ISOLATION AND CHARACTERIZATION OF D-*MYO*-INOSITOL-3-PHOSPHATE SYNTHASE GENE FAMILY MEMBERS IN SOYBEAN

LAURA LEE GOOD

Virginia Polytechnic Institute and State University

Candidate for the degree of
MASTERS IN LIFE SCIENCES
in
THE DEPARTMENT OF
PLANT PATHOLOGY, PHYSIOLOGY AND WEED SCIENCE

With a Major Emphasis In
Plant Physiology

COMMITTEE APPROVAL:
Dr. Elizabeth A. Grabau
Dr. John McDowell
Dr. Glenda Gillaspy

July 18, 2001
Blacksburg, Virginia

Keywords: MIPS, *myo*-inositol phosphate synthase, phytic acid, *Glycine max*, soybean seed development

ISOLATION AND CHARACTERIZATION OF D-*MYO*-INOSITOL-3-PHOSPHATE SYNTHASE GENE FAMILY MEMBERS IN SOYBEAN

by

Laura Lee Good

Dr. Elizabeth Grabau, Chairperson

Department of Plant Pathology, Physiology and Weed Science

Virginia Polytechnic Institute and State University

ABSTRACT

The objective of this research was to isolate genes encoding isoforms of the enzyme D-*myo*-inositol 3-phosphate synthase (MIPS, E.C. 5.5.1.4) from soybean and to characterize their expression, especially with respect to their involvement in phytic acid biosynthesis. A MIPS-homologous cDNA, designated GmMIPS1, was isolated via PCR using total RNA from developing seeds. Southern blot analysis and examination of MIPS-homologous soybean EST sequences suggested that GmMIPS1 is part of a multigene family of at least four similar members. The sequences of promoter and genomic regions of GmMIPS1 and GmMIPS2 revealed a high degree of sequence conservation. Northern and western blot analyses showed that MIPS transcript and protein are abundantly expressed early in seed development. Immunolocalization of MIPS protein in developing seeds confirmed expression of MIPS early in seed

development and correlated MIPS protein accumulation in soybean seed tissue with tissues in which phytic acid is known to accumulate. The promoter region of GmMIPS1 was isolated and analyzed for possible seed-specificity using promoter:GUS fusions. Two GmMIPS1 promoter fragments were capable of conferring GUS expression when bombarded directly into developing soybean seeds. However, preliminary bombardment experiments into soybean cell suspension culture indicated that both promoter fragments drove expression of GUS in undifferentiated tissue, indicating a potential lack of seed-specificity.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Introduction

Phytic acid in seed development

Phytic acid (*myo*-inositol hexa*kis*phosphate, $InsP_6$) is the major storage form of phosphorus in plant seeds. In soybean seeds, 60 to 80% of total phosphorus is present as phytic acid (Raboy and Dickenson, 1987). During seed germination, plants mobilize phosphorus and inositol from phytic acid for utilization in growth and development by specific expression of the phytic acid-degrading enzyme phytase (Gibson and Ullah, 1990). Non-ruminant animals are unable to cleave phosphorus from phytic acid molecules due to lack of appropriate digestive enzymes (Swick and Ivey, 1992). Total phosphorus (P) in seeds can be divided into two categories, available and unavailable. Available P consists of both cellular and inorganic phosphate reserves. Cellular P refers to organic phosphorus bound in biological molecules such as DNA, RNA, free nucleotides, sugar phosphates and phospholipids. Typically 10 to 20% of total phosphorus is partitioned into cellular P. Free inorganic phosphorus (P_i) represents a small amount of total phosphorus, typically 5%. Unavailable phosphorus, present as phytic acid, represents the remaining percentage of total phosphorus (Wilcox *et al.*, 2000; Larson *et al.*, 2000). Cellular and inorganic phosphate levels remain fairly constant during much of seed development and show limited response to fluctuations in phosphate levels in the soil environment. Phytic acid levels, however, can vary substantially from plant to plant, as well as from node to node within a plant. Phosphate levels in the environment limit total phytic acid levels. Fluctuations in soil P can result in dramatic reductions in total phytic acid levels (60-70%) without affecting seed germination. Phytic acid begins to accumulate in seeds after cellular phosphate levels have reached

maximum levels, and continues to increase linearly throughout seed development and seed filling (Raboy and Dickenson, 1987, Raboy *et al*., 1985).

Phytic acid in animal feed components

When poultry and swine are fed commercial feed containing corn and soybean meal, the phytic acid-bound phosphorus is nutritionally unavailable. Nonruminants lack phytase, the enzyme that cleaves phosphorus from phytic acid. Resident bacteria are presumed to be responsible for hydrolyzing phytic acid in ruminant animals. Feeds containing soybean and corn must be supplemented with free phosphorus, although total phosphorus levels in seeds are high enough to meet nutritional requirements if made available (Swick and Ivey, 1992).

Phytic acid acts as an antinutrient in the digestive tract of both ruminants and nonruminants, chelating mineral cations, including calcium, zinc, magnesium and iron. High levels of phytic acid can affect the bioavailability of these essential minerals. Phytic acid also has the potential to bind charged amino acid residues of proteins, with the concomitant reduction of protein availability (Ravindran *et al*., 1995). The antinutritional properties of phytic acid are especially important when cereals and legumes serve as the primary protein source in animal and human nutrition.

Phytic acid and complexed nutrients pass through the digestive tract of nonruminants and are excreted in their manure. Soil microorganisms cleave the phosphorus from phytic acid. Although soil naturally binds phosphorus, the phosphorus binding ability can become saturated with repeated applications of phytic acid containing animal waste as fertilizer (Ward, 1993). Runoff from fertilized fields transfers high

concentrations of phosphorus to surface water, leading to eutrophication of lakes, ponds and streams (Sharpley *et al.*, 1994). Elevated levels of phosphorus in runoff emptying into the Chesapeake Bay have been linked to excessive algae growth and increased fish kills (Ward, 1993). High phosphorus levels have been associated with *Pfiesteria piscicida* outbreaks in the Chesapeake Bay and connected waterways (Glasgow *et al*., 1995).

Reduction of phytic acid levels in soybean seeds could increase the value of soybean as a livestock feed component. Not only would the need to supplement livestock feed with phosphorus be eliminated, the nutritional quality of soybean meal could be improved. Lower phytic acid levels in seeds used for soybean meal would decrease phytic acid levels in animal waste, reducing overall phosphorus pollution in areas of intensive livestock production. Costlier alternatives to reducing phytic acid levels, such as direct addition of fungal phytase supplements to feed, could be eliminated with the introduction of low phytic acid feeds.

Approaches to improving phytic acid utilization: feeding studies

Until the isolation of low-phytic acid mutants, research in lowering phytic acid levels in animal diets had been directed at the utilization of phytase to break down phytic acid. Phytase cleaves inorganic phosphorus from phytic acid molecules. Feeding studies have shown that direct addition of microbial phytase to soybean meal increases the availability of phosphorus (Denbow *et al*., 1995; Ravindran *et al*., 1995b; Simons *et al*., 1990; Harper *et al*., 1997) and improves amino acid utilization in poultry (Biehl and Baker, 1997). Unfortunately, microbial phytase supplementation of livestock diets is

even more costly than direct addition of inorganic phosphorus to meet nutritional requirements. Engineering plants with microbial phytase is viewed as an alternative to direct enzymatic supplementation of animal diets. Broilers on a diet including tobacco seeds transformed with the phytase gene from *Aspergillus niger* showed increased growth rates over broilers on a diet supplemented with commercially available microbial phytase (Pen *et al*., 1993). This suggests that plants modified to express phytase in order to decrease seed phytic acid levels may be preferable than direct phytase supplementation of feed. Denbow *et al*. (1998) found increased phosphorus digestibility and reduced phosphorus excretion when broiler diets were supplemented with full-fat, milled transgenic soybeans expressing fungal phytase. Post-harvest processing of soybeans presents a difficulty for expression of active phytase. Soybeans are roasted at high temperatures to inactivate trypsin inhibitors and treated with organic solvents to extract soybean oil, an economically important commodity. In order to reduce phytate levels in feed, phytase in transformed soybeans must degrade seed phytic acid before processing, have thermostable properties or be protected within the cell from the harsh processing conditions.

Low phytic acid mutants

Low phytic acid (*lpa)* mutants have been identified in several important crop species, including rice (Larson *et al.*, 2000), corn (Larson *et al.*, 1999), barley (Larson *et al*., 1998) and soybean (Hitz and Sebastian, 1998; Wilcox *et al*., 2000). Possible mutants were first identified by screening seeds for high levels of free phosphorus. Mutants with reduced levels of phytic acid were then identified by HVPE (high voltage paper

electrophoresis). Screening for mutants with high inorganic phosphorus levels has identified low phytic acid (*lpa*) progeny that are otherwise phenotypically normal. In corn, barley and rice, *lpa* mutants can be classified into two discrete groups. *Lpa1* mutants show a gain in inorganic P levels equivalent to the decrease in seed phytic acid. *Lpa2* mutants accumulate lower inositol polyphosphates, primarily $InsP_5$, but also $InsP_4$ and $InsP_3$, with a limited increase in free phosphorus. No significant differences between non-mutant and *lpa1* or *lpa2* mutants seed total P (Raboy *et al*, 2000). While *lpa1* mutants appear phenotypically wildtype and germinate normally, *lpa2* mutants were recovered less frequently and germination rates are negatively affected in corn and barley. It is predicted that *lpa1* mutations affect phytic acid biosynthesis early in the pathway, possibly at the MIPS (*myo*-inositol 3-phosphate synthase) gene (Hitz and Sebastian, 1998) or in some aspect of MIPS gene regulation. *Lpa2* mutants are affected later in the biosynthetic pathway, possibly through disruption of one of the phosphoinositol kinases.

Feeding studies with *lpa* mutants have shown that phosphorus and mineral utilization are improved in swine (Veum *et al*., 1998; Spencer *et al*., 2000). Similar results were found when *lpa*-containing feed was fed to poultry (Huff *et al*., 1999; Douglas *et al*., 2000, Li *et al*., 2000,) and rainbow trout (Sugiura *et al.*, 1998)

MIPS expression

An alternative to phytase engineering for decreased seed phytic acid levels is the reduction of the expression of enzymes in the biosynthetic pathway of phytic acid. MIPS catalyzes the NADH-dependent conversion of glucose-6-P to *myo*-inositol-3-P, and is

critical in the *de novo* synthesis of inositol in higher organisms. One postulated pathway for the biosynthesis of phytic acid proceeds through the sequential phosphorylation of D-*myo*-inositol-3-phosphate (see Figure 1). MIPS genes have been isolated from a variety of microbial, plant and animal sources (reviewed in Majumder *et al.*, 1997), including yeast (Johnson and Henry, 1989), potato (Keller *et al.*, 1998), *Spirodela polyrrhiza* (Smart and Fleming, 1993), rice (Yoshida *et al*, 1999) and Arabidopsis (Johnson, 1994). Sequence conservation of MIPS is observed in widely divergent organisms, including protozoa, algae, plants and animals. This suggests the presence and functionality of MIPS before evolutionary divergence of eukaryotes and eubacteria (Majumder *et al.*, 1997). The molecular weight of individual MIPS subunits, both chloroplastic and cytosolic forms from a variety of plant sources, ranges between 58 and 67 kDa (Majumder *et al.*, 1997).

In addition to the phosphorus storage functions of MIPS-synthesized inositol derivatives, inositol and inositol phosphates are also critical in a variety of cellular functions (reviewed in Loewus and Murthy, 2000). Oxidation of *myo*-inositol to D-glucuronic acid 1-P provides polysaccharides for cell wall biogenesis. Galactinol, a product of inositol conjugation, serves as a galactosyl donor for the complex sugars raffinose, stachyose and verbascose, which in turn function in seed development and dessication, stress responses and phloem transport. Inositol conjugated to indole-3-acetic acid (IAA) serves as storage for the plant hormone, and may also act in IAA transport. Ononitol and pinitol, formed by methylation of inositol, are involved in stress protection, including salt tolerance (Ishitani *et al.*, 1996). Glycosylated lipids containing inositol moieties act to anchor proteins to cell membranes. These protein-lipid complexes have

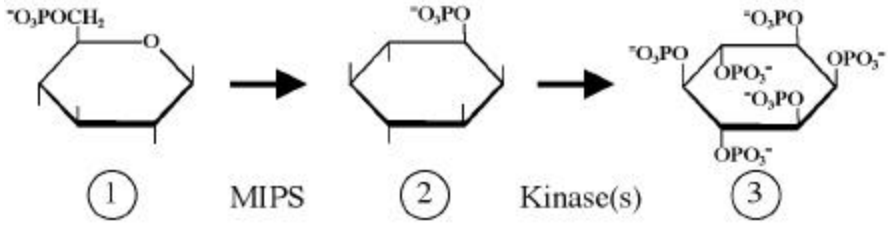


Figure 1. Phytate biosynthesis. D-glucose-6-phosphate (1) is converted to d-*myo*-inositol-3-phosphate (2) by the activity of D-*myo*-inositol 3-phosphate synthase (MIPS). D-*myo*-inositol-3-phosphate is further phosphorylated to yield *myo*-inositol hexakisphosphate (3) by several kinase steps.

been associated with detection of environmental stimuli within the cell. The role of phosphoinositides in cell signal transduction has been well documented in animal and plant cells (reviewed in Munnik *et al.,* 1998).

Induction of MIPS is associated with osmotic stress in the halophyte *Mesembryanthemum crystallinum* (common ice plant). Accumulation of the methylated inositol derivatives ononitol and pinitol was observed in iceplant, a salt and cold tolerant plant (Adams *et al.*, 1992; Vernon *et al*., 1993), after exposure to osmotic stress. These specialized sugars accumulated to high levels in salt-tolerant plants and are thought to function as osmoprotectants. Ishitani *et al*. (1996) observed a transcriptional induction of *Inps1* (MIPS homolog in *M. crystallinum*) after exposure to salt stress. *Arabidopsis,* a glycophyte, did not exhibit transcription induction of *Inps1* after salt stress. Salt-tolerant varieties of rice (*Oryza sativa* L.) exhibited higher levels of MIPS activity in cytosolic and chloroplastic fractions than salt-sensitive varieties (Raychaudhuri and Majumder, 1996). Although salt-inducible expression of MIPS is characteristic of halophytes, over-expression of MIPS in *Arabidopsis* was not sufficient to confer salt or cold tolerance (Smart and Flores, 1997). Although it seems clear that upregulation of MIPS during osmotic stress is characteristic of halophytes, creation of inositol flux is not the sole determinant of stress tolerance. It has been suggested that glycophytes lack the enzymes required to methylate inositol to form ononitol and pinitol (Vernon *et al*., 1993b). If glycophytes not only fail to upregulate inositol production during salt-stress but cannot methylate existing inositol, then engineering salt-tolerant plants would require transfer of the all the enzymes of the methyation pathway.

The S*pirodela polyrriza* MIPS homolog *tur1* was induced by the plant hormone abscisic acid (ABA) (Smart and Fleming, 1993). In *S. polyrrhiza*, ABA induces formation of turions, dormant overwintering structures. *Tur1* transcript and *myo*-inositol phosphate synthase activity was upregulated by ABA treatment, but not cold treatment, despite the fact that cold-treatment induces turion formation. The authors proposed that *tur1* upregulation was involved in ABA signaling. *In situ* hybridization of ABA-treated *S. polyrrhiza* showed localization of the *tur1* transcript to the stolon, the tissue connecting the mother frond and daughter frond. Flores and Smart (2000) showed that treatment of *S. polyrrhiza* with ABA not only increased levels of *myo*-inositol, but also levels of other soluble inositol polyphosphates ($InsP_3$, $InsP_4$, $InsP_5$). Phytic acid ($InsP_6$) accumulation was observed in vegetative tissue after ABA treatment. The purpose of phytic acid accumulation and its relation to ABA signal transduction is not known.

In *Arabidopsis,* five fold higher MIPS expression (and corresponding four fold increase in inositol levels) did not result in a discernable phenotype (Smart and Flores, 1997). In contrast, constitutive antisense reduction of MIPS activity in potato (*Solanum tuberosum*) resulted in an obvious phenotype (Keller *et al*., 1998). MIPS levels in transgenic lines were reduced to below 20% of wildtype levels. Plants exhibited reduced apical dominance, flowering was delayed or absent, leaf morphology was altered, tubers were smaller and more numerous, and leaves senesced earlier in transgenic plants than control plants. Inositol, galactinol and raffinose levels were reduced, while sucrose and starch levels increased. Constitutive reduction of *de novo* inositol synthesis results in a phenotype that drastically reduces plant productivity, an important consideration for agronomically important crops. However, soybean mutants with a point mutation in a

MIPS gene, reducing the specific activity of the enzyme to 10% of the wildtype, were phenotypically normal (Hitz and Sebastian, 1998). This indicates specific reduction of MIPS activity to decrease seed phytic acid levels would be more likely to result in productive plants than constitutive expression of an antisense MIPS construct.

Diurnal regulation of MIPS expression was found across plant species. Keller *et al*. (1998) noted increased MIPS transcript levels in photosynthetic tissue of potato and increased MIPS RNA levels in light compared to dark. Raychaudhuri and Majumder (1996) observed higher chloroplastic inositol synthase activity when rice plants were grown in the presence of light. Ishitani *et al.* (1996) observed diurnal shifts in the expression of the salt-inducible *Inps1*. Nelson *et al.* (1998) could not detect MIPS protein in chloroplastic fractions of *M. crystallinum*, but did note diurnal shifts in MIPS protein, although fluctuations in protein levels was less dramatic than diurnal shifts in MIPS transcript.

Johnson and Wang (1996) observed two distinct sizes of MIPS protein in western blot analysis of green bean (*Phaseolus vulgaris*). A 33 kD protein that cross reacted with a yeast MIPS polyclonal antibody was observed in the globular stage of embryogenesis, mature roots and in leaf extracts (after precipitation), while a 56 kD protein was observed in cotyledonary stage embryos and in the cotyledon of seeds 13 days after germination. After isolation and fractionation of chloroplasts, the 33 kD protein was found only in thylakoids.

Involvement of MIPS in *lpa* phenotype

Chromosome mapping of *lpa1* in maize demonstrated that the location of the mutation was on chromosome 1S near one of several regions of MIPS homology (Raboy *et al.*, 2000). Mutants with the low phytic acid phenotype had equivalent levels of total phosphorus when compared to wild type maize plants. However, more of the total phosphorus was present in forms other than phytic acid, and therefore more biologically available when maize was used as feed. Rice *lpa1* mutants mapped to chromosome 2L, whereas the single rice MIPS gene mapped to chromosome 3, demonstrating that the rice *lpa1* phenotype was not caused by a mutation in the MIPS gene. Larson *et al.* (2000) speculated that the similar *lpa1* phenotype could be caused by a mutation in a gene with MIPS activity that lacked homology to the identified MIPS gene. Alternatively, the *lpa* phenotype could be due to a mutation in a transcription factor involved in seed-specific MIPS activity. Linkage mapping in barley *lpa*1 mutants identified a *lpa1* locus on chromosome 2H. Comparative mapping studies indicate that this barley locus is not homologous to the maize *lpa1* locus associated with MIPS (Larson *et al*., 1998).

Soybean mutant lines with low raffinose saccharides developed by DuPont also exhibited low phytic acid levels (Hitz and Sebastian, 1998). Because of the poor digestibility of raffinose saccharides, soybean seeds with reduced levels of raffinose saccharides would be a valuable commodity. After chemical mutagenesis of soybean seeds, the authors screened for low raffinose saccharide content. A line, LR33, was isolated that exhibited not only the desirable phenotype of low raffinose saccharide content, but also contained lowered phytate levels and increased levels of free phosphorus and sucrose. Seeds from mutant lines contained less *myo*-inositol and lower

MIPS activity than wildtype soybean seeds. Free inositol and UDP-galactose are utilized to form galactinol. Sucrose is then combined with galactinol to form raffinose. The reduction in MIPS activity in LR33-derived soybean lines was attributed to a point mutation in the MIPS gene, changing amino acid residue 396 in wildtype MIPS from a lysine residue to an asparagine residue in the mutant plants. Soybeans containing the mutated allelle contained 8 to 9 times less raffinose saccharides and 40% less phytic acid than wildtype soybeans.

Soybean mutants similar to *lpa1* from corn, barley and rice were obtained by Wilcox *et al*. (2000). These mutants have lowered phytate levels and increased free inorganic phosphate compared to wildtype. The locations of the soybean *lpa1* mutants have not been mapped.

Yoshida *et al.* (1999) observed accumulation of rice MIPS transcript (*RINO1*) coincident with accumulation of phytic acid in globoid bodies. Through *in situ* hybridization of sections of developing rice seeds with antisense *RINO1* RNA probe, the authors visualized accumulation of *RINO1* transcript early in seed development. Transcript was first seen in the globular stage embryos of developing rice seeds at 2 DAA (days after anthesis). Signal was detected first in the embryo, then in scutellum, coleoptile, plumule, radicle, epiblast and aleurone, but not endosperm tissue. *RINO1* transcript accumulation was observed until 10 DAA. Globoid accumulation was observed beginning at 4 DAA in those tissues in which *RINO1* transcript accumulated. *RINO1* was observed in the vascular cambium of the developing embryo, but globoids were not observed to accumulate in this tissue. The accumulation of phytate-containing

globoids in tissues that expressed *RINO1* suggested a tissue specific expression of *RINO1* proceeding and correlating with phytate accumulation.

Soybean MIPS

The relationship between phytic acid biosynthesis and MIPS make the enzyme an attractive target for antisense regulation of MIPS transcript levels in soybean seed development. MIPS cDNA has been isolated from a variety of species, including *Arabidopsis* (Johnson, 1994), ice plant (Ishitani *et al.*, 1996), duckweed (Smart and Fleming, 1993) and green bean (Wang and Johnson, 1995). Hegeman *et al*. (2001) aligned these previously published sequences using Lasergene DNA analysis package from DNAStar, and oligonucleotide primers were designed from highly conserved regions. RNA from developing soybean cotyledons was used in RT-PCR (reverse transcription polymerase chain reaction) with MIPS consensus primers to amplify the soybean MIPS cDNA. Using RACE (Rapid Amplification of cDNA Ends), the complete soybean MIPS cDNA was isolated and sequenced. Results of Southern blot analysis with MIPS cDNA (GmMIPS1) showed cross reactivity of the probe with multiple genomic fragments under high stringency. Northern and western analysis of MIPS expression in developing soybean showed abundant expression of MIPS in early stages of soybean seed development (Hegeman *et al.*, 2001). These studies provided the basis for further research on the organization of the MIPS gene family in soybean and analyses of MIPS promoter regions for use in antisense MIPS constructs.

Goals and Objectives

The long-term goal of this research is to reduce the level of phytic acid accumulation in soybean seeds by reducing the amount of a key biosynthetic enzyme, *myo*-inositol-3-phosphate synthase, by antisense reduction of MIPS expression.

Reduction of gene expression through antisense occurs through the simultaneous expression of a target mRNA molecule with a complementary antisense mRNA molecule. The postulated mechanism is that sense and antisense mRNAs interact, resulting in a double-stranded RNA molecule, triggering RNA degradation (Mol *et al*., 1990). Reduced expression of native genes can occur through expression of complementary antisense mRNA ranging in size from 20 to several hundred base pairs (Klaff *et al*., 1996). The ability of small antisense primers to repress native gene expression indicates that regions of sequence identity, such as those found in conserved gene families, are susceptible to antisense suppression.

In order for an antisense construct to function in preventing MIPS expression, the construct must be transcribed at the same time as the native MIPS gene. The proposed antisense construct should be under the control of a MIPS promoter that is active in early seed development. Existing seed specific promoters, such as the soybean seed storage protein β-conglycinin (Chen *et al*., 1986), direct expression later in seed development than expression from the native MIPS promoter. The isolation and characterization of a soybean MIPS promoter active in early seed development is a crucial step in the generation of antisense MIPS constructs. Characterization of MIPS gene family members will provide information about isoform specific regulation necessary for proper control of antisense expression. Complete sequence of MIPS gene family members will

also aid in designing antisense constructs that are specific to a particular MIPS isoform in soybean.

Specific objectives include:

1. Isolation and sequencing of three MIPS gene family members.
2. Isolation of promoter regions of two MIPS genes through PCR and characterization of their specific expression patterns through bombardment of expression constructs into soybean tissue.
3. Characterization of MIPS protein expression in seeds and other plant tissues through western blot analysis.
4. Examination of MIPS protein accumulation in developing seeds through immunolocalization.

CHAPTER II

Expression of D-*myo*-inositol-3-phosphate synthase in soybean: implications for phytic acid biosynthesis

This chapter was published April 2001 in Plant Physiology (125: 1941-1948). Consent of the coauthors (C.E. Hegeman and E.A. Grabau) and the American Association of Plant Biologists to include this manuscript as a chapter is appreciated.

References listed in the published paper are included in the combined reference section following Chapter IV.

ABSTRACT

Phytic acid, a phosphorylated derivative of *myo*-inositol, functions as the major storage form of phosphorus in plant seeds. *Myo*-inositol phosphates, including phytic acid, play diverse roles in plants as signal transduction molecules, osmoprotectants, and cell wall constituents. D-*Myo*-inositol-3-phosphate synthase (MIPS, E.C. 5.5.1.4) catalyzes the first step in *de novo* synthesis of *myo*-inositol. A soybean (*Glycine max*) MIPS cDNA (*Gm*MIPS1) was isolated by RT-PCR using consensus primers designed from highly conserved regions in other plant MIPS sequences. Southern blot analysis and database searches indicated the presence of at least four MIPS genes in the soybean genome. Northern blot and immunoblot analyses indicated higher MIPS expression and accumulation in immature seeds than in other soybean tissues. MIPS was expressed early in the cotyledonary stage of seed development. The *Gm*MIPS1 expression pattern suggested that it encodes a MIPS isoform that functions in seeds to generate D-*myo*-inositol-3-phosphate as a substrate for phytic acid biosynthesis.

Nutrient reserves accumulate during seed development for re-mobilization during germination and early seedling growth. Phosphorus in seeds is stored primarily in the form of phytic acid (phytate, *myo*-inositol hexakisphosphate, $InsP_6$), which is a derivative of inositol. During seed development, phytic acid is deposited in spherical inclusions known as globoids or as complexes with seed storage proteins in protein bodies (Prattley and Stanley, 1982; Lott *et al.*, 1995). The stored phytate is hydrolyzed by the activity of phytase enzymes during germination to provide inorganic phosphate and *myo*-inositol to the growing seedling.

As a component of animal feed, phytate from seeds compromises the availability of dietary phosphorus because non-ruminants lack the digestive enzymes to hydrolyze phytic acid. The negatively-charged phytic acid molecule also chelates other mineral nutrients, which lowers their bioavailablity. In addition to the nutritional effects of phytate in animal diets, there are also environmental consequences. Undigested phytic acid is excreted in manure, which leads to elevated soil phosphorus levels when manure is applied repeatedly as fertilizer in areas of livestock production. High soil phosphorus levels, coupled with the potential for run-off, can lead to environmental phosphorus pollution and eutrophication (Sharpley *et al.*, 1994).

Myo-inositol is a precursor to compounds in plants that function not only in phosphorus storage, but also in signal transduction, stress protection, hormonal homeostasis, and cell wall biosynthesis (for reviews see Morré *et al.*, 1990; Loewus and Murthy, 2000). Production of the second messenger $InsP_3$, a derivative of inositol, leads to release of intracellular calcium in cellular signalling. Methylated forms of *myo*-inositol, ononitol and pinitol, accumulate in plants in response to salinity stress to

function in osmotic adjustment. Plant cells contain a number of conjugates of indole-3-acetic acid (IAA), including IAA-*myo*-inositol, which allow storage of excess amounts of auxin. Oxidation of *myo*-inositol to D-glucuronic acid plays a role in the biosynthesis of compounds that are constituents of plant cells walls.

The first step in the synthesis of *myo*-inositol is the conversion of D-glucose-6-phosphate to D-*myo*-inositol-3-phosphate by the isomerase D-*myo*-inositol-3-phosphate synthase (MIPS). The proposed phytic acid biosynthetic pathway (Loewus and Murthy, 2000) also includes phosphorylation steps catalyzed by one or more kinases that have not been well studied in plants (Fig. 1). The MIPS coding sequence has been cloned and characterized from a number of prokaryotic and eukaryotic sources. The first reported MIPS gene was isolated from *Saccharomyces cerevisiae* using a genetic complementation strategy (Johnson and Henry, 1989). A sequence similar to the yeast MIPS gene was identified in *Spirodela polyrrhiza*, an aquatic angiosperm (Smart and Fleming, 1993). MIPS sequences have been reported from *Arabidopsis thaliana* (Johnson, 1994), *Citrus paradisii* (Abu-Abied and Holland, 1994), *Mesembryanthemum crystallinum* (Ishitani *et al.*, 1996), and *Nicotiana tabacum* (Hara *et al.*, 2000), and are highly conserved at the nucleotide level.

MIPS sequences comprise gene families in some plant species. In maize, seven sequences hybridizing to a MIPS probe were mapped to different chromosomes (Larson and Raboy, 1999). In *Arabidopsis*, two distinct MIPS genes have been identified (Johnson, 1994; Johnson and Sussex, 1995; Johnson and Burk, 1995). The existence of multiple MIPS genes in plants may permit differential MIPS expression for many

different physiological functions, however, MIPS gene family expression has not been well characterized to date.

We report the isolation of a MIPS cDNA from developing soybean seeds. The genomic organization of soybean MIPS genes and the structure of the gene corresponding to the isolated MIPS cDNA were examined. At least four different MIPS sequences were identified from the soybean EST database and Southern hybridization results. Patterns of MIPS expression were analyzed in developing soybean seeds and other plant tissues. Together, these data suggest that we have isolated a member of a MIPS gene family that is abundantly expressed in developing soybean seeds to function in phytic acid synthesis.

RESULTS AND DISCUSSION

Isolation and analysis of a MIPS cDNA

A MIPS cDNA (*Gm*MIPS1) was isolated from developing soybean seeds using RT-PCR (Fig. 2). The 1791 bp cDNA sequence contained an 1533 bp open reading frame that could encode a protein of 510 amino acids. The putative soybean MIPS protein has a predicted molecular mass of 56.5 kD. BLAST database searches were performed with both the MIPS cDNA sequence and predicted protein sequence (Altschul *et al.*, 1997). The searches revealed a high degree of sequence identity between *Gm*MIPS1 and MIPS genes from plants and other organisms. The highest scoring match was to the *Nicotiana tabacum* MIPS sequence (Hara *et al.*, 2000), which showed 80.6% identity to *Gm*MIPS1 at the nucleotide level and 92.2% identity at the amino acid. Sequence identity to yeast was 40.2% at the nucleotide level and 41.5% at the amino acid level (Johnson and Henry,

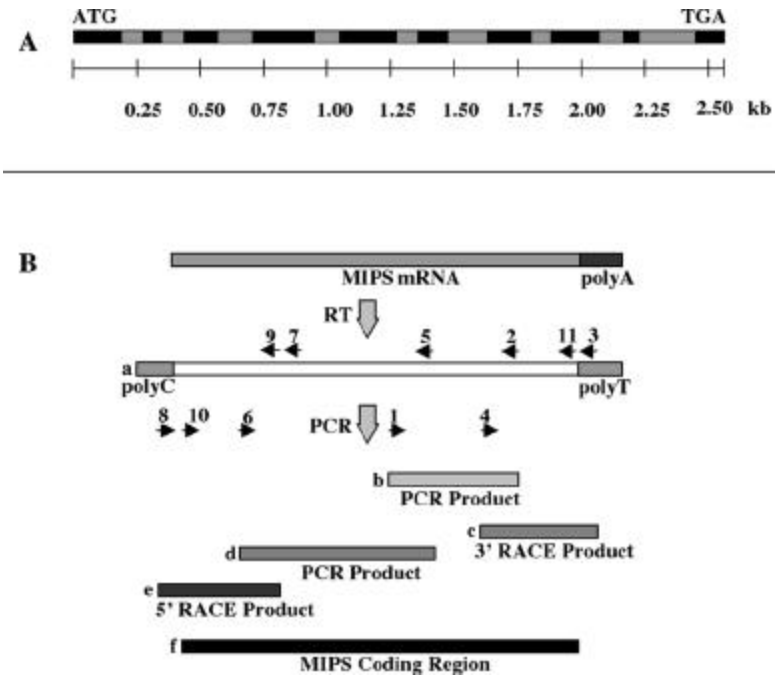


Figure 2. *Gm*MIPS1 gene structure and cDNA cloning strategy. (A) The sequence of a genomic PCR fragment was compared to the cloned MIPS cDNA sequence to define intron-exon borders. Introns are represented in gray and exons are represented in black. Scale indicates size in kb. (B) Cloning of *Gm*MIPS1 cDNA sequences. (a) Total RNA from developing seeds was used as template for cDNA synthesis. Locations of PCR primers on the MIPS sequence are indicated. (b) Primers 1 and 2 from conserved MIPS regions were used to amplify an internal fragment. (c) The 3' end of the MIPS cDNA was amplified by 3' RACE using anchored primer 3 and gene specific primer 4. (d) A second internal PCR product was amplified using primer 5 from a 5' conserved region in conjunction with a primer 6 from the first PCR product. (e) The 5' end of the coding sequence and untranslated sequence were obtained by using primer 7 for cDNA synthesis and adapter anchored primer 8 and gene specific primer 9 for amplification by 5' RACE. (f) A fragment containing the full-length coding sequence was amplified by high-fidelity PCR using primers 10 and 11.

1989), further confirming the high degree of conservation among MIPS sequences from different sources.

MIPS gene structure and organization

An estimate of MIPS sequence copy number was determined by Southern blot analysis. Under stringent conditions, multiple high molecular weight bands were detected using the full length cDNA as the MIPS hybridization probe (Fig. 3). The data indicated that the soybean genome contains four or more loci with MIPS homology.

A MIPS genomic sequence corresponding to *Gm*MIPS1 was generated by PCR amplification and was compared to the cDNA sequence for assignment of exons and introns. The region encompassing the start (ATG) to stop (TGA) codons in the genomic PCR product consisted of 2607 base pairs containing nine introns (Fig. 2A). Sequence analysis revealed that restriction sites for *Eco*RI, *Hind*III, and *Xba*I were absent in the genomic sequence. Genomic restriction fragments generated by the three enzymes were larger than the size of the genomic PCR product, consistent with the lack of these sites in the sequence (Fig. 3). Due to the high degree of sequence similarity among plant MIPS genes, high stringency hybridization did not reveal which fragment corresponded to the cloned *Gm*MIPS1 sequence.

Analysis of soybean cDNAs with homology to MIPS

The *Gm*MIPS1 cDNA sequence was compared to soybean expressed sequence tags (ESTs) that have been identified as putative MIPS sequences in a soybean database,

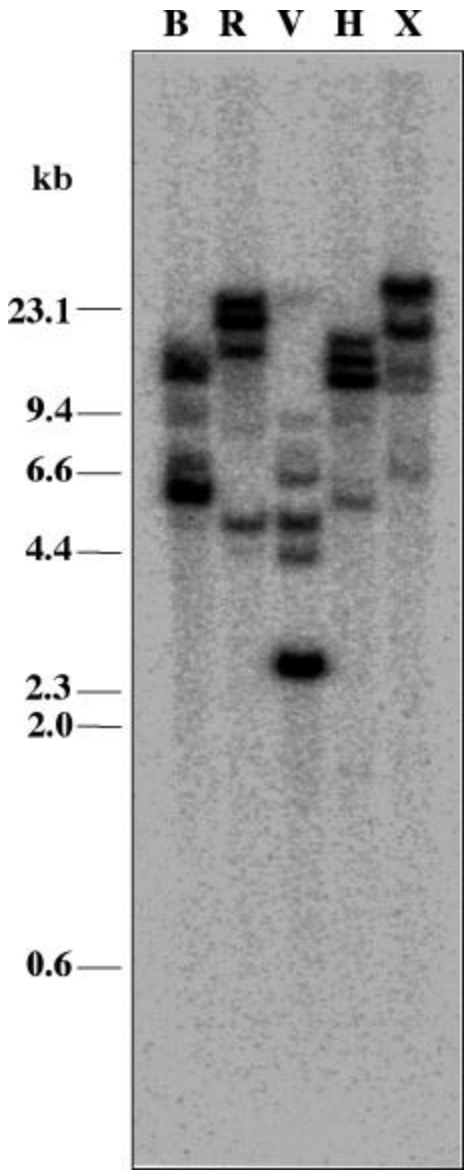


Figure 3. Southern blot analysis of soybean genomic DNA. A full-length MIPS cDNA fragment was [32P]-labeled to probe a blot of soybean genomic DNA at high stringency. DNA samples (10 mg) were digested with the following restriction endonucleases: *Bgl*II (B), *Eco*RI (R), *Eco*RV (V), *Hind*III (H), and *Xba*I (X), respectively. *Hind*III-digested bacteriophage

SoyBase (R. Shoemaker, USDA-ARS, Iowa State University). Analysis of the large number of soybean cDNA sequences with MIPS homology facilitated assignment of the ESTs to four highly similar, yet distinct MIPS sequences (*Gm*MIPS1 to *Gm*MIPS4) as shown in Table 1. (The complete *Gm*MIPS1 sequence and alignment with other MIPS family members from the EST database, as well as EST accession numbers are shown in on-line supplementary material. On-line supplementary material also found in Appendix A and Appendix B. EST accession numbers of GmMIPS3 and GmMIPS4 have been updated) The *Gm*MIPS1 cDNA isolated from immature seeds in these studies was identical to the EST sequences obtained from several immature cotyledon libraries. The majority of ESTs from *Gm*MIPS2 originated primarily from root libraries, but also included sequences from shoot, flower, and seed coat libraries. *Gm*MIPS1 and *Gm*MIPS2 showed 96% identity over a span of 1465 nucleotides and shared a very similar 3' untranslated region. The EST sequences corresponding to *Gm*MIPS3 were derived predominantly from flower libraries, but were also found in leaves, buds, and germinated cotyledons. Only two ESTs were assigned to *Gm*MIPS4 (from buds and young pods). *Gm*MIPS4 sequences were similar to *Gm*MIPS3 in both coding and 3' untranslated sequences. In addition to the soybean cDNAs that were assigned to *Gm*MIPS1 through *Gm*MIPS4, there were several individual cDNAs in the EST database that could not be placed with any other sequences.

The data from the soybean EST project facilitated the generation of an "electronic northern" which served as a predictive first step in determining the expression patterns of MIPS gene family members (Table 1). The frequency of *Gm*MIPS1 EST sequences and the derivation of these *Gm*MIPS1 sequences from immature cotyledons suggests that

Table 1. Classification of soybean ESTs with MIPS homology. Number of EST sequences and library source of cDNA for each MIPS family member are indicated.

Library source	Number of EST sequences encoding			
	*Gm*MIPS1	*Gm*MIPS2	*Gm*MIPS3	*Gm*MIPS4
Immature cotyledons	16	0	0	0
Roots	2	7	0	0
Immature seed coats	1	3	0	0
Shoot tips	0	1	0	0
Flowers	0	1	5	0
Immature flowers	0	1	3	0
Germinated cotyledons	0	0	2	0
Leaves	0	0	2	0
Buds	0	0	2	1
Young pods	0	0	0	1

*Gm*MIPS1 is abundantly expressed early in seed development, perhaps to serve in phytic acid biosynthesis.

Soybean MIPS RNA expression

Northern blot analyses were performed with total RNA from different soybean tissues and various stages of seed development, using the *Gm*MIPS1 cDNA as the hybridization probe (Fig 4). Steady state RNA levels were higher in developing seeds than in other soybean tissues, including flowers, leaves, roots, germinating cotyledons, and somatic embryos (Fig. 4A). MIPS transcript was observed in cotyledons at the earliest developmental stages analyzed, 0 to 2 mm (corresponding in size up to approximately stage C; Meinke *et al.*, 1981). Maximal levels of MIPS RNA were observed in 2-4 mm seeds (equivalent to stages D-F; Fig. 4B). As development progressed, MIPS transcript levels decreased to nearly undetectable levels by 8-10 mm seeds (stage J and beyond). Although differences in the relative abundance of MIPS RNA were easily detected using *Gm*MIPS1 as a probe, the high degree of sequence similarity between gene family members prevented us from concluding definitively that the observed patterns were gene specific. However, coupled with the relative abundance of *Gm*MIPS1 in the EST database and the origin of *Gm*MIPS1 sequences primarily from immature seeds, these results strongly suggest that the high level of MIPS transcript in soybean seeds may be primarily due to *Gm*MIPS1 expression.

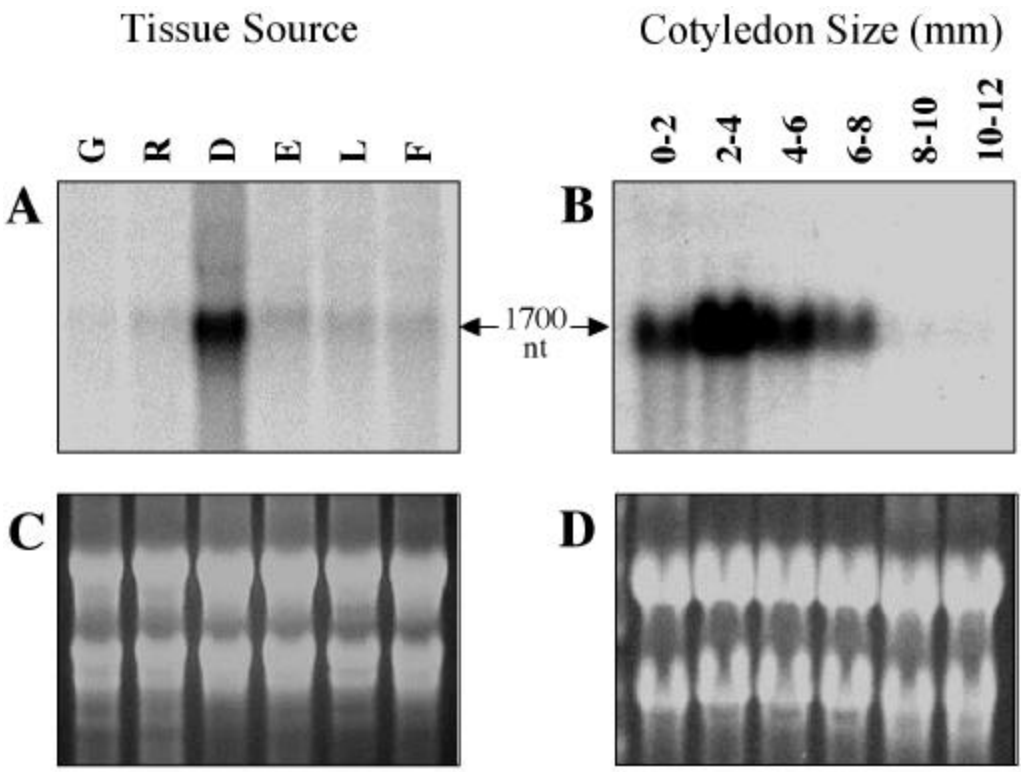


Figure 4. Analysis of MIPS expression patterns in soybean tissues. After transfer to membranes, soybean RNA samples were probed using a [32P]-labeled MIPS cDNA. Panels A and C, RNA samples from soybean tissues: G, germinated cotyledons; R, young roots; D, developing seeds; E, globular-stage somatic embryos; L, young leaves; F, flowers. Panels B and D, RNA samples from different stages of seed development: 0-2 mm, 2-4 mm, 4-6 mm, 6-8 mm, 8-10 mm, 10-12 mm. Panels A and B are autoradiographs; panels C and D are the corresponding ethidium bromide-stained samples to visualize RNA quantity and quality.

Immunodetection of soybean MIPS

Western blot analyses were performed to further corroborate the pattern of MIPS expression. To produce an antibody for immunodetection studies, a MIPS fusion protein construct was generated by insertion of a PCR product into an expression vector containing a His-tag. After demonstrating that the fusion protein was successfully expressed in bacterial cells, the vector was used for large scale protein expression and antibody production (Styer, 2000). The generation of a polyclonal antibody raised against the MIPS fusion protein has allowed the examination of MIPS protein expression and will provide a useful tool for further immunocytochemical studies. Immunoblot analysis of protein extracts from developing seeds, flowers, roots, somatic embryos, leaves, and germinated seeds indicated that developing seeds contained the highest levels of MIPS protein (Fig. 5A). The faint immunoreactive bands observed in other tissues may represent additional MIPS isoforms. The accumulation of the MIPS protein was observed in early seed development, at stages E, F, and G (2.6 to 5.4 mm), as seen from the developmental series of immature coyledons (Fig. 5B).

The pattern of MIPS expression observed in developing soybean seeds was consistent with the expectation that high levels of MIPS expression should precede and/or accompany synthesis and accumulation of phytic acid. In rice, recent studies demonstrated a clear relationship between patterns of MIPS expression and accumulation of phytic acid (Yoshida *et al.*, 1999). MIPS RNA expression was detected by *in situ* hybridization in globular stage rice embryos, prior to the formation of phytate-containing globoids. In maturing rice seeds, abundant levels of the MIPS transcript were observed in the scutellum and aleurone layer.

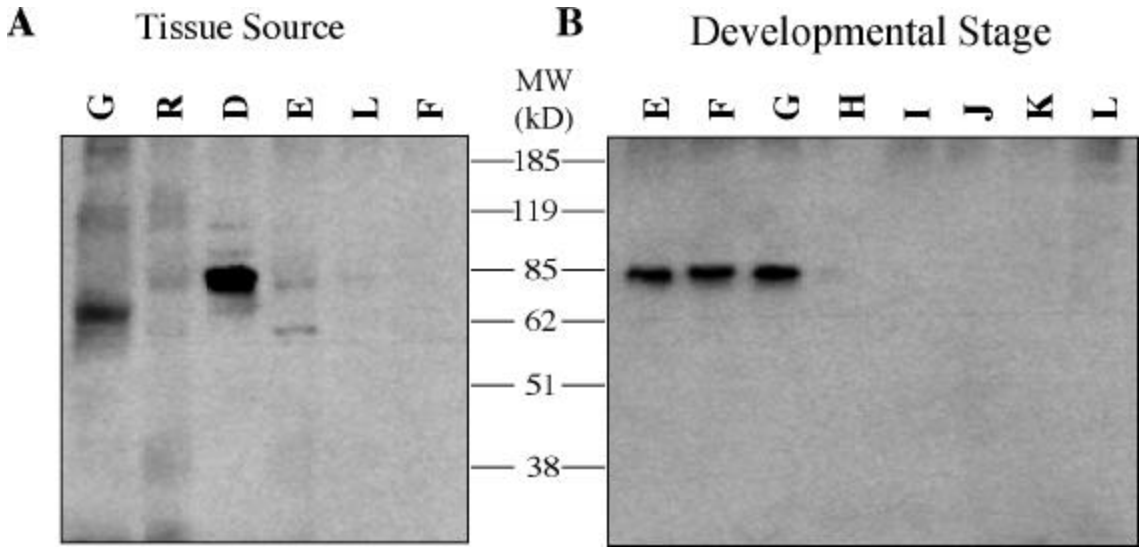


Figure 5. Immunodetection of soybean MIPS protein. Antibody produced against the soybean MIPS fusion protein was used to examine protein expression patterns in soybean protein extracts. Panel A, protein extracts from soybean tissues. G, germinating cotyledons; R, young roots; D, developing seeds; E, globular somatic embryos; L, young leaves; F, flowers. Panel B, lanes E through L represent protein extracts from soybean cotyledons at sequential stages of development as classified by Meinke et al. (1981). Size markers used were the BenchMark Prestained Protein Ladder (Life Technologies, Rockville, MD).

Raboy and Dickinson (1987) measured phytic acid and phosphate levels throughout soybean development. Separation of seed extracts using anion exchange chromatography showed that phytic acid levels increased steadily until late in seed maturation (18.7 to 33.6 µg per seed per day). At seed maturity, very little free phosphate remained in the seed because it had been incorporated into phytic acid. Although Raboy and Dickinson (1987) reported accumulation of phytic acid throughout seed maturation, our examination of MIPS protein levels during seed development indicated high level expression only in early cotyledonary stages. This suggests that conversion of D-glucose-6-P to D-*myo*-inositol-3-phosphate occurs earlier in seed development than accumulation of the final product, phytic acid. Identification of kinases and their expression patterns will provide a more detailed picture of the steps in phytic acid biosynthesis.

Potential for seed-specific MIPS isoforms

Further evidence for seed specific MIPS expression is provided by studies on low phytic acid (*lpa*) mutants in maize (*Zea mays*). Reduction in phytate levels by 50-95% has been achieved in *lpa* mutants following EMS mutagenesis (Larson and Raboy, 1999; Raboy *et al.*, 2000). The *lpa* mutants are viable and several appeared identical to wild type for traits other than the low phytic acid phenotype. One copy of the maize MIPS gene mapped to the same location as the *lpa-1* trait on chromosome 1S. It is possible that maize *lpa-1* represents a mutation affecting seed specific MIPS gene expression, resulting in decreased MIPS activity and lower phytate accumulation in seeds.

The association between the MIPS gene and the *lpa* phenotype is less clear in other cereals. The *lpa-1* mutations in barley (*Hordeum vulgare*) and rice (*Oryza sativa*)

mapped to chromosomes 2H and 2L, respectively (Larson *et al.*, 1998; Larson *et al.*, 2000). Only a single locus with identifiable MIPS homology has been located in both barley and rice, neither of which mapped to the same location as the *lpa* phenotype (Larson and Raboy, 1999; Larson *et al.*, 2000). This raises the possibility that barley and rice may contain other cryptic copies of genes encoding MIPS activity, or alternatively, that the *lpa* phenotype in barley and rice consists of mutations in regulatory genes affecting MIPS expression or enzyme activity.

In *Arabidopsis thaliana*, two distinct MIPS sequences were identified (Johnson, 1994; Johnson and Burk, 1995) and subsequently verified by their presence in EST and genome sequencing projects. No specific functions have been assigned to either gene in *Arabidopsis*. The two copies may simply be redundant or may perform different physiological functions within the plant. Additional copies of the MIPS gene in other plants may allow for further specialization of MIPS isozymes.

Compared to *A. thaliana*, soybean has a greater capacity for nutrient storage in the seed. The amount of phytate stored in soybean seeds may exceed levels required for normal seed function. Raboy *et al.* (1985) decreased the phytic acid content of soybean seeds to one-third the normal amount by growing maternal plants in low phosphate conditions. Germination and seedling viability were not compromised in the resulting low phytate soybeans. Our data suggest that large-seeded plants such as maize and soybean may utilize a seed-abundant isoform for high level accumulation of phytic acid. Future comparisons of MIPS function in *Glycine soja*, a small-seeded relative of soybean, and teosinte, a wild relative of maize, may address this hypothesis.

Alteration of the MIPS gene is a potential approach for the development of low phytic acid soybeans (Hitz and Sebastian, 1998). Mutants with decreased MIPS activity in soybean seeds contained significantly less seed phytic acid, as well as reduced levels of the inositol derivatives raffinose and stachyose. The MIPS sequence from the mutant soybeans contained a single substitution at nucleotide 1188 of the coding sequence (G to T), resulting in a change from lysine in the wild type to asparagine in the mutant sequence. Other strategies for decreasing MIPS gene expression, such as antisense suppression, may provide additional routes for the future development of soybeans with low phytic acid phenotypes to improve phosphorus availability.

Summary

Our MIPS RNA and protein expression data suggest that *Gm*MIPS1 expression is under the control of a strong, developmentally regulated promoter in immature seeds. This is further supported by the fact that the *Gm*MIPS1 ESTs are abundant and derived primarily from cDNA libraries from immature cotyledons. The observation of at least four bands by high stringency Southern blot analysis is consistent with the presence of four loci containing highly similar MIPS sequences in the soybean genome. It is likely that *Gm*MIPS1 is expressed most strongly in seeds and plays a critical role in phytic acid biosynthesis, while the other three soybean MIPS genes may be involved in other aspects of inositol metabolism. Differential regulation of multiple genes would allow specific MIPS expression during periods of high demand for inositol, such as would be expected during phytic acid accumulation in seed development. Closer examination of the

expression of these sequences in plant tissues using sensitive methods such as differential RT-PCR will aid in elucidating the multiple roles for MIPS in plant cells.

MATERIALS AND METHODS

Plant Material

Soybean plants (*Glycine max* L. Merr. cv. Williams 82), were grown in an environmental chamber (Conviron, Pembina, ND) in 24 h light until approximately 8 weeks of age and transferred to a 16 h light/8 h dark cycle to induce flowering. Somatic embryos were initiated from immature soybean cotyledons and maintained on D20 medium (Finer, 1988). Plant tissues were harvested and stored at -80°C.

DNA and RNA extraction

DNA for restriction endonuclease digestion and Southern blot analysis was isolated as described by Dellaporta *et al*. (1983). Total RNA was isolated from tissue samples (100 mg) using the RNAeasy kit (Qiagen, Valencia, CA). RNA used for RT-PCR amplification was isolated from seeds ranging in size from 2 mm to 10 mm, which were pooled since the precise timing of MIPS expression in seeds was unknown. For northern analysis, the seeds were divided into groups based on length as measured from apical end to basal end of the seed (0-2 mm, 2-4 mm, 4-6 mm, 6-8 mm, 8-10 mm, 10-12 mm).

Isolation and analysis of cDNA and genomic sequences

A soybean MIPS cDNA was isolated from developing seeds using a PCR approach that was based on conservation of MIPS coding sequences among plant species. A multiple sequence alignment of several previously reported plant MIPS cDNA sequences from Genbank (*Arabidopsis thaliana*, *Phaseolus vulgaris*, *Brassica napus*, *Spirodela polyrrhiza*, and *Mesembryanthemum crystallinum*) was performed using the Clustal method (Thompson *et al.*, 1994) in the MegAlign program (Lasergene Software, DNAstar, Madison, WI). Oligonucleotides for PCR amplification are included in on-line supplementary data (see Appendix C).

Oligo-dT primed cDNA synthesis was performed with 5 µg total RNA from pooled seeds using Superscript II reverse transcriptase (3' RACE system, Life Technologies, Rockville, MD). PCR reactions (50 µl) were assembled with Taq Master Mix according to manufacturer's recommendations (Qiagen, Valencia, CA) and amplification was performed with a Robocycler Gradient 40 thermocycler (Stratagene, La Jolla, CA). An initial 496 bp MIPS PCR product was cloned into the *Sma*I site of pTZ19R (Fermentas, Hanover, MD). Dideoxy sequencing was performed using the Sequitherm Excel II kit (Epicenter Technologies, Madison, WI).

Rapid amplification of cDNA ends (RACE) was used to amplify the 5' and 3' ends of the soybean MIPS cDNA sequence (Life Technologies, Rockville, MD). A 412 bp PCR product from the 3' end of the MIPS cDNA was cloned into the *Sma*I site of pTZ19R and sequenced as described above. The 5' end of the MIPS cDNA was amplified in two rounds of 5' RACE. PCR products of 836 bp and 295 bp were cloned into the *Sma*I site of pTZ19R and sequenced. The sequences from the overlapping MIPS

cDNA fragments were assembled into a contiguous sequence using the program SeqMan (DNAstar, Madison, WI). A 1551 bp product containing the full-length MIPS cDNA was amplified by high-fidelity PCR (*Pfu* polymerase; Strategene, La Jolla, CA) and cloned into the *Sma*I site of pTZ19R. The MIPS sequence was submitted to Genbank (Accession AF293970). DNA from soybean hypocotyls was utilized as a template to amplify a genomic product spanning the MIPS coding region and sequenced at the Virginia Tech Sequencing Facility. Alignment of MIPS sequences was performed using the Megalign program in DNAstar (Madison, WI).

DNA and RNA blot analyses

DNA samples (10 μg) were digested with restriction enzymes, separated by agarose gel electrophoresis, and transferred onto a nylon membrane (Schleicher & Schuell, Keene, NH). For use as a hybridization probe, a DNA fragment containing the full-length MIPS coding sequence was gel purified and labeled with α-[^{32}P]-dATP by random priming (Random Primers Kit, Life Technologies, Rockville, MD). High stringency DNA hybridizations and subsequent washes were performed according to the manufacturer's specifications (Ambion, Austin, TX). The membrane was exposed to a storage phosphor screen for 2 – 24 hours, which was scanned by a Storm 860 imager (Molecular Dynamics, Amersham, Piscataway, NJ).

The expression of the MIPS gene in developing seeds and other tissues was analyzed by northern blot hybridization. RNA (10 μg) from staged developing seeds and other soybean tissues (germinated cotyledons, young roots, somatic embryos, young leaves, and flowers) was separated by formaldehyde agarose gel electrophoresis (1.2%

agarose) prior to transfer to nylon membranes (Schleicher & Schuell, Keene, NH). Duplicate sets of lanes were loaded and one set was stained with ethidium bromide to confirm sample integrity and equal loading of samples. The blots were probed and washed as described for Southern hybridizations. Detection of hybridizing RNA bands was performed using the phosphorimager as described above.

Protein expression and immunoblot analyses

For high level bacterial expression and subsequent protein purification, a fragment containing the majority of the MIPS coding sequence was fused in frame to a 6X histidine sequence in the expression plasmid pET-32a (+) (Novagen, Madison, WI). A 1425 bp portion of the MIPS cDNA encoding amino acid residues 24 to 498 was amplified by high fidelity PCR from a MIPS cDNA and cloned into the pET-32a (+) vector. The resulting plasmid was introduced into the bacterial strain BL21 (DE3) pLysS (Novagen, Madison, WI). The constructs were provided to the laboratory of Dr. Glenda Gillaspy (Virginia Tech) for production and purification of the MIPS fusion protein (Styer, 2000) which was shipped to a commercial facility (Cocalico Biologicals, Reamstown, PA) for polyclonal antibody production.

Total protein was extracted from developing seeds, flowers, germinating cotyledons, leaves, roots, and somatic embryos. Immature seed samples were grouped by developmental stages as described previously (Meinke *et al.*, 1981). Stages collected included: E = 2.6-3.4 mm, F = 3.6-4.4 mm, G = 4.6-5.4 mm, H = 5.6-6.4 mm, I = 6.6-7.4 mm, J = 7.6-8.4 mm, K = 8.6-9.4 mm, and L = 9.6-10.4 mm. Protein was extracted with homogenization buffer (0.15M KCl, 50mM Tris-HCl, 5mM EDTA) in the presence of a

plant protease inhibitor cocktail (Sigma P-9599, St. Louis, MO) and 0.1% Triton X-100. Protein was quantified according to a BSA standard curve using a protein assay kit (BioRad, Hercules, CA).

For immunodetection of MIPS, protein extracts were subjected to SDS-polyacrylamide gel electrophoresis and subsequently transferred to membranes according to manufacturer's specifications (BioRad, Hercules, CA). Membranes were incubated with primary antibody and secondary (horseradish peroxidase conjugated goat-anti-rabbit; Sigma, St. Louis, MO) antibody at 1:10,000 dilutions. Protein bands were detected using the chemiluminescent ECL Plus Western Blotting Detection System according to the manufacturer's specifications (Amersham, Piscataway, NJ).

ACKNOWLEDGEMENTS

We thank Regina Hanlon for excellent technical assistance, and John McDowell and Craig Nessler for critical review of the manuscript.

For References see combined reference section following Chapter IV.

CHAPTER III

Isolation and characterization of a MIPS gene family in soybean

ABSTRACT

Myo-inositol 3-phosphate synthase (MIPS, E.C. 5.5.1.4) catalyzes the first step in *de novo* synthesis of inositol and phosphorylated derivatives such as phytic acid (*myo*-inositol hexa*kis*phosphate, $InsP_6$). Phytic acid serves as a storage molecule in plant seeds, where it typically accounts for 60-80% of total seed phosphorus. In soybean (*Glycine max*), MIPS is encoded by a highly similar gene family of at least four members (GmMIPS1 through GmMIPS4), at least one of which is highly expressed in developing seeds. GmMIPS1 and GmMIPS2 genomic and promoter regions were isolated via PCR. GmMIPS1 and GmMIPS2 sequences are highly similar, even within non-coding regions. A fragment of the GmMIPS1 promoter was found to drive expression of a promoter:GUS fusion construct in developing seeds and in soybean suspension cells, suggesting that the promoter may not be specific to seed tissue. Immunolocalization experiments showed that MIPS protein accumulated in developing seeds as early as the heart stage of embryogenesis and dramatically decreased by the cotyledonary stage, paralleling results previously observed by northern and western blot analysis.

INTRODUCTION

D-*Myo*-inositol 3-phosphate synthase (MIPS) catalyzes the conversion of D-glucose-6-phosphate to D-*myo*-inositol-3-phosphate, as the first step in the *de novo* synthesis of inositol. In plants, inositol and inositol derivatives are important in a variety of cellular functions, including signal transduction, cell wall synthesis, stress protection, and membrane biogenesis (reviewed in Loewus and Murthy, 2000). Phytic acid, *myo*-inositol hexa*kis*phosphate, acts as a phosphorus storage molecule in plant seeds. One proposed pathway for synthesis of phytic acid proceeds via sequential phosphorylation of *myo*-inositol-3-phosphate by phosphoinositol kinases. MIPS is thought to be the rate-limiting enzyme in the biosynthesis of phytic acid (reviewed in Raboy, 1990).

In the seeds of legumes and cereals, phytic acid represents a substantial portion of total seed phosphorus, ranging from 60% to 80% (Raboy, 1990). High levels of phytic acid have negative consequences in non-ruminant nutrition. Phytic acid acts as an antinutrient by chelating mineral cations and charged proteins, reducing overall digestibility of feedstuff. Nonruminants are unable to cleave phosphorus molecules from phytic acid, resulting in a requirement for additional sources of inorganic phosphorus for optimal growth. When manure containing undigested phytic acid is applied as fertilizer, soil microorganisms cleave phosphorus from phytic acid, leading to phosphorus loading in soil, and phosphorus pollution in neighboring watersheds. Reducing levels of phytic acid in the seeds of crop plants would increase nutritional value of the seeds and reduce the potential for phosphorus pollution (reviewed in Swick and Ivey, 1992).

One approach to reducing levels of phytic acid in seeds is to reduce flux through the biosynthetic pathway by downregulating or disrupting the activity of an enzyme in the

pathway. Because of its role in the initial synthesis of phytic acid precursor *myo*-inositol-3-phosphate, MIPS is a target for modification. However, due to the involvement of MIPS in other crucial plant functions, constitutive downregulation of MIPS is not desirable. Potato plants expressing a constitutive antisense MIPS were less vigorous than wildtype (Keller *et al*., 1998). An alternate approach to reducing seed phytate levels has been through screens of mutagenized seeds for high inorganic phosphorus and low phytic acid. This approach has been successful in a variety of species. *Lpa1* mutants isolated in rice, maize, barley and soybean have reduced levels of phytic acid, and increased levels of free inorganic phosphorus (Larson *et al*., 2000; Larson and Raboy, 1999, Wilcox *et al*., 2000). The accumulation of inorganic phosphorus without accumulation of phytic acid or less extensively phosphorylated inositol derivates indicates a mutation early in the biosynthesis of inositol, possibly at the level of the MIPS enzyme.

Soybean mutants developed by DuPont with low phytic acid and low levels of raffinose saccharides were found to have a mutation in a MIPS gene, leading to lowered levels of MIPS activity in developing mutant soybean seeds (Hitz and Sebastian, 1998). Although the phenotype of DuPont mutants is similar to that of *lpa1*, *lpa1* cannot be directly attribution to a mutation in the MIPS gene. The maize *lpa1* mutation maps to the same location on chromosome 1S as one of several MIPS genes, however sequencing of the 1S MIPS gene and promoter region have failed to identify any sequence changes (V. Raboy, personal communication). Rice and barley *lpa1* mutations do not map to a known MIPS gene. Because rice and barley have only one region with MIPS sequence homology, a mutation that would reduce or eliminate function of MIPS would likely

prove lethal. Therefore it is likely that the rice and barley *lpa1* phenotypes are due to mutations in a MIPS transcriptional regulator or MIPS associated protein.

Hegeman *et al*. (2001) observed high levels of MIPS expression in developing soybean seeds. A cDNA corresponding to a putative seed-specific MIPS RNA was isolated from developing seeds. This seed-derived member of the soybean MIPS gene family was designated GmMIPS1.

The purpose of this research was to isolate and characterize MIPS coding sequence and associated promoter regions. Previous research indicated that GmMIPS1 is likely the predominant isoform expressed in seeds. The GmMIPS1 gene is therefore a potential candidate for antisense regulation of MIPS expression to reduce phytic acid levels. In this study partial cDNA of GmMIPS3 as well as promoter and genomic regions of GmMIPS1 and GmMIPS2 were isolated and analyzed for sequence similarity. A fragment of the promoter region of GmMIPS1 was found to drive expression in developing soybean seeds and suspension cell culture. MIPS protein was observed to accumulate early in soybean seeds development and in those tissues where phytic acid is known to accumulate.

RESULTS AND DISCUSSION

Comparison of GmMIPS1, GmMIPS2 and GmMIPS3

The complexity of the MIPS gene family in soybean was determined though a combination of analysis of MIPS EST sequences and Southern blot analysis. The identification of multiple MIPS sequences in a soybean EST database search confirmed that at least four MIPS gene family members are transcribed in soybean (Hegeman *et al.*,

2001). For further characterization of the MIPS gene family, soybean cDNA and genomic MIPS sequences were isolated as described in Materials and Methods. EST data suggested that GmMIPS1 and GmMIPS2 isoforms shared the highest degree of sequence similarity among members of the MIPS gene family. Genomic sequence data confirmed this observation (Figure 6). When analyzed by the Clustal W alignment method (Thompson *et al*., 1994), GmMIPS1 and GmMIPS2 genomic sequences spanning the coding region exhibited 92% identity overall. Intron and exon sequences within GmMIPS2 genomic sequence were assigned based on GmMIPS1 cDNA and GmMIPS2 EST sequences. The number, size and sequence of predicted introns within GmMIPS2 were similar to GmMIPS1 (Figure 6). Sequence identity for individual exons ranged from 95 to 99%. Overall, the sequence of the coding regions of GmMIPS1 cDNA and GmMIPS2 predicted cDNA were 96% identical. Comparison of intron sequences between GmMIPS1 and GmMIPS2 revealed 64 to 90% identity. In comparison, genomic sequences spanning the coding region of two MIPS isoforms in *Arabidopsis* aligned by the Clustal W method share only 78% identity.

GmMIPS3 EST sequences were found in libraries from flowers, immature flowers, germinating cotyledons, leaves and buds (Hegeman *et al.*, 2001). GmMIPS3 EST sequences and 5'RACE sequence were utilized to predict the sequence of a full length GmMIPS3 cDNA. The sequence of the coding region of GmMIPS1 cDNA and GmMIPS3 cDNA were 90% identical (see Appendix J).

Primers designed to the 5' and 3' end of the GmMIPS3 cDNA sequence were utilized to amplify a putative genomic copy of GmMIPS3. A single product of approximately 1.7 kb was observed in several reactions under varying conditions (Figure

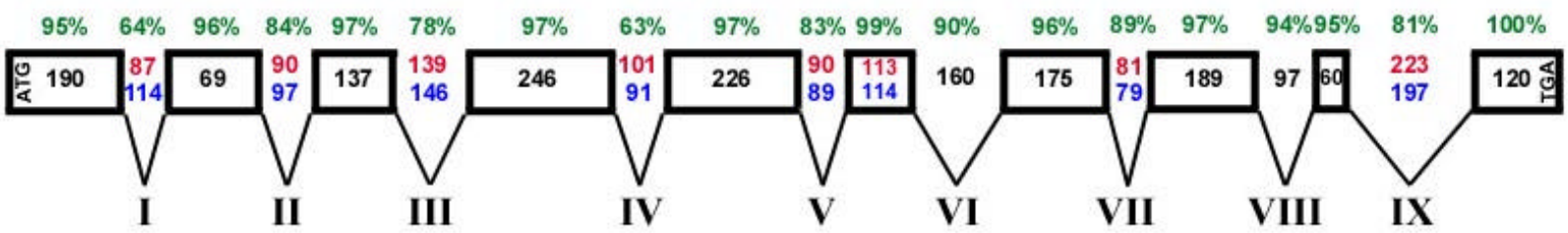


Figure 6. Comparison of genomic structure for GmMIPS1 and GmMIPS2..
Closed boxes represent exons. Introns are indicated by Roman numerals (I-IX). Size of exons (in bp) are indicated in boxes. Red numerals represent GmMIPS1 intron and exon sizes (in bp). Blue numerals represent GmMIPS2 intron and exon sizes (in bp). Black numerals represent intron and exon sizes that are identical between the two isoforms. Percentage nucleotide sequence identity between GmMIPS1 and GmMIPS2 is indicated in green above the corresponding region.

7). The size similarity of this GmMIPS3 PCR product to cDNA from GmMIPS1 (containing a 1.53 kb coding region) indicated that GmMIPS3 may lack introns or containing very small introns compared to GmMIPS1 and GmMIPS2. Precedent for an intronless genomic copy among related gene family members has been observed for inositol monophosphatase in tomato (G. Gillaspy, personal communication). A single clone of the PCR product was sequenced but did not contain MIPS homologous sequence.

GmMIPS1 cDNA was utilized to predict an amino acid sequence of a 56.5 kDA protein (Statistical Analysis of a Protein Sequence, Biology Workbench). The predicted protein sequences from GmMIPS2 and GmMIPS3 cDNAs were similar in size, 56.4 and 56.5 kDa, respectively. The predicted amino acid sequence of soybean MIPS isoforms were compared with other plant MIPS protein sequences with the CLUSTALTREE program, a phylogenetic analysis with Clustal W (Thompson *et al*., 1994). An unrooted phylogenetic tree was created through alignment of MIPS sequences from soybean and other plant sources (Figure 8). The evolutionary relationship of sequences was inferred from observation of the branching relationships within the tree. GmMIPS3 is divergent from the highly similar GmMIPS1 and GmMIPS2 sequences, as shown in Figure 8. MIPS proteins from soybean are more similar to each other than MIPS from any other plant (Figure 8). In comparison, two MIPS isoforms from *Arabidopsis thaliana*, AtMIPS1 and AtMIPS2, branch separately, inferring a greater divergence of the sequences from each other. This indicates that the sequence similarity between soybean MIPS gene family members is not characteristic of MIPS in all plants.

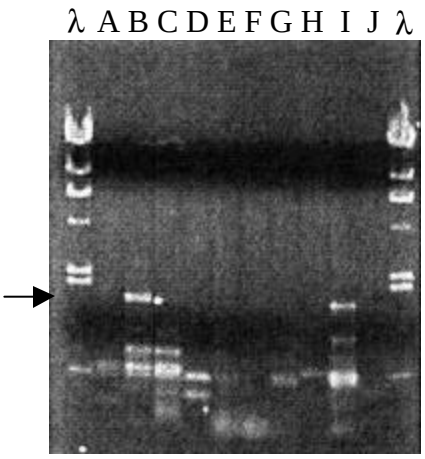


Figure 7. Genomic PCR amplification of GmMIPS3
A-J – PCR products resulting from reactions using Failsafe PCR Premix buffers A through J. PCR products were separated by electrophoresis in 0.8% agarose gels. ?- Lambda size marker (cut with *Hind*III). Arrow indicates a 1.7 kb product seen with Premix buffer B and I.

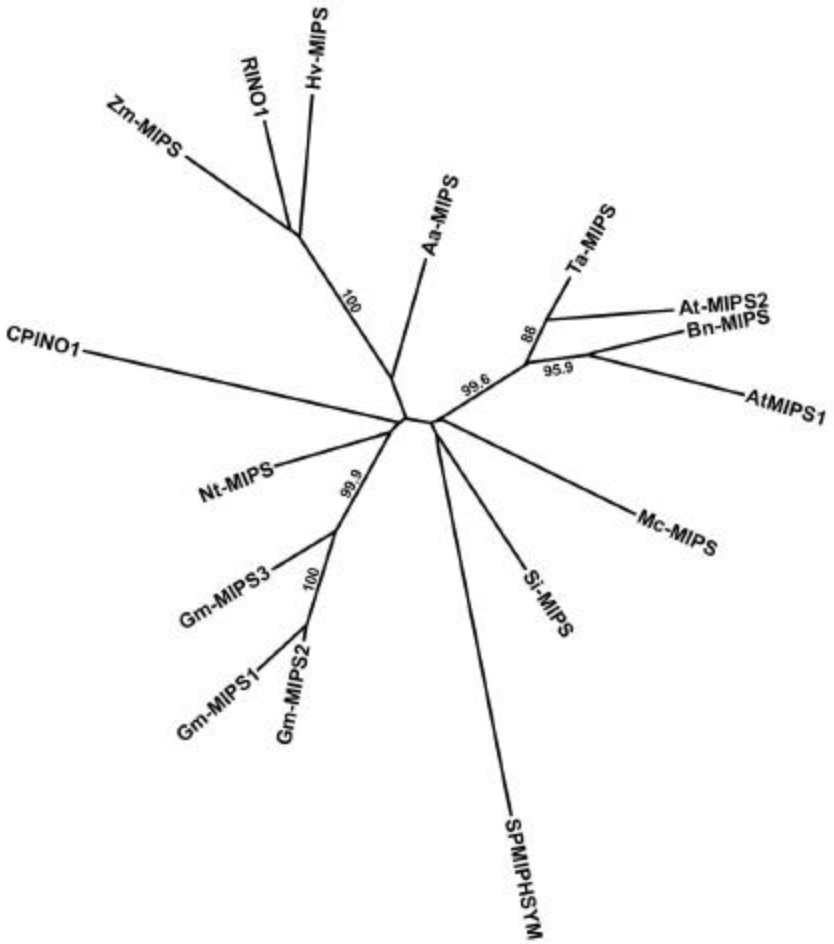


Figure 8. Unrooted phylogenetic tree of predicted MIPS protein sequences. Sequences were subjected to Clustal W multisequence alignment and PHYLIP (DRAWTREE, Biology Workbench, http://workbench.sdsc.edu). Confidence values at branch points were expressed as percentages of 1000 bootstrap replicates in the CLUSTALTREE program (Biology WorkBench)

Abbreviation	Name	Common Name	Accession
AaMIPS	*Actinidia arguta*	Kiwifruit	AY005128
AtMIPS1	*Arabidopsis thaliana*	Thale cress	U04876
AtMIPS2	*Arabidopsis thaliana*	Thale cress	U30250
BnMIPS	*Brassica napus*	Cabbage	U66307
CPINO1	*Citrus x paradisi*	Grapefruit	Z32632
GmMIPS1	*Glycine max*	Soybean	AF293970
GmMIPS2	*Glycine max*	Soybean	To be assigned
GmMIPS3	*Glycine max*	Soybean	To be assigned
HvMIPS	*Hordeum vulgare*	Barley	AF056325
McMIPS	*Mesembryanthemum crystallinum*	Iceplant	U32511
NtMIPS	*Nicotiana tabacum*	Tobacco	AB009881
RINO1	*Oryza sativa*	Rice	AB012107
SiMIPS	*Sesamum indicum*	Sesame	AF284065
SPMIPHSYM	*Spirodela polyrrhiza*	Duckweed	Z11693
TaMIPS	*Triticum aestivum*	Bread wheat	AF120146
ZmMIPS	*Zea mays*	Corn	AF056326

Isolation and sequence characterization of GmMIPS promoter regions

We previously speculated that GmMIPS1 may encode a MIPS isoform that is expressed in developing seeds to provide a supply of inositol for phytic acid synthesis. Antisense regulation of seed phytic acid synthesis would require that the antisense construct be transcribed synchronously with the seed-specific GmMIPS1 gene. To address the possibility of specific expression, promoter sequences were isolated for GmMIPS1 and GmMIPS2 by Inverse PCR (IPCR). Initial IPCR attempts amplified only a small 407 bp fragment of GmMIPS1 sequence upstream of the ATG. Subsequent attempts with different restriction enzymes and PCR conditions provided a total of 1241 bp of GmMIPS1 sequence upstream of the ATG. IPCR was also utilized to amplify a 2751 bp promoter fragment from GmMIPS2. Sequence comparison of the two promoter regions revealed surprising sequence similarity. When analyzed by the Clustal W alignment method, the 1241 bp upstream from the ATG (including promoter and untranslated region) shared 37% identity.

Previous work utilizing GmMIPS1 cDNA as a probe for Southern blot analysis indicated that the MIPS gene family in soybean was highly similar (Hegeman *et al.*, 2001, see also Figure 3 and Figure 9). Multiple hybridizing bands were visible even after high stringency hybridization. Promoter regions from GmMIPS1 and GmMIPS2 were utilized as probes in Southern blot analyses in an attempt to assign locations of different MIPS genes to individual bands. Southern blot analysis of soybean genomic DNA probed with GmMIPS1 upstream region (407 bp) under high stringency conditions indicated the presence of two fragments with several different restriction enzyme digests (Figure 9). The restriction sites utilized in the analysis were not present within the total

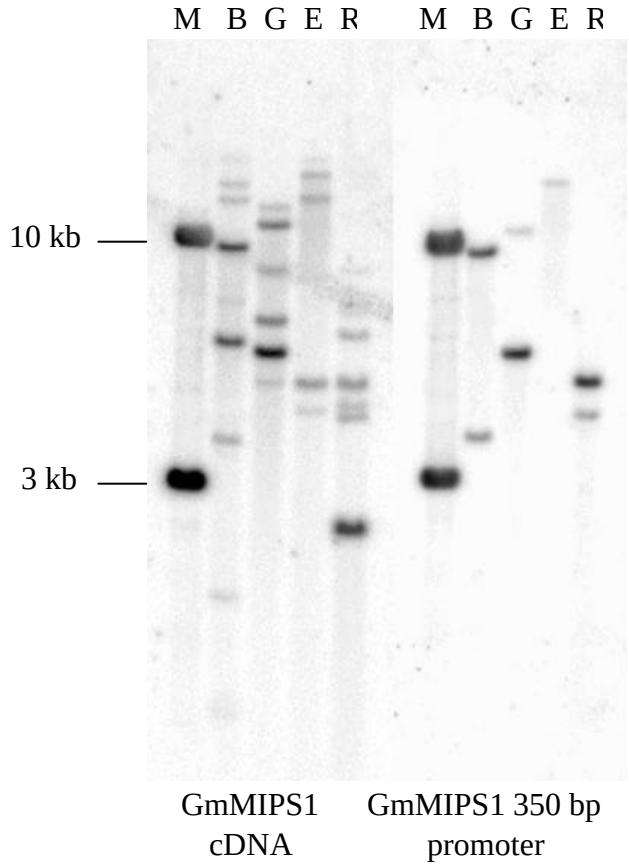


Figure 9. Southern blot analysis of soybean genomic DNA using GmMIPS1 cDNA and GmMIPS1 promoter probes. Genomic DNA was digested with appropriate restriction enzymes, separated by electrophoresis in 0.8% agarose gels and transferred by capillary blotting to nylon membranes. Blots were probed under high stringency hybridization conditions with either 1.7 kb GmMIPS1 cDNA or 350 bp GmMIPS1 promoter.

M-marker (Gibco BRL High DNA Mass Ladder), B-*Bam*HI, G-*Bgl*II, E-*Eco*RI, R-*Eco*RV

1241 bp GmMIPS1 upstream sequence. The band with strongest hybridization intensity with the GmMIPS1 promoter probe corresponded to the strongest hybridizing band observed in an identical blot probed with the full length GmMIPS1 cDNA. A second, fainter band observed with GmMIPS1 promoter corresponded to a band also observed in the cDNA-probed blot. The promoter probe from GmMIPS1 was designed to act as a gene specific probe under high stringency conditions. The presence of two hybridizing indicates that the GmMIPS1 promoter probe cross-hybridized with other MIPS gene family members, most likely GmMIPS2. The data indicated that sequence conservation in the MIPS gene family extends to include at least portions of promoter region.

Southern analysis of the promoter region of GmMIPS2 also revealed multiple hybridizing fragments, despite the fact that only *Eco*RV cuts within the 2751 bp upstream region (Figure 10). Some of the bands in the blot probed with GmMIPS2 promoter were similar in size to GmMIPS1 promoter fragments. The shared hybridizing bands in GmMIPS1 and GmMIPS 2 promoters under high stringency conditions reaffirms high levels of identity between the two promoter regions, and indicates that neither can be used as a gene-specific probe.

Southern analysis of the 717 bp 5' RACE product of GmMIPS3, which included 102 bp of untranslated region upstream of the ATG, resulting in a multiple banding pattern identical to that observed with GmMIPS1 cDNA (Figure 10), indicating that cDNA fragments from both GmMIPS1 and GmMIPS3 cross-hybridize with other members of the gene family.

GmMIPS1 and GmMIPS2 promoter sequences were subjected to analysis using PlantCARE, a plant *cis*-acting regulatory element database (Rombauts *et al.*, 1999).

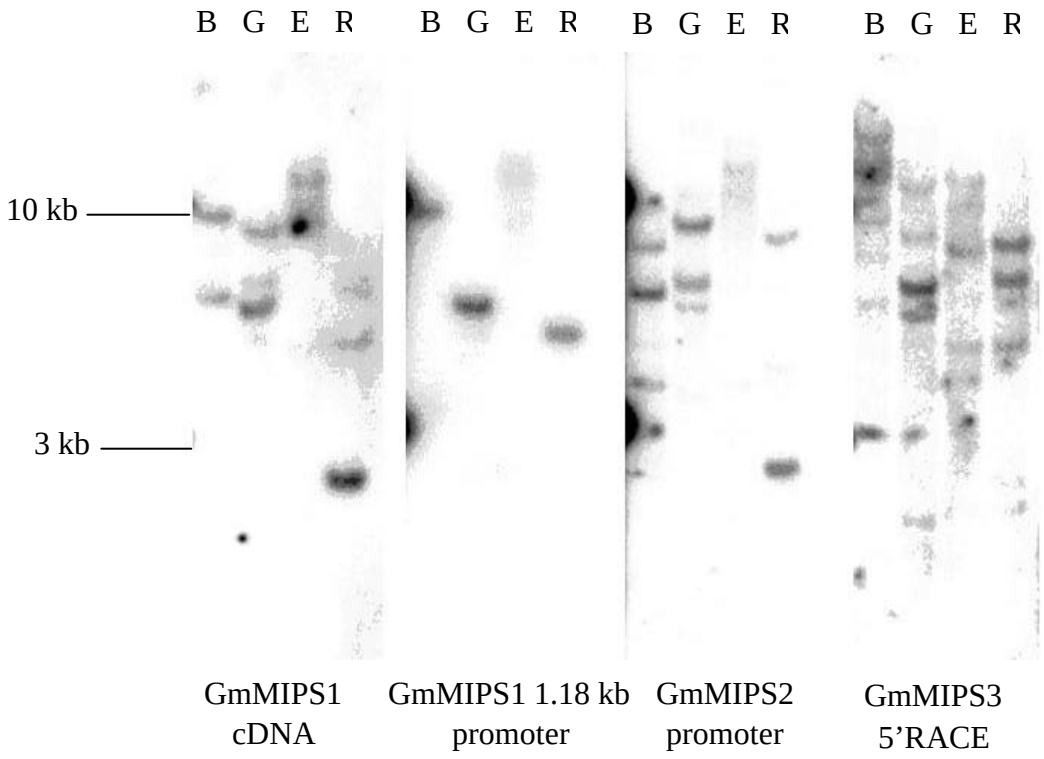


Figure 10. Southern blot analysis of soybean genomic DNA using GmMIPS1 cDNA, GmMIPS1 promoter, GmMIPS2 promoter and Gm-MIP3 5'RACE probes. Genomic DNA was digested with appropriate restriction enzymes, separated by electrophoresis in 0.8% agarose gels and transferred by capillary blotting to nylon membranes. Blots were probed under high stringency hybridization conditions with either 1.7 kb GmMIPS1 cDNA, 1.18 kb GmMIPS1 promoter, 2.75 kb GmMIPS2 promoter or 0.8 kb GmMIPS3 5'RACE product.

B-*Bam*HI, G-*Bgl*II, E-*Eco*RI, R-*Eco*RV

Putative elements in MIPS promoter sequences were identified by similarity to database sequences. The PlantCARE database contains elements of known function that interact with transcription factors to enhance or repress gene expression, elements with suspected function, and elements whose function is not known. Putative elements responsive to abcissic acid, methyl-jasmonate, auxin, gibberilic acid, salicyclic acid, wound and fungal elicitation were identified in GmMIPS1 promoter sequence (see Figure 11 and Table 2). Light responsive elements were observed most frequently. This correlates well with reports of strong diurnal regulation of MIPS transcript levels and enzyme activity in iceplant, green bean, potato and rice (Ishitani *et* al., 1996; Johnson and Wang, 1996; Keller *et al.*, 1998; Raychaudhuri and Majumder, 1996). Two OS-Prolamin boxes and a SKN1 motif involved in the endosperm-specific expression of the rice (*Oryza sativa*) glutelin Gt3 gene (Washida *et al*., 1999, and Wu *et al.*, 2000) are present in the promoter sequence. Glutelin functions as a seed-specific storage protein in rice.

Putative regulatory elements in the promoter region of GmMIPS2 were also identified using the PlantCARE database (see Figure 11). Elements identified were similar to those identified in the GmMIPS1 promoter, but their location and arrangement were not identical. Light responsive elements were also the most frequently observed, suggestive of transcriptional regulation leading to diurnal fluctuations of MIPS RNA levels observed by others. A OS-Prolamin box and two SKN1 motifs involved in the endosperm-specific expression of the rice (*Oryza sativa*) glutelin Gt3 gene are also present in GmMIPS2 promoter sequence. Putative elements responsive to auxin, anoxia, ABA, ethylene, cell cycle, gibberellin, methyl-jasmonate, heat shock and low temperature were also observed (see Table 3).

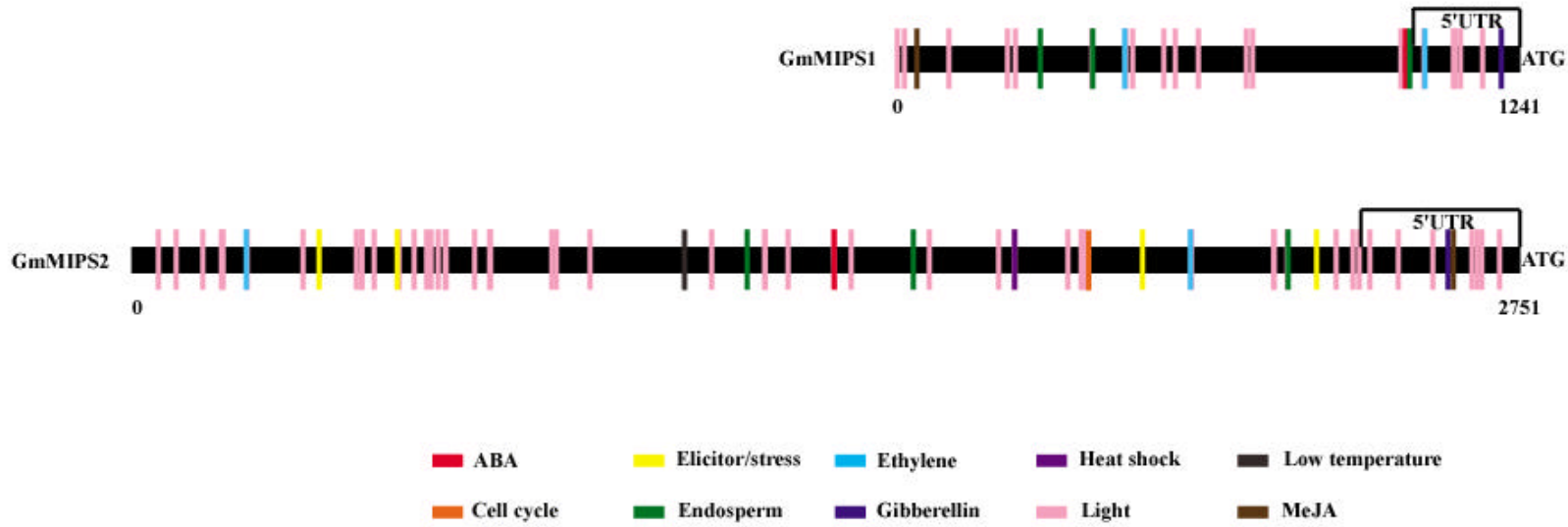


Figure 11. GmMIPS promoter region illustrating putative *cis*-acting regulatory elements. Elements identified using PlantCARE database (Rombauts *et al*., 1999).

Table 2. GmMIPS1 promoter and 5' UTR *cis*-acting elements				
Classification	Name	Position and Strand	Source Organism	Reference
ABA	ABRE	1110-	*Zea mays*	Polidoros and Scandalios, 1999
Endosperm	Prolamin box	390-,1019-	*Oryza sativa*	Washida *et al*., 1999
	Skn-1 motif	287-	*Oryza sativa*	Washida *et al*., 1999
Ethylene	ERE	453+, 1050-	*Dianthus caryophyllus*	Itzhaki *et al*., 1994
Gibberrillin	P box	1201+	*Oryza sativa*	Kim *et al*., 1992
Light	AE box	552+	*Arabidopsis thaliana*	Conley *et al*., 1994; Park *et al*., 1996
	AAGAA motif	389-	*Avena sativa*	Bruce *et al*., 1991
	ACE	1108+	*Pertroselinum crispum*	Feldbrugge *et al*., 1996
	Box 4	471	*Pertroselinum crispum*	Arguello-Astorga *et al*., 1996
	Box I	234+, 454+, 1050-	*Pertroselinum crispum*	Arguello-Astorga *et al*., 1996
	CATT motif	1121+	*Zea mays*	Arguello-Astorga *et al*., 1996
	G box	1110+	*Pisum sativum*	Arguello-Astorga *et al*., 1996
		1165-	*Hordeum vulgare*	Arguello-Astorga *et al*., 1996
	GA motif	1001+	*Glycine max*	Arguello-Astorga *et al*., 1996
	GAG motif	103+	*Spinacia oleracea*	Arguello-Astorga *et al*., 1996
	I Box	220-, 694+	*Pisum sativum*	Arguello-Astorga *et al*., 1996
		532-, 600-, 706-	*Solanum tuberosum*	Arguello-Astorga *et al*., 1996
	Lamp element	13-	*Spinacia oleracea*	Arguello-Astorga *et al*., 1996
	TCCC motif	2-	*Spinacia oleracea*	Arguello-Astorga *et al*., 1996
MeJA	CGTCA motif	40-	*Hordeum vulgare*	Rouster *et al*., 1997

Table 3. GmMIPS2 promoter and 5' UTR *cis*-acting elements				
Classification	Name	Position and Strand	Source Organism	Reference
ABA	ABRE	182-	*Zea mays*	Polidoros and Scandalios, 1999
		1396+	*Helianthus annus*	Prieto-Dapena *et al*., 1999
Cell cycle	MSA-like	1899+	*Pertroselinum crispum*	Ito *et al*., 1998
Elicitor/Stress	AT-rich sequence	369-	*Pisum sativum*	Seki *et al*., 1996
	ELI-box 3	2006+	*Brassica oleracea*	Pastuglia *et al*., 1997
	TC-rich repeats	527-, 2351-	*Nicotiana tabacum*	Klotz *et al*., 1996
Endosperm	Prolamin box	2295+	*Oryza sativa*	Washida *et al*., 1999
	Skn-1 motif	1223+, 1550-	*Oryza sativa*	Washida *et al*., 1999
Ethylene	ERE	230+, 2104+	*Dianthus caryophyllus*	Itzhaki *et al*., 1994
Gibberrillin	TATC box	2617+	*Oryza sativa*	Washida *et al*., 1999
Heat shock	HSE	1755+	*Brassica oleracea*	Pastuglia *et al*., 1997
Light	3-AF1 binding site	58+, 1888+	*Solanum tuberosum*	unpublished
	AAGAA motif	389-	*Avena sativa*	Bruce *et al*., 1991
	AE box	552+	*Arabidopsis thaliana*	Conley *et al*., 1994; Park *et al*., 1996
	AT1 motif	610+	*Solanum tuberosum*	unpublished
	ATC motif	2272-, 2587-, 2662+	*Arabidopsis thaliana*	Arguello-Astorga *et al*., 1996
	ATCC motif	2270-	*Pisum sativum*	Arguello-Astorga *et al*., 1996
	Box 4	469+, 588+, 682-, 711-, 2392+, 2517+	*Pertroselinum crispum*	Arguello-Astorga *et al*., 1996
	Box I	231+, 2105+	*Pertroselinum crispum*	Arguello-Astorga *et al*., 1996
	Box II	91+, 1259+, 1430+, 1722-	*Arabidopsis thaliana*	Arguello-Astorga *et al*., 1996
	CATT motif	2719+	*Zea mays*	Arguello-Astorga *et al*., 1996
	G box	182+, 1396-	*Pisum sativum*	Arguello-Astorga *et al*., 1996
		1586-	*Triticum aestivum*	Arguello-Astorga *et al*., 1996
		180+	*Brassica oleracea*	Pastuglia *et al*., 1997
		182+	*Nicotiana tabacum*	Klotz *et al*., 1996
		1260+, 1305-	*Zea mays*	Manjunath *et al*., 1997
		1395+	*Brassica napus*	Keddie *et al*., 1992
		1396+	*Helianthus annuus*	Prieto-Dapena *et al*., 1999
		2672-	*Hordeum vulgare*	Arguello-Astorga *et al*., 1996
	GA motif	142+, 1152+	*Glycine max*	Arguello-Astorga *et al*., 1996
		459-	*Arabidopsis thaliana*	Arguello-Astorga *et al*., 1996
	GT1 motif	598+, 834-	*Arabidopsis thaliana*	Vauterin *et al*., 1999
	Gap-box	2428+	*Arabidopsis thaliana*	Arguello-Astorga *et al*., 1996
	I Box	459-, 1890+	*Pisum sativum*	Arguello-Astorga *et al*., 1996
		341+, 483+, 834+, 626-, 714+, 843-, 1891-, 2460-, 2682+	*Solanum tuberosum*	Arguello-Astorga *et al*., 1996
		627+	*Arabidopsis thaliana*	Arguello-Astorga *et al*., 1996
		1862-	*Flaveria trinervia*	Arguello-Astorga *et al*., 1996
	MRE	911-	*Petroselinum crispum*	Feldbrugge *et al*., 1997
	TCT motif	563-, 2440-	*Arabidopsis thaliana*	Arguello-Astorga *et al*., 1996
Low temperature	LTR	1100+	*Hordeum vulgare*	Dunn *et al*., 1998
MeJA	CGTCA motif	2626+	*Hordeum vulgare*	Rouster *et al*., 1997

Promoter:GUS fusions

Promoter:GUS fusions were prepared to characterize promoter specificity of GmMIPS1. Promoter driven expression was assayed by bombardment of developing seeds and soybean root cell suspension culture with promoter:GUS constructs. Promoter fragments were used to generate different fusions containing a promoterless β-glucuronidase (GUS) gene fused to promoter fragments from GmMIPS1, the seed-storage protein β-conglycinin, or the constitutive cauliflower mosaic virus (CaMV) 35S (Figure 12). The promoter:GUS constructs were bombarded directly into immature soybean cotyledons. Promoter activity was assayed through observation of GUS staining after bombardment. As shown by preliminary results in Figure 13, a small fragment of GmMIPS1 upstream region (335 bp beginning 61 bp upstream of the ATG) was sufficient to drive GUS expression in early seed development (2-3 mm soybean seeds). Expression of GUS was also observed in early development using a larger 1180 bp GmMIPS1 promoter fragment. No expression was observed for 35S, β-conglycinin or no-promoter control constructs or in later stages of seed development (data not shown). For comparison to expression in developing seeds soybean cell suspension culture were bombarded with the promoter constructs. GUS staining was observed in suspension cells bombarded with GmMIPS1 promoter fragments (335 and 1180 bp) and the CaMV 35S promoter (Figure 14). No expression was observed for β-conglycinin or no-promoter control constructs(data not shown). Although the promoter fragments isolated from GmMIPS1 were sufficient to drive promoter activity, these preliminary results suggest that they were unable to confer seed specificity. The activity of the GmMIPS2 promoter in similar bombardment experiments has not yet been undertaken.

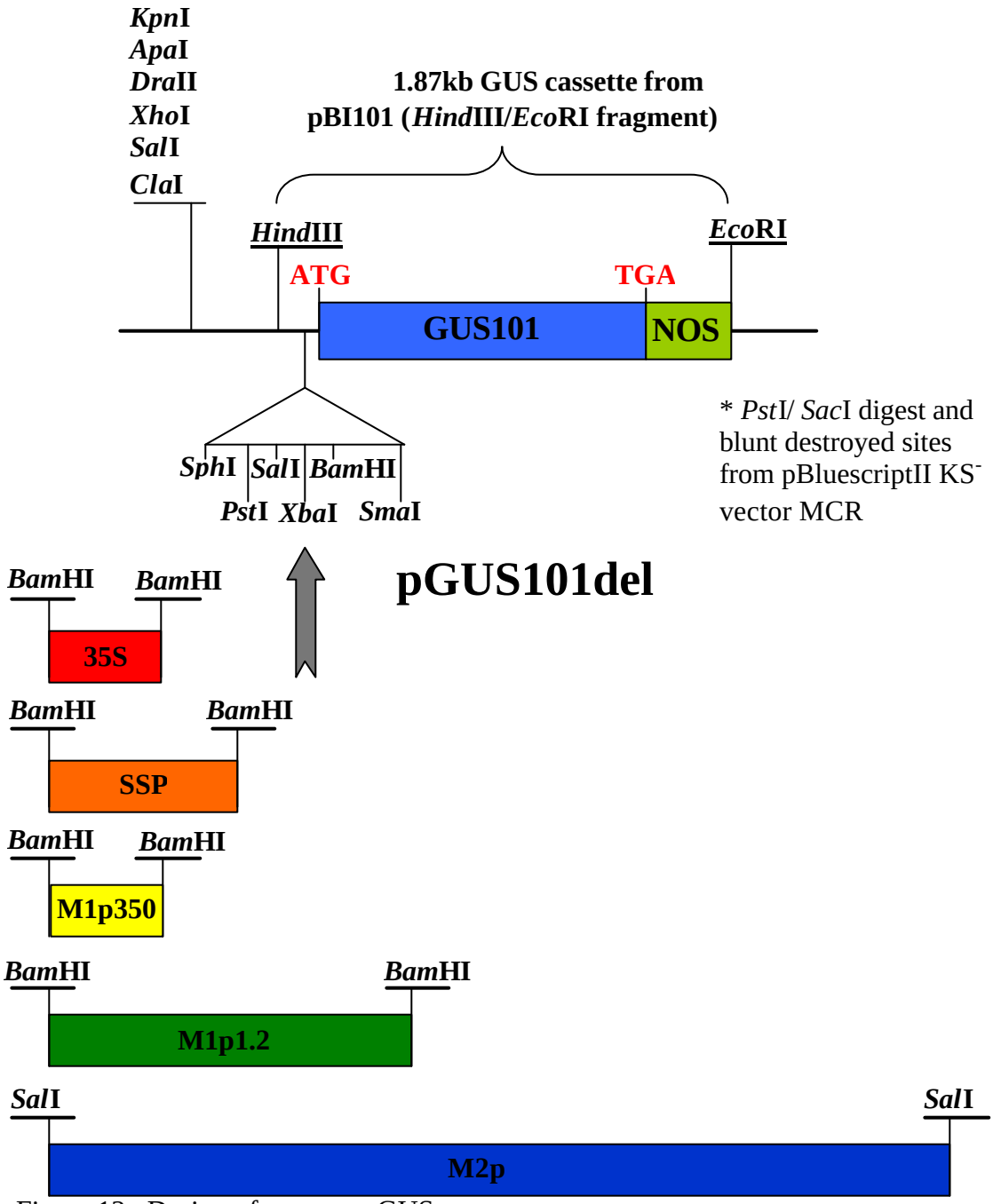


Figure 12. Design of promoter:GUS constructs.
Promoter fragments were amplified with *Taq* DNA Polymerase and cloned into an intermediate vector (pBluescriptII KS-). Fragments were removed from intermediate vector by digestion with *Bam*HI or *Sal*I and cloned into pGUS101del.
Abbreviations: 35S) Cauliflower Mosaic Virus (CaMV) 35S promoter, SSP) ß-conglycinin seed specific promoter, M1p350) GmMIPS1 335 bp promoter, M1p1.2) GmMIPS1 1180 bp promoter, M2p) GmMIPS2 2751 bp promoter

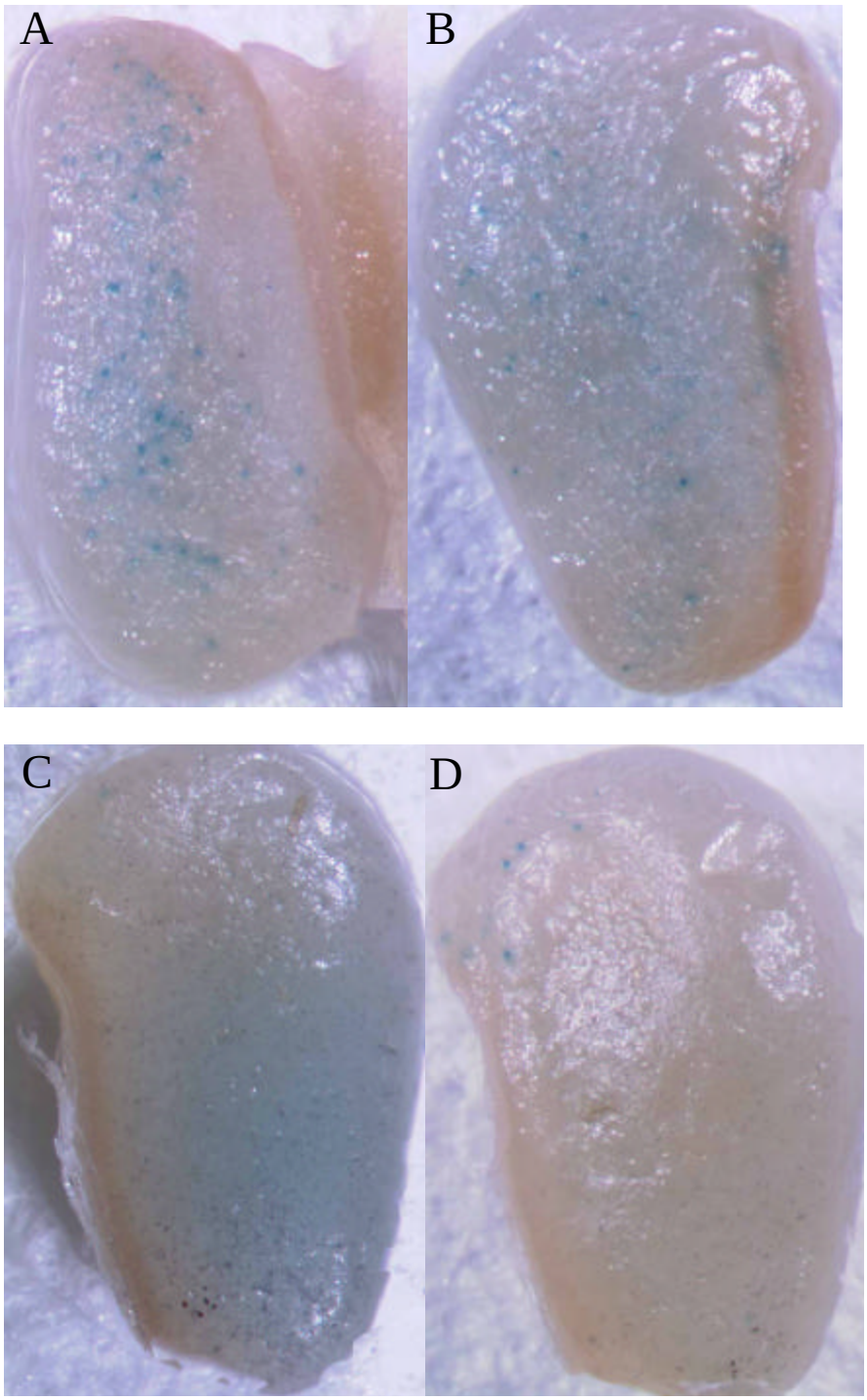


Figure 13. Expression of promoter:GUS constructs in bombarded developing soybean seeds.
Developing 'Essex' soybean seeds (2-3 mm) were bombarded with promoter:GUS DNA constructs. GUS expression was visualized by staining for ß-glucuronidase activity using X-Gluc substrate.

A & B. GmMIPS1 335 bp promoter

C & D. GmMIPS1 1180 bp promoter

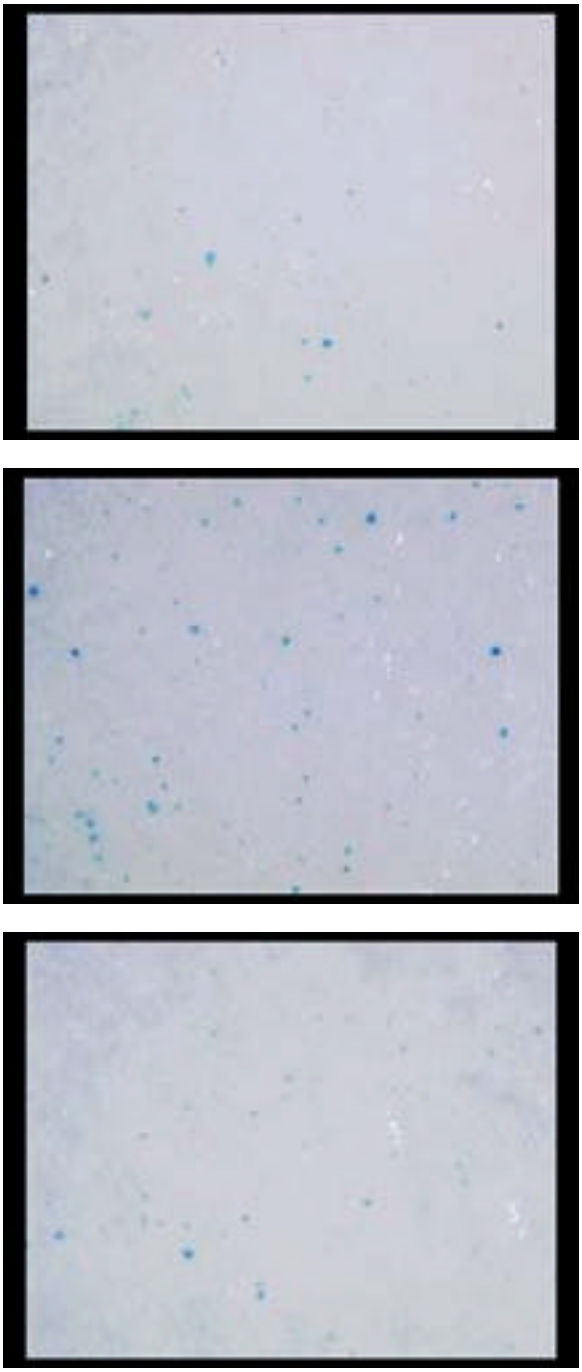


Figure 14. Expression of promoter:GUS constructs in soybean cell suspension cultures. 'Williams 82' soybean cell suspension cultures were bombarded with promoter:GUS DNA. GUS expression was visualized by staining for ß-glucuronidase activity using X-Gluc substrate.
A. CaMV 35S promoter
B. GmMIPS1 335 bp promoter
C. GmMIPS1 1180 bp promoter

Chromosomal mapping of GmMIPS1 and GmMIPS2

GmMIPS1 cDNA was utilized as a probe to map the MIPS1 locus in a segregating soybean population, consisting of a cross of the Virginia breeding line V71-370 with a *Glycine soja* introduction PI 407162. Multiple bands were observed when GmMIPS1 cDNA was used in RFLP mapping, similar to the multiple bands observed when the cDNA was utilized as a probe in Southern hybridization (Figure 3 and Figure 9). The darkest hybridizing band in cDNA probed blots was polymorphic and mapped to the end of linkage group G near marker Satt163. The 2751bp promoter region from GmMIPS2 was utilized for chromosomal mapping in the same population. GmMIPS2 promoter fragment also mapped to the end of linkage group G near marker Satt163 as was previously observed for GmMIPS1 cDNA. Although the GmMIPS1 335 bp promoter pattern originally appeared monomorphic in the mapping population, longer exposure indicated a faint banding pattern similar to that observed with GmMIPS2 promoter (S. Jeong and M.A. Saghai Maroof, personal communication). The polymorphism observed with the GmMIPS1 promoter after long exposure may be a result of its identity to GmMIPS2 promoter, rather than legitimate segregation. Nevertheless, the data suggest the possibility that GmMIPS1 and GmMIPS2 might be closely linked.

The possible location of one or more MIPS genes at the end of a chromosome may explain why multiple screening of a lambda and BAC (bacterial artificial chromosome) genomic libraries failed to yield clones that could be positively identified as GmMIPS1 and GmMIPS2. Two BAC libraries from the soybean cultivars 'Williams 82' (Kanazin *et al*., 1996) and 'Faribault' (Danesh *et al*., 1998) were probed independently with GmMIPS1 full-length cDNA. Forty-five BAC clones were purchased

for further characterization. Southern blot analysis of BAC DNA showed hybridizing bands with the GmMIPS1cDNA probe, but none of the bands corresponded to fragments observed on Southern blots of soybean genomic DNA. Localization to the end of a chromosome and lack of appropriate restriction enzyme sites may have precluded representation of certain MIPS sequences in the BAC library.

Immunolocalization

Immunolocalization of MIPS in developing soybean seeds was utilized to ascertain whether increases in MIPS protein were associated with tissues where phytic acid was known to accumulate. In rice, MIPS transcript accumulation was shown to be temporally and spatially associated with accumulation of phytic acid (Yoshida *et al*., 1999). Previous northern and western analysis of MIPS expression showed high accumulation of transcript and protein in developing soybean seeds (Hegeman *et al*., 2001). As shown by western analysis of developing soybeans seeds (Figure 5, panel B), MIPS expression was highest in early stages of seed development. MIPS expression was also significantly higher in developing seeds than in other tissues (Figure 5, panel A).

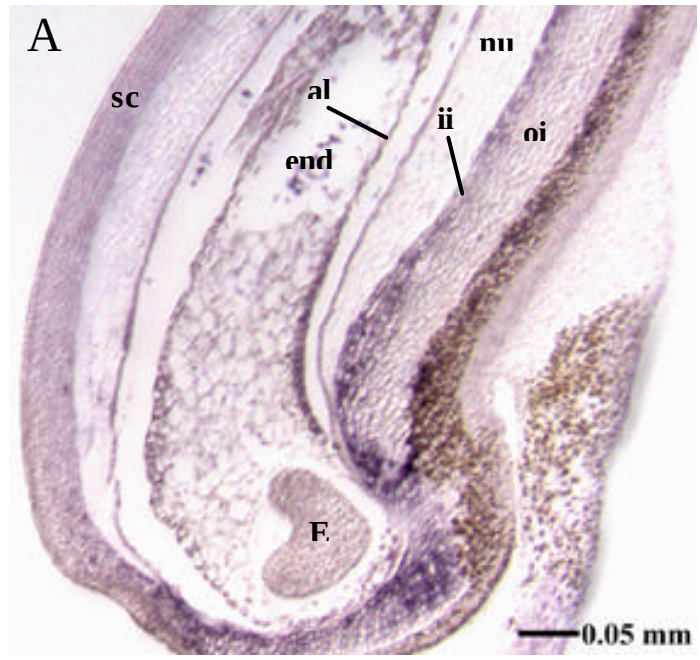
Although accumulation of MIPS was observed in developing soybean seeds, it is not known whether this was associated with phytate deposition. Immunolocalization of MIPS protein in developing soybean seeds was used to determine if MIPS accumulates in the endosperm, cotyledon and embryo, tissues were phytate deposits occur (Raboy, 1990). Affinity purified MIPS polyclonal antibody was used for immunolocalization of fixed sections of soybean seeds. Developing soybean seeds were stabilized in the coagulating acid fixative FAA (formalin-acetic acid-alcohol) before dehydration and

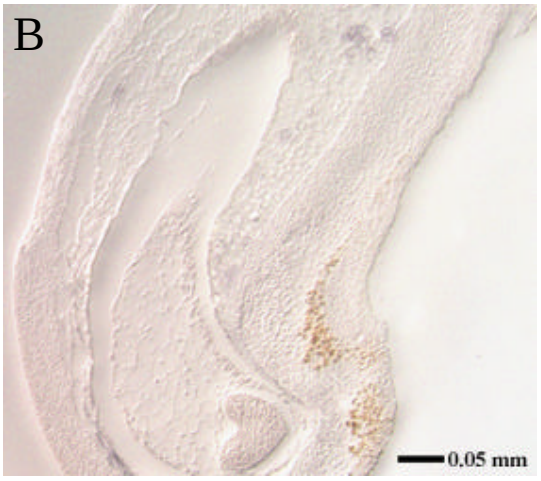
infiltration with paraffin for sectioning. Samples were incubated with antibody to MIPS and secondary conjugated antibody, with secondary antibody used alone as a background control. Samples were also stained for structure to observe morphology. The structural stain, Johansen's Safranin and Fast Green, differentially stains tissue dependent on composition with a regressive Safranin O dye and the progressive Fast Green. In this method, chromosomes, nuclei, lignified, suberized, or cutinized cell walls appear red, while cytoplasm and cellulosic cell walls appear green (Ruzin, 1999).

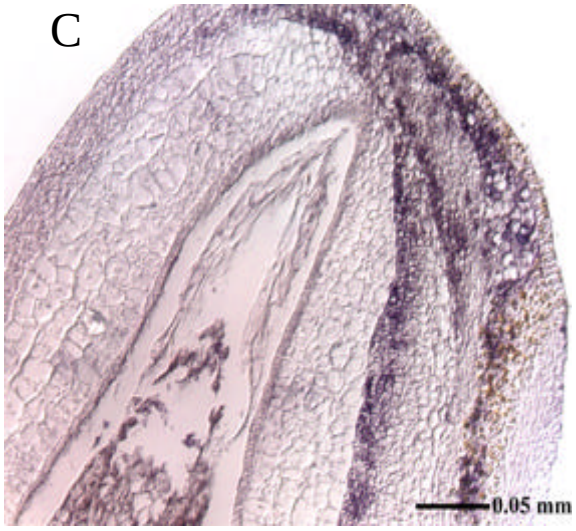
MIPS protein was detectable in developing soybean seeds as early as the heart stage of embryogenesis (Figure 15). Low-level expression of MIPS was found throughout the seed, with discrete regions of higher expression observed in the inner integument (bordering nucellus), and the aleurone. This developmental stage (C) occurs 6-8 days after anthesis (DAA) in 1.5-1.9 mm seeds (Meinke *et al.*, 1981). This stage precedes any of the samples examined by western analysis with MIPS antibody.

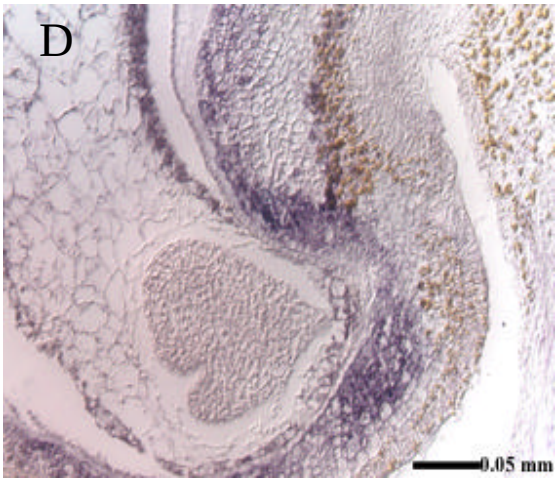
By the early cotyledon stage (8-10 DAA, 2.0-2.5 mm, stage D) MIPS expression dramatically increased, especially in endosperm and the developing embryo (Figure 16). MIPS was observed in the developing cotyledons. Expression of MIPS in the inner integument increased to encompass the entire layer of tissue. Outer integument tissue also showed expression throughout the developing seed. Perhaps most dramatic, the procambium in maternal tissue exhibited exceptionally high MIPS expression, exceeding that of the procambium tissue in the developing embryo. Aleurone cells surrounding the endosperm also showed high levels of MIPS protein.

MIPS protein levels significantly decreased by the cotyledon stage (10-12 DAA, 2.6-3.4 mm, stage E). Although western blot analysis showed high levels of expression


A
sc
al
ii
nu
oi
end
E
0.05 mm


B
0.05 mm


C
0.05 mm


D
0.05 mm

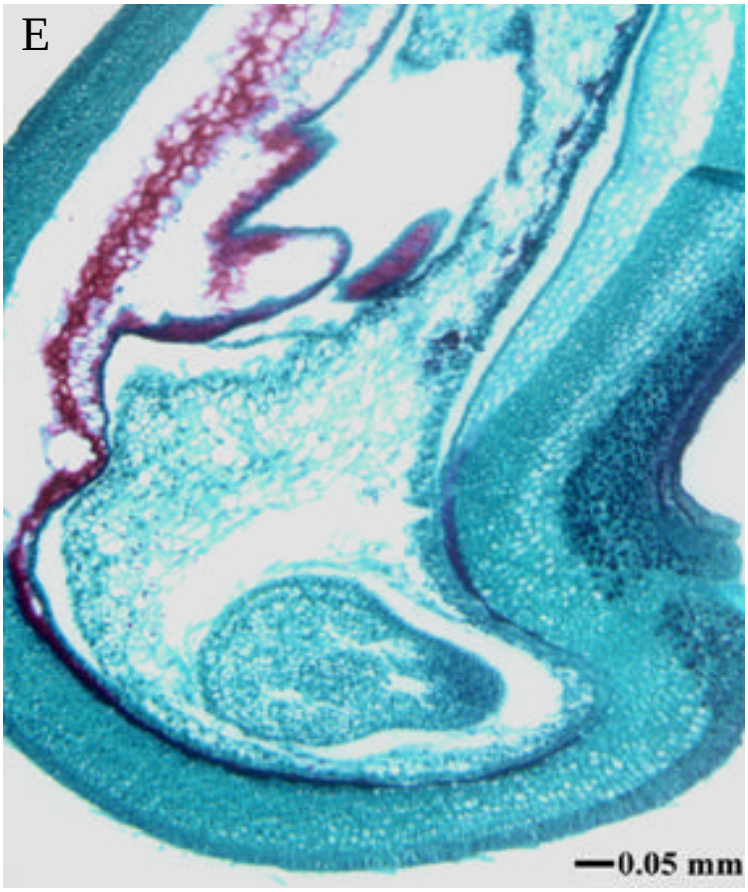


Figure 15. Immunolocalization of MIPS in heart stage soybean embryos

Developing seeds from field-grown soybeans were fixed in PFA and embedded in paraffin through a graded ethanol dehydration and CitriSolv infiltration. Embedded samples were sectioned longitudinally into 10 µm ribbons and affixed to slides. Paraffin was removed from the sections by clearing in CitriSolv and rehydrating in an ethanol series.

A. Developing V945153 seeds (0-2 mm) with a heart stage embryo were subjected to immunolocalization with anti-MIPS antisera. Protein was detected using a secondary antibody conjugated to AP (alkaline phosphatase) followed by incubation with NBT/BCIP (5-bromo-4-chloro-3-indolyl-phosphate/ nitro blue tetrazolium). (200X)
B. Control incubated only with secondary antibody and detected with NBT/BCIP. (200X)
C & D. Close up (300X) of anti-MIPS treated sections.
E. Structural stain - Johansen's Safranin and Fast Green. (200X)

Abbrievations: al-aleurone, E-embryo, end-endosperm, ii-inner integument, nu-nucellus, oi-outer integument and sc-seed coat.

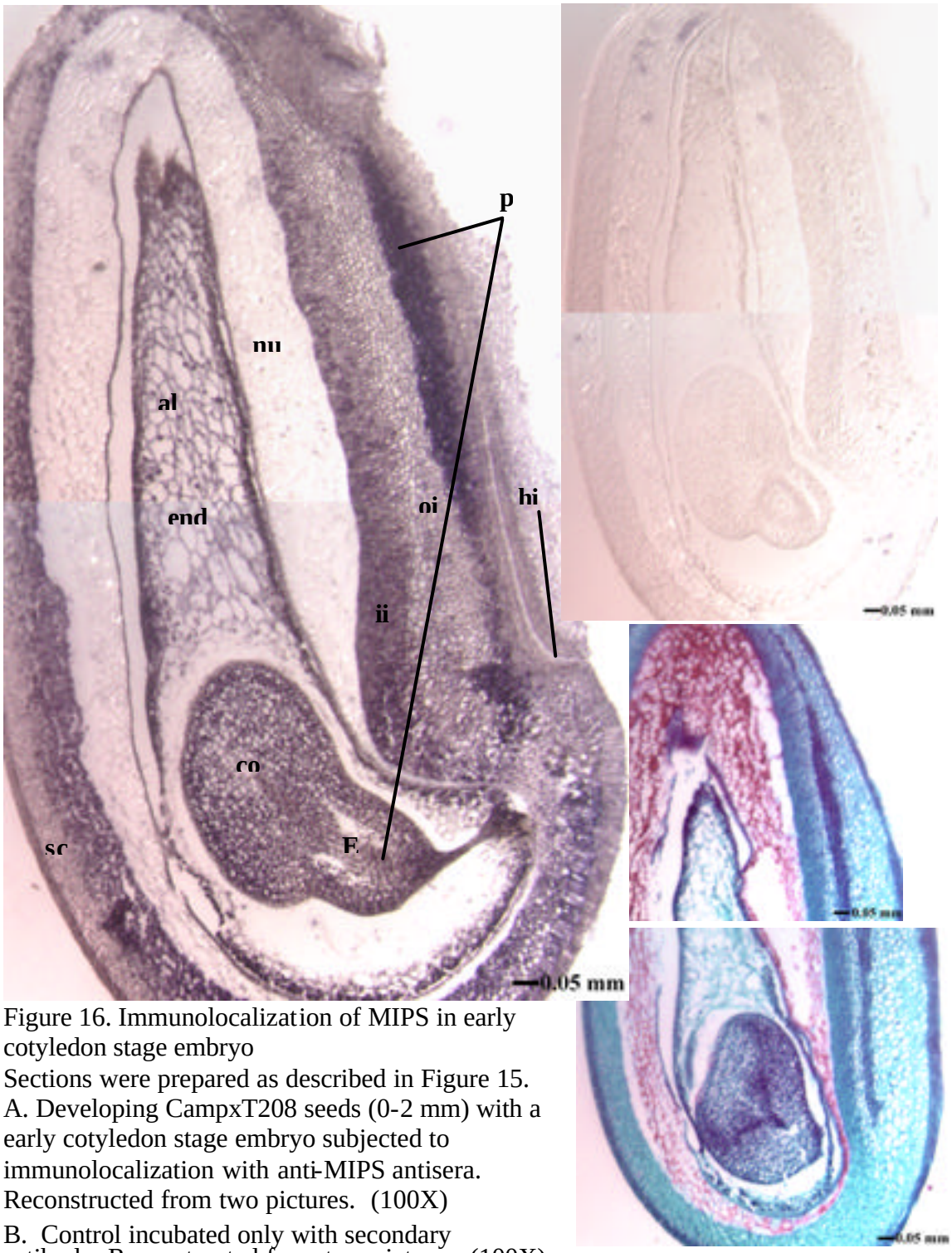


Figure 16. Immunolocalization of MIPS in early cotyledon stage embryo

Sections were prepared as described in Figure 15.
A. Developing CampxT208 seeds (0-2 mm) with a early cotyledon stage embryo subjected to immunolocalization with anti-MIPS antisera. Reconstructed from two pictures. (100X)

B. Control incubated only with secondary antibody. Reconstructed from two pictures. (100X)

C & D. Structural stain - Johansen's Safranin and Fast Green. (100X)

Abbrievations: al-aleurone, co-cotyledon, E-embryo, end-endosperm, hi-hilum, ii-inner integument, nu-nucellus, oi-outer integument, p-procambium and sc-seed coat.

(Figure 5, panel B), immunocytological observation showed an overall decrease in regions of MIPS expression (Figure 17) compared to early cotyledon stage (Figure 16). Expression in the cotyledon and embryo were minimal. Expression in the aleurone and inner integument decreased dramatically from higher levels in earlier stage embryos. High levels were still observed in maternal procambium.

By the late cotyledon stage embryos (17-19 DAA, 6.6-8.4 mm, stages I and J) high levels of MIPS expression was limited to developing maternal vascular tissue (Figure 18). At this stage, cotyledons are the largest structure visible in the developing seed, and minimal MIPS expression remains in the embryo and cotyledon. This observation corresponds with MIPS levels determined through western blot analysis. In Figure 5, panel B, only faint bands were observed in lane I, corresponding to late cotyledon stage seeds.

In green bean (*Phaseolus vulgaris*), MIPS protein was observed in globular-stage embryos and in cotyledon stage embryos (Johnson and Wang, 1996). The pattern of accumulation for soybean MIPS, which exhibited high expression early in development through the beginning of cotyledonary stage, differed from patterns of MIPS accumulation in green bean. Western analysis of MIPS protein in developing green been seeds showed protein bands of two distinct sizes (33 and 56 kDa) in globular and cotyledonary stage. No MIPS protein was observed in the green bean heart stage of embryogenesis.

Yoshida *et al.* (1999) examined the accumulation of the *RINO1*, a rice cDNA homologous to *myo*-inositol-3-phosphate synthase, in developing rice seeds. They found that *RINO1* accumulated in the scutellum (the single cotyledon in monocots) and

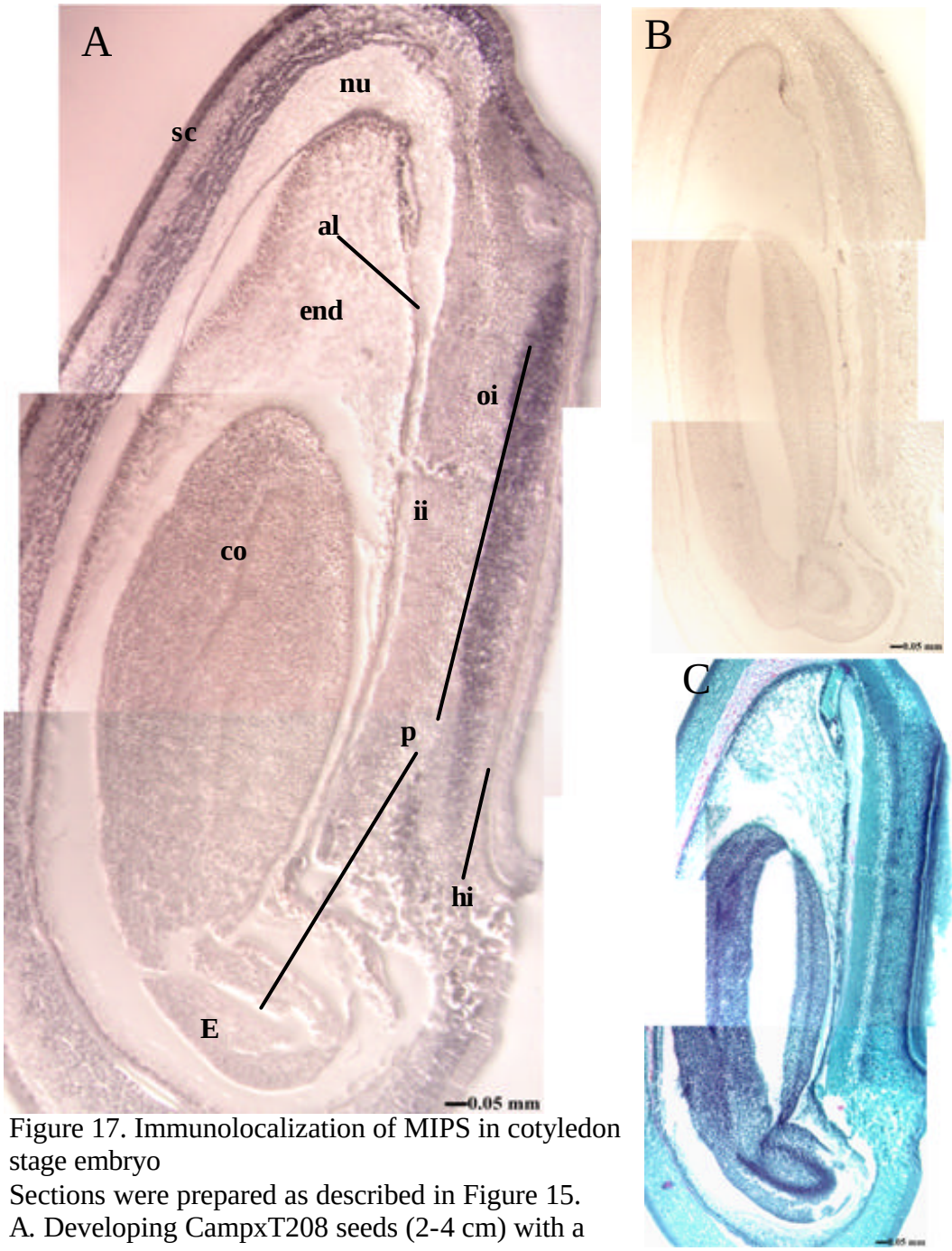


Figure 17. Immunolocalization of MIPS in cotyledon stage embryo

Sections were prepared as described in Figure 15.

A. Developing CampxT208 seeds (2-4 cm) with a cotyledon stage embryo subjected to immunolocalization with anti-MIPS antisera. Reconstructed from three pictures. (100X)

B. Control incubated only with secondary antibody. Reconstructed from three pictures. (100X)

C. Structural stain - Johansen's Safranin and Fast Green. Reconstructed from three pictures. (100X)

Abbrievations: al-aleurone, co-cotyledon, E-embryo, end-endosperm, hi-hilum, ii-inner integument, nu-nucellus, oi-outer integument, p-procambium and sc-seed coat.

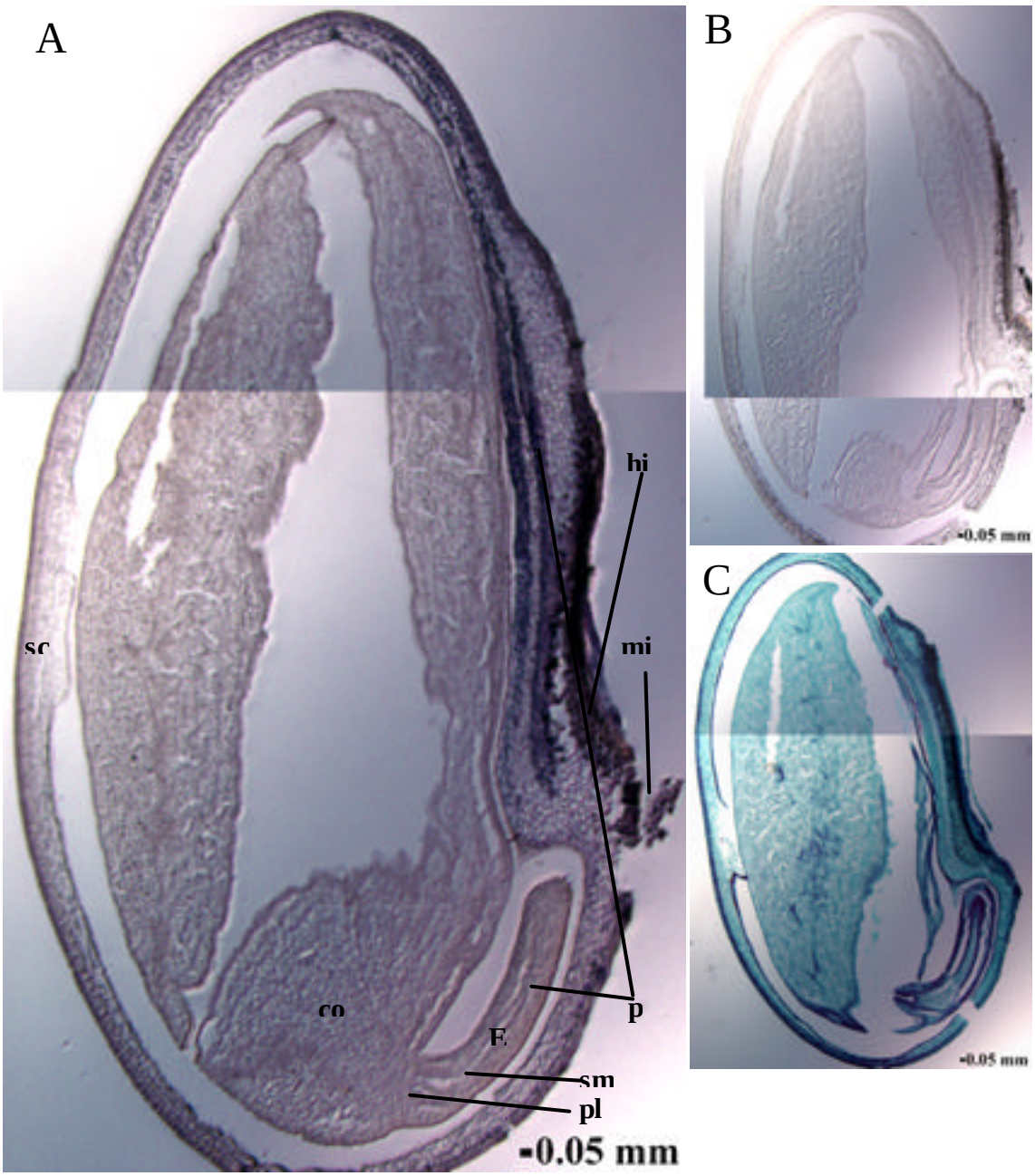


Figure 18. Immunolocalization of MIPS in late cotyledon stage embryo
Sections were prepared as described in Figure 15.
A. Developing V945153 seeds (6-8 mm) with a late cotyledon stage embryo subjected to immunolocalization with anti-MIPS antisera. Reconstructed from two pictures. (25X)
B. Control incubated only with secondary antibody. Reconstructed from two pictures. (25X)
C. Structural stain - Johansen's Safranin and Fast Green. Reconstructed from two pictures. (25X)
Abbrievations: co-cotyledon, E-embryo, end-endosperm, hi-hilum, mi-micropyle, p-procambium, pl-primary leaves, sc-seed coat and sm-shoot meristem.

aleurone preceding deposition of phytate in globoids in the same tissues. In rice, as in many other cereals, phytate deposits are limited to the aleurone and scutellum. In dicots, phytate deposits can be found in the endosperm, cotyledon and embryo of the developing seed (Raboy, 1990). Phytate begins to accumulate in soybean seeds 14 days after flowering, although only trace amounts of phytate are seen until 21 days after flowering. (Raboy and Dickenson, 1987). Observation of MIPS protein in early stages of developing seeds indicates that maximal expression of MIPS protein precedes phytate accumulation. MIPS protein is also seen in those tissues that later accumulate phytate, including the endosperm, cotyledon and embryo. High levels of MIPS protein in developing vascular tissue (procambium) of the seed does not appear to be associated with phytate deposition, both due to the extended time of expression and localization away from areas where phytate-containing globoids are typically found.

The observed pattern of MIPS accumulation at or near vascular tissue has been observed in other species. Nelson *et al*. (1998) observed in accumulation of MIPS protein in and near the phloem in leaves and roots of iceplant. After exposure to salt stress, an increase in MIPS was observed in mesophyll cells in the leaf and a decrease in root MIPS. Nelson *et al*. (1999) speculated that phloem-associated inositol metabolism is involved in salt-stress tolerance, with *myo*-inositol acting as a shoot to root signal for sodium uptake. To synthesize free inositol, the product of MIPS activity, inositol-3-phosphate, is dephosphorylated by the enzyme inositol monophosphatase (IMP). Gillaspy *et al*. (1995) observed accumulation of IMP protein in vascular tissue of light-grown tomato seedlings, particularly in phloem parenchyma and sieve tube member cells.

Conclusions

Sequence data obtained from GmMIPS1 and GmMIPS2 suggest that MIPS is highly similar in soybean, not only within the predicted protein sequence, but also within intron and promoter regions. GUS fusions with the GmMIPS1 promoter sequence indicate that the promoter fragments isolated thus far are sufficient to drive expression, but not limited to developing seeds. Isolation of additional upstream promoter sequence may provide a seed-specific promoter fragment that could be utilized to drive antisense expression of MIPS. Isolation and characterization of the two remaining MIPS isoforms, GmMIPS3 and GmMIPS4, was initiated with the sequencing of a GmMIPS3 5' RACE product. Immunolocalization analysis of MIPS in developing soybean seeds showed a pattern of protein accumulation early in seed development that supported northern and western blotting data. MIPS protein is found in those tissues of the developing soybean seed where phytic acid is known to accumulate. Characterization of the role of each MIPS isoform in soybean seed development, phytic acid synthesis and plant growth and development could be aided by utilization of sensitive, isoform-specific quantitative RT-PCR.

MATERIALS AND METHODS

GmMIPS gene sequence amplification

DNA extraction

Genomic DNA was isolated from two-day old soybean hypocotyls according to the method described by Dellaporta *et al*. (1983). DNA was extracted with phenol-

chloroform-isoamyl alcohol with Phase Lock Gel Tubes (5'→3', Boulder, CO) to aid in separation of aqueous phase before precipitation in ethanol. The DNA was treated with RNase Plus (5'→3') at a concentration of 1 µL RNase to 250 µL sample after resuspension.

Amplification and cloning of GmMIPS1 from genomic DNA

DNA from 'Williams' soybeans, isolated as described above, was utilized as a template for PCR amplification of the genomic coding region of a MIPS gene corresponding to the previously isolated GmMIPS1 cDNA. The gene-specific portion of the upstream primer "MIPS5'Kpn" (5'-G CTA **GGT ACC** ATG TTC ATC GAG AAT TTT AAG GT-3') included the translational start site (underlined). The gene specific portion of the downstream primer "MIPS3'Kpn" (5'-GC TT**G GTA CC**T ATT CTT CGG TCC CAT GCT TCA CT-3') ended at the translational stop codon (underlined). *Kpn*I restriction sites were designed at the 5'end of the primers (indicated in bold). PCR mixtures (50 µl) were assembled with Taq DNA Polymerase (Qiagen) according to manufacturer's recommendations. Reactions were denatured at 95°C for 2 minutes, followed by 35 cycles of denaturation at 95°C for 1 minute, annealing at 51°C for 1 minute, and extension at 72°C for 3 minutes in a Robocycler Gradient 40 thermocycler (Stratagene, La Jolla, CA). *Kpn*I restriction sites designed at the 5'end of the primers were utilized to clone the amplification product into the *Kpn*I site of pBluescriptII KS^-. The insert was sequence by automated cycle sequencing at the Virginia Tech DNA Sequencing Facility (See Appendix D for final GmMIPS1 genomic sequence including primer locations.).

Isolation of GmMIPS1 promoter by Inverse PCR

Self-ligated genomic DNA from 'Williams' soybeans was utilized as a template to amplify the promoter sequence upstream from the GmMIPS1 coding sequence. Soybean genomic DNA (1 µg) was digested for 3 hours with 10U *Eco*RV. The enzyme was heat inactivated at 65°C for 15 minutes, and the sample purified using a QIAquick spin column (Qiagen). The DNA sample was self-ligated in a 500 µL volume using 10 units of T4 DNA ligase. DNA from the ligation reaction was purified using a QIAquick spin column and used as template for inverse PCR. The primers "mIPSInv2-5'" (5'-GGA ATG GCA CCT ATC AG-3') and "mIPSInv2-3'" (5'-ACA AGT TCG GTG GTT TC-3') were complementary to the 5' end of the GmMIPS1 cDNA. 1X *Taq* extender buffer (Stratagene) was added to the reaction to facilitate synthesis of long PCR products. PCR cycling parameters were 94°C for 2 minutes, followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing at 39°C for 1 minute, and extension at 72°C for 5 minutes. The resulting 1145 bp PCR product (Inv5hyb) was purified using a QIAquick spin column. The purified product was treated with *Pfu* polymerase (Stratagene) in the presence of dNTPs to create blunt ends. The product was cloned into the *Sma*I site of pBluescript. The ligation mix was digested with *Sma*I to linearize plasmid self-ligations prior to transformation. *E. coli* cells were transformed using electroporation (BTX Electroporation System, BTX, San Diego, CA). The 1145 bp insert (Inv5hyb), containing 407 bp of sequence upstream of the ATG, was sequenced using dideoxy chain termination protocol (Sequitherm Excel II kit; Epicenter Technologies, Madison, WI).

Takara LA PCR Kit (Takara, Otsu, Shiga, Japan) and an alternate restriction enzyme were utilized to facilitate recovery of longer promoter sequence than those

obtained previously with the combination of *Taq* Extender and *Eco*RV digested genomic DNA. Soybean genomic DNA (2 µg) was digested overnight with 40 U *Pst*I. After heat inactivation of the restriction endonuclease (65°C for 15 min), samples were self-ligated with 20U T4 DNA Ligase in a total volume of 1000 µL for 18 hr at 25°C. DNA from the ligation reaction was purified using a QIAquick spin column. The primer "mIPSLAInvUP" (5'-TCT GTG TTG GTT GAT TTC CTT GTG GGG GCT GGT A-3') corresponded to the region directly 5' to intron VI. The primer "mIPSLApromdown" (5'-TTT GTT CAT TTG CCT TGT CTA CTA GTA TTG TTT TCA-3') corresponded to a region 290 bp upstream from the ATG, based on sequence information from Inv5hyb. PCR reactions were assembled according to manufacturer's instructions with 2 µL self-ligated template per 50 µL reaction. Reactions were denatured at 94°C for 2 minutes, followed by 30 cycles of denaturation at 94°C for 20 seconds and annealing/extension at 62°C for 15 minutes on an Eppendorf 96 Gradient thermocycler (Brinkmann Instruments, Inc., Westbury, NY). PCR products were purified using QIAquick spin columns and directly ligated into a pBluescriptII-KS$^-$ based T-vector (Marchuk *et al.*, 1990). The insert of approximately 1.8 kb (PstLAInv), containing 953 bp of promoter sequence upstream of the translational start site, was sequenced by automated cycle sequencing at the Virginia Tech DNA Sequencing Facility. (See Appendix E for final GmMIPS1 promoter sequence including primer locations.)

Isolation of GmMIPS2 promoter by Inverse PCR

GmMIPS2 promoter region was isolated by IPCR utilizing the Takara LA PCR Kit. Genomic self-ligated template was created as above, except 'Williams' genomic

DNA was digested with *Sca*I. The primer "mIPSLAInvUP" was utilized with the primer "mIPSLAinvDWN "(5'-TTT CCA CCC CAA CCC ACA AGC ATT ACC CTG AGA-3'), which corresponded to a region overlapping the 3'end of intron I. PCR reactions were assembled according to manufacturer's instructions with 5 µL self-ligated template per 50 µL reaction. Reactions were denatured at 94°C for 2 minutes, followed by 30 cycles of two-step denaturation at 94°C for 20 seconds and annealing/extension at 68°C for 15 minutes on a Robocycler Gradient 96 thermocycler. PCR products were purified using QIAquick spin columns and directly ligated into a T-vector, as described above. The insert of approximately 3.8 kb (ScaLAInv), containing 2751 bp of GmMIPS2 promoter and untranslated region, was sequenced by automated cycle sequencing at the Virginia Tech DNA Sequencing Facility. (See Appendix F for final GmMIPS2 promoter sequence including primers locations.)

Amplification and cloning of GmMIPS2 from genomic DNA

DNA from 'Williams' soybeans, isolated as described above, was utilized as a template for PCR amplification of the genomic coding region of a second MIPS gene. Primers were designed based on 5' coding sequence data from the GmMIPS2 IPCR product ScaLAInv and 3' sequence from GmMIPS2 ESTs. The upstream primer "MIPS2-UP" (5'-CAT TCT TAA TCG TTG GGA AAA ATG-3') corresponded to the region immediately upstream and including the translational start site (underlined). The downstream primer "MIPS2-DWN" (5'-TT ATT CTA TCC CAT GCG TCA CT-3') ended at the translational stop codon (underlined). PCR mixtures (50 µl) were assembled with Taq DNA Polymerase (Qiagen) according to manufacturer's recommendations. 5X

Q Solution was added to the reaction to facilitate synthesis of PCR products. Reactions were denatured at 98°C for 5 minutes, followed by 35 cycles of denaturation at 98°C for 1 minute, annealing at 42°C for 1 minute, and extension at 72°C for 3 minutes in a Robocycler Gradient 40 thermocycler. PCR products were purified using QIAquick spin columns and directly ligated into a T-vector. The 2610 bp insert (MIPS2full) was sequenced by automated cycle sequencing at the Virginia Tech DNA Sequencing Facility. (See Appendix G for final GmMIPS2 genomic sequence including primers locations.)

GmMIPS3 cDNA sequence amplification by 5' RACE

Previous classification of MIPS-homologous soybean EST sequences indicated a gene family of four potential members (Hegeman *et al*., 2001). Since ESTdata was generated through partial sequencing of the 3'end of cDNA clones, the complete sequence at the 5'end of GmMIPS3 was not available. 5' RACE (Rapid Amplification of cDNA Ends) was utilized to obtain 5' end sequence data from GmMIPS3. When MIPS homologous EST sequences from soybean were aligned, a majority of ESTs classified as GmMIPS3 were found in libraries derived from flower RNA (Hegeman *et al*. 2001). To obtain RNA containing GmMIPS3 transcript, RNA was extracted from flowers of field grown soybean cv. Jack, which had been frozen at -80°C immediately after harvesting. A QIAGEN RNeasy Plant Mini Kit was utilized according to manufacturer's specifications. Gene specific primers were designed based on sequence from GmMIPS3 ESTs. The gene-specific primer "MIPS3-5'RACE" (5'-TTG TTC TTT CTT TGT GCC CTT GAT-3') was utilized in first round cDNA synthesis from flower RNA template using the

5'RACE system (Gibco BRL, Rockville, MD). The cDNA was purified and tailed according to manufacturer's specifications prior to second round amplification with Abridged Anchor Primer (5'-CTA CTA CTA CTA GGC CAC GCG TCG ACT AGT-3') and gene specific "mIPS3-5'RACEnest" (5'-GTT GGC ACG GTC TCC TTG GTT GG-3'). Reactions were denatured at 94°C for 2 minutes, followed by 35 cycles of denaturation at 93°C for 30 seconds, annealing at 50°C for 30 seconds, and extension at 72°C for 1 minute in a Robocycler Gradient 96 thermocycler (Stratagene). An 717 bp product was gel purified utilizing the QIAGEN QIAquick Gel Extraction Kit. The gel purified PCR product was ligated into a T-vector. The 717 bp insert (MIPS35'RACElg) was sequence by automated cycle sequencing at the Virginia Tech DNA Sequencing Facility. Sequence information from the 5'RACE product was combined with GmMIPS3 EST sequences to infer the sequence of a full length GmMIPS3 cDNA. (See Appendix I for proposed GmMIPS3 sequence including primer locations.)

Attempts to amplify GmMIPS3 from genomic DNA

Primers designed from GmMIPS3 EST sequences and 5' RACE sequence data were utilized to amplify a product from genomic DNA. The upstream primer "MIPS3-UP" (5'-CTT GTT TCA TTT TCT CTC GAA GC-3') was designed from the untranslated region 7 bp upstream from the ATG (as determined from 5'RACE sequence). The downstream primer "MIPS3-DWN" (5'-CAA CTA CAC TAT CCC AAA CCG-3') was designed 20 bp downstream from the TGA translational stop codon (defined by GmMIPS3 EST consensus sequence). Reactions were prepared with the FailSafe PCR PreMix Selection Kit (Epicenter Technologies, Madison, WI) according to

manufacturer's specifications. 900 ng of genomic DNA from 'Williams' soybean, isolated as described above, was utilized as template for each 50 µL reaction. Reactions were denatured at 96°C for 5 minutes, followed by 40 cycles of denaturation at 96°C for 1 minute, annealing at 44°C for 1 minute, and extension at 72°C for 3 minutes in a Robocycler Gradient 40 thermocycler (Stratagene). When FailSafe PCR Premix B and I were utilized, a product of approximately 1.7 kb was observed (see Figure 8). A single clone containing the PCR product was sequenced from both ends, but did not appear to contain MIPS sequence when compared to soybean ESTs via BLAST.

Preparation of promoter:reporter gene constructs for bombardment

Preparation of a promoterless GUS construct (pGUS101del)

Constructs for characterization of promoter activity were created utilizing transcriptional fusions of promoter fragments to the β-glucuronidase (GUS) gene. A promoter-less GUS gene and nopaline synthase terminator (NOS-ter) were excised from pBI101 (Clontech, Palo Alto, CA) as a 1.87 kb *Hin*dIII/*Eco*RI fragment. This GUS cassette was ligated into pBluescriptII KS^-, which was previously digested with *Hin*dIII and *Eco*RI. The resulting clone (pGUS101) was digested with *Pst*I and *Sac*I. Overhanging ends resulting from the digestion were removed with *Pfu* Polymerase. The resulting blunted product was self-ligated, resulting in the 4.87 kb construct pGUS101del. The multiple cloning site (MCS) preceding the GUS gene was utilized for cloning of promoter fragments amplified by PCR (see Figure 12)

Amplification and cloning of a CaMV 35S promoter construct

The constitutive 35S promoter from cauliflower mosaic virus (CaMV) was utilized as a positive control for promoter expression after bombardment. The promoter region was amplified from a construct containing the dual enhanced CaMV 35S promoter utilizing the upstream primer "35SUP-Pst" (5'-TAA CTG CAG AAC ATG GTG GAG CAC GAC-3') and the downstream primer "35SDWN-Bam" (5'-ATA GGA TCC ATC ACA TCA ATC CAC TTG CTT TG-3'). PCR reactions (50 µl) were assembled with Taq DNA Polymerase (Qiagen) according to manufacturer's recommendations. 5X Q Solution was added to the reaction to facilitate synthesis of PCR products. Reactions were denatured at 98°C for 5 minutes, followed by 40 cycles of denaturation at 98°C for 1 minute, annealing at 39°C for 1 minute, and extension at 72°C for 1 minute in a Robocycler Gradient 40 thermocycler. PCR products were purified using QIAquick spin columns and directly ligated into a T-vector. The intermediate vector containing a 342 bp CaMV 35S promoter region (p35S-2) was digested with *Xba*I and *Xho*I to remove the insert. The insert was gel purified using the QIAGEN Gel Purification kit and digested with *Bam*HI. The bombardment construct pGUS101del was digested with *Bam*HI and treated with calf-intestinal alkaline phosphatase (CIAP) (Promega, Madison, WI) according to manufacturer's specifications. The CaMV 35S promoter region was ligated into pGUS101del, creating the constitutive control bombardment construct pGUS101del.35S-2 (see Figure 12).

Amplification and cloning of a SSP promoter construct

The seed-specific promoter (SSP) from the soybean seed-storage protein β-conglycinin α' subunit was utilized as a positive control for promoter expression in late seed development. The promoter region was amplified from a construct containing the SSP promoter utilizing the upstream primer "SSPUP-Pst" (5'-TAG CTG CAG AAG CTT TCG ACT CTA GAT CC-3') and the downstream primer "SSPDWN-Bam" (5'-ATA GGA TCC ATC TTC GAG CTC GCC G-3'). PCR reactions (50 µl) were assembled with Taq DNA Polymerase (Qiagen) according to manufacturer's recommendations. 5X Q Solution was added to the reaction to facilitate synthesis of PCR products. Reactions were denatured at 98°C for 5 minutes, followed by 40 cycles of denaturation at 98°C for 1 minute, annealing at 39°C for 1 minute, and extension at 72°C for 1 minute in a Robocycler Gradient 40 thermocycler. PCR products were purified using QIAquick spin columns. Products were treated with *Pfu* polymerase in the presence of dNTPs to produce blunt ends. The blunted PCR products were ligated into the *Sma*I site of pBluescriptII KS^-. The intermediate vector containing the 1 kb SSP promoter region (pSSP-1) was digested with *Xba*I and *Xho*I to remove the insert. The insert was gel purified using the QIAGEN Gel Purification kit and digested with *Bam*HI. The bombardment construct pGUS101del was digested with *Bam*HI and treated with CIAP according to manufacturer's specifications. The SSP promoter region was ligated into pGUS101del, creating the seed-specific control bombardment construct pGUS101del.SSP-1-42 (see Figure 12).

Amplification and cloning of a GmMIPS1p350 promoter construct

A 335 bp promoter region of GmMIPS1 was amplified to test whether it was of sufficient length to act as a promoter after bombardment. The upstream primer "MIPS1pUPsml-Pst" (5'-ATT CTG CAG CTA AGA TAA GTA TAC TCA TTG ATT G-3') was designed based on promoter sequence amplified from the IPCR product Inv5hyb. The downstream primer "MIPS1pDWN-Bam" (5'-CAT GGA TCC ACT ACA CAC ACT TGT GTT G-3') was designed to end at 61 bp upstream of the ATG. PCR reactions (50 µl) were assembled with *Pfu* Polymerase (Stratagene) according to manufacturer's recommendations. Genomic DNA from 'Williams' soybeans was utilized as template for PCR amplification. Reactions were denatured at 98°C for 5 minutes, followed by 40 cycles denaturation at 98°C for 1 minute, annealing at 38°C for 1 minute, and extension at 72°C for 1 minute in a Robocycler Gradient 40 thermocycler. PCR products were purified using QIAquick spin columns and directly ligated into a T-vector. The intermediate vector containing a 335 bp GmMIPS1 promoter region (pMIPS1p350-13) was digested with *Xba*I and *Xho*I to remove the insert. The insert was gel purified using the QIAGEN Gel Purification kit and digested with *Bam*HI. The bombardment construct pGUS101del was digested with *Bam*HI and treated with CIAP. The GmMIPS1 335 bp promoter region was ligated into pGUS101del, creating the bombardment construct pGUS101del.GmMIPS1p350-2-25 (see Figure 12).

Amplification and cloning of a GmMIPS1p1.2 promoter construct

A 1180 bp promoter region of GmMIPS1 was amplified to test promoter function in bombardment studies. The upstream primer "MIPS1pUP-Pst" (5'-TAA CTG CAG

GAG GGA GAA TGT TGG ATC -3') was designed based on promoter sequence amplified from the IPCR product PstLAInv. The downstream primer "MIPS1pDWN-Bam" (5'-CAT GGA TCC ACT ACA CAC ACT TGT GTT G-3') was designed to end 61 bp upstream of the ATG. PCR reactions (50 µl) were assembled with *Pfu* Polymerase (Stragene) according to manufacturer's recommendations. Genomic DNA from 'Williams' soybeans was utilized as template for PCR amplification. Reactions were denatured at 98°C for 5 minutes, followed by 40 cycles denaturation at 98°C for 1 minute, annealing at 38°C for 1 minute, and extension at 72°C for 1 minute in a Robocycler Gradient 40 thermocycler. PCR products were purified using QIAquick spin columns. Products were treated with *Pfu* polymerase in the presence of dNTPs to produce blunt ends. The blunted PCR products were ligated into the *Sma*I site of pBluescriptII KS. The intermediate vector containing a 1180 bp GmMIPS1p1.2 promoter region (pMIPS1p1.2-21) was digested with *Xba*I and *Xho*I to remove the insert. The insert was gel purified using the QIAGEN Gel Purification kit and digested with *Bam*HI. The bombardment construct pGUS101del was digested with *Bam*HI and treated with CIAP. The GmMIPS1 1180 bp promoter region was ligated into pGUS101del, creating the bombardment construct pGUS101del.GmMIPS1p1.2-4 (see Figure 12).

Amplification and cloning of a GmMIPS2 promoter construct

A 2751 bp promoter and untranslated region of GmMIPS2 was amplified to test promoter expression after bombardment. The upstream primer "MIPS2pUP-Sal" (5'-ATA GTC GAC TAA CTC TTG CAG GTT CCA CC-3') was designed based on maximum promoter region amplified from the IPCR product ScaLAInv. The gene-

specific portion of the downstream primer "MIPS2pDWN-Xba" (5'-GCC AGA TCT TTT TCC CAA CGA TTA AGA ATG AAA C-3') was designed directly upstream from the translational start site. PCR reactions (50 µl) were assembled with Taq DNA Polymerase (Qiagen) according to manufacturer's recommendations. 5X Q Solution was added to the reaction to facilitate synthesis of PCR products. The GmMIPS2 IPCR clone ScaLAInv was utilized as template for PCR amplification. Reactions were denatured at 98°C for 5 minutes, followed by 35 cycles denaturation at 98°C for 30 seconds, annealing at 42°C for 1 minute, and extension at 72°C for 3 minute in a Robocycler Gradient 40 thermocycler. PCR products were purified using QIAquick spin columns. Products were treated with *Pfu* polymerase in the presence of dNTPs to produce blunt ends. The blunted PCR products were ligated into the *Sma*I site of pBluescriptII KS. The intermediate vector containing a 2751 bp GmMIPS2p promoter region (pMIPS2p-9) was digested with *Kpn*I and *Sac*I to remove the insert. The insert was gel purified using the QIAGEN Gel Purification kit and digested with *Sal*I. The bombardment construct pGUS101del was digested with *Sal*I and treated with shrimp alkaline phosphatase (SAP) (Promega) according to manufacturer's specifications for standard dephosphorylation with precipitated digest. To date, the GmMIPS2 promoter has not successfully been cloned into the bombardment construct pGUS101del.

Microprojectile bombardment

Developing soybean seeds

Promoter:GUS fusion constructs, described above, were bombarded into developing soybean seeds to determine promoter activity. Developing 'Essex' soybean

seeds, ranging in size from 1 mm to 1 cm, were excised from pods. Seeds larger than 3 mm were sliced into sections. A range of seed sizes were placed on a Murashige and Skoog plate (basal MS, supplemented with 30g/L sucrose and 0.4 mg/L 2,4-D with 0.2% GelRite) in preparation for bombardment.

A BioRad biolistic PDS-1000/He Particle Delivery System (BioRad, Hercules, CA) was utilized to accelerate tungsten particles into the developing seeds. The microcarrier, M17 tungsten (BioRad), was washed and precipitated according to manufacturer's specifications and stored as frozen aliquots. One 50 µL aliquot of tungsten was used per bombardment. Plasmid DNA (5 µg) was precipitated onto the tungsten microcarrier with $CaCl_2$ and spermadine according to manufacturer's specifications. 10 µL of DNA-coated tungsten was dried on the macrocarrier. 1,100 psi rupture disks (BioRad) were used. Each plate containing soybean seeds was bombarded twice with the same construct.

Bombarded plates were allowed to recover 16 hours in the dark. Seeds were then transferred to 2 mL X-gluc (5-bromo-4-chloro-3-indolyl glucoronide) substrate (Stomp, 1992). Blue foci were photographed after six days of incubation in stain.

Soybean cell suspension cultures

Soybean cell suspension cultures (*Glycine max* [L.] Merr. cv. Williams 82, courtesy of Dr. J.M. Widholm, University of Illinois) were maintained in MS liquid media (Murashige and Skoog media; basal MS, supplemented with 30g/L sucrose and 0.4 mg/L 2,4-D) in darkness in an orbital shaker. Cells were subcultured four days prior to bombardment. Cells were packed by centrifugation and media was removed. Cells were

resuspended in an equal volume of fresh MS media. A 0.6 mL aliquot of the cells was spread on sterile #1 Whatman filters (Clifton, NJ) placed on MS plates.

Biolistic procedure was carried out as described above. Each plate was bombarded once. Cells were allowed to recover 16 hours in the dark. X-gluc substrate (1 mL) was added to plates. Blue foci were observed 24 hours after treatment.

Southern blot analysis

Soybean genomic DNA (10 µg), purified as described above, was digested with 90 U of the appropriate restriction endonuclease. Digested samples were purified utilizing QIAGEN QIAquick Purification columns. DNA was separated by agarose gel electrophoresis (0.8% agarose, 1X TAE). Separated DNA was dupurinated, denatured and neutralized according to MSI Southern blot protocol (Micron Separations Inc., Westborough, MA) before DNA was transferred to Nytran SuPerCharge nylon membrane (Schleicher and Schuell, Keene, NH) by capillary blotting in 10X SSC. Four DNA fragments were isolated from plasmids for use as probes. A GmMIPS1 cDNA probe consisted of a 1.7 kb *Sal*I fragment from the plasmid MIPSfull (containing full length GmMIPS1 cDNA) isolated by gel purification. The 1180 bp promoter region of GmMIPS1 was excised from pMIPS1p1.2-21 with the restriction enzymes *Xba*I and *Xho*I and gel purified to create a GmMIPS1 promoter probe. A GmMIPS2 promoter probe was isolated by digesting the intermediate vector pMIPS2p-9 with *Kpn*I and *Sac*I and gel purifying the 2751 bp promoter region. The GmMIPS3 5'RACE product was gel purified from MIPS35'RACElg after digestion with the restriction enzymes *Sal*I and *Xba*I, generating a 717 bp GmMIPS3 probe. The DNA probes (25 ng) were labeled with

α-^{32}P-dATP utilizing the Random Primers DNA labeling kit (Life Technologies, Rockville, MD). Probes were purified using a TE SELECT-D G50 spin column (5'→3'). Hybridizations were carried out in UltraHyb hybridization buffer (Ambion, Austin, TX), a high stringency formamide-based hybridization buffer, according to manufacturer's specifications. Membranes were washed twice in 2X SSC, 1% SDS for five minutes at 42°C, and twice in 0.2X SSC, 1% SDS for twenty minutes at 42°C. Membranes were exposed on phosphor screens for 2 to 48 hr and scanned by a Storm 860 Phosphoimager (Molecular Dynamics, Amersham Pharmacia Biotech, Piscataway, NJ).

Immunolocalization

Histological preparation of samples

Developing seeds from field grown soybeans were harvested and sorted according to size into four catagories; 0-2 mm, 2-4 mm, 4-6 mm and 6-8 mm. Whole seeds were fixed in PFA (4% paraformaldehyde [w/v] in 1X PBS) and stored at 4°C. The seeds were transferred to paraffin (TissuePrep2, Fisher, Pittsburg, PA) through a graded ethanol dehydration and CitriSolv (Fisher) infiltration procedure as described by in Plant Microtechnique and Microscopy (Ruzin, 1999). Fixed samples were dehydrated in 10%, 20%, 30% and 50% ethanol for 30 minutes each, before being transferred to 70% ethanol for an overnight incubation. Samples were then further dehydrated in 90%, 95%, 100% (twice) ethanol for 20 minutes each. Samples were infiltrated with CitriSolv by transitioning through series of 100% ethanol:CitriSolv mixtures (3:1, 1:1, 1:3) for 30 minutes each before two 30 minute changes of 100% CitriSolv. Samples were infiltrated

with melted paraffin (58°C) over a period of 48 hours. Samples embedded in paraffin were transferred to small aluminum boats and allowed to cool overnight at 4°C. Individual seeds were then mounted on wooden chucks for longitudinal sections. Sectioning of samples to yield 10 µm ribbons was performed with a microtome (Microm HM 300, Microm International, Walldorf, Germany). Sections were affixed to Probe-On Plus slides (Fisher).

Immunocytochemistry

Slides with mounted sections were deparrafinized by clearing in CitriSolv and rehydrated in ethanol. Slides were treated with two changes of 100% CitriSolv for 10 minutes, then transferred to a 1:1 mixture of CitriSolv:100% ethanol for two minutes. Slides were subsequently rehydrated in an ethanol series (100%, 90%, 70% and 40%). After equilibration in 1X PBS, slides were incubated with blocking solution (1X PBS, 0.1% BSA) for 10 minutes. Sections were incubated with 100 µL of a primary anti-MIPS antibody at a dilution of 1:250 in blocking solution or a control of blocking solution alone for 6 hr. MIPS antibody was affinity purified (Styer, 2000). Slides were washed three times with PBS and incubated with 100 µL of a 1:1000 dilution of secondary antibody, goat anti-rabbit IgG-AP $F(ab')_2$ (Jackson ImmunoResearch, West Grove, PA) for one hour. After washing with PBS, slides were washed briefly with alkaline phosphatase buffer (100 mM Tris pH 9.5, 100 mM NaCl, 5 mM $MgCl_2$). Slides were incubated with the colorametric alkaline phosphatase substrates NBT/BCIP (5-bromo-4-chloro-3-indolyl-phosphate/ nitro blue tetrazolium) (Promega, Madison, WI) in the presence of the eukaryotic AP inhibitor levamisole (0.24 mg/mL in alkaline

phosphatase buffer) for three hours. Slides were mounted with Crystal/Mount (Fisher) and visualized with a Ziess Axioscope 2 compound microscope (Carl Ziess Inc, Thornwood, NY). Images were recorded with a Spot Digital Camera (SPOT model 1.4.0. Diagnostic Instruments, Inc., Sterling Heights, MI). Structural features were determined by comparison with published images (Perez-Grau and Goldberd, 1989; Carlson and Lersten, 1987).

Histological Staining

Slides were stained by Johansen's Safranin and Fast Green method, as described by Ruzin (1999). Slides were deparrafinized and brought to 70% ethanol as described above. Slides were stained for 2 hours in Safranin O staining solution (1% Safranin, 1% NaOAc, 2% formalin in 2:1:1 methyl cellosolv, 95% EtOH, water). Excess stain was removed with water. Samples were dehydrated for 10 seconds in 95% ethanol, 0.5% picric acid, then washed for 10 seconds in 95% ethanol and 2 drops 14.8 N NH_4OH (per 50 mL). After additional dehydration for 10 seconds in 100% ethanol, slides were counterstained for 10-20 seconds in Fast Green FCF (0.15% Fast Green FCF in 1:1: methyl cellosolv, 100% EtOH, clove oil). Excess stain was removed by dipping in "used" clearing solution (2:1:1 clearing solution, 100% EtOH, CitriSolv). Slides were then dipped for 10 seconds in clearing solution (2:1:1, methyl salicylate: 100% EtOH: CitriSolv). Clearing solution was removed by dipping for 10 seconds in CitriSolv and 2-3 drops 100% EtOH (per 50 mL). Slides were cleared in 100% CitriSolv for two changes. Coverslips were mounted over samples with Permount (Fisher).

Computer Analysis of MIPS sequences

The sequence of GmMIPS1 and GmMIPS2 genomic regions were compared using the Clustal W alignment method (Thompson *et al*., 1994). The regions extending from the ATG start codon to the TGA stop codon of GmMIPS1 and GmMIPS2 were compared for sequence identity. Complete genomic sequences of *Arabidopsis thaliana* MIPS sequences were obtained using AtMIPS cDNA sequences to search the TIGR *Arabidopsis thaliana* database (http://www.tigr.org/tdb/e2k1/ath1/). *Arabidopsis* MIPS genomic sequences identified in the TIGR database as At2g22240 and At4g39800 were identical to AtMIPS1 and AtMIPS2, respectively (U04876 and U30250).

Predicted protein sequences of plant MIPS, including GmMIPS1, GmMIPS2 and GmMIPS3, were aligned using the Clustal W alignment method in Biology Workbench. An unrooted phylogenetic tree was then created using the DRAWTREE (Biology Workbench). Confidence values at major branch points were detemined by 1000 bootstrap replicates of the alignment in the CLUSTALTREE program (Biology WorkBench). Values were expressed as a percentage.

GmMIPS1 and GmMIPS2 promoter and 5'UTR (untranslated region) sequences were analyzed for putative cis-acting promoter motifs by comparison of sequence with known elements in the PlantCARE database (Rombauts *et al*., 1999). The 1241 bp GmMIPS1 sequence analyzed contained promoter sequence from the IPCR product PstLAInv, and 5'UTR sequence from full length MIPS cDNA (AF293970). The 2751 bp sequence of GmMIPS2 contained all sequence upstream of the ATG, promoter and 5'UTR inclusive.

Putative cis-acting elements were identified in the GmMIPS1 and GmMIPS2 sequences through a search of the PlantCARE database (http://sphinx.rug.ac.be:8080/PlantCARE/index.htm). Elements of interest were separated from less conserved elements by ranking according to similarity. Core similarity of 1 and a matrix similarity greater than 0.95 were arbitrarily chosen as cutoffs for elements of interest, regardless of the origin of the element. Elements that met these requirements are diagrammed in Figure 11 according to their function. A comprehensive list of elements identified in GmMIPS1 and GmMIPS2 promoter and 5'UTR are shown in Table 2 and Table 3.

The sequence comprising the 5' UTR, shown boxed in Figure 11, was determined by comparison of the 3' end of promoter sequence to *Glycine max* EST sequences, identified with BLAST (Altschul et al., 1997). The extent of the 5' UTR was defined as that region represented by at least 1 cDNA sequence. The soybean EST AI496426 extended 211 bp upstream of the ATG of GmMIPS1. EST AI442850 extended 317 bp upstream of the ATG of GmMIPS2.

CHAPTER IV

Summary and Future Directions

A gene family of at least four members, designated GmMIPS1 through GmMIPS4, encodes *myo*-inositol 3-phosphate synthase in soybean. MIPS appears to be highly expressed in developing seeds. Accumulation of MIPS protein in developing seeds was observed as early as the heart stage of embryogenesis. MIPS protein was observed in the endosperm, aleurone and embryo of the developing seeds through the cotyledon stage of embryogenesis, but was dramatically decreased by late cotyledon stage. This pattern of expression coincides with the spacial and temporal accumulation of phytate in protein bodies observed by other researchers. MIPS protein was also observed localized around vascular tissue. Vascularly localized accumulation of MIPS did not appear to be downregulated to the same extent as MIPS accumulation in other tissues of the developing seed.

GmMIPS1 cDNA, corresponding genomic DNA and promoter region have been isolated and sequenced. Low phytic acid soybean mutants developed by DuPont (Hitz and Sebastian, 1998) were found to have a mutation in a MIPS gene whose sequence is identical to that of GmMIPS1. GmMIPS1 therefore may be the isoform solely responsible for seed specific expression of MIPS produced for the biosynthesis of phytate. The ability of GmMIPS1 1180 bp promoter region to drive expression of GUS was tested by bombardment of a promoter construct into developing soybean seeds and soybean cell suspension culture. GUS expression was observed with both promoters (335 bp and 1180 bp), but preliminary results suggest that the GmMIPS1 promoter fragment is not specific to seed tissue.

Use of the 1180 bp GmMIPS1 promoter in an antisense MIPS construct would likely be detrimental to the health of transgenic plants. The promoter fragment isolated

thus far lacks tissue specificity and appears to drive strong expression; an antisense construct with the GmMIPS1 1180 bp promoter would act more like a constitutively expressed antisense construct, reducing MIPS activity throughout the plant. Transgenic potatoes with a constitutively expressed MIPS antisense construct exhibited drastically altered morphology, including reduced apical dominance and a delayed or absent flowering (Keller *et al*., 1998). The value of soybeans as a commodity requires that any potential transgenic plant retain similar yield to non-transgenic cultivars, therefore an overall decrease in MIPS activity throughout the plant, and the resulting phenotypic changes, would be undesirable. Seed-specific expression of GmMIPS1 promoter region may be obtained through promoter fragments larger than 1180 bp.

Although all the soybean MIPS isoforms appear to be highly similar, GmMIPS1 and GmMIPS2 share a higher degree of sequence identity than other plant MIPS gene families, even within intron sequences,. GmMIPS3 sequence is divergent from GmMIPS1 and GmMIPS2. A true comparison of the sequence conservation between GmMIPS1, GmMIPS2 and GmMIPS3 must wait for the successful amplification and sequencing of a GmMIPS3 genomic PCR product. Sequence information from all four GmMIPS isoforms will allow design of an antisense construct specific to GmMIPS1.

A 2751 bp region upstream of the translational start codon of GmMIPS2 was isolated via PCR. The activity of this promoter has not yet been characterized. Many of the putative *cis*-acting elements observed in the promoter region of GmMIPS1, including elements responsive to light, plant hormones and stress, were also observed in the promoter region of GmMIPS2. Characterization of the activity of this promoter region will be dependant upon additional bombardment experiments.

Future Directions

The creation of a MIPS antisense construct for the reduction of phytate levels in soybean seeds is dependant upon a promoter that not only is active the same time as MIPS in the developing seed, but is limited to the seed. While in theory the GmMIPS1 promoter is ideal for this task, the fragment isolated thus far does not demonstrate tissue specific activity. Additional bombardment experiments into differentiated soybean tissue may provide additional information about promoter activity, apart from the vagarities of undifferentiated suspension cell culture. Additional upstream sequence from GmMIPS1 might be neccessary in order to obtain a promoter fragment with restricted expression. Inverse PCR primers designed to the promoter region of GmMIPS1 will be utilized to amplify promoter region further upstream. Fragments larger than 1180 bp will then be tested for seed specific expression through bombardment into developing soybean seeds and other soybean tissues.

In order to confirm the predicted location of exon-intron boundries in the genomic sequence of GmMIPS2, a GmMIPS2 cDNA will be isolated via RT-PCR from root RNA. Sequencing of the GmMIPS2 cDNA will allow an unequivocal assignment of intron sequences in GmMIPS2.

Genomic sequence of GmMIPS3 is still unknown to date. Primers designed to the predicted GmMIPS3 cDNA failed to amplify a product of the expected size. The product amplified failed to contain MIPS homologous sequence. Single base pair errors at the 3'end of the primers could be responsible for lack of specific priming in the PCR reaction. Analysis of additional EST sequences recently deposited to GenBank may provide additional sequence data, which can be utilized to design primers more specific

to GmMIPS3. Complete GmMIPS3 sequence data will be utilized to design IPCR primers to amplify promoter regions of GmMIPS3. This promoter region will then be utilized for isoform-specific chromosomal mapping of GmMIPS3.

Due to the sequence conservation of MIPS among the soybean isoforms, determination of tissue and special specificity of the isoforms by Northern analysis is not possible. Primers designed to single nucleotide polymorphisms between the isofoms will be utilized to perform real-time quantitative RT-PCR on soybean RNA. This will allow the contribution of each isoform to MIPS accumulation in specific tissues and development stages to be determined.

The sequence information for GmMIPS4 is based only on two EST sequences. Additional analysis of EST sequences recently deposited to the database may provide more EST sequences with GmMIPS4, enabling design of primers to amplify a GmMIPS4 genomic copy.

LITERATURE CITED

Abu-Abied M, Holland, D (1994) The gene *c-ino1* from *Citrus paradisi* is highly homologous to *tur1* and *ino1* from yeast and Spirodela encoding for myo-inositol phosphate synthase. Plant Physiol 106: 1689

Adams P, Vernon DM, Thomas JC, Bohnert HJ, Jensen RG (1992) Distinct cellular and organismic responses to salt stress. Plant Cell Physiol 33:1215-1223

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl Acids Res 25: 3389-3402

Arguello-Astorga GR, Herrera-Estrella LR (1996) Ancestral multipartite units in light-responsive plant promoters have structural features correlating with specific phototransduction pathways. Plant Physiol 112:1151-1166

Biehl RR, Baker DH (1997) Microbial phytase improves amino acid utilization in young chicks fed diets based on soybean meal but not diets based on peanut meal. Poult Sci 76: 355-360

Bruce WB, Deng XW, Quail PH (1991) A negatively acting DNA sequence element mediates phytochrome-directed repression of *phyA* gene transcription. EMBO J 10: 3015-24

Carlson JB, Lersten NR (1987) Reproductive Morphology. In Wilcox JR Soybean: Improvement, Production, and Uses. Second Edition. Agronomy 16: 95-134

Chen Z-L, Schuler MA, Beachy RN (1986) Functional analysis of regulatory elements in a plant embryo-specific gene. Proc Natl Acad Sci USA 83: 8560-8564

Conley TR, Park SC, Kwon HB, Peng HP, Shih MC (1994) Characterization of *cis*-acting elements in light regulation of the nuclear gene encoding the A subunit of chloroplast isozymes of glyceraldehyde-3-phosphate dehydrogenase from Arabidopsis thaliana. Mol Cell Biol 14:2525-2533

Danesh D, Penuela S, Mudge J, Denny RL, Nordstrom H, Martinez JP, Young ND. (1998) A bacterial artificial chromosome library for soybean and identification of clones near a major cyst nematode resistance gene. Theor Appl Genet 96: 196-202

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep 1: 19-21

Denbow DM, Grabau EA, Lacy GH, Kornegay ET, Russell DR, Umbeck PF (1998) Soybeans transformed with a fungal phytase gene improve phosphorus availability for broilers. Poult Sci 77:878-871

Denbow DM, Ravindran V, Kornegay ET, Yi Z, Hulet RM (1995) Improving phosphorus availability in soybean meal for broilers by supplemental phytase. Poult Sci 74: 1831-1842

Douglas MW, Peter CM, Boling SD, Parsons CM, Baker DH (2000) Nutritional Evaluation of Low Phytate and High Protein Corns. Poult Sci 79:1586-1591

Dunn MA, White AJ, Vural S, Hughes MA (1998) Identification of promoter elements in a low-temperature-responsive gene (*blt*4.9) from barley (*Hordeum vulgare* L.). Plant Mol Biol 38: 551-564

Feldbrugge M, Hahlbrock K, Weisshaar B (1996) The transcriptional regulator *CPRF1*: expression analysis and gene structure. Mol Gen Genet 251:619-627

Feldbrugge M, Sprenger M, Hahlbrock K, Weisshaar B (1997) *Pc*MYB1, a novel plant protein containing a DNA-binding domain with one MYB repeat, interacts *in vivo* with a light-regulatory promoter unit. Plant J 11: 1079-1093

Finer JJ (1988) Apical proliferation of embryogenic tissue of soybean [*Glycine max* (L.) Merrill]. Plant Cell Rep 7: 238-241

Flores S, Smart CC (2000) Abscisic acid-induced changes in inositol metabolism in *Spirodela polyrrhiza*. Planta 211:823-832

Gibson DM, Ullah AHJ (1990) Phytases and their action on phytic acid. In Morre DJ, Boss WF, Loewus FA. Inositol Metabolism in Plants. Wiley-Liss, Inc., New York, 77-92

Gillaspy GE, Keddie JS, Oda K, Gruissem W (1995) Plant Inositol Monophosphatase Is a Lithium-Sensitive Enzyme Encoded by a Multigene Family. Plant Cell 7: 2175-2185

Glasgow HB, Burkholder JM, Schemechel DE, Tester PA, Rublee PA (1995) Insidious effects of a toxic estuarine dinoflagellate on fish survival and human health. J toxicol Env Health 46: 501-522

Hara K, Yagi M, Koizumi N, Kusano T, and Sano H (2000) Screening of wound-responsive genes identifies an immediate-early expressed gene encoding a highly charged protein in mechanically wounded tobacco plants. Plant Cell Physiol 41: 684-691

Harper AF, Kornegay ET, Schell TC (1997) Phytase supplementation of low-phosphorus growing-finishing pig diets improves performance, phosphorus digestibility, and bone mineralization and reduces phosphorus excretion. J Anim Sci 75:3174-3186

Hegeman CE, Good LL, Grabau EA (2001) Expression of D-*myo*-Inositol-3-Phosphate Synthase in Soybean: Implications for Phytic Acid Biosynthesis. Plant Physiol 125: 1941-1948

Hitz WD, Sebastian SA (1998) Soybean plant producing seeds with reduced levels of raffinose saccharides and phytic acid. International patent application WO 98/45448, pp 1-57

Huff WE, Moore PA, Waldroup PW, Waldroup AL, Balog M, Huff GR, Rath NC, Danial C and Raboy (1998) Effect of Dietary phytase and high available phosphorus corn on Broiler chicken performance. Poult Sci : 77:1899-1901

Ishitani M, Majumder AL, Bornhouser A, Michalowshi CB, Jensen RG, Bohnert HJ (1996) Coordinate transcriptional induction of *myo*-inositol metabolism during environmental stress. Plant J 9: 537-548

Ito M, Iwase M, Kodama H, Lavisse P, Komanine A, Nishihama R, Machida Y, Watanabe A (1998) A Novel *cis*-Acting Element in Promoters of Plant B-Type Cyclin Genes Activates M Phase-Specific Transcription. Plant Cell 10: 331-341

Itzhaki H, Maxson JM, Woodson WR (1994) An ethylene-responsive enhancer element is involved in the senescence-related expression of the carnation glutathione-*S*-transferase (*GST1*) gene. Proc Natl Acad Sci 91: 8925-8929

Johnson MD (1994) The *Arabidopisis thaliana myo*-inositol 1-phosphate synthase (EC 5.5.1.4). Plant Physiol 105: 1023-1024

Johnson MD, Burk D (1995) Isozyme of 1L-*myo*-inositol-1-phosphate synthase from Arabidopsis (Accession No. U30250) (PGR95-067). Plant Physiol 109: 721

Johnson MD, Henry SA (1989) Biosynthesis of inositol in yeast: primary structure of myo-inositol-phosphate synthase (EC 5.5.1.4) and functional analysis of its structural gene, the *INO1* locus. J Biol Chem 264: 1274-1283.

Johnson, MD, Sussex IM (1995) 1-L-*myo*-inositol 1-phosphate synthase from *Arabidopsis thaliana*. Plant Physiol 107: 613-619

Johnson MD, Wang X (1996) Differentially Expressed Forms of 1-L-*myo*-Inositol-1-Phosphate Synthase (EC 5.5.1.4) in *Phaseolus vulgaris*. J Biol Chem 271: 17215-17218

Kanazin V, Marek LF, Shoemaker RC (1996) Resistance gene analogs are conserved and clustered in soybean. Proc. Natl. Acad. Sci. USA 93:117466-11750

Keddie JS, Hubner G, Slocombe SP, Jarvis RP, Cummins I, Edwards EW, Shaw CH, Murphy DJ (1992) Cloning and characterization of an oleosin gene from *Brassica napus*. Plant Mol Biol 19: 443-453

Keller R, Brearley CA, Trethewey RN, Müller-Röber B (1998) Reduced inositol content and altered morphology in transgenic potato plants inhibited for 1D-*myo*-inositol 3-phosphate synthase. Plant J 16: 403-410

Kim JK, Cao J, Wu R (1992) Regulation and interaction of multiple protein factors with the proximal promoter regions of a rice high pI a-amylase gene. Mol Gen Genet 232: 383-393

Klaff P, Riesner D, Steger G (1996) RNA structure and the regulation of gene expression. Plant Mol Biol 32: 89-106

Klotz KL, Lagrimini LM (1996) Phytohormone control of the tobacco anionic peroxidase promoter. Plant Mol Biol 31: 565-573

Larson SR, Rutger JN, Young KA, Raboy V (2000) Isolation and genetic mapping on a non-lethal rice *low phytic acid I* mutation. Crop Sci 40:1397-1405

Larson SR, Raboy V (1999) Linkage mapping of maize and barley *myo*-inositol 1-phosphate synthase DNA sequences: correspondence with a *low phytic acid* mutation. Theor Appl Genetics 99:27-36

Larson SR, Young KA, Cook A, Blake TK, Raboy V (1998) Linkage mapping of two mutations that reduce phytic acid content of barley grain. Theor Appl Genetics 97:141-146

Li YC, Ledoux DR, Veum TL, Raboy V, Ertl DS (2000) Effects of Low Phytic Acid Corn on Phosphorus Utilization, Performance and Bone Mineralization in Broiler Chicks. Poult Sci 79:1444-1450

Loewus FA, Murthy PPN (2000) *Myo*-Inositol metabolism in plants. Plant Science 150: 1-19

Lott JNA, Greenwood JS, Batten GD (1995) Mechanisms and regulation of mineral nutrient storage during seed development. In J Kigel, G Galili, eds, Seed Development and Germination. Marcel Dekker, New York, 215-235

Majumder AL, Johnson MD, Henry SA (1997) 1L-*myo*-Inositol-1-phosphate synthase. Biochimica et Biophysica Acta 1348: 245-256

Manjunath S, Sachs MM (1997) Molecular characterization and promoter analysis of the maize cytosolic glyceraldehyde 3-phosphate dehydrogenase gene family and its expression during anoxia. Plant Mol Biol 33: 97-112

Marchuk D, Drumm M, Saulino A, Collins FS (1990) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Research 19:1154

Meinke DW, Chen J, Beachy RN (1981) Expression of storage-protein genes during soybean seed development. Planta 153: 130-139

Mol JNM, van der Krol AJ, van Tunen R, van Blokland R, de Lange P, Stuitje AR (1990) Regulation of plant gene expression by antisense RNA. FEBS Letters 268: 427-430

Morré DJ, Boss WF, Loewus FA (1990) Inositol Metabolism in Plants. Wiley-Liss, Inc., New York , 393 p

Munnik T, Irvine RF, Musgrave A (1998) Phospholipid signaling in plants. Biochim Biophys Acta 1389:222-272

Nelson DE, Koukoumanos M, Bohnert HJ (1999) *Myo*-Inositol-Dependent Sodium Uptake in Ice Plant. Plant Physiol 119: 165-172

Nelson DE, Rammesmayer G, Bohnert HJ (1998) Regulation of Cell-Specific Inositol metabolism and Transport in Plant Salinity Tolerance. Plant Cell 10:753-764

Park SC, Kwon HB, Shih MC (1996) *Cis*-acting elements essential for light regulation of the nuclear gene encoding the A subunit of chloroplast glyceraldehyde 3-phosphate dehydrogenase in *Arabidopsis thaliana*. Plant Physiol 112:1563-1571

Pastuglia M, Roby D, Dumas C, Cock JM (1997) Rapid induction by wounding and bacterial infection of an *S* gene family receptor-like kinase gene in *Brassica oleracea*. Plant Cell 9: 49-60

Pen J, Verwoerd TC, van Paridon PA, Beudeker RF, van den Elzen PJM, Geerse K, van der Klis JD, Versteegh HAJ, van Ooyen AJJ, Hoekema A (1993) Phytase-containing transgenic seeds as a novel feed additive for improved phosphorus utilization. Bio/Tech 11: 811-814

Perez-Grau L, Goldberd RB (1989) Soybean Seed Protein Genes Are Regulated Spatially During Embryogenesis. Plant Cell 1: 1095-1109

Polidoros AN, Scandalios JG (1999) Role of hydrogen peroxide and different classes of antioxidants in the regulation of catalase and glutathione S-transferase gene expression in maize (*Zea mays* L.). Physiologia Plantarum 106: 112-120

Prattley CA, Stanley DW (1982) Protein-phytate interactions in soybeans: I. Localization of phytate in protein bodies and globoids, J. Food Biochem. 6: 243-253

Prieto-Dapena P, Almoguera C, Rojas A, Jordano J (1999) Seed-specific expression patterns and regulation by ABI3 of an unusual late embryogenesis-abundant gene in sunflower. Plant Mol Biol 39: 615-627

Raboy V (1990) Biochemistry and Genetics of Phytic Acid Synthesis. In: Morre DJ, Boss WF, Loewus FA Inositol Metabolism in Plants. Wiley-Liss, Inc., New York, 55-76

Raboy V, Dickenson DB (1987) The Timing and Rate of Phytic Acid Accumulation in Developing Soybean Seeds. Plant Physiol 85: 841-844

Raboy V, Gerbasi P, Young KA, Stoneberg S, Pickett SG, Bauman AT, Murthy PPN, Sheridan WF, Ertl DS (2000) Origin and seed phenotype of maize *low phytic acid 1-1* and *low phytic acid 2-1*. Plant Physiol 124: 355-368

Raboy V, Hudson SJ, Dickinson DB (1985) Reduced phytic acid content does not have an adverse effect on germination of soybean seeds. Plant Physiol. 79: 323-325

Ravindran V, Bryden WL, Kornegay ET (1995) Phytic acids: occurrence, bioavailability and implications in poultry nutrition. Poultry and Avian Biology Reviews 6: 125-143

Ravindran V, Kornegay ET, Denbow DM, Yi Z, Hulet RM (1995b) Response of turkey poults to tiered levels of Natuphos phytase added to soybean meal-based semi-purified diets containing three levels of non-phytate phosphorus. Poult Sci 74:1843

Raychaudhuri A, Majumder AL (1996) Salinity-induced enhancement of L-myo-inositol 1-phosphate synthase in rice (*Oryza sativa* L.). Plant Cell Environ 19:1437-1442

Rombauts S, Déhais P, Van Montagu M, Rouzé P (1999) PlantCARE, a plant *cis*-acting regulatory element database. Nucleic Acids Res. 27: 295-296

Rouster J, Leah R, Mundy J, Cameron-Mills V (1997) Identification of a methyl jasmonate-responsive region in the promoter of a lipoxygenase 1 gene expressed in barley grain. Plant J 11:513-523

Ruzin SE (1999) Plant Microtechnique and Microscopy. Oxford, New York, NY. 322p

Seki H, Ichinose Y, Kato H, Shiraishi T, Yamada T (1996) Analysis of *cis*-regulatory elements involved in the activation of a member of chalcone synthase gene family (*PsChs1*) in pea. Plant Mol Biol 31: 479-491

Sharpley AN, Chapra SC, Wedepohl R, Sims JT, Daniel TC, Reddy KR (1994) Managing agricultural phosphorus for protection of surface waters: issues and options. J. Environ. Qual. 23: 437-451

Simons RCM, Versteegh AJ, Jongbloed AW, Kemme PA, Slump P, Bos KD, Wolters GE, Buedecker RF, Vershoor GJ (1990) Improvement of phosphorus availability by microbial phytase in broilers and pigs. Brit J Nutr 64:525-540

Smart CC, Fleming AJ (1993) A plant gene with homology to D-*myo*-inositol-3-phosphate synthase is rapidly and spatially up-regulated during an abscisic-acid-induced morphogenic response in *Spirodela polyrrhiza*. Plant J 4:279-293

Smart CC, Flores S (1997) Overexpression of D-*myo*-inositol-3-phosphate synthase leads to elevated levels of inositol in *Arabidopsis*. Plant Mol Biol 33:811-820

Spencer JD, Allee GL, Sauber TE (2000) Phosphorus bioavailability of normal and genetically modified low-phytate corn for pigs. J Anim Sci 78: 675-681

Stomp AM (1992) Histochemical Localization of β-Glucuronidase. In: Gallager SR GUS Protocols: using the GUS gene as a reporter of gene expression. Academic Press, Inc., San Diego, 103-113

Styer JC (2000) Regulating Inositol Biosynthesis in Plants: *Myo*-Inositol Phosphate Synthase and *Myo*-Inositol Monophosphatase. Masters thesis Virginia Polytechnic Institute and State University

Sugiura SH, Raboy V, Young KE, Do ng FM, Hardy R W (1999) Availablility of phosphorus and trace elements in low-phytate varieties of barley and corn for rainbow trout (*Oncorhynchus mykiss)*. Aquaculture 170:856-296

Swick RA, Ivey FJ (1992) The value of improving phosphorus retention. Feed Management 43: 8-17

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673-4680

Vauterin M, Frankard V, Jacobs M (1999) The *Arabidopsis thaliana dhdps* gene encoding dihydrodipicolinate synthase, key enzyme of lysine biosynthesis, is expressed in a cell-specific manner. Plant Mol Biol 39: 695-708

Vernon DM, Ostrem JA, Bohnert HJ (1993) Stress perception and response in a facultative halophyte - the regulation of salinity-induced genes in *Mesembryanthemum crystallinum*. Plant Cell Environ 16:437-444

Vernon DM, Tarczynski MC, Jensen RG, Bohnert HJ (1993b) Cyclitol production in transgenic tobacco. Plant J 4:199-205

Veum T, Roboy V, Ertl D, Ledoux D (1998) Low phytic acid corn improves calcium and phosphorous utilization for growing pigs. J Anim Sci 76(Suppl)1:177

Wang X, Johnson MD (1995) An Isoform of 1L-myo-Inositol 1-Phosphate Synthase (EC 5.5.1.4) from *Phaseolus vulgaris*. Plant Physiol 110:336

Ward NE (1993) Phytase in nutritional and waste management. Poultry Digest 52:10

Washida H, Wu CY, Suzuki A, Yamanouchi U, Akihama T, Harada K, Takaiwa F (1999) Identification of *cis*-regulatory elements required for endosperm expression of the rice storage protein glutelin gene *GluB-1*. Plant Mol Biol 40: 1-12

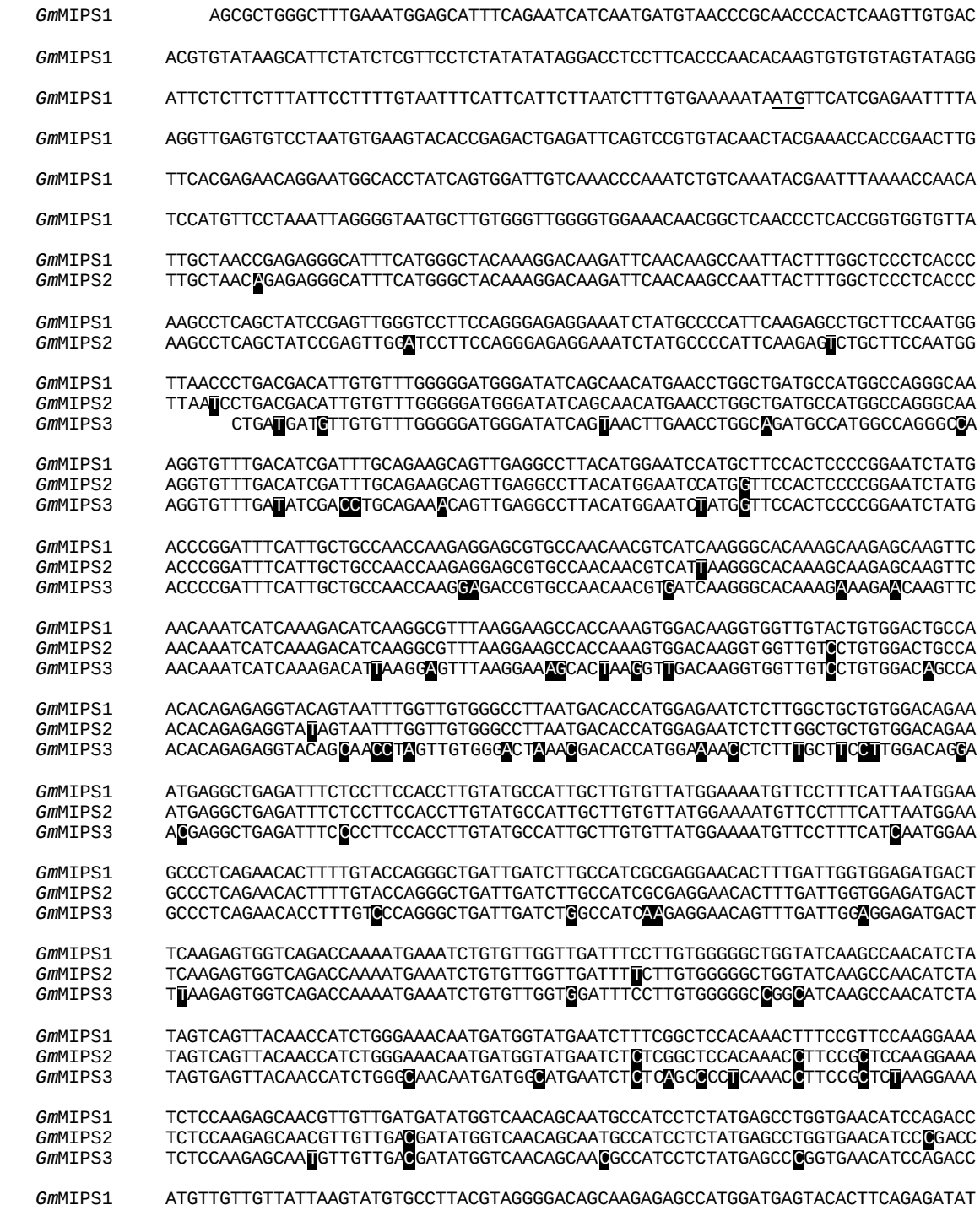
Wilcox JR, Premachandra GS, Young KA, Raboy V (2000) Isolation of High Seed Inorganic P, Low-Phytic acid Soybean Mutants. Crop Sci 40:1601-1605

Wu C, Washida H, Onodera Y, Harada K, Takaiwa F (2000) Quantitative nature of the Prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal

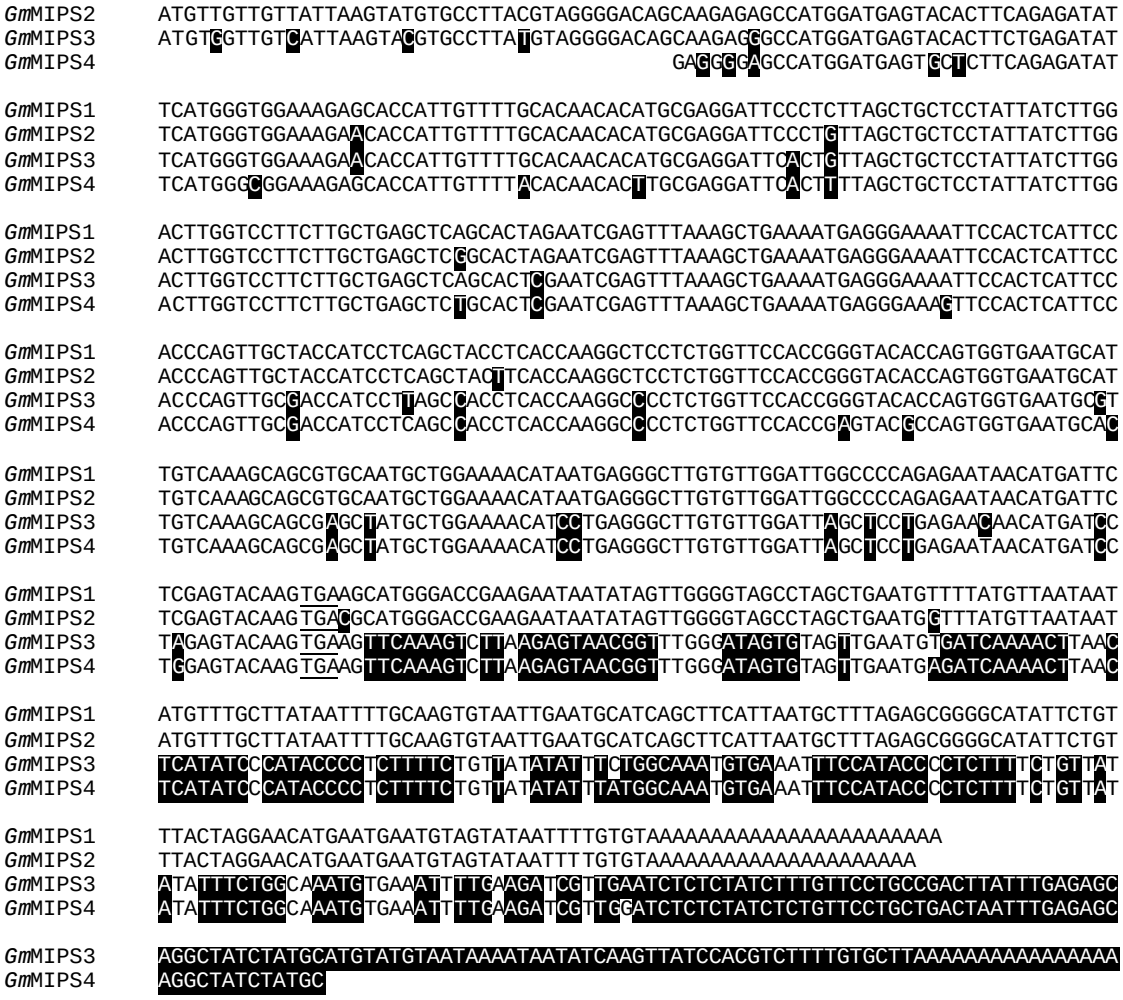
cis-element requirements for endosperm-specific gene expression. Plant J 23: 415-421

Yoshida KT, Wada T Koyama H, Mizobuchi-Fukuoka R, Naito S (1999) Temporal and Spatial Patterns of Accumulation of the Transcript of *Myo*-Inositol-1-Phosphate Synthase and Phytin Containing Particles during Seed Development in Rice. Plant Physiol 119:65-72

APPENDICES



A. On-line supplementary data. Alignment of GmMIPS1 with other MIPS-homologous sequences from the soybean EST database. Shaded boxes indicate differences from the GmMIPS1 sequence. Start and stop codons are underlined

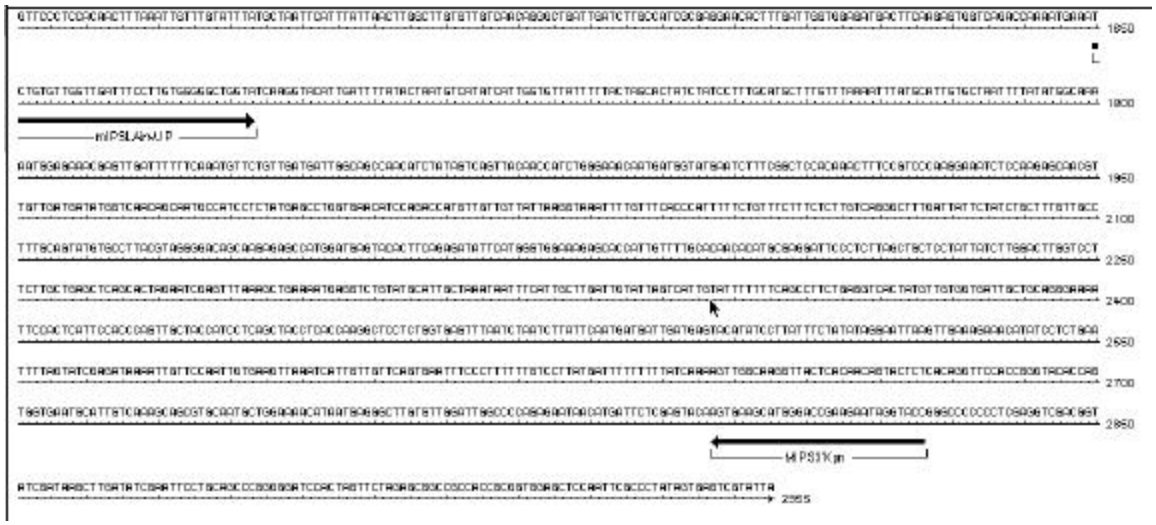

GmMIPS2 ATGTTGTTGTTATTAAGTATGTGCCTTACGTAGGGGACAGCAAGAGAGCCATGGATGAGTACACTTCAGAGATAT
GmMIPS3 ATGTCGTTGTCATTAAGTACGTGCCTTATGTAGGGGACAGCAAGAGGGCCATGGATGAGTACACTTCTGAGATAT
GmMIPS4 GAGGGGAGCCATGGATGAGTGCCTCTTCAGAGATAT
GmMIPS1 TCATGGGTGGAAAGAGCACCATTGTTTTGCACAACACATGCGAGGATTCCCTCTTAGCTGCTCCTATTATCTTGG
GmMIPS2 TCATGGGTGGAAAGAACACCATTGTTTTGCACAACACATGCGAGGATTCCCTGTTAGCTGCTCCTATTATCTTGG
GmMIPS3 TCATGGGTGGAAAGAACACCATTGTTTTGCACAACACATGCGAGGATTCACTGTTAGCTGCTCCTATTATCTTGG
GmMIPS4 TCATGGGCGGAAAGAGCACCATTGTTTTACACAACACTTGCGAGGATTCACTTTTAGCTGCTCCTATTATCTTGG
GmMIPS1 ACTTGGTCCTTCTTGCTGAGCTCAGCACTAGAATCGAGTTTAAAGCTGAAAATGAGGGAAAATTCCACTCATTCC
GmMIPS2 ACTTGGTCCTTCTTGCTGAGCTCGGCACTAGAATCGAGTTTAAAGCTGAAAATGAGGGAAAATTCCACTCATTCC
GmMIPS3 ACTTGGTCCTTCTTGCTGAGCTCAGCACTCGAATCGAGTTTAAAGCTGAAAATGAGGGAAAATTCCACTCATTCC
GmMIPS4 ACTTGGTCCTTCTTGCTGAGCTCTGCACTCGAATCGAGTTTAAAGCTGAAAATGAGGGAAACTTCCACTCATTCC
GmMIPS1 ACCCAGTTGCTACCATCCTCAGCTACCTCACCAAGGCTCCTCTGGTTCCACCGGGTACACCAGTGGTGAATGCAT
GmMIPS2 ACCCAGTTGCTACCATCCTCAGCTACTTCACCAAGGCTCCTCTGGTTCCACCGGGTACACCAGTGGTGAATGCAT
GmMIPS3 ACCCAGTTGCGACCATCCTTAGCCACCTCACCAAGGCCCCTCTGGTTCCACCGGGTACACCAGTGGTGAATGCGT
GmMIPS4 ACCCAGTTGCGACCATCCTCAGCCACCTCACCAAGGCCCCTCTGGTTCCACCGAGTACGCCAGTGGTGAATGCAC
GmMIPS1 TGTCAAAGCAGCGTGCAATGCTGGAAAACATAATGAGGGCTTGTGTTGGATTGGCCCCAGAGAATAACATGATTC
GmMIPS2 TGTCAAAGCAGCGTGCAATGCTGGAAAACATAATGAGGGCTTGTGTTGGATTGGCCCCAGAGAATAACATGATTC
GmMIPS3 TGTCAAAGCAGCGAGCTATGCTGGAAAACATCCTGAGGGCTTGTGTTGGATTAGCTCCTGAGAACAACATGATCC
GmMIPS4 TGTCAAAGCAGCGAGCTATGCTGGAAAACATCCTGAGGGCTTGTGTTGGATTAGCTCCTGAGAATAACATGATCC
GmMIPS1 TCGAGTACAAGTGAAGCATGGGACCGAAGAATAATATAGTTGGGGTAGCCTAGCTGAATGTTTTATGTTAATAAT
GmMIPS2 TCGAGTACAAGTGACGCATGGGACCGAAGAATAATATAGTTGGGGTAGCCTAGCTGAATGGTTTATGTTAATAAT
GmMIPS3 TAGAGTACAAGTGAAGTTCAAAGTCTTAAGAGTAACGGTTTGGGATAGTGTAGTTGAATGTGATCAAAACTTAAC
GmMIPS4 TCGAGTACAAGTGAAGTTCAAAGTCTTAAGAGTAACGGTTTGGGATAGTGTAGTTGAATGAGATCAAAACTTAAC
GmMIPS1 ATGTTTGCTTATAATTTTGCAAGTGTAATTGAATGCATCAGCTTCATTAATGCTTTAGAGCGGGGCATATTCTGT
GmMIPS2 ATGTTTGCTTATAATTTTGCAAGTGTAATTGAATGCATCAGCTTCATTAATGCTTTAGAGCGGGGCATATTCTGT
GmMIPS3 TCATATCCCATACCCCTCTTTTCTGTTATATATTTTCTGGCAAATGTGAAATTTCCATACCCCTCTTTTCTGTTAT
GmMIPS4 TCATATCCCATACCCCTCTTTTCTGTTATATATTTATGGCAAATGTGAAATTTCCATACCCCTCTTTTCTGTTAT
GmMIPS1 TTACTAGGAACATGAATGAATGTAGTATAATTTTGTGTAAAAAAAAAAAAAAAAAAAAAAAA
GmMIPS2 TTACTAGGAACATGAATGAATGTAGTATAATTTTGTGTAAAAAAAAAAAAAAAAAAAAAA
GmMIPS3 ATATTTTCTGGCAAATGTGAAATTTTGAAGATCGTTGAATCTCTCTATCTTTGTTCCTGCCGACTTATTTGAGAGC
GmMIPS4 ATATTTTCTGGCAAATGTGAAATTTTGAAGATCGTTGGATCTCTCTATCTCTGTTCCTGCTGACTAATTTGAGAGC
GmMIPS3 AGGCTATCTATGCATGTATGTAATAAAATAATATCAAGTTATCCACGTCTTTTGTGCTTAAAAAAAAAAAAAAAA
GmMIPS4 AGGCTATCTATGC

MIPS Gene	Accession Numbers
*Gm*MIPS1	AI496426, AI794750, AW317927, AW348857, AW397453, AW397496, AW398011, AW460108, AW472088, AW472315, AW472457, AW508269, AW508280, AW509380, AW567912, AW568795, AW705757
*Gm*MIPS2	AI442485, AI442850, AI960904, AW099866, AW153049, AW184796, AW203517, AW279066, AW310369, AW348136, AW432728, AW596973, AW620996
*Gm*MIPS3	AI458285, AI939184, AI938770, AI939288, AI965652, AW119754, AW306404, AW308846, AW394671, AW395224, AW432887, AW458285, AW458924, AW570129, AW595925, AW734281, BE191041, BE191260, BE191264, BE330770, BE331096, BE331133, BE346146, BE347713, BE348137, BE610499, BE611186, BE660319, BE660321, BE800100, BE800112, BE800119, BE800271, BE800837, BE801159, BE803215BE803317, BE806314, BE820883, BE820432, BF068026, BF069873, BG154723, BG154733, BG155232, BG155304, BG155329, BG405299, BG406262, BG651304, BG652576, BG653100, BG725477, BG726561, BG882566
*Gm*MIPS4	AI939184, AI939288, AW100674, AW394570, BE058023, BE347770, BE347771, BE611790, BE657641, BE800525, BE800974, BE800995, BE802849, BF066650, BF067760, BF071073, BG839877, BG839891, BG726578, BG726751, BG882738

B. On-line supplementary data. EST accession numbers with MIPS homology. Additional GmMIPS3 and GmMIPS4 EST sequences.

Primer	DNA sequence
1	5'- GTGGAGATGATTTCAAGAGTGGTC-3'
2	5'- CTTGGTGAGGTAGCTGAGTATGGT -3'
3 (RACE adapter primer)	5’- GGCCACGCGTCGACTAGTAC$[T]_{17}$-3’
4	5'- TTAAAGCTGAAAATGAGGGAAAAT-3'
5	5’- CCACAAGGAAATCAACCAACA -3’
6	5’- AAGCCAAAGACTGTCAAATACGAT -3’
7	5'- ATGAAATGCCCTCTCGGTTAGC-3'
8 (Abridged Anchor Primer)	5’- CTACTACTACTAGGCCACGCGTCGACTAGT-3’
9	5'- GGTGAGGGTTGAGCCGTTGTTT-3
10	5’- GCTAGTCGACATGTTCATCGAGAATTTTAAGGT-3’
11	5’- GCTTGGTACCTATTCTTCGTCCCATGCTTCACT-3’
Upstream genomic primer with KpnI site	5’- GCTAGGT ACC**ATG**TTCATCGAGAATTTTAAGGT-3’
Downstream genomic primer with KpnI site	5’- GCTTGGTACCTATTCTTCGGTCCCATGCTTCACT-3’
Upstream primer for expression vector	5’- GCAAGGATCCTACAACTACGAAACCACCGAACT-3'
Downstream primer for expression vector	5’- CCATAAGCTTCATCCAACACAAGCCCTCATTAT-3’

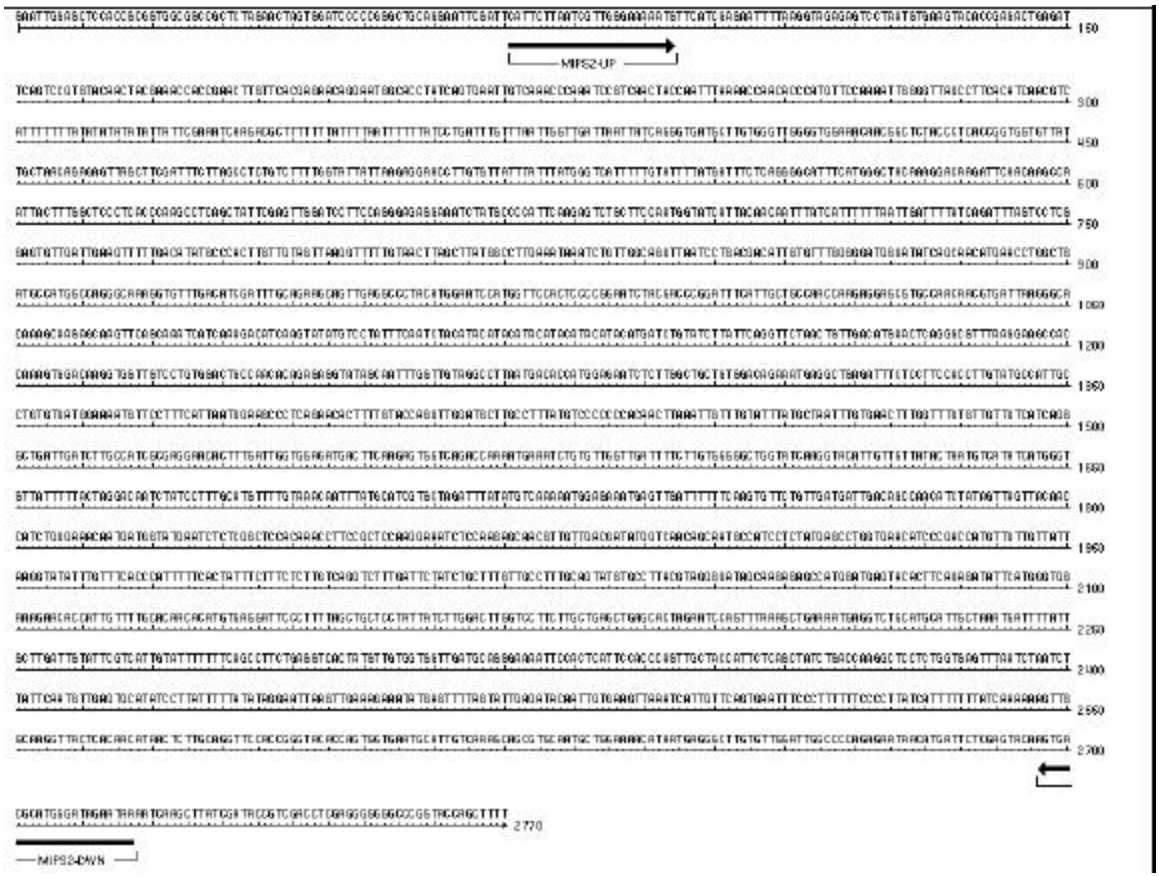
C. On-line supplementary data. Table of primers for PCR amplification.



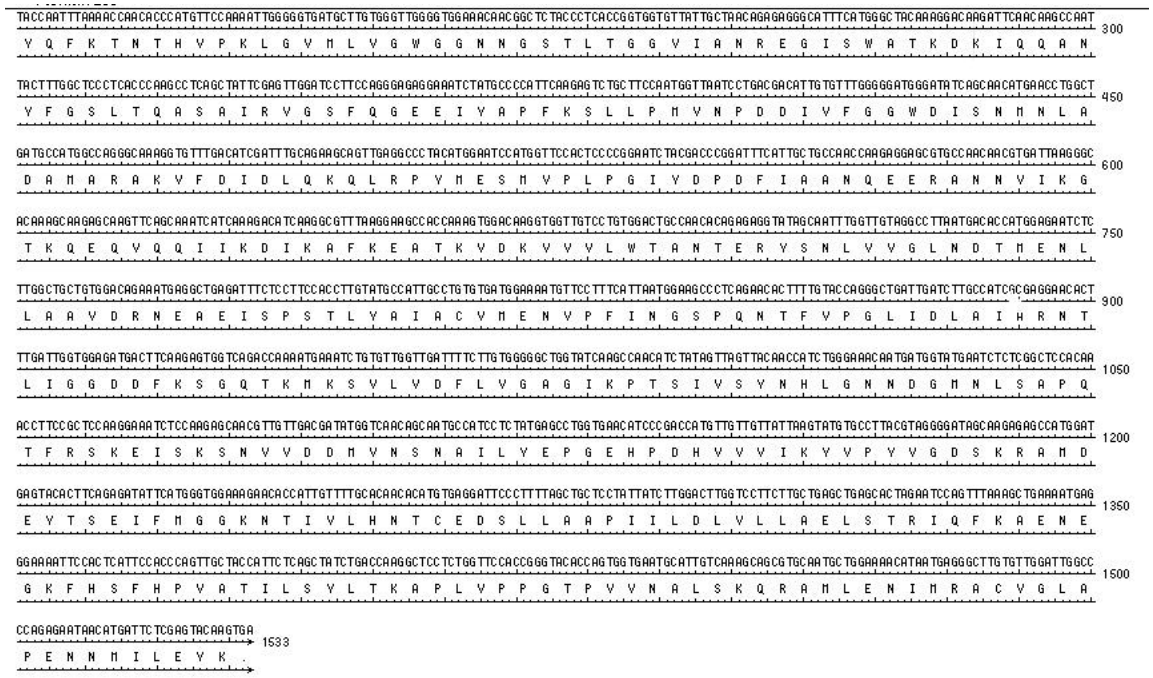
D. GmMIPS1 genomic sequence and location of PCR primers

E. GmMIPS1 promoter sequence and location of PCR primers

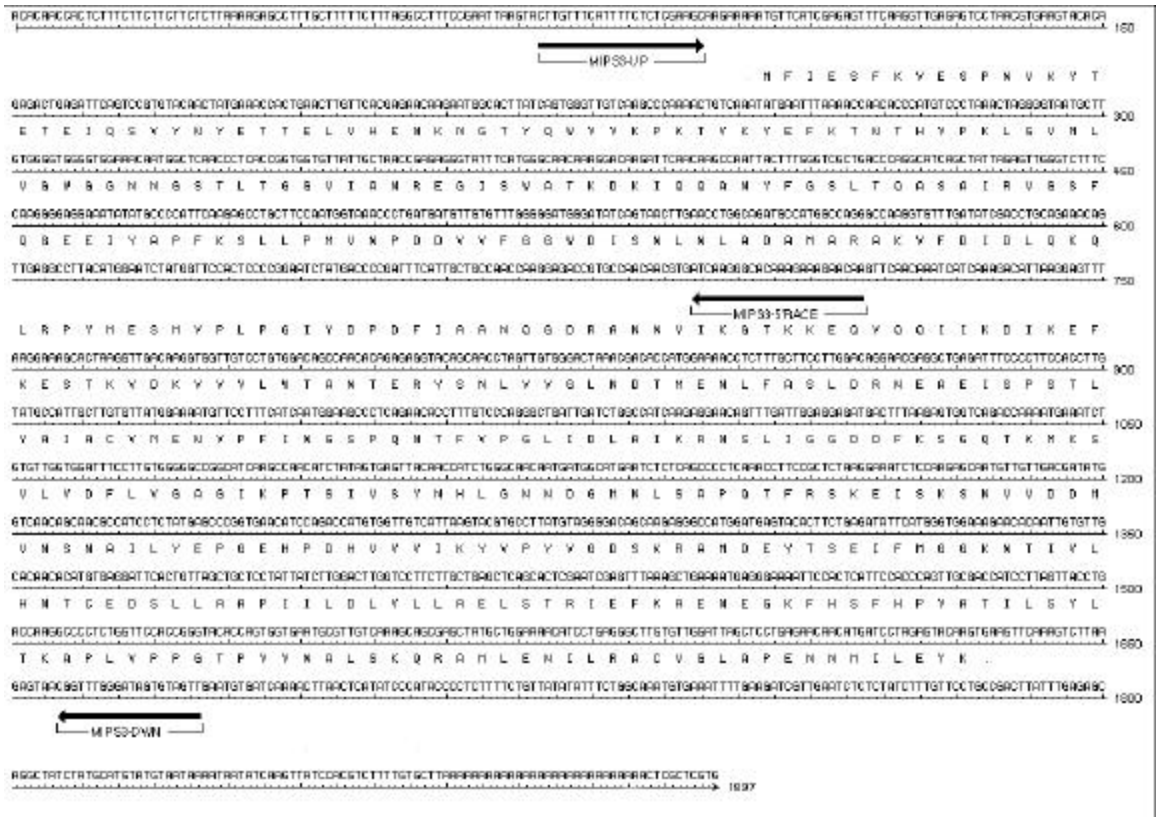
F. GmMIPS2 promoter sequence and location of PCR primers



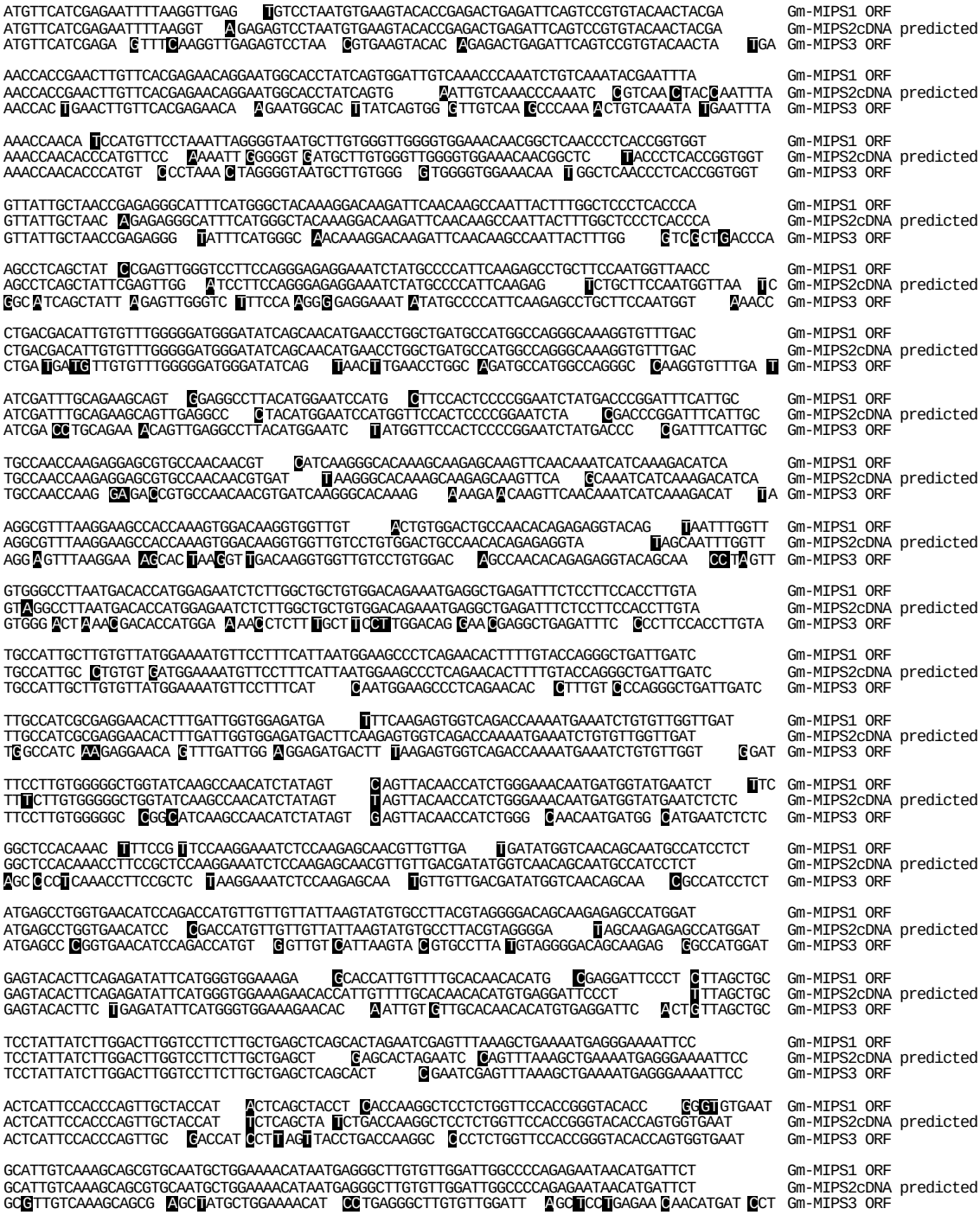
G. GmMIPS2 genomic sequence and location of PCR primers



H. Predicted GmMIPS2 cDNA (ATG to TGA) and protein sequence



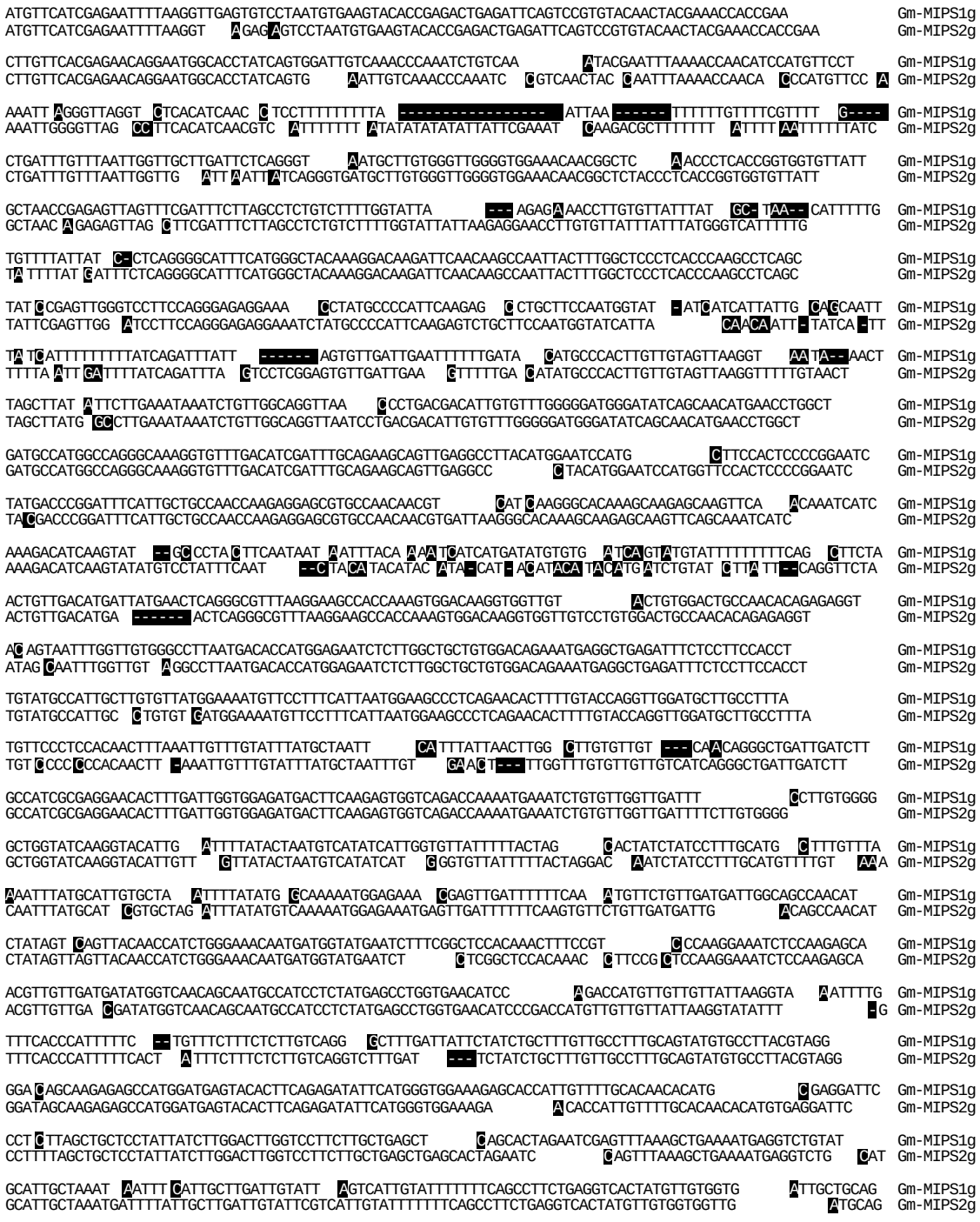
I. GmMIPS3 predicted cDNA sequence, primer locations and protein sequence



J. Clustal W alignment of coding region of GmMIPS1 cDNA (AF293970) with GmMIPS2 and GmMIPS3 predicted cDNA

```
CGAGTACAAGTGA                                                    Gm-MIPS1 ORF
CGAGTACAAGTGA                                                    Gm-MIPS2cDNA predicted
AGAGTACAAGTGA                                                    Gm-MIPS3 ORF
```

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.



K. Clustal W alignment of GmMIPS1 genomic with GmMIPS2 genomic (ATG to TGA)

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

VITAE

Laura Good was born September 9, 1977 in Charlottesville, Virginia. In 1995 she graduated from Broadway High School, located in the Shenandoah Valley, Virginia. She enrolled in Virginia Polytechnic Institute and State University, where she graduated *In Honors* from the Biochemistry Department. Continuing her undergraduate research in the laboratory of Dr. Elizabeth Grabau, she enrolled as a Masters student in the Department of Plant Pathology, Physiology and Weed Science, completing her degree in the summer of 2001.