Reporter Gene Expression Patterns Regulated by an Ara h 2 Promoter Differ in Homologous Versus Heterologous Systems[1]

A. Bhattacharya[2], M.L. Ramos[3], P. Faustinelli[4], and P. Ozias-Akins*

ABSTRACT

Peanut (*Arachis hypogaea* L.) is a globally important crop whose seeds are widely used in food products. Peanut seeds contain proteins that serve a nutrient reservoir function and that also are major allergens. As part of an investigation to determine the effect of reducing/eliminating the peanut allergen Ara h 2 from seeds, gene sequence including upstream regulatory regions was characterized. The ability of regions upstream of the translation initiation site to regulate seed-specific expression of reporter genes was tested in peanut and *Arabidopsis*. Two independent transgenic peanut lines biolistically transformed with 1kb of DNA upstream of the Ara h 2.02 (B-genome) coding sequence controlling a Green Fluorescent Protein – **β**-glucuronidase (*Gfp-Gus*) fusion were obtained. All T_1, T_2 and T_3 generations of transgenic plants showed the expression of GFP and GUS restricted to seeds and near background levels in vegetative tissues. However, constitutive GUS expression was observed in *Arabidopsis* transgenic lines, a heterologous system. It is possible that trans-acting factors regulating seed specificity in peanut are too divergent in *Arabidopsis* to enable the seed specific response. Thus, the promoter described in this paper may have potential use for expression of transgenes in peanut where seed-specificity is desired, but expression patterns should be tested in heterologous systems prior to off-the-shelf adoption.

Key Words: Peanut allergen, Ara h 2.02 promoter, reporter gene, seed specificity, motif analysis, peanut, *Arabidopsis*, transgenic.

Peanut (*Arachis hypogaea* L.) is an allotetraploid species, $2n = 4x = 40$, having an AABB genome derived from its putative progenitor species, *Arachis duranensis* (AA) and *Arachis ipaensis* (BB) (Burow *et al.*, 2009; Kochert *et al.*, 1996; Seijo *et al.*, 2004), though other wild progenitor species have been proposed (Raina and Mukai, 1999; Smartt *et al.*, 1978). Thus, at least two genes, one from each sub-genome, are likely to encode each protein in peanut. The seed storage proteins of peanut, which are largely 2S albumin (Ara h 2, 6, and 7), 7S globulin (Ara h 1), and 11S globulin (Ara h 3, 4), are allergenic (Ozias-Akins *et al.*, 2006). Of these seed storage proteins, 11S globulins are the most abundant comprising on average 30% of total protein (Kang *et al.*, 2007a), whereas Ara h 2 is the most potent allergen (Flinterman *et al.*, 2007). In the course of our study to assign expressed allergen genes to their sub-genomes (Ramos *et al.*, 2006), we cloned the regulatory and coding regions of *Ara h 2* genes. This study revealed that *Ara h 2* has two copies, one from the A genome designated *Ara h 2.01* and the other from B genome, *Ara h 2.02*. The two protein isoforms mainly differ by a 12-amino acid insertion in *Ara h 2.02*. Ara h 2 expression was shown to begin in immature seeds (stages 1 and 2) and peak as seeds matured (stages 3–4) (Kang *et al.*, 2007b). Also, the expression of Ara h 2 was reduced in embryonic axes compared to cotyledons in the post germination (seedling) growth stages. The *Ara h 2* transcript was absent from the vegetative parts of the plant as confirmed by Northern blots and RT-PCR analysis (Kang *et al.*, 2007b).

Accurate prediction of promoter function in eukaryotic systems is difficult, since promoters may contain many different motifs, some present in multiple copies, that can interact with transcription factors (Priest *et al.*, 2009). Core promoter regions typically extend 40 bp upstream of a start codon, although upstream promoter regions may encompass more than 200 bp farther upstream (Potenza *et al.*, 2004). Studies on promoters controlling seed specific expression have shown that essential motifs are present within 500 bp upstream of the ATG start codon (Priest *et al.*, 2009).

The primary objective of this study was to identify the promoter region responsible for tissue-specific expression patterns resulting in accumulation of Ara h 2 allergen protein in peanut seeds. Seed-specific storage protein accumulation is regulated by hormonal and metabolic regimes (Priest

[1]Research location and current address of P. Ozias-Akins: Department of Horticulture and NESPAL, The University of Georgia Tifton Campus, Tifton, GA 31793-5766.

[2]Present address: Bench Biotechnology, Vapi, Gujarat, India.

[3]Present address: NIDERA S.A., Departamento de Biotecnologia, Venado Tuerto, Santa Fe CP2600, Argentina.

[4]Present address: Faculty of Agricultural Sciences, Catholic University of Cordoba, Camino a Alta Gracia km 7 1/2 (5017), Cordoba, Argentina.

*Corresponding author email: pozias@uga.edu

Table 1. List of primers used in the present study.

Primer ID	Target Fragment	Sequence
91	*Ara h 2* coding	5′-GAAGATCCAACGTGACGA-3′
189	*Ara h 2* coding	5′-GTCCAAATCGCAACGCTG-3′
400	*Adf* coding	5′-GCAGCATCAGGTATGGCAGT-3′
401	*Adf* coding	5′-TTCCATCAAGTTCCCCTCTT-3′
697	*Ara h 2* 1 kb promoter	5′-CCCAAGCTTGGGAACAGCTGAAACAAAAAGA-3′
698	*Ara h 2* promoter	5′-CATGCCATGTTGTTGTGGTGGTTAT-3′
703	*Ara h 2* 2 kb promoter	5′-CCCAAGCTTCACTATCTTTTCTCCTCGTCACACTA-3′
969	*Gfp* coding	5′-CTCTTGAAGAAGTCGTGCCGCTT-3′
1013	*Ara h 2* 2.5 kb promoter	5′- CCCAAGCTTGATACTTTTATTAAATACAGCCAT -3′
1621	*Gus* coding	5′-AGTGTACGTATCACCGTTTGTGTGAAC-3′
1622	*Gus* coding	5′-ATCGCCGCTTTGGACATACCATCCGTA-3′

et al., 2009; Verdier and Thompson, 2008); therefore, *in silico* analysis of the sequence upstream from the start codon of *Ara h 2* was carried out to identify motifs known to play a role in seed-specific expression. Since upstream regulatory regions of arbitrary length often are used to test for promoter function, an ~ 1 kb (989 bp) region was selected for fusion with reporter genes, green fluorescent protein (*Gfp*) and ß-glucuronidase (*Gus*), and expression analysis in peanut, whereas 989 bp, 1927 bp, and 2517 bp regions were tested in *Arabidopsis*.

Materials and Methods

Motif analysis

The 3595 bp region upstream of the start codon in *Ara h 2.02* was analyzed for different motifs, particularly those known to confer seed specificity, using PLACE (Higo *et al.*, 1999) and PLANT-CARE (Lescot *et al.*, 2002) databases and search tools.

Plasmid construct

The *Ara h 2.02*pro:*Gfp:Gus* fusion was constructed from pCambia1304 where the CaMV 35S promoter driving the *Gfp:Gus* cassette had been excised by digestion with *Nco*I and *Hin*dIII. The promoter region of *Ara h 2.02* was amplified from the plasmid clone 8A2a (GI 148613180, EF 609644) using specific primer combinations, 697 sense / 698 anti-sense (989 bp), 703 sense / 698 anti-sense (1927 bp), or 1013 sense / 698 anti-sense (2517 bp) that included appropriate restriction sites (Table 1). PCR amplification was carried out according to the following conditions: initial denaturation at 94 C for 5 min followed by 35 cycles of 94 C for 30 s, variable annealing temperatures and extension times at 72 C (primers 697 / 698: 48 C for 30 sec, 1-min extension; 703/698: 50 C for 30 sec, 2-min extension; 1013/698: 50 C for 30 sec and 3.5-min extension) with a final extension of 7 min at 72 C and a 4 C hold. The PCR products were digested with *Nco*I and *Hin*dIII and gel purified using a Qiagen kit as per the manufacturer's protocol (Qiagen). After ligation of the digested vector and PCR products, the plasmids were transformed into *E. coli* DH5α (NEB) competent cells, extracted using the Qiagen Miniprep Kit (Qiagen), and inserts were sequenced. Plasmids were transformed into *Agrobacterium tumefaciens* strain GV 3101 by electroporation.

Plant tissue culture

Somatic embryos (SE) were obtained by initiating cultures from mature peanut seeds of 'Georgia Green' in the SE induction medium which consisted of FN lite macro salts (Samoylov *et al.*, 1998), MS micronutrients (Murashige and Skoog, 1962), B5 vitamins (Gamborg *et al.*, 1968), 1 g/L L-asparagine (Fisher Scientific), 3 mg/L picloram (Dow Agrosciences), 30 g/L sucrose (J. T. Baker, Inc.), and 8 g/L agar (Sigma). The pH was adjusted to 5.7–5.8 prior to autoclaving at 121 C for 20 min. The medium was cooled below 65 C and then filter sterilized L-glutamine solution, 1 g/L (Acros Organics), was added as a supplement prior to dispensing to plates. Seeds were surface sterilized in 50% solution of commercial bleach (Clorox) for 15 min on a shaker, followed by three rinses with sterile deionized water. The plumule region of the embryo axis was carefully excised under a stereomicroscope. Embryogenic cultures were grown in the dark at 26 ± 2 C for up to 9 mo with periodic subculture every 4 wk. Bombardment was conducted 12–14 d after subculture.

Microprojectile bombardment, selection, and regeneration of peanut plants

The Qiagen Plasmid Maxi/Midi Kit was used to extract plasmid DNA from an overnight culture of *E. coli* cells. DNA quantification was carried out with the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen). Microprojectile bombardment and selection were conducted as described previously (Chu *et al.*, 2008). Hygromycin-resistant embryogenic

tissues were regenerated on MS medium supplemented with 1 mg/L thidiazuron (NOR-AM Chemical Co.) along with 20 mg/L hygromycin (Sigma), 30 g/L sucrose, and 8 g/L agar under a daily photoperiod of 16 h light (50 μmol/m^2/s) followed by 8 h of darkness for at least 7–8 wk. Somatic embryos undergoing conversion were transferred to fresh MS medium containing 20 mg/L hygromycin. Embryos displaying well-developed shoot-root axes were subsequently transferred to ventilated vessels containing MS medium with 0.5 mg/L kinetin (Sigma) and 0.25 mg/L 6-benzylaminopurine (Sigma). Well-developed shoots were selected for rooting in Magenta vessels on MS medium supplemented with 0.1 mg/L 1-naphthaleneacetic acid (Sigma). Plants with well-developed roots were transferred to soil, acclimatized, and grown in the greenhouse.

Agrobacterium-mediated _Arabidopsis_ transformation and screening

Arabidopsis transformation was carried out by the standard floral dip procedure (Clough and Bent, 1998). Seeds (T_1) from treated inflorescences and subsequent generations (T_2–T_3) were sterilized in 70% alcohol for 30 sec, 50% Clorox/0.02% Triton X-100 for 6 min, and washed three times with sterile deionized water. The putative transgenic seeds were selected in ½-strength MS medium containing hygromycin (50 mg/L final concentration) and 8 g/L agar. Seeds were plated and stratified at 4 C for 4–7 d before being transferred to a growth chamber at 22 ± 2 C under a 16 h light: 8 h dark photoperiod. Hygromycin-resistant seedlings were transferred to soil and analyzed after 2 wk by PCR and GUS histochemistry to confirm positive transformants. From each T_2 line showing 3:1 Mendelian segregation for hygromycin resistance, at least 8 transgenic plants were chosen for bulk harvest and T_3 seeds were screened for hygromycin resistance. Plates showing 100% survival and growth were considered to contain homozygous lines.

PCR analysis

Genomic DNA was extracted as described previously (Singsit *et al.*, 1997). Putative transgenic *Arabidopsis* plants were screened by amplifying the 989 bp, 1927 bp, or 2517 bp promoter region by the same set of primers used initially for cloning (Table 1). The *Gus* gene was detected by using primers 1621 sense/1622 antisense, which generated a product of 1061 bp. Actin depolymerizing factor, a housekeeping gene, was detected by primers 400 sense / 401 antisense producing a 450 bp product, and *Ara h 2* by primers 91 sense / 189 antisense, for a PCR product size of 337 bp. Putative transgenic peanut plants were screened by using the 1 kb promoter forward primer 697 and *Gfp* reverse primer 969, product size of 1270 bp. PCR amplification was carried out according to the following conditions: initial denaturation at 94 C for 5 min followed by 35 cycles of 94 C for 30 s, variable annealing temperatures and extension times (annealing temperatures: 697 / 969, 59.3 C; 400 / 401, 53 C; 91 / 189, 54 C; 1621 / 1622, 71 C for 60 s; extension times: 697 / 969, 120 s; 400 / 401, 60 s; 91 / 189, 60 s; 1621 / 1622, 75 s at 72 C) with a final extension of 7 min at 72 C and a 4 C hold. PCR products (8 μl) were separated in 1% (w/v) agarose gels (70V for 1 h) containing ethidium bromide (0.1 μg/ml) with a 1 kb Hi-Low ladder (Minnesota Molecular) as marker. Gels were visualized using a UV transilluminator.

Southern blot hybridization

Genomic DNA was extracted with the Qiagen Plant DNA extraction kit according to the manufacturer's protocol and quantified by a Nanodrop1000 (Nanodrop). Approximately 20 μg DNA was digested overnight at 37°C with 50 units of *Hin*d III restriction enzyme in a 20 μl final reaction volume. Digested DNA fragments were separated overnight by electrophoresis on a 1% agarose gel in 1× TBE buffer at 25V. DNA was blotted onto Genescreen Plus nylon membrane (NEN Life Sciences) by capillary transfer. The membrane was pre-hybridized in 30 ml hybridization buffer (6×SSC, 1% SDS, 100 μg/ml salmon sperm DNA) at 65 C overnight. Probes were labeled with α-^{32}P-dCTP following the protocol described in the PCR labeling kit (Sigma) and cleaned with Sephadex® G-50 (Amersham). Labeled probe was denatured at 95 C for 10 min and placed immediately on ice before adding into the hybridization buffer. Hybridization was conducted at 65 C overnight followed by four washes at the same temperature for 15 min each with the following buffers: 1) 2×SSC, 0.1% SDS; 2) 1× SSC, 0.1% SDS; 3) 0.5× SSC, 0.1% SDS; and 4) 0.1 ×SSC, 0.1% SDS. Signal was detected with the Storm Phosphorimager system (Amersham Biosciences).

Semi-quantitative reverse transcriptase PCR (RT-PCR)

RNA was extracted from 100 mg leaf samples / seed cotyledons (stage 3) (Paik-Ro *et al.*, 2002) of transformed and non-transformed control plants using an RNeasy® Plant Mini Kit (Qiagen). Two rounds of DNase treatment were employed, one at the column purification stage (Qiagen) and another post purification (DNase from Invitrogen) according to the manufacturers' protocols. RT-PCR employed a One Step RT-PCR Kit (Qiagen) according to the manufacturer's instructions. Subsequent conditions were as above for PCR analysis. A set

Table 2. Cis-regulatory elements present in 3595 bp upstream sequence from Arah 2.02 according to PLACE and PlantCARE promoter databases.

Motif ID	Position upstream from ATG: −1 to −989 bp	−990 to −1927 bp	−1928 to −2517 bp	−2518 to −3595	Motif sequence	Function	References
−300ELEMENT (PROLAMIN BOX)	−320, −725	np[a]	np	−2764, −3507	TGHAAARK	seed specific expression	(Colot *et al.*, 1987; Kreis *et al.*, 1986)
2SSEEDPROTBANAPA	−155	np	np	np	CAAACAC	seed storage protein expression	(Stalberg *et al.*, 1996)
AACACOREOSGLUB1	np	−1527	np	np	AACAAAC	seed specific expression	(Wu *et al.*, 2000)
ABRELATERD1	−183	np	−2048	np	ACGTG	ABA response	(Simpson *et al.*, 2003)
ACGTTBOX (T- Box)	np	np	np	−2917	AACGTT	seed specific expression	(Fauteux and Stromvik, 2009; Foster *et al.*, 1994)
AMYBOX2	np	np	np	−2908	TATCCAT	seed protein expression	(Huang *et al.*, 1990; Hwang *et al.*, 1998)
ARFAT	np	np	np	−2900	TGTCTC	auxin response	(Nag *et al.*, 2005)
ARR1AT	−262	np	−2147	−2639	NGATT	growth regulator response	(Ross *et al.*, 2004)
	−370		−2177	−2782			
	−380			−2834			
	−481			−2937			
	−670			−2980			
				−3039			
				−3341			
ATHB6COREAT	np	np	−2441	np	CAATTATTA	growth regulator response	(Himmelbach *et al.*, 2002)
BIHD10S	np	np	−1950	np	TGTCA	homeodomain transcription factor binding, seed storage protein expression	(Fauteux and Stromvik, 2009; Luo *et al.*, 2005)
CAATBOX1	−31	−1623	−2133	−2610	CAAT	seed protein expression	(Fauteux and Stromvik, 2009)
	−167	−1662	−2441	−2795			
	−790	−1688	−2581	−2867			
	−813			−3564			
CANBNNAPA	−155	np	np	np	CNAACAC	seed specific expression	(Ellerstrom *et al.*, 1996)
CATATGGMSAUR	np	−1435	np	−2824	CATATG	auxin response	(Xu *et al.*, 1997)
DPBFCOREDCDC3	−688	np	np	np	ACACNNG	ABA response	(Lopez-Molina and Chua, 2000)
EBOXBNNAPA	−105	−1095	−2117	−2824	CANNTG	seed storage protein expression	(Hartmann *et al.*, 2005)
	−891	−1435		−2904			
	−985						
ERELEE4	−627	np	np	np	AWTTCAAA	ethylene response	(Rawat *et al.*, 2005)
MYB1AT	−17	np	−2121	np	WAACCA	ABA response	(Abe *et al.*, 2003)
MYCATRD22	−105	np	−2117	np	CACATG	ABA response	(Busk and Pages, 1998)
	−891						

Table 2. Continued. Cis-regulatory elements present in 3595 bp upstream sequence from Arah 2.02 according to PLACE and PlantCARE promoter databases.

Motif ID	Position upstream from ATG				Motif sequence	Function	References
	−1 to −989 bp	−990 to −1927 bp	−1928 to −2517 bp	−2518 to −3595			
NAPINMOTIFBN	−28	np	np	np	TACACAT	seed storage protein expression	(Ericson *et al.*, 1991)
RYREPEATBNNAPA / RY- Elements	−103	−1462	np	np	CATGCA	seed storage protein expression	(Ezcurra *et al.*, 2000; Fujiwara and Beachy, 1994; Lelievre *et al.*, 1992)
SEF1MOTIF	−429	np	np	−2526 −3592	ATATTTAWW	seed storage protein expression	(Lessard *et al.*, 1991)
SEF3MOTIFGM	np	−1537	np	np	AACCCA	seed storage protein expression	(Lessard *et al.*, 1991)
SEF4MOTIFGM7S	np	−1504	−2381 −2482	−2637 −3307 −3545	RTTTTTR	seed storage protein expression	(Lessard *et al.*, 1991)
T/GBOXATPIN2	−184	np	np	np	AACGTG	jasmonate response	(Boter *et al.*, 2004)

[a]np -not present.

of negative RT controls was also included to test for genomic DNA contamination. Actin depolymerizing factor (*Adf*) served as a control housekeeping gene and was amplified for 24 cycles. *Gus* was amplified for 30 cycles and *Ara h 2* for 28 cycles for semi-quantitative RT-PCR purposes. Image quantification (ImageQuant TL, Amersham Biosciences) was used to estimate amounts of RT-PCR products by densitometry.

GUS expression analysis

Four seeds were selected from the T_1 peanut lines and each seed was divided into two cotyledons, one containing the embryo axis. The half containing the embryo axis was transferred to soil. The other half was further divided into two pieces, one of which was used for GUS assay and the other for GFP analysis. Leaf tissues (2 and 8 wk post planting) were tested for GUS activity. Tissues were immersed in a GUS reagent solution (Jefferson *et al.*, 1987) and incubated overnight at 37 C in the dark with appropriate controls, i.e. non-transformed or 35S transformed plants. An indigo blue color confirmed GUS expression.

GUS quantification

GUS quantification was carried out by grinding 100 mg of sample (seeds/leaf tissues) in 200 µl of extraction buffer (100 µl 0.1 M NaH_2PO_4, pH 7, 0.014 µl ß-mercaptoethanol, 0.4 µl 0.5 M Na_2EDTA, 0.2 mg Sarkosyl, 0.2 µl Triton X-100 and 99.38 µl of dH_2O). Cell debris was removed by centrifugation at 12,000 rpm for 15 min at 4 C. The supernatant was transferred to a new tube and protein was estimated by the Bradford assay (Bio-Rad Laboratories). Homogenate (20 µl) containing 5 µg protein was mixed with 80 µl of GUS assay buffer prepared by adding 0.88 mg MUG (4-methylumbelliferyl ß-D-glucuronide; Sigma) in each ml of extraction buffer. A 1 mM MU (sodium methylumbelliferone; Sigma) stock solution was used to prepare dilutions for generation of a standard curve. After 30 min, 475 µl of 200 mM Na_2CO_3, pH 11.2, stop-buffer was added to 25 µl of each reaction mixture or standard dilution. Duplicate samples of 200 µl for each of three plants per line were loaded in a microtiter plate and fluorescence was determined at excitation and emission wavelengths of 365 nm and 444 nm, respectively. To determine endogenous background GUS activity, plant extract from non-transgenic plants was taken as a control. GUS activity was calculated as pmoles MU/mg protein/min.

GFP signal detection

GFP was detected by placing the plant material (seeds or leaf tissues, as appropriate) directly under blue light and observing with a SV11 epifluorescence stereomicroscope (Zeiss) equipped with a

100W mercury lamp as the light source with a 480 ± 30 nm excitation filter and a 515 nm long-pass emission filter (Chroma Technology). Images were captured with an Axiocam digital camera (Zeiss).

Statistical analysis

Data were subjected to Student's t-test using standard MINITAB version 15, statistical software (Minitab, Inc.).

Results

Motif analysis of *Ara h 2.02* upstream sequence

A detailed analysis for short conserved putative *cis*-regulatory sequences was undertaken for 3595 bp upstream of the translational start site for *Ara h 2.02* using PLACE and PLANTCARE databases and search tools with the sequence information of EF 609644, GI148613180. The analysis revealed *cis*-regulatory elements known to confer seed-specific expression and response to growth regulators (Table 2). Since numerous *cis*-regulatory elements responsible for seed-specific expression were present in the 1 kb promoter region (Table 2), both peanut and *Arabidopsis* transformation were conducted with this promoter-reporter construct.

Transgenic peanut

Copy number determination by Southern blot hybridization

Two fertile transgenic lines, designated BP 1.4 and BP 1.8, were obtained via hygromycin selection of peanut embryogenic cultures bombarded with a plasmid carrying the *Ara h 2.02*pro:*Gfp:Gus* fusion. Southern blot hybridization results showed that both the lines were homozygous in the T_3 generation although BP 1.4 showed one band and BP 1.8 showed four bands hybridizing with the *Gus* gene (Fig. 1). Transgene integration at a single locus seems likely given the similar hybridization pattern among all T_3 progeny from each parental line.

GUS and GFP expression in seeds

GUS and GFP expression was restricted to the cotyledons of seeds and essentially absent from vegetative tissues (Fig. 2). Semi-quantitative RT-PCR results showed that *Ara h 2* gene expression in transgenic stage 3 peanut seeds was similar to the control (Fig. 2B), particularly when expressed as a fraction of *Adf* (*Adf* : *Ara h 2* ratio of 1.32 in transgenic vs. 1.44 in control). Between the two transgenic lines BP 1.4 and BP 1.8, *Gus* transcripts were not significantly different at $P<0.01$ (Fig. 2B). Even though BP 1.8 carried a higher copy number of *Ara h 2.02*pro:*Gfp:Gus* than BP 1.4, GUS expression levels as measured in pmol MU/mg protein/min were not significantly different between the two lines for both tissues tested, although GUS expression was shown to be several orders of magnitude higher in stage 3 seeds compared with 2-wk-old vegetative tissues of transgenic peanut lines (Fig. 2C, D).

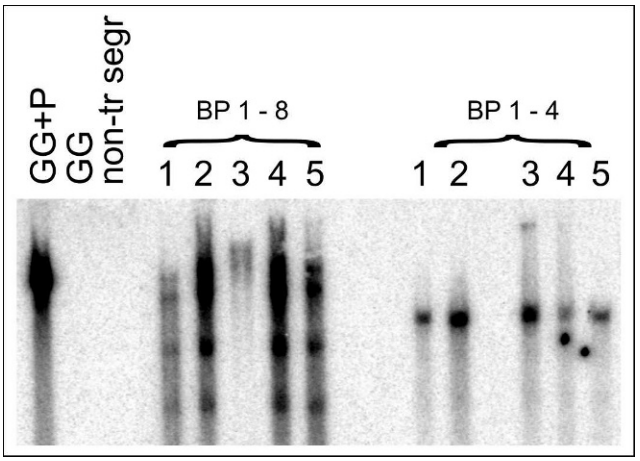

Fig. 1. Southern blot hybridization analysis from two *Ara h 2.02*pro:*Gfp:Gus* transgenic lines, BP 1.4 and BP 1.8, in the T_3 generation (5 plants from each individual line numbered 1 through to 5). Upon digestion with *Hin*dIII, BP 1.4 showed one band whereas BP 1.8 showed four bands (except for lane 3 which was incompletely digested) when probed with *Gus*. Controls include a non-transgenic segregant from BP 1.8 (non-tr segr), non-transformed 'Georgia Green' (GG), and non-transformed 'Georgia Green' DNA mixed with *Hin*dIII-digested plasmid DNA (GG+P).

GUS expression in vegetative tissues declines post germination

Faint GUS expression was observed in vegetative parts of young seedlings 2 wk after germination and GUS expression was undetectable in 8-wk-old leaves. The presence of transcript (*Gfp*:*Gus* cassette) was confirmed by RT-PCR in 2-wk-old tissues from T_2 individuals (Fig. 3A) and also by semi-quantitative RT-PCR in the T_3 generation (*Gus* transcript only, data not shown). There was no significant difference between the two transgenic lines (BP 1.4 and BP 1.8) in the level of *Gus* transcripts from 2-wk-old vegetative tissues (data not shown). Quantitative GUS analysis showed that expression in line BP 1.4 was similar to BP 1.8 after a 30-min assay (Fig. 2D). Negative controls consisted of two lines, which were transgene nulls segregating from the transformed lines, i.e., lacking the transgene cassette (confirmed by PCR and Southern blot). Neither GUS enzyme activity nor *Gus* transcripts was detectable in 8-wk-old plants from either transgenic line at any of the three generations.

Transgenic *Arabidopsis*

Six transgenic *Arabidopsis* T_1 lines containing the 1 kb *Ara h 2.02*pro:*Gfp:Gus* cassette, five with the 2 kb promoter, and six with the 2.5 kb promoter were recovered. Five transgenic lines

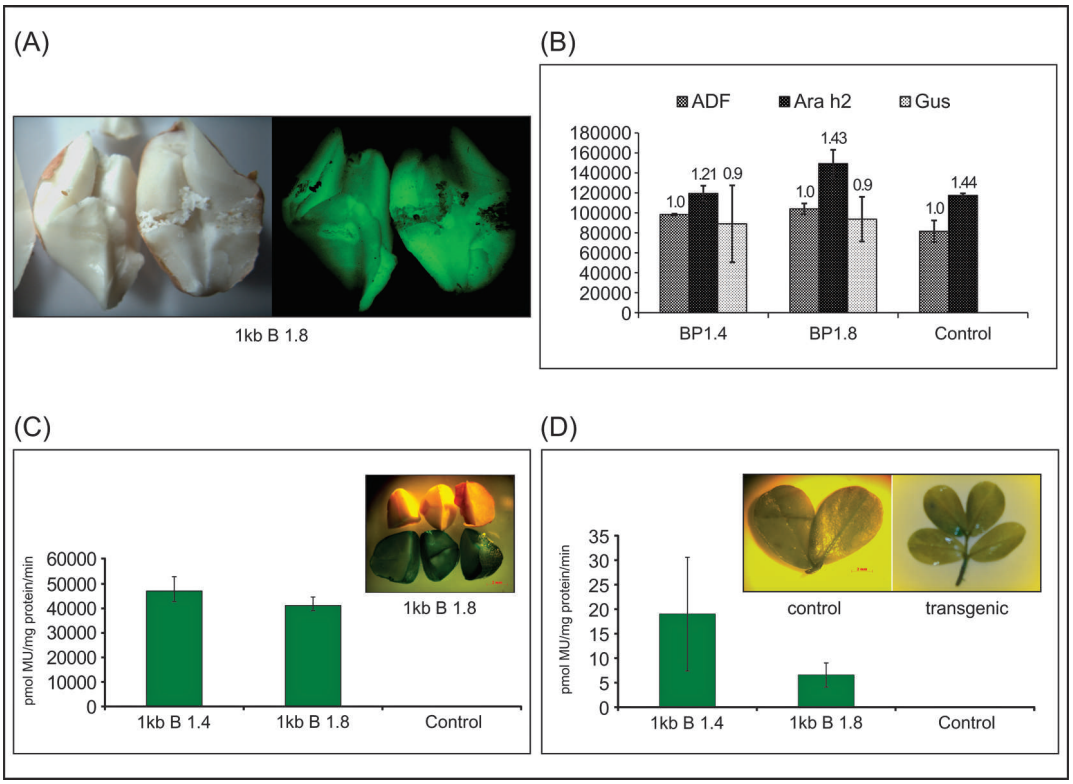

Fig. 2. Transgenic 1 kb *Ara h 2.02*pro:*Gfp:Gus* peanut lines (T_2 generation) (A) GFP expression (right) in transgenic cotyledons. Left panel shows half-seeds under white light. (B) Densitometric quantification of *Adf* (Actin depolymerizing factor), *Ara h 2* and *Gus* transcripts. Error bars indicate standard error of mean. Numbers above bars are ratio of respective gene transcript level to *Adf* transcript level. (C) Quantitative GUS assay with transgenic (B1.4 and B1.8) and non-transgenic control seeds. Inset - GUS expression in peanut cotyledons (top – non-transgenic; bottom – transgenic). (D) Quantitative GUS assay with 2-wk-old leaves from transgenic (B1.4 and B1.8) and non-transgenic control plants. Inset - leaflets stained for GUS 2 wk post-planting; non-transgenic (left panel), transgenic (right panel).

containing the reporter genes controlled by the CaMV 35S promoter also were obtained. Single locus lines for each construct were identified, namely line numbers 58 (1 kb); 67, 68 (2 kb) and 37, 38 (CaMV 35S). Transgenic lines containing the 2.5 kb promoter showed poor seed set. Similar constitutive GUS expression patterns from all three *Ara h 2.02* promoter regions, and the CaMV 35S promoter, were observed in transgenic *Arabidopsis* plants compared to no expression in the non-transformed controls (Fig. 3).

Discussion

The present study investigates the potential of using promoter sequence from *Ara h 2.02* as a seed specific promoter (SSP). Transcriptional regulation of any gene in plants is controlled by a complex interaction of transcription factors and *cis*-regulatory elements (Dare *et al.*, 2008). Several important seed specific regulatory motifs found in the *Ara h 2.02* promoter (PROLAMINBOX, 2SSEEDPROTBANAPA) were identified for high level of activity of the *napA* promoter from *Brassica napus*. Deletion of the region containing these motifs from the *napA* promoter decreased GUS expression in transgenic tobacco plants (Stalberg *et al.*, 1993). A RY motif (CATGCA / CACGT) is commonly found in seed storage protein genes (Fauteux and Stromvik, 2009), including *Ara h 1* (Viquez *et al.*, 2003) and *Ara h 2* (this study), and acts by enhancing seed specific gene expression (Dickinson *et al.*, 1988). Certain regulatory elements are fundamental to growth regulator signaling pathways which are important in seed development. The ARR1AT element found in the *Ara h 2.02* promoter has been shown in other studies to bind ARR1 proteins which are response regulators, act as transcription factors (Kuriakose *et al.*, 2009), and are involved in cytokinin signaling. Abscisic acid (ABA) also plays an important role in expression of seed-storage protein genes and embryogenesis related genes (Luo *et al.*, 2008). Plant *et al.* (1994) showed that ABA positively regulated expression of GUS in transgenic *Brassica* containing the oleosin promoter from *Arabidopsis* fused to the reporter gene. Rowley and Herman (Rowley and Herman, 1997) found that soybean oleosin genes contain regulatory elements in the upstream region which also were found in vacuolar storage protein genes. These included the −300 element (also present in the *Ara h 2.02* promoter) that moreover

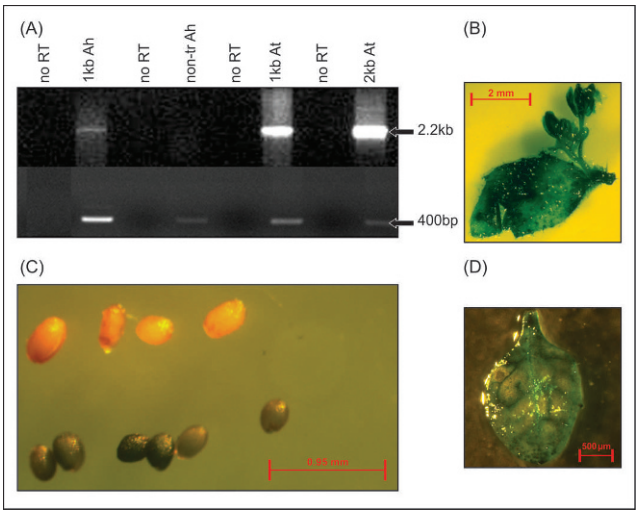

Fig. 3. (A) Top panel - RT-PCR with primers for detecting *Gfp:Gus* transcripts in 2-wk-old leaves of transgenic BP 1.8 (1 kb Ah) and non-transgenic (non-tr Ah) peanut, and leaves from transgenic *Arabidopsis* plants carrying 1 kb (1 kb At) or 2 kb (2 kb At) promoter regions. Each sample included a corresponding negative control (no RT) with all reaction components except reverse transcriptase enzyme. Bottom panel - actin depolymerizing factor (*Adf*) primers were also used to confirm integrity of the cDNA. (B) GUS expression in *Arabidopsis* vegetative tissue transformed with the 1 kb promoter (line no. 58). (C) GUS expression in *Arabidopsis* seeds from plants transformed with the 2 kb promoter (line no. 67); top row of seeds is non-transgenic control. (D) GUS expression in *Arabidopsis* vegetative tissue transformed with the 2.5 kb promoter (line no. 1).

quantitatively regulates prolamin expression, an endosperm protein in cereals (Colot *et al.*, 1987). Qu *et al.* (2008) found prolamin and endosperm-specific motifs in all but one rice glutelin gene promoter which controlled seed specific expression.

Our results for reporter gene expression in peanut under the control of the *Ara h 2* promoter largely conform with the results obtained by Kang *et al.* (2007b) who showed that Ara h 2 polypeptides were abundant in mature seeds and diminished after germination during seedling growth. An unexpected result was the lack of reporter gene expression in embryo axes, although expression in cotyledons was strong. Furthermore, we did not detect *Ara h 2* transcripts in either 2-wk- or 8–wk-old leaf tissues even though reporter gene transcripts could be observed as faint RT-PCR products in the 2-wk-old samples, data which also are consistent with Kang *et al.* (2007b) who could not detect *Ara h 2* transcripts in leaves by Northern blot analysis. The present study also shows that a higher copy number of the transgene did not result in a difference in the steady state level of transcripts. Reporter gene expression pattern largely paralleled *Ara h 2* expression pattern with the minor exceptions described above.

Therefore, the 1 kb promoter region reported in this study is sufficient for seed specific reporter gene expression, although it only confers seed-specificity in peanut and not in *Arabidopsis*. In addition to *cis*-regulatory elements, transcription factors that bind to a particular element and interact with other recruited transcription factors ultimately control transcriptional activation or suppression (Verdier and Thompson, 2008). The lack of seed-specificity in *Arabidopsis* could be due to the absence, divergence, or temporal difference in expression of trans-acting factors, which are present and functioning in peanut. There is precedent for altered expression patterns from promoters introduced into a heterologous system. For example, a −765 bp alfalfa isoflavone reductase promoter fused with *Gus* conferred expression in alfalfa root meristems, cortex and nodules, following the pattern of endogenous isoflavone reductase expression, but in transgenic tobacco, GUS expression was observed in additional vegetative tissues, including leaf and stem, as well as in reproductive tissues (Oommen *et al.*, 1994). While multiple motifs with evidence for seed-specific regulation and growth regulator response may contribute to the expression pattern for *Ara h 2* in peanut seeds, further promoter dissection with transgenic peanut would be laborious because of its low transformation efficiency. However, from a comparative perspective, the set of seed-specific motifs present in the *Ara h 2.02* promoter reported during this work are very attractive for future *in vitro* or *in vivo* analyses in other legume systems to understand evolutionary differences.

Summary and Conclusions

The present study describes a novel promoter from a peanut Ara h 2.02 protein with a seed specific expression pattern. However, the promoter-regulated reporter gene expression patterns differed in homologous versus heterologous (*Arabidopsis*) systems. Although an approximately 1 kb fragment upstream from the translational start codon in peanut was sufficient to confer seed-specific expression of GUS and GFP in peanut, even up to ~2.5 kb of upstream sequence was not sufficient in *Arabidopsis*. Thus, this promoter is a valuable molecular tool accessible for seed-specific expression of transgenes for genetic improvement of peanut and testing in heterologous legumes.

Acknowledgements

This work was supported by the Georgia Seed Development Commission/University of Georgia Research Foundation, the Peanut Foundation, and USDA Specific Cooperative Agreement 58-6435-6-050.

Literature Cited

Abe, H., T. Urao, T. Ito, M. Seki, K. Shinozaki, and K. Yamaguchi-Shinozaki. 2003. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. Plant Cell 15:63-78.

Boter, M., O. Ruiz-Rivero, A. Abdeen, and S. Prat. 2004. Conserved MYC transcription factors play a key role in jasmonate signaling both in tomato and *Arabidopsis*. Genes & Development 18:1577-1591.

Burow, M.D., C.E. Simpson, M.W. Faries, J.L. Starr, and A.H. Paterson. 2009. Molecular biogeographic study of recently described B- and A-genome *Arachis* species, also providing new insights into the origins of cultivated peanut. Genome 52:107-119.

Busk, P.K. and M. Pages. 1998. Regulation of abscisic acid-induced transcription. Plant Mol. Biol. 37:425-435.

Chu, Y., P. Faustinelli, M.L. Ramos, M. Hajduch, S. Stevenson, J.J. Thelen, S.J. Maleki, H.P. Cheng, and P. Ozias-Akins. 2008. Reduction of IgE binding and nonpromotion of *Aspergillus flavus* fungal growth by simultaneously silencing Ara h 2 and Ara h 6 in peanut. J. Agric. Food Chem. 56:11225-11233.

Clough, S.J. and A.F. Bent. 1998. Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* Plant J. 16:735-743.

Colot, V., L.S. Robert, T.A. Kavanagh, M.W. Bevan, and R.D. Thompson. 1987. Localization of sequences in wheat endosperm protein genes which confer tissue-specific expression in tobacco. EMBO J. 6:3559-3564.

Dare, A.P., R.J. Schaffer, K. Lin-Wang, A.C. Allan, and R.P. Hellens. 2008. Identification of a cis-regulatory element by transient analysis of co-ordinately regulated genes. Plant Methods 4:17.

Dickinson, C.D., R.P. Evans, and N.C. Nielsen. 1988. RY repeats are conserved in the 5′-flanking regions of legume seed-protein genes. Nucl. Acids Res. 16:371-371.

Ellerstrom, M., K. Stalberg, I. Ezcurra, and L. Rask. 1996. Functional dissection of a napin gene promoter: Identification of promoter elements required for embryo and endosperm-specific transcription. Plant Mol. Biol. 32:1019-1027.

Ericson, M.L., E. Muren, H.-O. Gustavsson, L.-G. Josefsson, and L. Rask. 1991. Analysis of the promoter region of napin genes from *Brassica napus* demonstrates binding of nuclear protein invitro to a conserved sequence motif. Eur. J. Bioch. 197:741-746.

Ezcurra, I., P. Wycliffe, L. Nehlin, M. Ellerstrom, and L. Rask. 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 Journal 24:57-66.

Fauteux, F. and M.V. Stromvik. 2009. Seed storage protein gene promoters contain conserved DNA motifs in Brassicaceae, Fabaceae and Poaceae. BMC Plant Biol. 9:126.

Flinterman, A.E., E. Van Hoffen, C.F. Den Hartog Jager, S. Koppelman, S.G. Pasmans, M.O. Hoekstra, C.A. Bruijnzeel-Koomen, A.C. Knulst, and E.F. Knol. 2007. Children with peanut allergy recognize predominantly Ara h 2 and Ara h 6, which remains stable over time. Clin. Exp. Allergy 37:1221-1228.

Foster, R., T. Izawa, and N.H. Chua. 1994. Plant bZIP proteins gather at ACGT elements. FASEB J. 8:192-200.

Fujiwara, T. and R.N. Beachy. 1994. Tissue-specific and temporal regulation of a beta-conglycinin gene - roles of the RY repeat and other cis-acting elements. Plant Mol. Biol. 24:261-272.

Gamborg, O.L., R.A. Miller, and K. Ojima. 1968. Nutrient requirements of suspension cultures of soybean root cells. Exp. Cell Rev. 50:151-158.

Hartmann, U., M. Sagasser, F. Mehrtens, R. Stracke, and B. Weisshaar. 2005. Differential combinatorial interactions of cis-acting elements recognized by R2R3-MYB, bZIP, and bHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. Plant Mol. Biol. 57:155-171.

Higo, K., Y. Ugawa, M. Iwamoto, and T. Korenaga. 1999. Plant cis-acting regulatory DNA elements (PLACE) database. Nucl. Acids Res. 27:297-300.

Himmelbach, A., T. Hoffmann, M. Leube, B. Hohener, and E. Grill. 2002. Homeodomain protein AthB6 is a target of the protein phosphatase ABI1 and regulates hormone responses in *Arabidopsis*. EMBO J. 21:3029-3038.

Huang, N., T.D. Sutliffe, J.C. Litts, and R.L. Rodriguez. 1990. Classification and characterization of the rice alpha-amylase multigene family. Plant Molec. Biol. 14:655-668.

Hwang, Y.S., E.E. Karrer, B.R. Thomas, L. Chen, and R.L. Rodriguez. 1998. Three cis-elements required for rice alpha-amylase Amy3D expression during sugar starvation. Plant Mol. Biol. 36:331-341.

Jefferson, R.A., T.A. Kavanagh, and M.W. Bevan. 1987. Gus fusions: β-glucusonidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. 6:3901-3907.

Kang, I.H., M. Gallo, and B.L. Tillman. 2007a. Distribution of allergen composition in peanut (*Arachis hypogaea* L.) and wild progenitor (*Arachis*) species. Crop Sci. 47:997-1003.

Kang, I.H., P. Srivastava, P. Ozias-Akins, and M. Gallo. 2007b. Temporal and spatial expression of the major allergens in developing and germinating peanut seed. Plant Physiol. 144:836-845.

Kochert, G., H.T. Stalker, M. Gimenes, L. Galgaro, C.R. Lopes, and K. Moore. 1996. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). Am. J. Bot. 83:1282-1291.

Kreis, M., M.S. Williamson, J. Forde, D. Schmutz, J. Clark, B. Buxton, J. Pywell, C. Marris, J. Henderson, N. Harris, P.R. Shewry, B.G. Forde, and B.J. Miflin. 1986. Differential gene-expression in the developing barley endosperm. Phil. Trans. Royal Soc. London Ser. B-Biol. Sci. 314:355-365.

Kuriakose, B., V. Arun, S.S. Gnanamanickam, and G. Thomas. 2009. Tissue-specific expression in transgenic rice and *Arabidopsis thaliana* plants of GUS gene driven by the 5′ regulatory sequences of an anther specific rice gene YY2. Plant Sci. 177:390-397.

Lelievre, J.M., L.O. Oliveira, and N.C. Nielsen. 1992. 5′-CATGCAT-3′ elements modulate the expression of glycinin genes. Plant Physiol. 98:387-391.

Lescot, M., P. Dehais, G. Thijs, K. Marchal, Y. Moreau, Y. Van De Peer, P. Rouze, and S. Rombauts. 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucl. Acids Res. 30:325-327.

Lessard, P.A., R.D. Allen, F. Bernier, J.D. Crispino, T. Fujiwara, and R.N. Beachy. 1991. Multiple nuclear factors interact with upstream sequences of differentially regulated beta-conglycinin genes. Plant Mol. Biol. 16:397-413.

Lopez-Molina, L. and N.H. Chua. 2000. A null mutation in a bZIP factor confers ABA-insensitivity in *Arabidopsis thaliana*. Plant Cell Physiol. 41:541-547.

Luo, H., F. Song, R.M. Goodman, and Z. Zheng. 2005. Up-regulation of OsBIHD1, a rice gene encoding BELL homeodomain transcriptional factor, in disease resistance responses. Plant Biol. 7:10.

Luo, K.M., G.F. Zhang, W. Deng, F.T. Luo, K. Qiu, and Y. Pei. 2008. Functional characterization of a cotton late embryogenesis-abundant d113 gene promoter in transgenic tobacco. Plant Cell Rep. 27:707-717.

Murashige, T. and F. Skoog. 1962. A revised medium for rapid growth and bioassays with tobacco cultures. Physiol. Plant. 15:473-497.

Nag, R., M.K. Maity, and M. Dasgupta. 2005. Dual DNA binding property of ABA insensitive 3 like factors targeted to promoters responsive to ABA and auxin. Plant Mol. Biol. 59:821-838.

Oommen, A., R.A. Dixon, and N.L. Paiva. 1994. The elicitor-inducible alfalfa isoflavone reductase promoter confers different patterns of developmental expression in homologous and heterologous transgenic plants. Plant Cell 6:1789-1803.

Ozias-Akins, P., M.L. Ramos, and Y. Chu. 2006. Hypoallergenic foods beyond infant formulas, in: S. Maleki, *et al.* (Eds.), Food allergy. ASM Press, Washington, D.C. pp. 287-–306.

Paik-Ro, O.G., J.C. Seib, and R.L. Smith. 2002. Seed-specific, developmentally regulated genes of peanut. Theor. Appl. Genet. 104:236-240.

Plant, A.L., G.J.H. van Rooijen, C.P. Anderson, and M.M. Moloney. 1994. Regulation of an *Arabidopsis* oleosin gene promoter in transgenic *Brassica napus*. Plant Mol. Biol. 25:193-205.

Potenza, C., L. Aleman, and C. Sengupta-Gopalan. 2004. Targeting transgene expression in research, agricultural, and environmental applications: Promoters used in plant transformation. In Vitro Cell. Devel. Biol.-Plant 40:1-22.

Priest, H.D., S.A. Filichkin, and T.C. Mockler. 2009. Cis-regulatory elements in plant cell signaling. Curr. Opin. Plant Biol. 12:643-649.

Qu, L.Q., Y.P. Xing, W.X. Liu, X.P. Xu, and Y.R. Song. 2008. Expression pattern and activity of six glutelin gene promoters in transgenic rice. J. Exp. Bot. 59:2417-2424.

Raina, S.N. and Y. Mukai. 1999. Genomic in situ hybridization in *Arachis* (Fabaceae) identifies the diploid wild progenitors of cultivated (*A. hypogaea*) and related wild (*A. monticola*) peanut species. Plant Syst. Evol. 214:251-262.

Ramos, M., G. Fleming, Y. Chu, Y. Akiyama, M. Gallo, and P. Ozias-Akins. 2006. Chromosomal and phylogenetic context for conglutin genes in *Arachis* based on genomic sequence. Mol. Genet. Genomics 275:578-592.

Rawat, R., Z.F. Xu, K.M. Yao, and M.L. Chye. 2005. Identification of cis-elements for ethylene and circadian regulation of the *Solanum melongena* gene encoding cysteine proteinase. Plant Mol. Biol. 57:629-643.

Ross, E.J.H., J.M. Stone, C.G. Elowsky, R. Arredondo-Peter, R.V. Klucas, and G. Sarath. 2004. Activation of the Oryza sativa non-symbiotic haemoglobin-2 promoter by the cytokinin-regulated transcription factor, ARR1. J. Exp. Bot. 55:1721-1731.

Rowley, D.L. and E.M. Herman. 1997. The upstream domain of soybean oleosin genes contains regulatory elements similar to those of legume storage proteins. Biochim. Biophys. Acta -Lipids and Lipid Metabolism 1345:1-4.

Samoylov, V.M., D.M. Tucker, and W.A. Parrott. 1998. Soybean [*Glycine max* (L.) Merill] embryogenic cultures: The role of sucrose and total nitrogen content on proliferation. In Vitro Cell. Dev. Biol. 34:8-13.

Seijo, J.G., G.I. Lavia, A. Fernandez, A. Krapovickas, D. Ducasse, and E.A. Moscone. 2004. Physical mapping of the 5s and 18s–25s rRNA genes by FISH as evidence that *Arachis duranensis* and *A. ipaensis* are the wild diploid progenitors of *A. hypogaea* (Leguminosae). Am. J. Bot. 91:1294-1303.

Simpson, S.D., K. Nakashima, Y. Narusaka, M. Seki, K. Shinozaki, and K. Yamaguchi-Shinozaki. 2003. Two different novel cis-acting elements of erd1, a clpa homologous *Arabidopsis* gene function in induction by dehydration stress and dark-induced senescence. Plant J. 33:259-270.

Singsit, C., M.J. Adang, R.E. Lynch, W.F. Anderson, A. Wang, G. Cardineau, and P. Ozias-Akins. 1997. Expression of a *Bacillus thuringiensis* CryIA(c) gene in transgenic peanut plants and its efficacy against lesser cornstalk borer. Transgen. Res. 6:169-176.

Smartt, J., W.C. Gregory, and M.P. Gregory. 1978. The genomes of *Arachis hypogaea*. 2. The implications in interspecific breeding. Euphytica 27:677-680.

Stalberg, K., M. Ellerstrom, L.G. Josefsson, and L. Rask. 1993. Deletion analysis of a 2S seed storage protein promoter of *Brassica napus* in transgenic tobacco. Plant Mol. Biol. 23:671-683.

Stalberg, K., M. Ellerstom, I. Ezcurra, S. Ablov, and L. Rask. 1996. Disruption of an overlapping E-box/ABRE motif abolished high transcription of the napA storage-protein promoter in transgenic *Brassica napus* seeds. Planta 199:515-519.

Verdier, J. and R.D. Thompson. 2008. Transcriptional regulation of storage protein synthesis during dicotyledon seed filling. Plant Cell Physiol. 49:1263-1271.

Viquez, O.M., K.N. Konan, and H.W. Dodo. 2003. Structure and organization of the genomic clone of a major peanut allergen gene, Ara h 1. Mol. Immunol. 40:565-571.

Wu, C.Y., H. Washida, Y. Onodera, K. Harada, and F. Takaiwa. 2000. Quantitative nature of the prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: Minimal cis-element requirements for endosperm-specific gene expression. Plant J. 23:415-421.

Xu, N.F., G. Hagen, and T. Guilfoyle. 1997. Multiple auxin response modules in the soybean SAUR 15A promoter. Plant Sci. 126:193-201.