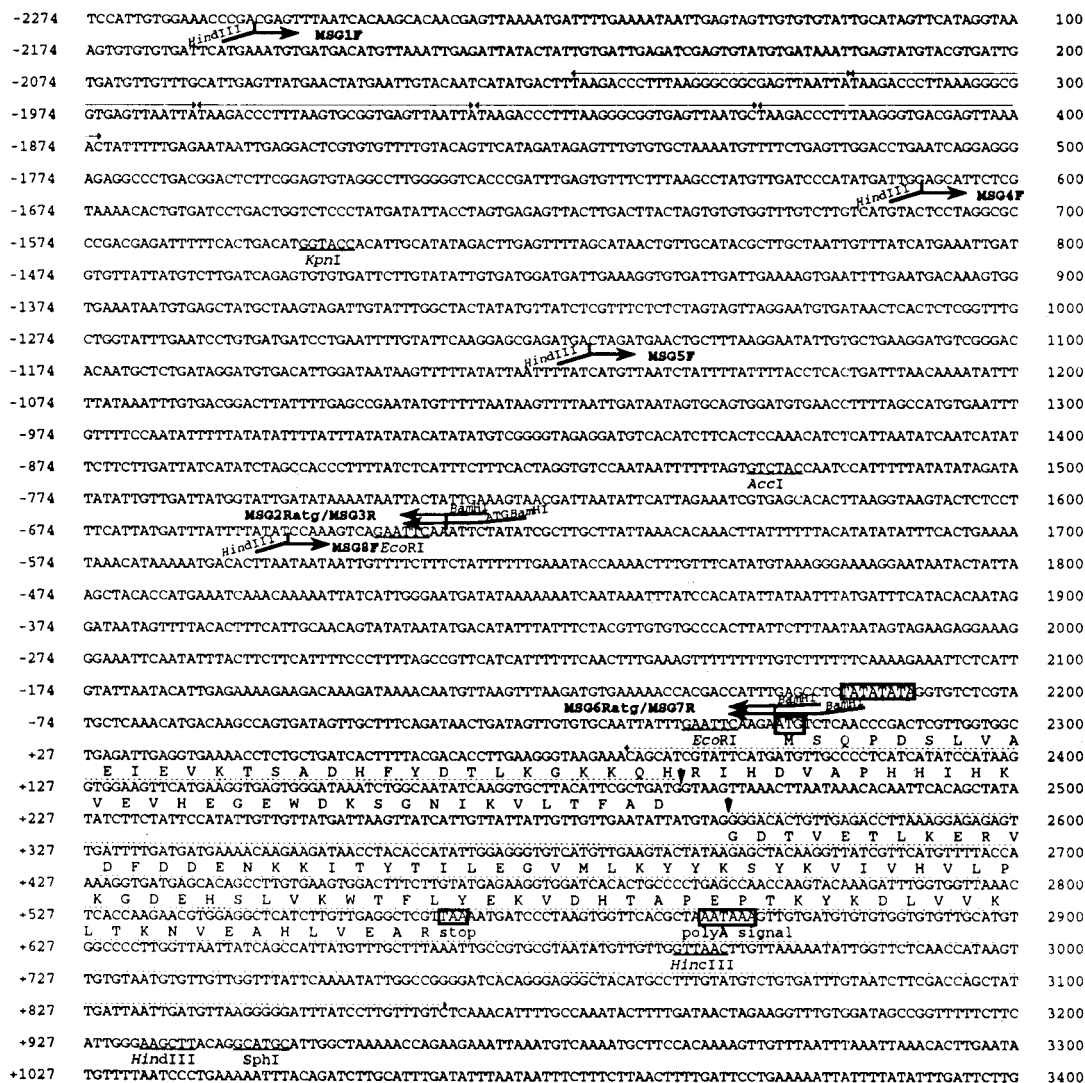


Figure 3. The nucleotide sequence of *Msg* and flanking regions. The primers used to obtain the PCR fragments for the promoter analysis are indicated by an arrow, with the overhang indicated by the name of the restriction site. Restriction sites are underlined and the splicing sites of the intron are marked by vertical arrowheads. The putative TATA box, start, stop and polyadenylation signals are boxed. The original cDNA sequence (VS-107) is indicated by a dashed overline, starting and ending with arrows. The protein sequence is given below the coding nucleotide sequence. Between -1871 and -2026, five highly conserved repeats are located (marked with thin lines and arrows).

are also sequences in the databases for *Arabidopsis thaliana*, rape seed (*Brassica napus*), raspberry (*Rubus idaeus*) and common ice plant (*Mesembryanthemum crystallinum*). The function of the *MSG* protein and its family members are unknown but they are all expressed in fruits and/or related organs such as flower parts and in general they are of low copy number. They are 16.8–18.1 kDa in size and pI's range between 5.5 and 6.7. These sequences also show significant homology to pathogenesis-related plant proteins (PR proteins) such as ribonuclease of ginseng (Moi-

seyev *et al.*, 1997) and the parsley PR1 protein (Rush-ton *et al.*, 1996) and to the related allergens such as *Api g I* of celery (Breiteneder *et al.*, 1995) and especially to the major tree pollen allergen proteins such as *Cor a I* of hazel (Breiteneder *et al.*, 1993) and *Car b I* of hornbeam (Larsen *et al.*, 1992).

Figure 4 presents an alignment of the gene products of the published major latex homologues. In the figure, only those amino acid residues that are identical matches to the soybean *Msg* are boxed. Some regions are highly conserved, for example a glycine

				1	15	16	30	31	45
1. <i>Capsicum annuum</i> (Sn-1)	-----	MGVKGKL	IASVEVKCEGNLIHE	LFHIHA-	HHVPNISE				
2. <i>Nicotiana tabacum</i> (FB7-4)	-----	MGLKGKL	IAQIEMKCAGDLLHE	HFKNP-	HQTSTMS				
3. <i>Glycine max</i> (MSG)	-----	MSQPDSL	VAEIEVKTSAHDFD	TLFGKKQRIHDVAP					
4. <i>Cucumis melo</i> (MEL7)	-----	MSLIGKL	VSELEINAAAEKFYE	IFKDQC-	FQVFNITP				
5. <i>Papaver somniferum</i> (MLP 22)	MAEHHHTISGLVGKL	VTELEVNCNADEVYK	IFKHH--	EDLPNAIP					

	46	60	61	75	76	90	91	105	106	120
1. NFINHFEIHEGETVK	VG--SVVSWSYNEAG	QKRYMKQ	LIEDIDPD	MKLIRWKA	IEGDVLE	SYNSFT	IIVT----	S		
2. NKITNFTLHEGQLGS	TG--SVVSWKFLVGG	KERHAKQ	VLH-IDDA	KKSITFN	FVEGDMNE	LYKSMT	TATL----	T		
3. HIIHIVVEVHEGEWDK	SGN--IKVLTFADGD	TVETLKERV	DFDDE	NKKITYTILEGVMLK	YKSYKVIVHVLPGK					
4. RCIQQVEIHGTNWDG	HCHGSIKSWYITIDG	KAEVFKERV	EFHDD	KLLIVLDG	VGGDVFK	NYKSF	KPAYQFVPKD			
5. HIYRGVKAVEGDRI	SGF--IKEWYIIEG	KPLTCKERTT	YEDE	ARTIHHST	VEGVLLD	DYK	KFDATLVNPKAD			

	121	135	136	150	151	165
1. SEHEWTTWTIEYEKK	TEGTPEPLVLLGLVL	DMTKDIEAHL	LKK--			147
2. AEGNWMTWTFVYEKL	NENIPEPLDIFELAI	CLLKDEPHHV	GK--			146
3. DEHSLVKWTFVYEKV	DHTAPEPTKYKDLV	KLTKNVEAHL	VEAR			153
4. RNHCQAILLSIEYEKL	HGSPDPHKYIDLMI	GIINDIGSH	K---			151
5. GHGSIVTWIVEYEKI	NEDSPVRISYLTFFH	KIIEDLNTYL	CASD			158

Figure 4. Sequence alignment of the published proteins belonging to the major latex protein homologue family. The MAP, Multiple Sequence Alignment function at BCM Search Launcher was used for the alignment. Only those amino acids that perfectly match with MSG have been boxed. The MSG (*Glycine max*) is marked in bold. *Capsicum annuum* (Sn-1), accession number X79230; *Nicotiana tabacum*, (FB7-4), S44872; *Glycine max* (MSG); *Cucumis melo* (MEL7), Z70522; *Papaver somniferum* (MLP 22), 320544.

at position 62 in the figure, a lysine at position 81, an isoleucine at 94, a glutamic acid and a glycine at 100–101, a tyrosine, a lysine and a serine at 107–109, a tyrosine, a glutamic acid and a lysine at 132–134, proline-glutamic acid-proline at 140–142 and, finally, a histidine at 159. The melon MEL7 is 31.6% identical to the *Msg* protein, the poppy MLP22 is 29.6% identical, the tobacco FB7-4 protein 29.4% identical, and the bell pepper Sn-1 protein is 28.6% identical to the *Msg* protein. In general, these proteins have about 30% overall identity to each other and as a group they show about 25% identity to the IPRs (Osmark *et al.*, 1998).

The 5'-flanking region of the *Msg* gene is very A/T-rich (68% overall in the region investigated) and has a region at -711 to -767 that shows high homology (70% identity) with a promoter region from a soybean heat shock protein (Schöffl *et al.*, 1984).

Only a few sequences in the soybean genome hybridize to *Msg*

Figure 5 shows the restriction fragment patterns of genomic DNA when hybridized to the VS-107 cDNA clone. There are possibly one or two additional copies of genes related to *Msg* based on both the *EcoRI* and *HindIII* restriction patterns. Two *EcoRI* fragments of 10.0 kb and 6.5 kb are found although there are no *EcoRI* sites internal to the probe (the VS-107 cDNA

clone). Three bands are detected with *HindIII* digestion, the longest of which is about 8.5 kb. This fragment is consistent with a *HindIII*-*HindIII* fragment of the MSGb map (Figure 2). The other two fragments in the *HindIII* genomic blot are about 4 and 2.3 kb or 4 and 3.1 kb for cv. Williams (lane 3). Each of these two fragments could represent additional copies of genes related to *Msg*. Alternatively, they might represent one copy containing a *HindIII* polymorphism that interrupts the coding region of that gene. The two smaller bands are also fainter than the 8.5 kb band, which could be due to sequence divergence with a second or third copy of the *Msg* sequence. Figure 5 also shows that there is a polymorphism detected in one of the *HindIII* fragments in cv. Williams.

Design of *Msg* promoter:GUS fusions

Fourteen PCR-generated deletions of the *Msg* promoter were fused to the *uidA* gene in the pBI101 binary vector (Bevan, 1984; Jefferson *et al.*, 1987). Figure 3 shows the positions of the primers and Figure 6 shows the names and sizes of the constructs. Series G305atg-N353 represents a conventional deletion series in which successively longer regions of the 5'-flanking promoter are eliminated as shown in Figure 6A. Series A92atg-F219 represents the same three distal 5'-deletion fragments as G305atg-L318 except that the 650 bp proximal to the start codon of the *Msg*

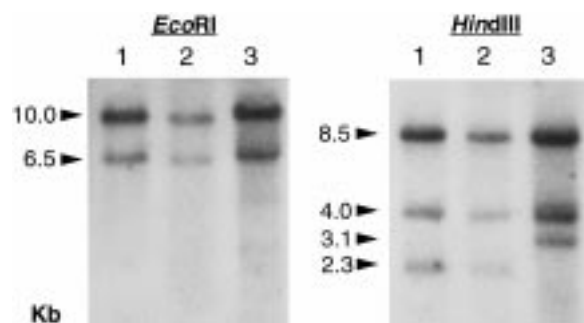


Figure 5. Genomic DNA blots indicate that there are at least two *Msg* genes in soybean. Genomic DNA (10 μ g) was digested with *Eco*RI or *Hind*III, electrophoresed, and probed with the cDNA clone VS-107. Lane 1 is cv. Richland, lane 2 is T157, and lane 3 is cv. Williams. Clark has the same polymorphism as Williams (not shown).

gene have also been deleted as shown in Figure 6B. In essence, A92atg–F219 represent the deletion of the 650 bp *Eco*RI fragment of the promoter (Figure 3) that contains the TATA box and all of the 5'-untranslated region of the *Msg* gene. Initially, this deletion was a mistake in that the close proximity of the *Eco*RI site to the start site of the gene leads to accidental loss of that fragment in subcloning of the promoter region.

The PCR primers were designed to anneal at 2.26, 1.68, 1.22 and 0.65 kb from the translational start codon of the gene. The reverse primer used to amplify the *Msg* promoter was designed to either include or exclude a start codon in the overhang. The A92atg, C211atg and E178atg constructs contain an atg provided by the primer. The practical difference between the 'atg constructs' and the others is that in the atg constructs 10 extra amino acid residues (9 without the methionine) from the polylinker cloning region will be fused to the GUS protein. In the B103, D162, and F219 constructs, on the other hand, 30 extra base pairs will be part of the untranslated region. The inserts of all 14 constructs were confirmed by PCR and by Southern blot after *Bam*HI and *Hind*III restriction digests of plasmid preparations.

Stable transformation of *Arabidopsis* reveals complexity of the tissue-specific promoter

The 14 *Msg* promoter constructs were transformed into *A. tumefaciens* strain GV3101(pMP90) (Koncz and Schell, 1986) which were then used for vacuum infiltration or floral dipping-mediated transformation (Bechtold *et al.*, 1993; Clough and Bent, 1998) of *Arabidopsis thaliana* ecotype Columbia. Bolting plants were subjected to vacuum infiltration of *Agrobac-*

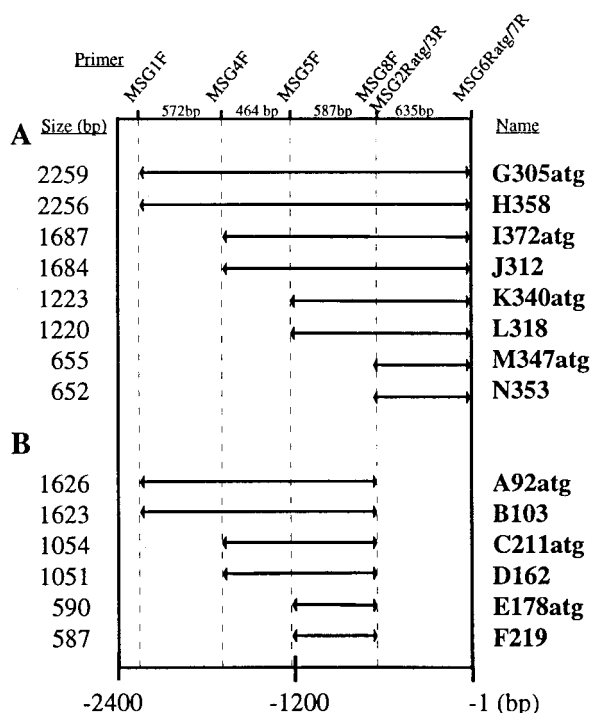


Figure 6. Deletion constructs of the *Msg* promoter. A. The promoter of *Msg* was successively 5'-deleted as indicated in the G305atg–N353 constructs and the resulting regions were fused to the *uidA* (GUS) reporter gene. B. Each of the A92atg–F219 constructs also lack 635 bp adjacent to the start codon of the *Msg* gene that contains the TATA box. (Note: because the primers used to amplify the M347atg and N353 constructs contains 17 bases of *Msg* sequence, the total lengths of the M347atg and N353 constructs are 655 and 652 bp, respectively, whereas the exact length of the deletion in the A92atg–F219 constructs is 635 bp; these exact sizes are shown in the figure; however, for simplicity throughout the paper, we have referred to the proximal region as 650 bp.) The PCR primers indicated in the figure were custom-made with restriction enzyme overhangs (*Hind*III and *Bam*HI) to facilitate cloning of the fourteen different promoter PCR fragments into the promoterless binary vector pBI101. For each construct there are two versions; one contains the start codon (atg) incorporated into the primer that amplifies the *Msg* promoter fragment and the other contains the start codon from the *uidA* gene. The 'atg' in the vector name designates the construct in which the atg derives from the primer used to amplify the *Msg* promoter. The 'atg' constructs contain 9 extra amino acids from the polylinker region fused to GUS whereas the other constructs contain these 27 bases as part of the untranslated region.

terium and the seeds were harvested and selected on plates containing kanamycin. T₁ transformants were transferred from the plates to soil and a histochemical GUS assay was performed on flowers, leaves, and siliques at different stages of development. Southern blots confirm that all constructs have been successfully transformed into *Arabidopsis* (data not shown). Figure 7A–F show GUS expression in *Arabidopsis* trans-

formants. The longest construct (G305atg, 2.26 kb) drives GUS expression in the short style, the nectaries and in the guard cells of the silique and the pedicel but not in the leaves. In general, fewer expressing cell types and less intense expression was correlated with successive deletion of the 5' region as illustrated in Figure 7F. In addition, developing siliques generally showed higher expression than mature siliques. The short style expressed GUS at early silique development but not in mature siliques. The guard cells and nectaries also showed the highest expression during flowering and in the early stage of silique development. The T₂ and T₃ generations show similar patterns of gene expression (data not shown).

There was no major difference found between the expression pattern and the position of the atg in the different constructs. Therefore, these have been grouped together in Figure 8, which presents a graphical representation of the expression patterns found in the 14 constructs. The largest number of cell types that express GUS in *Arabidopsis* were obtained with the largest promoter fragments (2.26 kb, constructs G305atg and H358). In these cases, GUS is expressed in nodes, sepals, short style, nectaries, and in guard cells of the silique and the pedicel. Sometimes guard cells of the cauline leaves expressed GUS though no expression was ever observed in guard cells of mature leaves. Within the same raceme, GUS was sometimes present in some pedicels but not in others. Interestingly, if the 650 bp proximal to the start of translation is removed (constructs A92atg and B103), the expression in short style and sepals is generally weakened but is otherwise similar to that of the G305atg and H358 transformants.

When the 570 bp most distal to start site are removed (C211atg, D162, I372atg and J312 constructs), expression is lost in the guard cells, sepals, short style and nectary lobes and only nectary base cells and nodes show some expression. If another 460 bp are deleted, only nodes (E178atg and K340atg constructs) and occasionally nectary base cells (L318 but not F219 construct) show expression.

All of the A92atg–F219 transformants, lacking the proximal 650 bp with the TATA box, show similar patterns as the corresponding 5'-deletion transformants (G305atg–L318), although with somewhat weaker levels of expression. Thus, the 650 bp proximal deletion that contains the TATA box does not seem to alter the tissue-specific expression pattern. The constructs that contain only the 650 bp proximal to the start site of the *Msg* gene are not expressed at all in

Arabidopsis, although they are expressed transiently in soybean pods as are all of the constructs (see below). It is clear that the most distal 5' sequences of the *Msg* promoter are required for adding tissue-specific expression in *Arabidopsis* with or without the proximal 650 bp containing the TATA box.

High transient expression of GUS is observed in pods of soybeans and green beans after particle bombardment

Particle bombardment (Klein *et al.*, 1987; Finer *et al.*, 1992) of developing soybean and green bean pods were performed to test for transient expression of the constructs. All fourteen *Msg* promoter:GUS constructs, a CaMV35S:GUS-positive control construct (pZA300) and the pBI101 promoterless negative control were precipitated on gold particles. Soybean and green bean pods were surface-sterilized and the pods were shot on the inside surface after the seeds had been removed. Two days after bombardment, a histochemical GUS assay was performed. Figure 7 (panels G, H, and I) illustrate the typical results. In soybean pods, visual examination indicates that the G305atg–N353 constructs express GUS stronger than the control CaMV35S:GUS construct. The A92atg–F219 constructs are almost as strong as the CaMV35S control for both soybean and green bean pods. No expression was seen when bombardment was done with the promoterless plasmid vector (pBI101).

Typically after bombardment of developing soybean pods, the sclerenchyma cells and inner epidermis cells in the seed cavity were transformed leading to the conclusion that the promoter is transiently active in both of these cell types. Similarly, in green bean pods the inner epidermis and the immediately underlying tissue in the seed cavity express GUS. The CaMV35S:GUS positive control construct also gave the same cell type pattern. Bombardment of soybean cotyledons, seed coats and leaves showed a maximum of 1 or 2 dots per bombardment, which could be due to contamination. Though transient expression of the *Msg* promoter is strong in the pods, it is not detected in the leaves and this result is consistent with the general pattern seen in RNA blots (Figure 1).

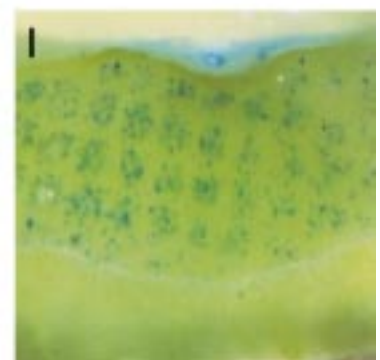
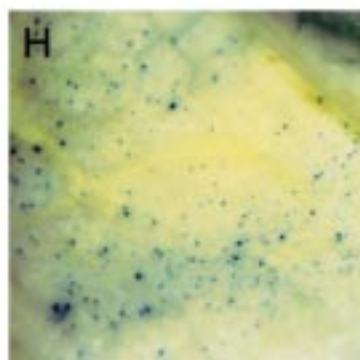
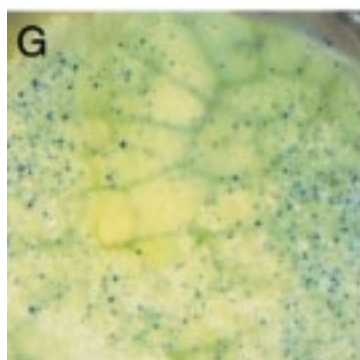
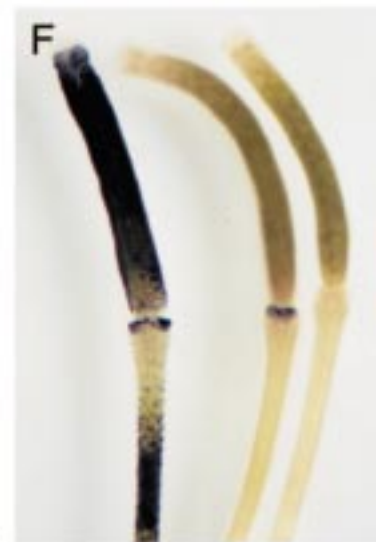
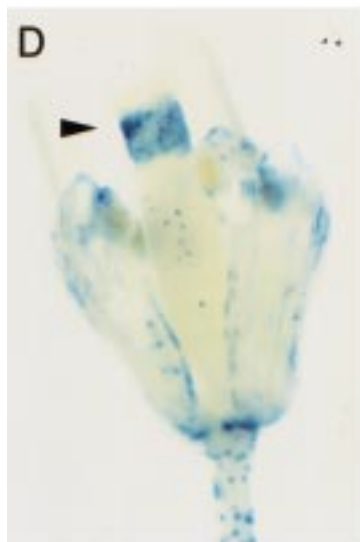
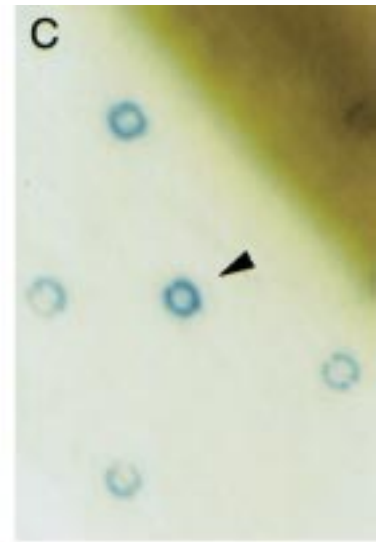
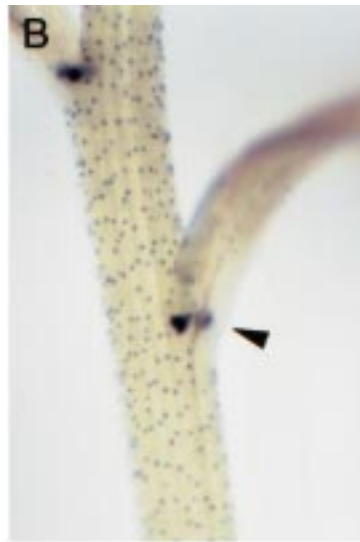
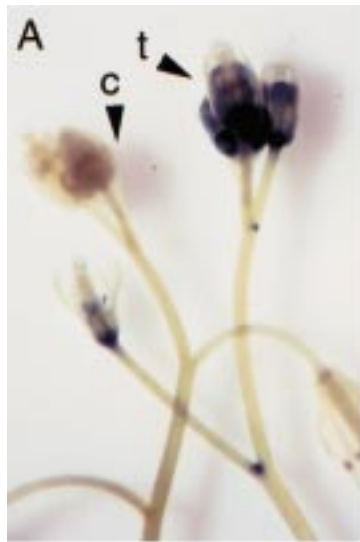


Figure 7. Transient and stable expression patterns of *Msg* promoter:GUS fusions. A–F. Stable expression of GUS in *Arabidopsis* under control of *Msg* promoter deletions. A. A transformant (t) and a control non-transformant plant (c) ($\times 8$). B. A group of strongly expressing cells are situated in the node by the base of cauline leaves and pedicels. The guard cells of the upper stem and pedicels are also GUS-positive ($\times 30$). C. Guard cells of a silique expressing GUS ($\times 190$). D. GUS expression in nectaries at the base of the flower, petals, sepals, and short style (arrow) of the pistil/premature silique ($\times 20$). E. GUS expression in nectary lobes and adjacent cells of the nectary at the base of the flower or silique are expressing GUS ($\times 30$). F. Siliques of T_1 plants transformed with (from left to right): G305atg (2.26 kb), J312 (1.68 kb), K340atg (1.22 kb). G–I. Transient GUS expression in pods of soybean (*Glycine max*), and green bean (*Phaseolus vulgaris*), after particle bombardment ($\times 5$). G. Soybean pod bombarded with pBMGL318 (1.22 kb). H. Soybean pod bombarded with control 35S-GUS construct. I. Green bean pod bombarded with pBMGB103 (1.62 kb, and lacking the proximal 635 bp).

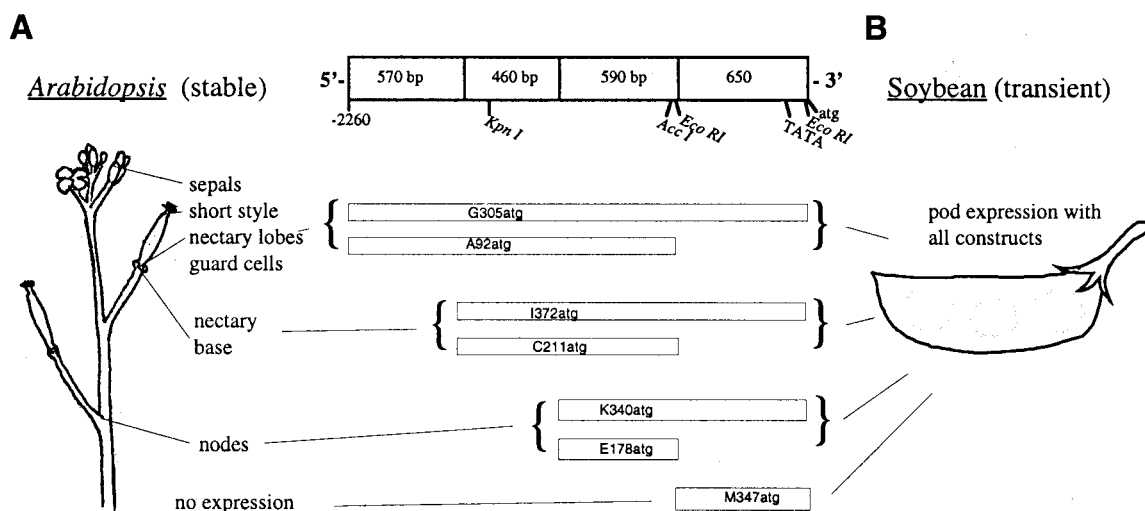


Figure 8. 5'-Deletions of the *Msg* promoter lead to loss of expression in various *Arabidopsis* cell types. A. Tissue-specific expression in *Arabidopsis* plants transformed with the indicated deletions. B. All constructs tested for transient expression in soybean activate GUS in pods but not in leaves.

Discussion

Soybean contains both MLP (major latex protein) and IPR (intracellular pathogenesis-related protein) gene families

We describe the isolation and characterization of a soybean gene, *Msg*, that shows a high level of expression in soybean pod tissue. We have shown that the *Msg* gene has significant homology with the major latex protein (MLP) homologues, a family of flower- and fruit-specific genes for which the function of the gene products are unknown. The first major latex protein reported was from opium poppy, *Papaver somniferum*, where it is present at high levels in the latex (Griffing and Nessler, 1989; Nessler and Von der Haar, 1990; Nessler *et al.*, 1990; Nessler and Burnett, 1992; Nessler, 1994), a substance believed to have a protective function to the plant. Homologues are also reported in tobacco (Meeks-Wagner *et al.*, 1989; Neale *et al.*, 1990), bell pepper (Pozueta-Romero *et al.*, 1995), melon (Aggelis *et al.*, 1997), and

wild strawberry (Nam *et al.*, 1999). The name may be misleading since latex is not a feature of every plant for which the genes have been found.

These homologues are similar in size and location of the intron and no obvious signal peptides have been found in any of them. In opium poppy, however, the proteins have been localized to membrane-bound vesicles in the laticifers (Griffing and Nessler, 1989) and the *Sn-1* of bell pepper has been shown to be associated with small vacuoles close to the plasmalemma (Pozueta-Romero *et al.*, 1995). The tobacco *FB7-4* was initially reported to be expressed during floral initiation (Meeks-Wagner *et al.*, 1989) and was later shown to be highly expressed in the basal internodes of the stem (Neale *et al.*, 1990). The melon *MEL7* gene and the strawberry 3.1.R4 gene are highly expressed in ripening fruits but also to some extent in root and stem (Aggelis *et al.*, 1997; Nam *et al.*, 1999). In poppy, the major latex proteins are reported to be highly expressed in flower buds and weakly in roots in addition to the laticifers (Nessler and Burnett, 1992) which are present throughout the poppy plant.

The major latex protein homologues also show significant homology to major pollen allergen proteins and other pathogenesis-related (PR) proteins, i.e. *Cor a 1* of hazel (Breiteneder *et al.*, 1993), *Car b1* of hornbeam (Larsen *et al.*, 1992), ribonuclease of ginseng (Moiseyev *et al.*, 1997), *Api g 1* of celery (Breiteneder *et al.*, 1995), and parsley *PR1* (Rushton *et al.*, 1996). Especially the intracellular pathogenesis-related proteins, the IPR-PRs, seem to have an interesting relation to the major latex protein homologues as recently reviewed by (Osmark *et al.*, 1998). The homology between the two gene families is not more than about 25% but they have similar structures, sizes and pI's and their functions are unknown (Osmark *et al.*, 1998). There is an IPR protein reported for soybean, SAM22 (Crowell *et al.*, 1992), which makes our report the first report of a plant that has representative genes from both the MLP family and the IPR family. SAM22 is highly expressed in roots but is also expressed in leaves after wounding (Crowell *et al.*, 1992). The soybean SAM22 and MSG proteins show 25.3% sequence identity.

Msg is highly expressed in pods and maintains tissue-specific expression in transient assays in soybean

We have shown by transient expression assays that the promoter from the *Msg* gene can function as a strong activator of GUS in soybean and green bean pods after particle bombardment (Figure 7). Control bombardment of leaves and cotyledons did not result in GUS expression. The transient experiments agree well with the expression of *Msg* transcripts as determined by RNA blots (Figure 1) in which there is strong expression in the pods but not in the expanded leaves or in the developing cotyledons. Some expression is detected in nodes but it is not clear if this is due to *Msg* or the second related gene that is detected by DNA blotting.

The Msg promoter directs a complex pattern of regulation in transgenic Arabidopsis including guard cell expression

In contrast to the transient soybean transformation, maximal tissue-specific expression in stable *Arabidopsis* transformants required 2.26 kb of the *Msg* promoter. The expression was localized to nodes between pedicel and stem, or cauline leaf and stem, to nectaries, short style and to guard cells of the stem, pedicel and silique. GUS expression was never seen

in rosette leaves or roots. Guard cell-specific expression is reported for genes involved in different tasks. Guard cells of transgenic tobacco are reported to express GUS driven by a 2753 bp fragment of the tomato *tas14* promoter (del Mar Parra *et al.*, 1996). The *tas 14* gene is an inducible dehydrin involved in water stress but it is also developmentally expressed in flowers and pollen. The promoter of an *Arabidopsis* acidic chitinase is also proved to activate GUS in guard cells of transgenic *Arabidopsis* (Samac and Shah, 1991). Stomates are natural openings through which pathogens can intrude and chitinase is involved in the plants' primary defense against pathogens. Expression due to the chitinase promoter was also seen in anthers (Samac and Shah, 1991). Expression of GUS in guard cells of different floral organs and the receptacle was observed when the GUS gene was driven by the promoter region of the *rha1* gene from *Arabidopsis* (Terry *et al.*, 1993). The *rha1* gene encodes a small GTP-binding protein involved in secretion and vesicle-mediated transport.

The relationship of the major latex homologues to the PR proteins and allergens and the activity of the *Msg* promoter in guard cells, fruit and floral organs may be a clue to a putative defense or stress function of the *Msg* gene product.

The Msg promoter appears to have multiple tissue-specific elements in the distal region of the promoter

Tissue-specific expression is assumed to be the result of combinatorial regulation due to the specific set of regions present in the promoter (Singh, 1998). Many previously reported promoters are shown to have distal quantitative regions and proximal tissue-specific regions (Stougaard *et al.*, 1987; Bustos *et al.*, 1989; Zhao *et al.*, 1994; Hamilton *et al.*, 1998; Ruíz-Rivero and Prat, 1998). However, this may not be a strict rule. In the case of two fruit ripening-specific genes (*E4* and *E8*) from tomato, sites flanking the genes were shown to contain protein-binding sites. The genes are simultaneously expressed but the positions of the binding sites in the 5' regions of the two genes differed by 900 bp (Cordes *et al.*, 1989; Deikman *et al.*, 1998). Our results also show that the tissue specificity is not necessarily confined to the most proximal region of a promoter but might be much farther out. The *Msg* promoter appears to gain expression in specific tissues with the addition of increasing amounts of 5' sequence (Figures 7 and 8).

The proximal 650 bp of the Msg promoter that contains the TATA box is not required for expression

Surprisingly, our stable transformation experiments show that the proximal 650 bp alone does not induce gene expression at all in *Arabidopsis* as illustrated in Figure 8, although it can transiently express GUS in soybean pods by particle bombardment. In fact, including this region in the promoter does not alter the pattern of organ/tissue specificity either. The transformants of the constructs including the 650 bp show the same pattern as do the transformants of the counterpart constructs without the 650 bp. From both the transient expression experiment in soybean and the stable transformation of *Arabidopsis*, it appears that including the 650 bp enhances the expression of GUS. These results suggest three things: (1) the A/T-rich proximal 650 bp might contain transcription enhancer elements; (2) the 590 bp region just upstream of the proximal 650 bp contains regions that can take over the basic promotion functions (i.e. TATA box, etc.); and (3) the conserved repeats found in the most distal 570 bp (2 kb upstream of the start codon) might be responsible for the specificity of expression in the style, nectaries and guard cells in *Arabidopsis*.

Small repeated motifs have been reported from other regulating domains in tissue-specific promoters such as the β -phaseolin storage protein of bean (Bustos *et al.*, 1991). For the same β -phaseolin promoter, A/T-rich domains have been reported to enhance transcription (Bustos *et al.*, 1989). A/T-rich regions of the tissue-specific pea plastocyanin (*PetE*) gene promoter have recently been shown to enhance transcription, whether fused upstream or downstream of a cauliflower 35S minimal promoter (Sandhu *et al.*, 1998). There are also reports of A/T-rich enhancing promoter regions in soybean heat shock proteins (Czarnecka *et al.*, 1992). Interestingly, regions about –500 upstream of transcription start in a soybean heat shock promoter (Schöffl *et al.*, 1984) show 70% homology to a region –700 to –800 upstream of the translational start of the *Msg* promoter.

Regions upstream of the proximal 650 bp can take over the function of the true TATA box as we have shown with the 650 bp proximal deletion constructs. These upstream regions are relatively AT-rich and there is a possibility that alternative TATA-box-like elements reside there. However, there is also the possibility that there are other protein binding elements that can help initiate transcription as in TATA-less promot-

ers. Such exciting systems have been reported from other organisms (Wieczorek *et al.*, 1998).

The *Msg* promoter behaves differently in stably transformed *Arabidopsis* and transiently in soybean. This indicates the importance of well-defined regulating sequences useful for other species than model plants. Tissue-specific regulation in different plants might occur in different tissues due to differences in set of organs and function of the organ. Soybean does not have nectaries and the function of the gene product is apparently carried out in the pod though the cell type in the pod has yet to be determined. However, based on the expression pattern and the high level of expression, the *Msg* promoter will be a useful promoter to drive transgene expression in pods of not only soybean but also other legumes and plants with a similar fruit anatomy.

In summary, we have identified a soybean equivalent of the major latex protein homologues and hereby reported the first plant to have both a major latex protein (MLP) homologue and an intracellular pathogenesis-related protein (IPR-P). The function of the gene product is unknown but it is highly expressed in soybean pods. We have shown that the promoter contains different regions important for maximal expression of GUS in a heterologous plant but that, surprisingly, the 650 bp adjacent to the start codon does not affect the pattern of expression.

Acknowledgements

We are thankful to Dr Steve Clough and Dr Andrew Bent for providing the vacuum infiltration device and to Ms Deborah Petrik for helping in screening the cDNA library. This project was supported by the Illinois Soybean Program Operating Board.

References

- Aggelis, A., John, I., Karvouni, Z. and Grierson, D. 1997. Characterization of two cDNA clones for mRNAs expressed during ripening of melon (*Cucumis melo* L.) fruits. *Plant Mol. Biol.* 33: 313–322.
- Bechtold, N., Ellis, J. and Pelletier, G. 1993. *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C.R. Acad. Sci. Paris, Life Sci.* 316: 1194–1199.
- Bevan, M.W. 1984. Binary *Agrobacterium* vectors for plant transformation. *Nucl. Acids Res.* 12: 8711–8721.
- Breiteneder, H., Ferreira, F., Hoffman-Sommergruber, K., Ebner, C., Breitenbach, M., Rumpold, H., Kraft, D. and Scheiner, O. 1993. Four recombinant isoforms of *Cor a 1*, the major allergen

- of hazel pollen, show different IgE-binding properties. *Eur. J. Biochem.* 212: 355–362.
- Breiteneder, H., Hoffman-Sommergruber, K., O’Riordain, G., Susani, M., Ahorn, H., Ebner, C., Kraft, D. and Scheiner, O. 1995. Molecular characterization of *Api g1*, the major allergen of celery (*Apium graveolens*), and its immunological and structural relationships to a group of 17 kDa tree pollen allergens. *Eur. J. Biochem.* 233: 484–489.
- Bustos, M.M., Guiltinan, M.J., Jordano, J., Begum, D., Kalkan, A. and Hall, T.C. 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, M.M., Begum, D., Kalkan, F.A., Battraw, M.J. and Hall, T.C. 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.
- Chen, Z.-L., Naito, S., Nakamura, I. and Beachy, R.N. 1989. Regulated expression of genes encoding soybean beta-conglycinins in transgenic plants. *Dev. Genet.* 10: 112–122.
- Cho, M.-J., Widholm, J.M. and Vodkin, L.O. 1995. Cassettes for seed-specific expression tested in transformed embryogenic cultures of soybean. *Plant Mol. Biol. Rep.* 13: 255–269.
- Clough, S.J. and Bent, A.F. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16: 735–743.
- Cordes, S., Deikman, J., Margossian, L.J. and Fischer, R.L. 1989. Interaction of a developmentally regulated DNA-binding factor with sites flanking two different fruit-ripening genes from tomato. *Plant Cell* 1: 1025–1034.
- Crowell, D.N., John, M.E., Russell, D. and Amasino, R.M. 1992. Characterization of a stress-induced, developmentally regulated gene family from soybean. *Plant Mol. Biol.* 18: 459–466.
- Czarnecka, E., Ingersoll, J.C. and Gurley, W.B. 1992. AT-rich promoter elements of soybean heat shock gene *GMHSP17.5E* bind two distinct sets of nuclear proteins in vitro. *Plant Mol. Biol.* 19: 985–1000.
- Deikman, J., Xu, R., Kneissl, M.L., Ciardi, J.A., Kim, K.-N. and Pelah, D. 1998. Separation of *cis* elements responsive to ethylene, fruit development, and ripening in the 5′-flanking region of the ripening-related *E8* gene. *Plant Mol. Biol.* 37: 1001–1011.
- del Mar Parra, M., del Pozo, O., Luna, R., Godoy, J.A. and Pintor-Toro, J.A. 1996. Structure of the dehydrin *tas14* gene of tomato and its developmental and environmental regulation in transgenic tobacco. *Plant Mol. Biol.* 32: 453–460.
- Feinberg, A. and Vogelstein, B. 1982. A technique for radiolabeling DNA restriction enzyme fragments to high specific activity. *Anal. Biochem.* 132: 6–13.
- Finer, J.J., Vain, P., Jones, M.W. and McMullen, M.D. 1992. Development of the particle inflow gun for DNA delivery to plant cells. *Plant Cell Rep.* 11: 323–328.
- Griffing, L.R., Nessler, C.L. 1989. Immunolocalization of the major latex proteins in developing laticifers of opium poppy (*Papaver somniferum* L.). *J. Plant Physiol.* 134: 357–363.
- Hamilton, D.A., Schwarz, Y.H. and Mascarenhas, J.P. 1998. A monocot pollen-specific promoter contains separable pollen-specific and quantitative elements. *Plant Mol. Biol.* 38: 663–669.
- Holsters, M., De Waele, D., Depicker, A., Messens, E., Van Montagu, M. and Schell, J. 1978. Transfection and transformation of *Agrobacterium tumefaciens*. *Mol. Gen. Genet.* 163: 181–187.
- Jefferson, R.A. 1987. Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol. Biol. Rep.* 5: 387–405.
- Jefferson, R.A., Kavanagh, T.A., Bevan, M.W. 1987. GUS fusions: β -glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* 6: 3901–3907.
- Klein, T.M., Wolf, E.D., Wu, R. and Sanford, J.C. 1987. High-velocity microprojectiles for delivering nucleic acids into living cells. *Nature* 327: 70–73.
- Koncz, C. and Schell, J. 1986. The promoter of the T_L-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* 204: 383–396.
- Kulik, M. 1984. Symptomless infection, persistence, and production of pycnidia on host and non-host plants by *Phomopsis batatae*, *Phomopsis phaseoli*, and *Phomopsis sojae*, and the taxonomic implications. *Mycologia* 76: 274–291.
- Larsen, J.N., Stroman, P. and Ipsen, H. 1992. PCR based cloning and sequencing of isogenes encoding the tree pollen major allergen *Car b1* from *Carpinus betulus*, hornbeam. *Mol. Immunol.* 29: 703–711.
- Lauridsen, P., Franssen, H., Stougaard, J., Bisseling, T. and Markcker, K.A. 1993. Conserved regulation of the soybean early nodulin *ENOD2* gene promoter in determinate and indeterminate transgenic root nodules. *Plant J.* 3: 483–492.
- Lindstrom, J.T., Vodkin, L.O., Harding, R.W. and Goeken, R.M. 1990. Expression of soybean lectin gene deletions in tobacco. *Dev. Genet.* 11: 160–167.
- Lu, Z., Templer, M. and Nielsen, B. 1994. Rapid method for recovery of DNA from agarose gels. *BioTechniques* 16: 400–402.
- Meeks-Wagner, D.R., Dennis, E.S., Tran T.V.K. and Peacock, W.J. 1989. Tobacco genes expressed during in vitro floral initiation and their expression during normal plant development. *Plant Cell* 1: 25–35.
- Moiseyev, G.P., Fedoreyeva, L.I., Zhuravlev, Y.N., Yasnetskaya, E., Jekel, P.A. and Beintema, J.J. 1997. Primary structures of two ribonucleases from ginseng calluses. New members of the PR-10 family of intercellular pathogenesis-related plant proteins. *FEBS Lett.* 407: 207–210.
- Nam, Y.W., Tichit, L., Leperlier, M., Cuerq, B., Marty, I. and Lelièvre, J.M. 1999. Isolation and characterization of mRNAs differentially expressed during ripening of wild strawberry (*Fragaria vesca* L.) fruits. *Plant Mol. Biol.* 39: 629–636.
- Neale, A.D., Wahleithner, J.A., Lund, M., Bonnett, H., Kelly, A., Meeks-Wagner, D.R., Peacock, J.W. and Dennis, E.S. 1990. Chitinase, β -1,3-glucanase, osmotin and extensin are expressed in tobacco explants during flower formation. *Plant Cell* 2: 673–684.
- Nessler, C.L. 1994. Sequence analysis of two new members of the major latex protein gene family supports the triploid-hybrid origin of the opium poppy. *Gene* 139: 207–209.
- Nessler, C.L. and Burnett, R.J. 1992. Organization of the major latex protein gene family in opium poppy. *Plant Mol. Biol.* 20: 749–752.
- Nessler, C.L. and Von der Haar, R.A. 1990. Cloning and expression analysis of DNA sequences for the major latex protein of opium poppy. *Planta* 180: 487–491.
- Nessler, C.L., Kurz, W.G.W. and Pelcher, L.E. 1990. Isolation and analysis of the major latex protein genes of opium poppy. *Plant Mol. Biol.* 15: 951–953.
- Osmark, P., Boyle, B. and Brisson, N. 1998. Sequential and structural homology between intracellular pathogenesis-related proteins and a group of latex proteins. *Plant Mol. Biol.* 38: 1243–1246.
- Pozueta-Romero, J., Klein, M., Houlné, G., Schantz, M.-L., Meyer, B. and Schantz, R. 1995. Characterization of a family of genes encoding a fruit-specific wound-stimulated protein of

- bell pepper (*Capsicum annuum*): identification of a new family of transposable elements. *Plant Mol. Biol.* 28: 1011–1025.
- Ruiz-Rivero, O. and Prat, S. 1998. A –308 deletion of the tomato LAP promoters is able to direct flower-specific and MeJA-induced expression in transgenic plants. *Plant Mol. Biol.* 36: 639–648.
- Rushton, P.J., Torres, J.T., Parniske, M., Wernert, P., Hahlbrock, K. and Somssich, I.E. 1996. Interactions of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. *EMBO J.* 15: 5690–5700.
- Samac, D.A. and Shah, P.M. 1991. Developmental and pathogen-induced activation of the *Arabidopsis* acidic chitinase promoter. *Plant Cell* 3: 1063–1072.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. *Molecular cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sandhu, J.S., Webster, C.I. and Gray, J.C. 1998. A/T-rich sequences act as quantitative enhancers of gene expression in transgenic tobacco and potato plants. *Plant Mol. Biol.* 37: 885–896.
- Schneider, R.W., Dhingra, O.D., Nicholson, J.F. and Sinclair, J.B. 1974. *Colletotrichum truncatum* borne within the seed coat of soybean. *Phytopathology* 64: 154–155.
- Schöffl, F., Raschke, E. and Nagao, R.T. 1984. The DNA sequence analysis of soybean heat-shock genes and identification of possible regulatory promoter elements. *EMBO J.* 3: 2491–2497.
- Singh, K.B. 1998. Transcriptional regulation in plants: the importance of combinatorial control. *Plant Physiol.* 118: 1111–1120.
- Stomp, A.-M. 1992. Histochemical localization of β -glucuronidase. In: S.R. Gallagher (Ed.), *GUS Protocols: Using the GUS Gene as a Reporter of Gene Expression*, Academic Press, San Diego, CA, pp. 103–113.
- Stougaard, J., Sandal, N.N., Grøn, A., Kühle, A. and Marcker, K.A. 1987. 5' Analysis of the soybean leghaemoglobin *lbc₃* gene: regulatory elements required for promoter activity and organ specificity. *EMBO J.* 6: 3565–3569.
- Terryn, N., Brito Arias, M., Engler, G., Tiré, C., Villarroel, R., Van Montagu, M. and Inzé, D. 1993. *rha1*, a gene encoding a small GTP binding protein from *Arabidopsis*, is expressed primarily in developing guard cells. *Plant Cell* 5: 1761–1769.
- Todd, J.J. and Vodkin, L.O. 1996. Duplications that suppress and deletions that restore expression from a soybean multigene family. *Plant Cell* 8: 687–699.
- Wieczorek, E., Brand, M., Jacq, X. and Tora, L. 1998. Function of TAFII-containing complex without TBP in transcription by RNA polymerase II. *Nature* 393: 187–191.
- Zhao, Y., Leisy, D.J. and Okita, T.W. 1994. Tissue-specific expression and temporal regulation of the rice glutelin Gt3 gene are conferred by at least two spatially separated *cis*-regulatory elements. *Plant Mol. Biol.* 25: 429–436.
- Zhou, C., Yang, Y. and Jong, A.Y. 1990. Mini-prep in ten minutes. *BioTechniques* 8: 172–173.