GM Crops

ISSN: 1938-1999 (Print) 1938-2006 (Online) Journal homepage: https://www.tandfonline.com/loi/kgmc19

Identification of maize embryo-preferred promoters suitable for high-level heterologous protein production

Stephen J. Streatfield, Jeffrey Bray, Robert T. Love, Michael E. Horn, Jeffrey R. Lane, Carol F. Drees, Erin M. Egelkrout & John A. Howard

To cite this article: Stephen J. Streatfield, Jeffrey Bray, Robert T. Love, Michael E. Horn, Jeffrey R. Lane, Carol F. Drees, Erin M. Egelkrout & John A. Howard (2010) Identification of maize embryo-preferred promoters suitable for high-level heterologous protein production, GM Crops, 1:3, 162-172

To link to this article: https://doi.org/10.4161/gmcr.1.3.12816

Copyright © 2010 Landes Bioscience

Published online: 01 May 2010.

Submit your article to this journal

Article views: 450

View related articles

Citing articles: 13 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=kgmc20

GM Crops 1:3, 162-172; May/June 2010 © 2010 Landes Bioscience

Identification of maize embryo-preferred promoters suitable for high-level heterologous protein production

Stephen J. Streatfield,[1] Jeffrey Bray,[1] Robert T. Love,[1] Michael E. Horn,[1] Jeffrey R. Lane,[1] Carol F. Drees,[1] Erin M. Egelkrout[2] and John A. Howard[2,*]

[1]Formerly of ProdiGene Inc.; College Station, TX USA; [2]Applied Biotechnology Institute; California Polytechnic Institute; San Luis Obispo, CA USA

Key words: molecular pharming, protein accumulation, seed expression, promoters, tissue specific promoters, transgene expression, maize

The production of heterologous proteins in plants at levels consistent with commercialization of protein products requires molecular tools to ensure high-level transgene expression. The identification of strong promoters, preferably specific to the target expression tissue, is a focus for improving foreign protein yields using transgenic cereals as a production system. Thus, there is a requirement for strong embryo preferred monocot promoters. We obtained the sequences of 500 randomly selected maize cDNA clones to determine gene expression profiles in embryo tissues at multiple stages during development. Promoters corresponding to the most abundant clones were identified and isolated. These promoters were fused to the β-glucuronidase reporter and their tissue specificity and developmental expression characteristics assessed in transgenic maize. All of the isolated promoters tested drove transgene expression predominantly in the embryo and were most active late in embryogenesis during storage protein deposition. One of the most active promoters assessed by transgene expression was associated with the *globulin-1* protein. Sequence identified here extended approximately 1.6 kb distal to the previously identified extent of the *globulin-1* promoter, and this additional sequence boosted expression over two-fold. The extended *globulin-1* promoter sequence isolated in this study has the potential for driving transgene expression at higher levels than those previously reported for cereals. Also, other highly active embryo promoters identified here offer opportunities to express multiple foreign proteins simultaneously at high levels in embryo tissues, while avoiding concerns over gene silencing due to the repeated use of a single promoter.

Introduction

Promoters are vital molecular tools that are utilized widely in plant biotechnology to control the expression of introduced genes. The applications for promoters in driving gene expression in plant tissues include the synthesis of scoreable and selectable markers to identify transgenic plants[1,2] and the overexpression of control point enzymes to modify metabolic flux through key pathways, thereby affecting the yield of important plant products.[3,4] Other uses of plant promoters include the expression of genes conferring resistance to pests, thus conferring protection,[5] and the expression of non-native enzymes to facilitate the production of foreign metabolites in particular plant species.[6,7] A further application of plant promoters is to overexpress controlling regulatory genes affecting aspects of plant physiology such as flowering time and so modify plant growth characteristics.[8] Promoters are also used to repress the expression of specific genes by driving the synthesis of interfering RNA species,[9] thus affecting plant metabolic and developmental pathways.[10] Although high levels of expression may not be necessary for all of the above applications, there is clearly a need for promoters showing activity in specific plant tissues.

Apart from these and other applications of promoters to modify plant traits, promoters are also required for plants to function as production systems for heterologous proteins. Plants have been used to produce a wide range of recombinant proteins of potential economic and/or medicinal importance. These include research chemicals,[11,12] processing enzymes that are used, for example, in the pharmaceutical industry,[13] industrial enzymes that are deployed in large-scale processing operations such as bleaching,[14,15] candidate vaccine antigens for animal or plant disease prevention[16-19] and therapeutic pharmaceuticals including antibodies.[20,21] The expressed proteins may either be purified from plant tissues[11,13] or, if as with orally administered vaccines the final application allows, the recombinant plant material may be processed into a suitable form for use.[22,23] For these and other protein products to be produced in plant systems it is necessary that promoters drive sufficiently high levels of expression to ensure commercial viability.

*Correspondence to: John A. Howard; Email: jhoward@appliedbiotech.org
Submitted: 06/03/10; Revised: 07/08/10; Accepted: 07/09/10
Previously published online: www.landesbioscience.com/journals/gmcrops/article/12963

Spatial and temporal control is also often important in driving gene expression in plants. For example, selectable and scoreable markers must be expressed at a suitable time in appropriate tissues to allow for screening while controlling enzymes and regulatory factors must be produced in metabolically active and physiologically responsive tissues, respectively. Similarly, genes conferring host protection must be expressed in the target tissues for the pathogen or pest, and plant produced protein products should be expressed in tissues suitable for protein accumulation and storage. Furthermore, since certain protein products may have detrimental effects on plant health and yield when expressed in metabolically active plant tissues that are essential for survival and growth, promoters may be favored that are active in plant storage tissues but show low or no activity in other, non-storage tissues.

Several promoters of plant and plant pathogen (bacterial and viral) origin have been used to direct transgene expression in plants. Prominent examples include the French bean β-*phaseolin* promoter,[24] the mannopine synthase promoter of *Agrobacterium tumefaciens*[25] and the *35S* promoter of cauliflower mosaic virus.[26] These and several other promoters in widespread use in plants were originally developed and utilized in dicot species. However, the cereals comprise particularly important crops and there is therefore a pressing need for promoters that have high activity and/or tissue specificity in monocots. The *globulin-1* promoter has been characterized and previously used in transgenic maize seeds.[27-30] Since the nutritional value of cereals is in their seeds that are also well suited for recombinant protein accumulation, storage promoters besides *globulin-1* that are also active in cereal seed tissues are especially useful.

In the case of expression in maize seed, two broad classes of promoters have been deployed: constitutive and seed preferred. Constitutive promoters, such as maize *polyubiquitin-1* drive expression in the seed but also in other tissues.[31] A drawback with such constitutive promoters is that expression in tissues other than seed storage tissues may result in plant health being compromised, for example if an enzymatic protein is expressed in metabolically active tissues required for germination or growth.[15] Furthermore, constitutive expression may result in the expressed foreign protein being synthesized in pollen grains and thus being difficult to bio-contain. By contrast, seed preferred promoters limit all or the bulk of transgene expression to seed tissues, so avoiding such concerns.

The principal tissue types in maize seeds are the embryo, the endosperm including a surrounding aleurone cell layer and the maternally derived pericarp. Of these, the endosperm and to a lesser extent the embryo, comprise most of the volume of the seed. Thus, endosperm and embryo promoters are particularly important for modifying seed characteristics and content. The relatively high soluble protein content of the embryo makes it a particularly good target for foreign protein expression. The proximal 1.1 kb of the maize *27 kD γ-zein* promoter[32] and the proximal 1.4 kb of the maize *globulin-1* promoter (GenBank accession L22344)[33] are prominent examples of seed preferred promoters that have been used to express transgenes in endosperm and embryo tissues, respectively, of monocots. However, despite these examples, there is currently a very limited repertoire of promoters for expressing foreign proteins in seed tissues of cereals. There is a need for further promoters that express transgenes at similar or higher levels to those currently deployed and with similar or improved tissue specificity. Such promoters may facilitate the expression of foreign proteins in seeds at higher levels than are currently achieved, while restricting expression specifically or predominantly to seed tissues. Also, a range of new promoters would allow the expression of multiple copies of a single transgene in seeds without the need to repeatedly use the same promoter. This should reduce silencing phenomena associated with promoter methylation,[34] and thereby should also serve to boost expression. Similarly, multiple distinct transgenes could be simultaneously expressed in seed tissues, allowing more complex traits and foreign protein products to be introduced into cereal seeds. Embryo specific promoters may be particularly useful for modifying seed oil content and amino acid composition.

Here we describe a genomics approach to identify further sequences that can drive high levels of transgene expression specifically in maize embryo tissues, targeting at least 1% of soluble embryo protein. We also describe the cloning of 3 kb of proximal promoter sequence of five maize genes: *globulin-1*, *globulin-2*, *16 kD oleosin*, an *abscisic acid inducible* like (*ABA-inducible*) gene and a *novel (a)* peptide gene. The *globulin-1* clone extends by approximately 1.6 kb the previously isolated proximal 1.4 kb of the maize *globulin-1* promoter, which has previously been used to express transgenes in maize seeds.[13,15] We demonstrate that the 3 kb extent of the maize *globulin-1* promoter cloned here has significantly increased activity over the previously cloned 1.4 kb region, while retaining a highly embryo preferred expression pattern. Three of the other isolated promoters also show high levels of transgene expression specific to the embryo in maize seeds.

©2010 Landes Bioscience.
Do not distribute.

Results

Identification of highly represented cDNA clones in maize embryo libraries. In order to identify genes with promoters capable of driving high levels of heterologous gene expression in maize embryos, highly represented cDNAs were identified in libraries prepared from this target tissue. Five independent cDNA libraries were prepared from embryo tissues dissected out of developing seeds at 10 to 12, 19, 28, 37 or 46 days post-pollination, the last time point corresponding to seed maturity, just prior to harvest. These libraries provided a source of representative clones from each of five time points during seed development. Each cDNA library was prepared from a pool of embryo tissues supplied equally from four distinct maize lines: a representative Lancaster line (SN103), a representative Stiff Stalk line (DA502), a high protein line (ILHP90) and a high oil line (G30B3). These four lines were chosen to yield information on highly expressed genes that should be valid for more than one variety of maize. Lancaster and Stiff Stalk lines are the most commonly used classes of lines used in field production of grain, and are hence relevant to large scale commercial production of heterologous proteins. Also, high protein and high oil lines are potentially promising germplasm

Table 1A. Number of transcript sequences per 100 sequences indicating highly expressed genes (representing >0.5% of total transcripts) obtained by sequencing cDNA libraries prepared from embryos at five distinct time points post-pollination

Sequence identity	10–12 dpp[1]	19 dpp	28 dpp	37 dpp	46 dpp	Number and % of total transcripts
ABA-inducible sequence	0	0	7	2	3	12 (2.4%)
Globulin-2	0	1	5	2	3	11 (2.2%)
16 kD oleosin	2	3	0	4	0	9 (1.8%)
18 kD oleosin	0	2	2	3	0	7 (1.4%)
Barley per1 like regulator sequence	0	1	4	2	0	7 (1.4%)
Globulin-1	0	1	1	3	2	7 (1.4%)
Histone H2B	4	1	0	0	0	5 (1%)
Glycine rich protein	1	0	1	2	1	5 (1%)
Dehydrin	0	0	0	1	4	5 (1%)
Histone H3C3	3	1	0	0	0	4 (0.8%)
ABA inducible glycine rich	0	2	0	1	1	4 (0.8%)
17 kD oleosin	0	1	0	3	0	4 (0.8%)
Novel (a)	0	0	1	3	0	4 (0.8%)
Cytochrome P450	3	0	0	0	0	3 (0.6%)
Histone H4C14	2	1	0	0	0	3 (0.6%)
Anaerobically regulated fructose bisphosphate aldolase	1	1	1	0	0	3 (0.6%)
Gibberellin 2-oxidase/2 β-hydroxylase	0	1	2	0	0	3 (0.6%)
ADP-ribosylation factor	0	1	1	1	0	3 (0.6%)
Novel (b)	0	1	1	1	0	3 (0.6%)
Novel (c)	0	1	1	0	1	3 (0.6%)
Gamma thionin	0	0	1	1	1	3 (0.6%)
	101	101	110	118	115	545

[1]Days post-pollination.

for heterologous protein production, with high oil lines having large embryos and thus favoring protein accumulation in this tissue.

For each of the five libraries, nucleotide sequences of at least 100 randomly selected cDNA clones were determined. These sequences were assessed to give a rough overview of the gene expression profile of developing maize embryos and to identify the most highly represented sequences corresponding to the most strongly expressed genes (**Table 1A**). Sequence identification information and GenBank accession numbers of selected clones are provided in **Table 1B.** Individual cDNAs representing over 2% of the total sequenced clones for a particular library were considered as highly represented. On this basis, at 10 to 12 days post-pollination sequences encoding two classes of histone and a cytochrome P450 were highly represented, while at 19 days post-pollination sequence encoding *16 kD oleosin* predominated. At 28 days post-pollination, sequence with similarity to an *abscisic acid inducible* gene (*ABA-inducible*), *globulin-2* sequence and sequence encoding a protein with similarity to a per1 like regulator were most highly represented, with the *ABA-inducible* sequence alone accounting for over 6% of the clones. By 37 days post-pollination, sequences encoding 16 kD, 17 kD and *18 kD oleosin*, *globulin-1* and a novel peptide [*novel (a)*] predominated, and finally at 46 days post-pollination, just before harvest, sequences encoding a *dehydrin*, *globulin-2* and the *ABA-inducible* sequence were most highly represented.

Table 1B. GenBank accession and ABI identification numbers

Sequence identity	Identification number	GenBank accession number
ABA-inducible sequence	Pr26	EA076965.1
Globulin-2	Pr23	AR947679.1
16 kD oleosin	Pr29	U13701.1
Globulin-1	Pr36	EA061987.1
Novel (a)	Pr36	HM6308

Considering the frequency of specific sequences in the five libraries taken together (**Table 1A**, last column), two distinct cDNAs were represented at over 2% of the total sequences, with a further four cDNAs at over 1%, and fifteen more cDNAs at over 0.5%. The most highly represented sequence was the *ABA-inducible* sequence. These observations were very encouraging for identifying promoters capable of expressing foreign proteins at high levels in embryo tissues.

Analysis of isolated cDNA clones. Sequences identified as representing over 0.5% of the total clones and/or over 2% of the clones in one post-pollination embryo library were scanned against maize databases. This allowed an initial screen to identify

Table 2. Representation (%) of selected clones from four cDNA libraries when screened against pooled phage from these libraries

Sequence	% representation
globulin-2	1.4
globulin-1	1.3
16 kD oleosin	0.7
ABA inducible	0.6
Novel (a) protein	0.6
17 kD oleosin	0.5
18 kD oleosin	0.4

Table 3. Clone representation in the maize genome determined by DNA hybridization

Sequence	Bands with EcoRI digest	Bands with HindIII digest
globulin-2	3	3
globulin-1	4	3
16 kD oleosin	3	3
ABA-inducible	5	4
novel (a) protein	8	6
17 kD oleosin	7	9
18 kD oleosin	8	6

sequences corresponding to expressed sequence tags (ESTs) isolated from cDNA libraries prepared from specific tissues. On this basis only five sequences, those encoding the embryonic *ABA-inducible*, *globulin-2*, two novel peptides including *novel (a)* peptide, and *17 kD oleosin* were identified as being embryo specific, with a further three sequences, those encoding *globulin-1*, *18 kD oleosin* and *16 kD oleosin*, being identified as highly embryo preferred, with in the case of *globulin-1* and *18 kD oleosin* the only other recorded expression being in endosperm tissue.

Since the nucleotide sequences of only a little over 500 clones were determined in order to identify highly expressed genes in embryo tissues, a hybridization screening approach was then followed to confirm high level expression for seven of the eight preferred sequences (*ABA-inducible*, *goblulin-1*, *globulin-2*, *16 kD oleosin*, *17 kD oleosin*, *18 kD oleosin* and *novel (a)* peptide), with the lowest represented, a second novel sequence, not pursued further. To provide a greater level of confidence, a sequence of each selected clone was screened against 20,000 to 40,000 phage plaques. The phage were derived from an equally represented combination of all five post-pollination time points for the embryo libraries that were themselves derived from all four lines of maize. Since the plaque hybridization has a 40 to 80-fold greater level of sampling than the cDNA sequencing approach, it should more accurately reflect expression across all the libraries. In general, sequence representation determined by plaque hybridization was lower than that estimated by nucleotide sequence analysis (**Table 2**). This was most apparent for the *ABA-inducible* sequence, which was four-fold more highly represented among sequenced clones than hybridizing phage. By contrast, *globulin-1* sequence representation was very similar between the two approaches.

A complication with identifying highly expressed genes on the basis of partial DNA sequence analysis, and especially with confirming high level expression by phage hybridization, is that very similar sequences derived from distinct genomic loci may be misinterpreted as representing a single sequence. This would result in an overestimation of a particular clone's representation. To determine whether this could be a serious limitation with sequences identified here as being highly expressed in embryos, each selected sequence was assessed for approximate copy number in the maize genome by high stringency DNA hybridization analysis of Missouri-13 genomic DNA. To simplify hybridization patterns, fairly short sequences lacking sites for restriction enzymes used to digest the genomic DNA were used as probes, although these sites could be present in intervening introns. Hybridization conducted with EcoRI or HindIII digested genomic DNA gave similar results, with the banding pattern being considered to represent the number of similar clones in each case. From three to seven similar sequences were identified for the different clones (**Table 3**). Thus, the frequency of these or similar sequences in the genome is fairly low, particularly since some of these hybridizing sequences may represent pseudogenes. However, the DNA hybridization data do indicate that the phage hybridization results may arise from multiple similar sequences, and thus may overestimate expression for single genes.

To provide an initial assessment of specificity of expression of the seven selected sequences within seed, RNA was prepared from 28 day post-pollination embryos isolated from each of the four maize lines that were used to prepare the embryo cDNA libraries, and also from 28 day post-pollination embryos and endosperm isolated from Missouri-13, a representative of a highly used maize line. Since sampling across all time points was prohibitive, 28 days post-pollination was selected as a central time point during embryogenesis in maize. The selected clones (*ABA-inducible*, *globulin-1*, *globulin-2*, *novel (a)*, *16 kD oleosin*, *17 kD oleosin*, *18 kD oleosin*) were used as probes for hybridization to these RNA preparations. For all seven selected clones (partial data shown), strong signals were observed with RNA prepared from embryos of each of the five lines, with the exception that no message was observed for *globulin-1* sequence in ILHP90, which is a known *globulin-1* null mutant line (**Fig. 1**). By contrast no signal was observed with RNA prepared from Missouri-13 endosperm tissue when *ABA-inducible* sequence, *globulin-1*, *globulin-2* or *18 kD oleosin* were used as probes, and only very weak signals were observed for *16 kD oleosin*, *17 kD oleosin* and *novel (a)* sequence (data not shown).

Expression of the seven selected sequences in non-seed tissues was also assessed by RNA hybridization analysis using RNA prepared from tissues pooled from the four maize lines used to prepare the developing embryo cDNA libraries. Expression was assessed in leaf, stem, root, tassel, anther, pollen, husk, silk, immature ear and cob tissues of maize plants. *Globulin-1* probe gave very weak signals with stem, tassel and silk tissues, and *globulin-2* probe gave a weak signal with stem tissues only. None of the other five probes gave a signal with any of the RNA samples (data not shown). Thus, on the basis of RNA hybridization analysis, all seven selected sequences appear to be seed specific or highly seed

©2009 Landes Bioscience.
Do not distribute.

preferred, and within the seed all of the selected sequences are embryo specific or embryo preferred.

Isolation and activity of genomic sequences 5' to the selected embryo expressed cDNAs. In order to isolate the most promising of the identified promoters, genomic sequences located upstream of the *ABA-inducible* sequence, *novel (a)*, *globulin-1*, *globulin-2* and *16 kD oleosin* were cloned by a combination of plaque hybridization and polymerase chain reaction based gene-walking approaches. A library of genomic sequences prepared from leaf tissue of Missouri-13 provided a template. DNA sequence extending approximately 3 kb upstream of the predicted translation start codon was cloned and nucleotide sequence determined for each of the five genes selected on the basis of clone representation in the cDNA libraries and phage hybridization.

Expression constructs were prepared by fusing each of the cloned approximately 3 kb regions of 5' sequences to sequence encoding the β-glucuronidase (GUS) reporter gene (*uidA*). As a comparative positive control the 1.4 kb extent of previously identified maize globulin-1 5' sequence was also fused to *uidA*. For all constructs, the potato *protease inhibitor II* (*PinII*) terminator sequence was introduced downstream of *uidA* coding sequence. In addition the vectors carried the *phosphinothricin Phosphinothricin* gene (*pat*) of *Streptomyces viridochromogenes* to confer herbicide resistance for selection of transgenic plants. Expression of the *pat* marker was controlled by the cauliflower mosaic virus *35S* promoter and terminator sequences. The vectors also contained border sequences flanking these transcription units to allow the transformation of vector DNA enclosed within the borders into the target plant's genome.

Agrobacterium strains containing the expression cassettes were assembled as described previously[15] and used to initiate plant transformation of a Hi-II/elite inbred line to generate transgenic events. For each reporter construct, plants were regenerated from callus, grown to maturity and fertilized using pollen from the elite inbred line. From 9 to 22 independent transformation events were isolated for each construct, and between 2 and 13 plants were regenerated for each transformation event. Soluble protein was extracted from seed harvested from each plant and GUS activity was determined. For each plant the assay was conducted on six randomly selected seeds, and for each construct, mean expression across all transformation events was determined as a percentage of total soluble protein (**Table 4**). All plants tested were T_1 seed that had been crossed to an elite non-transgenic line and showed the expected segregation pattern. Only the positive seed were used for statistical analysis between various constructs.

The previously isolated *globulin-1* sequence has been used to express high levels of up to approximately 15% of total soluble protein for foreign proteins in the seed of transgenic maize[15,39] dependent on the sub-cellular targeting site and the specific protein. In this case, all constructs encoding GUS were targeted to the cytoplasm and expected expression levels are low in comparison to some reports for other proteins, but consistent with *uidA* expression in other reports. The analysis performed here on individual seed gave 0.454% of total soluble protein as the highest level of GUS observed with any of the 5' sequences, and this was recorded with the *novel (a)* promoter sequence. For comparison purposes, mean values are given for plants derived from multiple transgenic events (**Table 4**), and expression of transgenes typically varies greatly between transgenic events, for example,[40] and it is not surprising that standard deviations are relatively large.

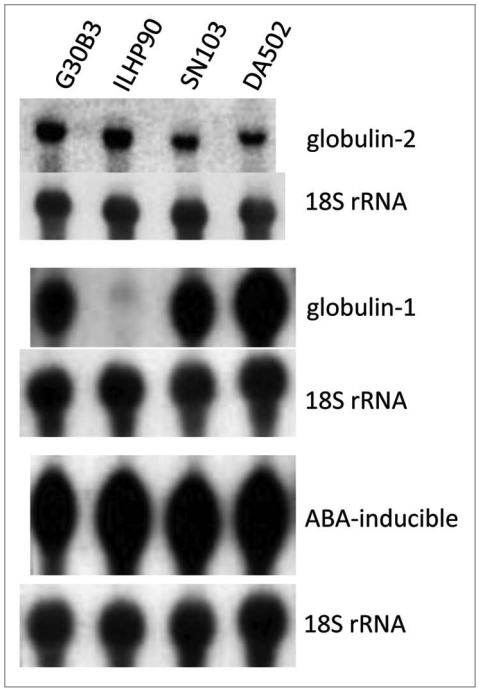

Figure 1. RNA hybridization analysis of selected highly expressed sequences globulin-2, globulin-1, ABA-inducible in embryo tissues of four maize lines, with each blot subsequently probed with 18S rRNA sequence to demonstrate equivalent sample loading between maize lines.

©2010 Landes Bioscience. Do not distribute

An analysis of variance does allow 5' sequences to be grouped according to activity. On the basis of single seed analysis, the 3 kb 5' *globulin-1* sequence and the *novel (a)* sequence appear to be stronger promoters than the 1.4 kb *globulin-1* positive control, whereas 5' sequences of the *ABA-inducible* gene and *globulin-2* have a similar activity to the control, although the *ABA-inducible* sequence is more active than *globulin-2*. The *16 kD oleosin* sequence appears to be weaker than the control.

Tissue specificity of cloned 5' sequences fused to a reporter gene. The tissue specificity of expression driven by the cloned 5' sequences was assessed across a range of non-seed plant tissues: leaf at 21 days post-germination and at 12 days post-pollination, stem, root and silk at 12 days post-pollination, husk and cob at 19 days post-pollination, and pollen and anther at pollen shed. For each construct, analysis was performed on three lines derived from separate transformation events that showed high level GUS expression in T_1 seed. The 5' sequences cloned here for the *globulin-1*, *globulin-2*, *16 kD oleosin*, *novel (a)* and *ABA-inducible* genes did not drive GUS expression in any of the non-seed tissues tested except for the cob. The *globulin-1* positive control gave the same results (data not shown).

Table 4. Mean expression of GUS for each reporter construct, allowing for 3 unequal promoter variance groupings

Sequence	Single seed mean[1] (%TSP[3])	Single seed s.d.[2] (%TSP[3])	Single seed sample size	Single seed group[4]
novel (a) (PMR)	0.062	0.047	11	A
globulin-1 (PMH)	0.049	0.031	18	AB
ABA-inducible (PMJ)	0.023	0.010	13	BC
globulin-1 control (PME)	0.020	0.013	7	CD
globulin-2 (PMF)	0.008	0.009	22	DE
16 kD oleosin (PMM)	0.006	0.004	9	E

[1]Mean expression for each construct calculated from means of expression for each event, themselves calculated from means of expression for positive plants only regenerated from each event and derived from positive seed data only. [2]Standard deviation. [3]Total soluble protein. [4]Analysis of variance. (Tukey-Kramer method at 0.05 experiment-wise error rate).

Table 5. Temporal expression of GUS within the developing embryo driven by cloned 5' sequences

Sequence	12 dpp[1]	19 dpp[1]	28 dpp[1]	37 dpp[1]	21 dph[2]
globulin-2	+/-	+	+	+	+
globulin-1	+	+	+	+	+
16 kD oleosin	-	+	+/-	+	+
ABA-inducible	+/-	+	+	+	+
novel (a)	+	+	+	+	+
globulin-1 control	+	+	+	+	+

Expression evidenced by blue staining across all three tested lines is indicated by +, whereas no expression in any line is indicated by -, and expression in one or two of the three lines is indicated by +/-. [1]Days post-pollination. [2]Days post-harvest.

©Landes Bioscience. Do not distribute.

An assessment of temporal expression in embryo tissue was assessed in developing T_2 seed for the same three transgenic lines for each construct (**Table 5 and Fig. 2**). Since GUS is a stable reporter, detection of GUS activity in later stages of development may result from expression at earlier stages. However, GUS staining does provide an indication of the onset of expression in developing tissues. Sequences isolated here and located 5' of *globulin-1*, *globulin-2*, *novel (a)* and the *ABA-inducible* gene, and sequence located 5' of the previously cloned *globulin-1* control, all appear to drive embryo expression from at least as early as 12 days post-pollination onwards, whereas sequences located 5' of *16 kD oleosin* only start to drive expression between 12 and 19 days post-pollination, and expression in mature seed is predominantly in aleurone tissue rather than embryo tissue. Within the embryo, the *globulin-1*, *globulin-2*, *novel (a)* and *ABA-inducible* promoters all drive expression predominantly in scutellum tissue. For each of these promoter sequences, expression was also assessed in endosperm tissue at 21 days post-harvest. No GUS activity was observed in endosperm tissue of any line, except for lines in which the *16 kD oleosin* sequence promotes GUS activity in the aleurone layer (data not shown).

Activity of the proximal 3 kb of *globulin-1* promoter sequence relative to the 1.4 kb sequence confirmed with a second reporter. In order to assess whether the 3 kb extent of the *globulin-1* promoter and leader would also show increased activity over the 1.4 kb extent of sequence with a second reporter, transcription units were made in which approximately 3 kb or 1.4 kb of sequence 5' to the translation start codon of *globulin-1* were fused to DNA encoding the *Bos taurus trypsinogen* coding sequence. DNA encoding the barley *alpha-amylase* signal sequence was also included between the *globulin-1* leader and *trypsinogen* coding sequence. This sequence has been shown to target the protein product to the cell surface.[15,41] As for the *uidA* fusions, the *PinII* terminator sequence was present downstream of the *trypsinogen* reporter, and the transcription units were included in a plant transformation vector. The promoter-reporter fusions were stably introduced into the maize genome by Agrobacterium mediated transformation and plants were regenerated from transformation events obtained using each vector. A total of 146 plants were regenerated from twenty one independent transformation events obtained using the 3 kb promoter-*trypsinogen* fusion, and 140 plants were regenerated from nineteen independent transformation events using the 1.4 kb promoter fusion. Seed was harvested, soluble protein was extracted and for each plant the level of trypsin was determined in each of six randomly selected seeds.

Expression data for the transgenic seed are summarized in **Table 6**. As with the *uidA* reporter, the promixal 3 kb of maize *globulin-1* promoter sequence drives a significantly higher level of *trypsinogen* reporter gene expression than the proximal 1.4 kb of promoter sequence. Analysis of variance demonstrates that the 3 kb promoter and the 1.4 kb promoter give distinct sets of expression data, with the 3 kb promoter being the more active in driving *trypsinogen* gene expression.

As a guide to the expression potential of each extent of the globulin-1 promoter with *trypsinogen*, the highest expressing single seed was noted for each promoter. With the 3 kb extent of promoter the highest recorded expression level was in excess of 5% of total soluble protein, whereas with the 1.4 kb extent of promoter the highest recorded expression level was 1.32% of total soluble protein (data not shown). Compared to the average seed expression levels of 0.34% for the 3 kb extent and 0.20% for the 1.4 kb extent of promoters (**Table 6**), these levels indicate considerable potential for increasing expression through selection. Very similar numbers of plants and transformation events were assessed for each promoter. Thus, the 3 kb extent of the *globulin-1* promoter appears to have a greater expression potential than the 1.4 kb extent of promoter.

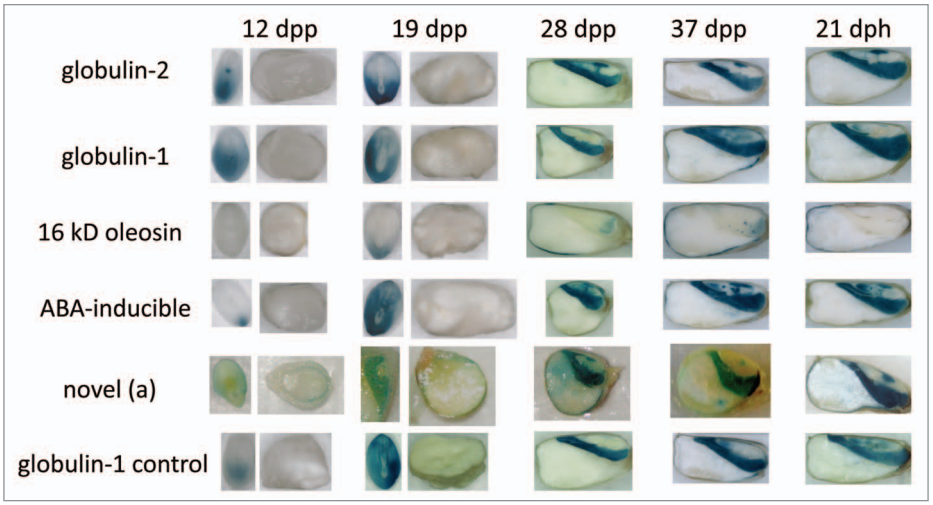

Figure 2. Accumulation of the β-glucuronidase reporter in embryos and endosperm when driven by selected upstream sequences of genes highly expressed in embryo tissues. For 12 and 19 days post-pollination, embryo and endosperm tissue are shown separately, with embryos positioned to the left of endosperm tissue. dpp, days post-pollination; dph, days post-harvest.

Discussion

Promoters that are active in plant tissues are vital tools in implementing a range of strategies to engineer plant characteristics. However, overexpression of transgenes throughout the plant can have undesired effects and consequences. Tissue specific promoters are important for restricting the expression of selected transgenes to particular parts of the plant, thereby eliminating deleterious effects that might arise from constitutive expression. Promoters well suited to expressing transgenes specifically in target tissues are most obviously identified as those that drive expression of native genes in those tissues. In the case of cereals, including maize, seed tissues are of particular interest for crop improvement and for acting as a repository for protein accumulation. Thus, promoters that are active in seed tissues are of considerable value for crop development and for innovations pertaining to seeds.

Cereal seeds consist principally of two tissue types, the embryonic plant or embryo and the surrounding endosperm. The endosperm is comprised almost entirely of nutritional reserves, primarily of complex carbohydrate and aqueous insoluble protein, but the embryo also contains considerable stores, mainly of oils and aqueous soluble proteins. Globulin-1 is one of the most abundant proteins in maize embryo tissue. It is largely limited to this tissue and becomes particularly concentrated in the scutellum late in embryo development. Given the high concentration of this protein observed in embryo tissues the maize, the globulin-1 promoter was previously identified as being a good candidate to direct high levels of transgene expression in the embryo. An approximately 1.4 kb extent of this promoter has been cloned (Genbank accession L22344)[33] and used to drive high levels of transgene expression specifically in maize seed.[13,15] However, still more active promoters are very desirable for some applications, such as the expression of highly cost sensitive foreign proteins in cereal seeds. Here we report the isolation of a number of highly active embryo-specific promoters by sampling cDNA libraries representing embryo tissues from diverse lines at different stages of development.

Table 6. Mean expression of trypsinogen for each reporter construct

Extent of 5' *globulin-1* sequence (kb)	Single seed mean[1] (%TSP[3])	Single seed s.d.[2] (%TSP[3])	Single seed group[4]
3	0.34	0.15	A
1.4	0.20	0.10	B

[1]Mean expression for each construct calculated from means of expression for each event, themselves calculated from means of expression for all plants regenerated from each event and derived from positive seed data only. [2]Standard deviation. [3]Total soluble protein. [4]Analysis of variance. (Duncan's multiple range test at 95% confidence).

©2010 Landes Bioscience. Do not distribute

One of the promoters isolated was that of *globulin-1*, confirming that it is one of the most highly expressed genes in maize embryos. *Globulin-1* sequences appeared to comprise over 1% of the total clones, a proportion confirmed by a nucleic acid hybridization approach. Furthermore, expression appeared specific to embryo tissues. Promoters from a number of other genes including *globulin-2*, a *16 kD oleosin* and two novel sequences including *novel (a)*, were also isolated. The identification of promoters driving *globulin-2* and *globulin-2* is perhaps not surprising in that they are highly expressed storage proteins. The function of the other two proteins is not known.

Each of the promoters was examined by RNA hybridization for tissue specificity and their expression was found to be largely seed specific and embryo-preferred. The 3 kb *globulin-1* promoter cloned here appears to have a similar high specificity to the previously cloned 1.4 kb extent of *globulin-1* promoter, with expression being seed specific apart from some minor activity in the cob, though even this may represent a staining artifact of seed activity. The 3 kb and 1.4 kb extents of the *globulin-1* promoter also have similar specificities within seed tissues during development. Activity is evident in the embryo as early as 12 days after

pollination and appears increased by 19 days after pollination. However, based on the abundance of cDNAs in the developmental seed libraries, the native full length *globulin-1* promoter appears not to be as active in young developing embryo tissue as in maturing tissue, at least relative to other promoters. Both extents of the *globulin-1* promoter appear to continue to be active throughout seed development, but with *uidA* as the reporter gene the presence of GUS activity in late stage embryos may reflect upon previously synthesized protein rather than active transcription and translation. During the later stages of seed development promoter activity within the embryo is strongest in the scutellum and coleoptile but is also evident in plumule and radicle tissues. Expression is also observed in the aleurone, and at a very low level in the endosperm.

Expression of the *globulin-2*, *ABA-inducible* and *novel (a)* promoters is also apparent as early as 12 days after pollination, with a similar expression pattern at later time points to the *globulin-1* promoter. Expression of the *16 kD oleosin* promoter is delayed until 19 days and is generally much lower than for the other promoters at later time points.

From genomic clones extending upstream of the translation start codon, approximately 3 kb of each promoter plus untranslated leader was isolated and the DNA sequence determined. Each promoter was fused to the *uidA* reporter gene and transformed back into maize. The promoters showed mean GUS expression levels of 0.006% tsp to 0.062% tsp (**Table 4**). As we had previously worked with *globulin-1* promoter, we were particularly interested to find that the 3 kb *globulin-1* promoter sequence appears over twice as active as the previously cloned 1.4 kb extent of the *globulin-1* promoter. Also, when fused to the *Bos taurus trypsinogen* reporter gene, the 3 kb promoter is again more active than the 1.4 kb promoter. While the GUS protein was synthesized in the cytoplasm, bovine *trypsinogen* was targeted by the presence of the barley alpha amylase signal to the cell surface. Targeting to the cell surface is a strategy that is often used to boost expression of particular proteins in plants.[15,42] Thus, the 3 kb extent of the *globulin-1* promoter boosts expression over the 1.4 kb promoter with both of the reporters tested in two different intercellular locations, and results in increased protein levels whether reporter proteins remain in the cytoplasm or are targeted to the cell surface.

In addition to the 3 kb *globulin-1* promoter, promoters with homology to gene sequences from *globulin-2*, *16 kD oleosin*, an *ABA-inducible* gene and a novel gene also showed strong embryo preferred expression (**Table 4**). These were all confirmed by using the GUS reporter gene in transgenic tissue. Several of these new promoter sequences are likely to be excellent choices for seed specific expression in maize, and likely in other cereals. The increased activity of the 3 kb extent of the *globulin-1* promoter over the 1.4 kb promoter, makes it an especially promising candidate along with the *novel (a)* promoter, which also shows very high seed expression. Higher levels of transgenes should be achievable without having to resort to using multiple copies of the same promoter, a strategy which can lead to transcriptional gene silencing, and thus muted expression.

This new group of promoters provides a battery of strong embryo preferred promoters that should be of use in the overexpression of proteins in germ tissue. These may be useful when trying to overexpress several different proteins in the same seed by using different promoters for different proteins. This may also be used to provide higher expression of a single protein by making multiple cassettes with different promoters upstream of the target gene coding sequence. In either case, the different promoters should decrease the potential for gene silencing or recombination, and therefore may result in higher expression.

Materials and Methods

Construction of cDNA libraries. *Zea mays* plants representing four lines, SN103, DA502, ILHP90 and G30B3, were grown in soil from seed and were self-pollinated. Developing or mature seed were harvested from each line at 10 to 12, 19, 28, 37 and 46 days post-pollination. Embryo tissue was separated from other seed tissues, and equal weights of embryos for each of the four lines were mixed together for each time point, and these pooled line embryos were snap frozen with liquid nitrogen, for each of the five time points cDNA libraries were constructed. Total RNA was isolated from the pooled embryo tissues using a phenol-based method,[35] and poly-A message was then prepared from this RNA using Poly(A) Quik mRNA isolation columns (Stratagene; La Jolla, CA). The poly-A RNA samples were used to prepare cDNA libraries, each representative of all four maize lines and each corresponding to a different time point of embryo development. The libraries were constructed in the lambda ZAP II vector (Stratagene; La Jolla, CA).

©2010 Landes Bioscience.
Do not distribute.

DNA sequence analysis of cDNA clones. Phagemids were excised from the phage vector for each of the five libraries representing five developmental time points post-pollination. Approximately 100 clones were randomly selected to represent each library and the nucleotide sequences of the cDNA inserts were determined at the 'DNA Sequencing and Synthesis Facility' at Iowa State University (Ames, IA). Nucleotide sequences of clones were compared using the 'Sequencher' package (Gene Codes Corporation; Ann Arbor, MI).

Plaque hybridization. Equal aliquots of each of the five embryo developmental time point cDNA libraries were pooled, and the pooled phage infected onto the bacterial strain XL1-Blue MRF' (Stratagene; La Jolla, CA) to generate approximately 20,000 to 40,000 plaques upon plating. Phage DNA was transferred onto charged nylon filters (Amersham; Piscataway, NJ) and cross-linked to the filters by exposure to ultraviolet light. Radionucleotide (^{32}P) was incorporated into the selected cDNA sequences by random prime labeling using the 'High Prime' reagent mix (Roche Diagnostics GmbH; Mannheim, Germany) and the filters were incubated with this probe. Filters were washed under conditions of high stringency (15 mM NaCl, 1.5 mM sodium citrate, 0.1% sodium dodecyl sulfate, 65°C) and exposed to BioMax MS film (Kodak; Rochester, NY), to reveal clones homologous to the cDNAs.

Preparation of genomic DNA and genome analysis of highly expressed sequences. DNA from *Z. mays* Missouri-13 leaves was prepared using a hexadecyltrimethyl-ammonium bromide based method.[36] DNA (15 μg samples) was digested with EcoRI or

HindIII and DNA fragments were size separated on 0.7% agarose gels. The DNA was transferred onto charged nylon filters and probed with seven ^{32}P radiolabelled cDNA sequences representing *ABA-inducible*, *globulin-1*, globulin-2, *16 kD oleosin*, *17 kD oleosin*, *18 kD oleosin* and the *novel (a)* peptide sequences at high stringency as above. The filters were exposed to film, as above.

Preparation of RNA and analysis of message levels. Plants of each of the four lines used to generate the cDNA libraries, SN103, DA502, ILHP90 and G30B3, were grown under greenhouse conditions and self-pollinated. Missouri-13 plants were similarly grown and self-pollinated. Embryos were isolated from each of the four lines used to generate cDNA libraries and embryos and endosperm were isolated from Missouri-13. Also, leaf, stem, root, tassel, anther, pollen, husk, silk, immature ear and cob tissues were pooled from each of the four lines. All tissues were frozen under liquid nitrogen. Total RNA was isolated from plant tissues using a phenol-based method.[35] RNA (20 μg samples) was size separated on agarose/formaldehyde gels, transferred onto charged nylon filters and probed with the six selected ^{32}P radiolabelled cDNA sequences at high stringency as above. After exposure to film, DNA probes were stripped from filters by washing with near-boiling 0.1% sodium dodecyl sulfate, and were re-probed with maize 18S rRNA gene sequences as a loading control.

Cloning of promoter sequences. Genomic sequences 5' to the translational start codons of the *ABA-inducible* gene sequence, *globulin-1*, *globulin-2*, *16 kD oleosin* and the *novel (a)* peptide gene sequence were cloned by screening a commercial *Z. mays* Missouri-13 genomic phage library in the lambda FIXII vector, where the plant DNA was prepared from leaf tissue (Stratagene; La Jolla, CA), using cloned coding sequences as probes for plaque hybridization. The phage library was infected onto the bacterial strain XL1-Blue MRA (Stratagene; La Jolla, CA) and plated to generate plaques. For plaque hybridization, DNA was transferred onto charged nylon filters. Homologous clones were recovered and the phage inserts mapped by comparing restriction endonuclease digests of the clones following size fractionation using agarose gel electrophoresis. The nucleotide sequence of DNA identified as extending approximately 3 kb 5' of the selected open reading frames was determined at the 'DNA Sequencing Facility' of Iowa State University (Ames, IA).

In cases where at least 3 kb of 5' sequence was not immediately identified, cloned genomic DNA served as template for the Universal GenomeWalker Kit (Clontech Laboratories, Inc.; Palo Alto, CA). Separate genomic DNA samples were each digested with one of five restriction endonucleases and each aliquot of digested DNA was ligated to adaptors to generate a genomic library. Oligonucleotide primer pairs were then used in two rounds of polymerase chain reactions (PCRs) to isolate sequences extending upstream of the sequences identified by plaque hybridization. In each round of reactions, one primer, provided by the supplier, was identical to sequence in the adaptor and was directed downstream into novel newly cloned promoter sequence, and the other corresponded to already cloned promoter sequence proximal to coding sequence and was directed upstream. The primers for the second round of PCRs were nested within those for the first round. Following agarose gel electrophoresis of the reaction products, DNA fragments of defined sizes were isolated, sub-cloned into the PCR product cloning vector pCR2.1 (Invitrogen; Carlsbad, CA) and their DNA sequences determined at the 'DNA Sequencing and Synthesis Facility' of Iowa State University (Ames, IA). Further upstream promoter sequences were isolated by repeating the above PCR based procedure on the genomic library, using oligonucleotide primers corresponding to newly cloned promoter sequences along with the adaptor based primers. When the series of cloned products extended over 3 kb upstream of the translation start site, a full-length clone was isolated from undigested genomic DNA using primers specific to the cloned promoter sequences. The product was also sub-cloned into the PCR product cloning vector pCR2.1 and its nucleotide sequence was verified as matching that of the individual GenomeWalker clones.

Generation of promoter-reporter constructs. Five DNA constructs were generated, each with approximately 3 kb of sequence immediately 5' to the translation start codon of a selected gene identified as being highly expressed in the embryo. Genomic sequences upstream of the *ABA-inducible* gene, *globulin-1*, *globulin-2*, *16 kD oleosin* and the *novel (a)* peptide gene were fused to sequence encoding the β-glucuronidase (GUS) reporter gene (*uidA*). To provide a comparative control, a previously identified and commonly used 1.4 kb 5' extent of *Z. mays globulin-1* sequence was similarly positioned upstream of *uidA* in a sixth construct. For all six constructs, the potato *protease inhibitor II* (*PinII*) terminator sequence was positioned downstream of *uidA*. All constructs also carried a selectable marker cassette to allow for selection of transgenic plants on the herbicide bialaphos, and comprising the cauliflower mosaic virus (CaMV) *35S* promoter, *phosphinothricin Phosphinothricin* (*pat*) of *Streptomyces viridochromogenes*, and the CaMV terminator. Right and left border sequences of an *A. tumefaciens* Ti plasmid flanked sequence encompassing the reporter and selectable marker cassettes. Outside of these borders sequences, the *streptomycin/spectinomycin nucleotydyltransferase* (*aadA*) gene was located, to allow for selection in *Escherichia coli* during plasmid construction and maintenance.

In the case of the 3 kb *globulin-1* promoter and the 1.4 kb globulin-1 control promoter, further fusions were made to a second reporter, the *trypsinogen* gene of *Bos taurus*. In this case, DNA encoding the barley *alpha-amylase* signal sequence was included immediately 5' to reporter sequence.

Maize transformation. Promoter-reporter fusions were introduced by a modified procedure for Agrobacterium-mediated transformation of *Z. mays*.[37] Immature zygotic embryos of Hi-II/elite inbred kernels were transformed with *A. tumefaciens* strain EHA101 containing the relevant isolated upstream sequence-*gusA* reporter fusions. T_0 plants were regenerated from tissue culture and pollinated using pollen from the elite inbred line to produce T_1 seeds.

Tissue staining. Callus tissue and leaf, stem, root, husk, cob, silk, pollen, anther and seed tissues of T_1 generation transgenic *Z. mays* were stained for 24 hours with 0.5 mgml^{-1} 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid: cyclohexylammonium

©2010 Landes Bioscience. Do not distribute

salt (Inalco; Milan, Italy), and were subsequently transferred to 70% ethanol. Blue staining indicated the presence of GUS activity. Prior to screening, T_1 transgenic plants were identified by screening for bialaphos resistance.

Preparation of plant extracts and β-glucuronidase activity assays. For seed extracts, six dry seeds from each ear were individually pulverized and extracted with 1 ml of lysis buffer (50 mM sodium phosphate pH 7.0, 1 mM EDTA, 10 mM β-mercaptoethanol or 1 mM DTT). Samples were placed in extraction tubes held in a rack, with a ball bearing added to each tube, and were then homogenized in a high-speed shaker for 20 s. Samples were centrifuged, and the supernatants recovered and stored on ice prior to analysis. Assays were performed in triplicate to determine GUS activity resulting from expression of the *gusA* reporter gene.[1] Total soluble protein (1 µg) was incubated in 100 µl of lysis buffer and the reaction was initiated with 25 µl of 5 mM 4-methylumbelliferyl β-D-glucuronide (Sigma; St. Louis, MO). The reaction was incubated for up to 30 minutes at 37°C. At specific time points 25 µl volumes of the reaction mixture were transferred to PolySorp 96-well plates (Nalge Nunc International; Rochester, NY) that had 175 µl of stop buffer (0.2 M Na_2CO_3) per well. Fluorescence was measured at an excitation wavelength of 360 nm and an emission wavelength of 460 nm on a Microplate Fluorometer (Molecular Devices; Sunnyvale, CA). GUS protein levels were then calculated by comparison to a standard curve of 1 to 100 µM 4-methylumbelliferone (Sigma; St. Louis, MO). Protein concentrations were determined in duplicate using a dye-binding assay.[38]

Statistical analysis. The analysis of variance was carried out using the mixed procedure in SAS system software version 9.1 (SAS Institute; Cary, NC). Promoters were placed into three groups such that promoters within groups had approximately equal standard deviations and promoters between groups had unequal standard deviations. Specifically, the promoters *ABA-inducible* (PMJ; s.d. = 0.010), *globulin-1* control (PME; s.d. = 0.013) and *globulin-2* (PMF; s.d. = 0.009) were placed into one group, promoters *novel (a)* peptide (PMR; s.d. = 0.047) and globulin-1 (PMH; s.d. = 0.031) were placed into another group and *16 kD oleosin* (PMM; s.d. = 0.004) was in a group by itself. Pair wise comparisons of promoter means between standard deviations groups were done using the Tukey-Kramer method, allowing for unequal standard deviations. The experiment-wise error rate was controlled at 0.05.

Trypsin activity assays. Extracts were prepared as for β-*glucuronidase* assays. Trypsin levels were determined as described by[13] using thiobenzyl benzyloxycarbonyl-L-lysinate (Sigma; St. Louis, MO) as a substrate and 5,5'-dithiobis-(2-nitrobenzoic) acid (Sigma; St. Louis, MO) to monitor the product. Assays were performed in triplicate and samples were measured against a bovine trypsin standard curve that was spiked with maize seed protein. Product formation was monitored using a SpectraMax Plus[384] plate reader (Molecular Devices; Sunnyvale, CA) at 412 nm over a period of 15 minutes. Protein concentrations were determined in duplicate using a dye-binding assay.[38] Seed expression data was assessed using SAS system software version 8 (SAS Institute; Cary, NC).

©2010 Landes Bioscience.
Do not distribute.

Acknowledgements

We would like to thank Dr. V. Rajan for her help in preparation of the manuscript and Dr. K. McGaughey and Dr. J. Walker for their help in the statistical analysis.

References

1. Jefferson R, Kavanagh T, Bevan M. GUS fusions: beta-*glucuronidase* as a sensitive and versatile gene fusion marker in higher plants. EMBO J 1987; 6:3901-7.
2. Wohlleben W, Arnold W, Broer I, Hillemann D, Strauch E, Punier A. Nucleotide sequence of the *phosphinothricin Phosphinothricin* gene from *Streptomyces viridochromogenes* Tü494 and its expression in *Nicotiana tabacum*. Gene 1988; 70:25-37.
3. Nessler C. Metabolic engineering of plant secondary products. Transgenic Res 1994; 3:109-15.
4. Lessard P, Kulaveerasingam H, York G, Strong A, Sinskey A. Manipulating gene expression for the metabolic engineering of plants. Metab Eng 2002; 4:67-79.
5. Estruch J, Carozzi N, Desai N, Duck N, Warren G, Koziel M. Transgenic plants: an emerging approach to pest control. Nat Biotechnol 1997; 15:137-41.
6. Poirier Y, Nawrath C, Somerville C. Production of polyhydroxyalkanoates, a family of biodegradable plastics and elastomers, in bacteria and plants. Nat Biotechnol 1995; 13:142-50.
7. Ye X, Al-Babili S, Klöti A, Zhang J, Lucca P, Beyer P, et al. Engineering the provitamin A (-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science 2000; 287:303.
8. Weigel D, Nilsson O. A developmental switch sufficient for flower initiation in diverse plants. Nature 1995; 377:495-500.
9. Waterhouse P, Wang M, Lough T. Gene silencing as an adaptive defence against viruses. Nature 2001; 411:834-42.
10. Yu H, Kumar P. Post-transcriptional gene silencing in plants by RNA. Plant Cell Rep 2003; 22:167-74.
11. Hood E, Witcher D, Maddock S, Meyer T, Baszczynski C, Bailey M, et al. Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. Mol Breed 1997; 3:291-306.
12. Zhong G, Peterson D, Delaney D, Bailey M, Witcher D, Register Iii J, et al. Commercial production of aprotinin in transgenic maize seeds. Mol Breed 1999; 5:345-56.
13. Woodard S, Mayor J, Bailey M, Barker D, Love R, Lane J, et al. Maize (*Zea mays*)-derived bovine trypsin: characterization of the first large-scale, commercial protein product from transgenic plants. Biotechnol Appl Biochem 2003; 38:123-30.
14. Bailey M, Woodard S, Callaway E, Beifuss K, Magallanes-Lundback M, Lane J, et al. Improved recovery of active recombinant laccase from maize seed. Appl Microbiol Biotechnol 2004; 63:390-7.
15. Hood E, Bailey M, Beifuss K, Magallanes-Lundback M, Horn M, Callaway E, et al. Criteria for high-level expression of a fungal laccase gene in transgenic maize. Plant Biotechnol 2003; 1:129-40.
16. Mason H, Lam D, Arntzen C. Expression of hepatitis B surface antigen in transgenic plants. Proc Natl Acad Sci USA 1992; 89:11745-49.
17. Haq T, Mason H, Clements J, Arntzen C. Oral immunization with a recombinant bacterial antigen produced in transgenic plants. Science 1995; 268:714-6.
18. Carrillo C, Wigdorovitz A, Oliveros J, Zamorano P, Sadir A, Gomez N, et al. Protective immune response to foot-and-mouth disease virus with VP1 expressed in transgenic plants. J Virol 1998; 72:1688-90.
19. Streatfield S, Jilka J, Hood E, Turner D, Bailey M, Mayor J, et al. Plant-based vaccines: unique advantages. Vaccine 2001; 19:2742-8.
20. Daniell H, Streatfield S, Wycoff K. Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. Trends Plant Sci 2001; 6:219-26.
21. Hood E, Woodard S, Horn M. Monoclonal antibody manufacturing in transgenic plants—myths and realities. Curr Opin Biotechnol 2002; 13:630-5.
22. Streatfield S, Mayor J, Barker D, Brooks C, Lamphear B, Woodard S, et al. Development of an edible subunit vaccine in corn against enterotoxigenic strains of *Escherichia coli*. In Vitro Cell Dev-Pl 2002; 38:11-7.
23. Lamphear B, Streatfield S, Jilka J, Brooks C, Barker D, Turner D, et al. Delivery of subunit vaccines in maize seed. J Control Release 2002; 85:169-80.
24. Bustos M, Guiltinan M, Jordano J, Begum D, Kalkan F, Hall T. Regulation of [beta]-*glucuronidase* Expression in Transgenic Tobacco Plants by an A/T-Rich, cis-Acting Sequence Found Upstream of a French Bean [beta]-*Phaseolin* Gene. Plant Cell 1989; 1:839-53.
25. Leung J, Fukuda H, Wing D, Schell J, Masterson R. Functional analysis of cis-elements, auxin response and early developmental profiles of the mannopine synthase bidirectional promoter. Mol Gen Genet 1991; 230:463-74.
26. Guilley H, Dudley R, Jonard G, Balàzs E, Richards K. Transcription of Cauliflower mosaic virus DNA: detection of promoter sequences and characterization of transcripts. Cell 1982; 30:763-73.
27. Liu S, Kriz A, Duncan D, Widholm J. Abscisic acid-regulated Glb1 transient expression in cultured maize P3377 cells. Plant Cell Rep 1998; 17:650-5.
28. Shen L, Petolino J. Pigmented maize seed via tissue-specific expression of anthocyanin pathway gene transcription factors. Mol Breed 2006; 18:57-67.

29. Shepherd C, Scott M. Construction and evaluation of a maize (*Zea mays*) chimaeric promoter with activity in kernel endosperm and embryo. Biotechnol Appl Biochem 2009; 52:233-43.
30. Shepherd C, Vignaux N, Peterson J, Johnson L, Scott M. Green fluorescent protein as a tissue marker in transgenic maize seed. Cereal Chem 2008; 85:188-95.
31. Christensen A, Sharrock R, Quail P. Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing and promoter activity following transfer to protoplasts by electroporation. Plant Mol Biol 1992; 18:675-89.
32. Russell D, Fromm M. Tissue-specific expression in transgenic maize of four endosperm promoters from maize and rice. Transgenic Res 1997; 6:157-68.
33. Belanger F, Kriz A. Molecular basis for allelic polymorphism of the maize Globulin-1 gene. Genetics 1991; 129:863-72.
34. De Wilde C, Van Houdt H, De Buck S, Angenon G, De Jaeger G, Depicker A. Plants as bioreactors for protein production: avoiding the problem of transgene silencing. Plant Mol Biol 2000; 43:347-59.
35. Chatterjee M, Sparvoli S, Edmunds C, Garosi P, Findlay K, Martin C. DAG, a gene required for chloroplast differentiation and palisade development in *Antirrhinum majus*. EMBO J 1996; 15:4194-4207.
36. Stacey J, Isaac P. Issac P, ed. Isolation of DNA from plants. Protocols for nucleic acid analysis by nonradioactive probes 1994; 9.
37. Ishida Y, Saito H, Ohta S, Hiei Y, Komari T, Kumashiro T. High efficiency transformation of maize (*Zea mays* L.) mediated by *Agrobacterium tumefaciens*. Nat Biotechnol 1996; 14:745-50.
38. Bradford M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 1976; 72:248-54.
39. Clough R, Pappu K, Thompson K, Beifuss K, Lane J, Delaney D, et al. Manganese peroxidase from the white-rot fungus *Phanerochaete chrysosporium* is enzymatically active and accumulates to high levels in transgenic maize seed. Plant Biotechnol J 2005; 4:53-62.
40. Streatfield S, Magallanes-Lundback M, Beifuss K, Brooks C, Harkey R, Love R, et al. Analysis of the maize *polyubiquitin-1* promoter heat shock elements and generation of promoter variants with modified expression characteristics. Transgenic Res 2004; 13:299-312.
41. Mølhøj M, Degan F. Leader sequences are not signal peptides. Nat Biotechnol 2004; 22:1502.
42. Streatfield S, Lane J, Brooks C, Barker D, Poage M, Mayor J, et al. Corn as a production system for human and animal vaccines. Vaccine 2003; 21:812-5.

©2010 Landes Bioscience.
Do not distribute.