

## ABSTRACT

Title of Thesis: CHARACTERIZATION OF THE *MYO*-INOSITOL (3) PHOSPHATE SYNTHASE GENE (MIPS) AND MAPPING OF A LPA MUTANT IN SOYBEAN (*GLYCINE MAX (L.) MERRILL*).

Katherine Salmon, Masters of Science, 2004.

Thesis Directed By: Dr. William Kenworthy, Natural Resource Science and Landscape Architecture.

Low phytic acid (LPA) is a mutation causing phosphorus to be stored as unbound phosphorus in the seed. LPA mutants show a high inorganic phosphorus (HIP) phenotype. Previous studies had indicated that LPA might be linked to the *myo*-inositol (3) phosphate synthase (MIPS) gene; this research attempted to associate a soybean HIP mutant with the MIPS gene.

The parental and the F<sub>2</sub> genotypes were tested in four ways: 1) SNP detection using the LCR protocol; 2) polymorphism detection with PCR; 3) high inorganic phosphorus (HIP) phenotype detection; and 4) oil and protein concentration.

The two parental genotypes could not be differentiated in the LCR study. A PCR-based polymorphism was heritable in the F<sub>2</sub> genotypes. HIP assay indicated multiple genes control the LPA mutant. A polymorphism was associated to the HIP phenotype. The three types of HIP phenotypes were not statistically different in oil and protein concentrations allowing implementation into a breeding program.

CHARACTERIZATION OF THE *MYO*-INOSITOL (3) PHOSPHATE SYNTHASE  
GENE (MIPS) AND MAPPING OF A LPA MUTANT IN SOYBEAN (*GLYCINE*  
*MAX (L.) MERRILL*).

By

Katherine Diane Salmon

Thesis submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Masters of Sciences  
2004

Advisory Committee:  
Professor William J. Kenworthy, Chair  
Associate Professor Jose M. Costa  
Dr. Perry Cregan

## Dedication

This thesis is dedicated to my mother, Patsy Anne Hayes, and my grandmother, Mildred Katherine Hayes, both who taught me to persevere through difficult times. Their love and support has been instrumental in achieving all my successes.

## Acknowledgements

The completion of this dissertation was made possible through the assistance and support extended by several persons. I would like to express my thanks to my advisors, Dr. William Kenworthy and Dr. Jose Costa, who shared with me their expertise and advice throughout my graduate studies. Thanks to Dr. Costa for the chance to work as a lab technician in your breeding lab and your further help when I decided to pursue my masters' degree. Thanks to Dr. Perry Cregan for sharing his knowledge as valuable member of my committee.

I would also like to thank Justin Kerns. Without his support I would not have gotten through my masters program.

## Table of Contents

Dedication.....	ii
Acknowledgements.....	iii
Table of Contents.....	iv
List of Tables.....	vi
List of Figures.....	vii
List of Figures.....	vii
Introduction.....	1
Literature Review.....	3
<u>Soybean Importance in the Poultry Industry</u> .....	3
<u>Phytic Acid</u> .....	3
Phytic Acid: Storage.....	3
Phytic Acid: Soil Interaction.....	5
Phytic Acid: Nutrient Interactions.....	6
Phytic Acid: Phytase.....	8
<u>Myo-inositol biosynthesis and MIPS gene</u> .....	8
<u>Nutritional Issues Concerning Phytate</u> .....	10
Phytic Acid: Human Issues.....	10
Phytic Acid: Zinc.....	11
Phytic Acid: Iron.....	11
Phytic Acid: Calcium.....	12
Phytic Acid: Diet.....	12
<u>Poultry Issues</u> .....	13
<u>Pollution Concerns</u> .....	14
<u>Solutions for the Poultry Industry for Phytate</u> .....	15
<u>Low Phytate Acid Mutations</u> .....	17
Generation of Mutations.....	17
Germination of Mutations.....	21
Mutant Genotypes.....	21
LPA1 Mutants.....	22
LPA2 Mutants.....	22
<u>High Inorganic Phosphorus (HIP) Phenotype</u> .....	23
<u>LPA mutants: solutions for feed nutrition and pollution problems?</u> .....	23
<u>LPA mutations associated with the MIPS gene</u> .....	24
<u>Single Nucleotide Polymorphism</u> .....	25
<u>Ligase Chain Reaction (LCR)</u> .....	25
<u>Genomic Mapping</u> .....	27
Materials and Methods.....	29
<u>DNA Population</u> .....	29
<u>HIP Phenotype Assay</u> .....	29
<u>DNA Extraction</u> .....	31
<u>MIPS-Based Primer Design</u> .....	33
<u>MIPS Polymerase Chain Reaction (PCR) Amplification</u> .....	33
<u>Nested PCR</u> .....	34
<u>Cloning of MIPS</u> .....	34

<u>Alignment of MIPS sequence</u> .....	35
<u>Ligase Chain Reaction (LCR)</u> .....	36
Results and Discussion .....	38
<u>High Inorganic Phenotype Assay</u> .....	38
<u>Oil and Protein Concentration</u> .....	41
<u>PCR and Nested PCR for Cloning</u> .....	43
<u>Cloning and sequence alignment</u> .....	53
<u>LCR Primer Design and Reaction</u> .....	64
<u>Marker Design and F<sub>2</sub> Analysis</u> .....	68
Conclusions .....	72
Literature Cited .....	74

## List of Tables

Table 1: Phenotypic classification of F <sub>2</sub> genotypes by the HIP soybean assay conducted on F <sub>3</sub> seeds.....	38
Table 2: Primers designed based on the published MIPS sequence using Primer3 program for use in PCR-based cloning.....	44
Table 3: LCR primers designed for SNP detection.....	65
Table 4: Presence/absence of 700bp fragment from a PCR reaction with primer pair 4.....	69
Table 5: Polymorphism with primer 4, observed and expected HIP phenotype in the F <sub>2</sub> population derived from a cross of a HIP mutant and a wild-type.....	70

## List of Figures

Figure 1: Basic Structure of the stable bound form of phosphorus, phytic acid .....	4
Figure 2: Phytic Acid Interactions .....	7
Figure 3: Ethylmethane sulfonate is an alkylating agent that induces a point mutation. .....	19
Figure 4: Generation of M153 and M766, both LPA putative mutants.....	20
Figure 5: Using LCR for SNP detection creates a positive/negative product on an agarose gel. ....	27
Figure 6: High Inorganic Phosphorus Phenotype (HIP) assays were preformed on six seeds for each genotype. ....	31
Figure 7: MIPS soybean cDNA sequence from NCBI Blast website.....	37
Figure 8: Oil and Protein Concentrations of the two parental genotypes and the three HIP types.....	42
Figure 9: Three types of readable results were produced on an agarose gel with the primer pairs designed. ....	50
Figure 10: Sequenced fragment of the MIPS gene composed of the two contigs which were homologous to the Hegeman et al. (2001) published cDNA sequence.....	57
Figure 11: Alignment of contig and three published sequences using MultAlin. ....	60
Figure 12: A PCR fragment was designed from the sequenced MIPS fragment in which all SNPs between the two parental genotypes were included between the two primer pairs. ....	67
Figure 13: Agarose gel electrophoresis showing polymorphism observed using primer 4.....	68



## Introduction

Soybean (*Glycine max*) is an important crop for the economy of the United States as well as for the local economy of Maryland. The main use of soybeans is in the production of a protein supplement used in feed for agrarian animals. The increased production of poultry in recent years has increased the amount of soybean grain used in Maryland. One of the wastes produced by using soybean feed in the diet of monogastric animals is phosphorus-rich manure. Phosphorus in soybean grain is found primarily in a bound form, phytic acid, which is indigestible to monogastric animals such as poultry. The majority of phosphorus ends up in the manure of monogastrics, when the manure is continually applied to fields near poultry production farms, excess phosphorus in the manure on these farms can build up in the soil and run-off into nearby water systems. This pollutant is the major cause of algae blooms. These blooms have been linked to both an extinction of aquatic life and human environment-linked health problems.

To reduce excess phosphorus in poultry manure, a mutation in a soybean genotype was created; phytic acid in these mutants is not the main form of phosphorus in soybean seed. Mineral elements normally bound to phytic acid are available in the low phytic acid mutant genotype. To help understand how the low phytic acid (LPA) mutation is inherited, the location of the mutation in the soybean genome needs to be determined. The knowledge of the genome position would also be helpful in marker assisted breeding.

LPA mutants similar to the one in the soybean genotype have been linked to a gene coding for *myo*-inositol (3) phosphate synthase (MIPS); a major component of

the phytic acid synthesis pathway. To determine a relationship between LPA and the MIPS gene a marker needs to be designed that links the two.

An SNP (single nucleotide polymorphism) mutation on the MIPS gene might stop the production of phytic acid. A fragment of the gene was sequenced from a wild-type genotype and an LPA mutant genotype and a population derived from the cross of the two was compared for an SNP. A primer designed from the published cDNA sequence of MIPS showed a polymorphic difference between the two parental genotypes. This difference was tested on F<sub>2</sub> genotypes to determine heritability. The polymorphism was a dominant marker in the F<sub>2</sub> genotypes fitting the 3:1 ratio.

There were four objectives: 1) test the phenotype, 2) clone the MIPS gene, 3) create a marker that was associated the HIP phenotype with a genotype difference between two parental genotypes, and 4) determine the effect of HIP type on protein and oil content. The phenotype of each genotype was determined using a HIP assay, which classified the genotype as HIP positive, HIP negative, or HIP intermediate. The MIPS gene was PCR-cloned from both the wild-type parent and the LPA parent. The sequence information from the cloning was used to test SNPs between the two parental genotypes using the ligase chain reaction (LCR). A primer that showed a polymorphism between the two parental genotypes was also used to determine the association of HIP type to the MIPS gene. Protein and oil content of each type of HIP was compared to each other in order to determine whether the mutation produced a negative effect on these two agriculturally important traits in the three HIP classes. These four objectives were used to characterize the MIPS gene and map the LPA mutant in soybean.

## **Literature Review**

### **Soybean Importance in the Poultry Industry**

Five thousand years ago soybeans, along with rice (*Oryza sativa*), wheat (*Triticum sativum*), barley (*Hordeum vulgare*), and millet (*Pennisetum squamulatum*), were named as the five sacred plants of China. Since this illustrious beginning, soybeans have become a major agricultural crop. The first commercial crop of soybeans was planted in the United States in 1929 and it is a vital crop in the U.S. agricultural economy.

How important are soybeans to the American economy? Over thirty million hectares of soybeans were planted in the United States in the year 2000 with a total crop value of 12 billion U.S. dollars (American Soybean Association). One of the main uses for soybeans in the United States is as a protein source for the poultry industry. Phosphorus in soybeans has remained an untapped resource because of the chemical form in which phosphorus stored in the soybean grain. Farmers not only have to supplement feed with phosphorus, but they have to be concerned with the effect of the unused phosphorus in the manure on the environment.

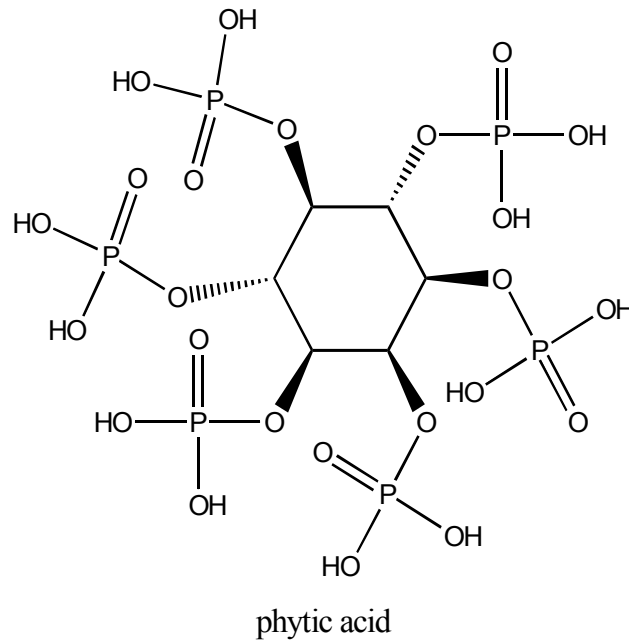
### **Phytic Acid**

#### **Phytic Acid: Storage**

Phosphorus in soybean seed is stored primarily as phytic acid (*myo*-inositol 1,2,3,4,5,6-hexakisphosphate). Phytic acid (Figure 1), identified in 1855, is an

inositol ring with six phosphate groups attached and is the most stable form of phosphorus (Oatway et al., 2001). Phytic acid is found in cereal grains, legumes, nuts, oilseeds, tubers, pollen, spores, and organic soils (Oatway et al., 2001).

**Figure 1: Basic Structure of the stable bound form of phosphorus, phytic acid (ChemDraw ultra 8.0, Adept Scientific Inc, Bethesda).**



The phytic acid as a percent total phosphorus in cereal grains, oilseeds, and legumes ranges from 60 to 97%, whereas phytic acid only represents 21 to 25% of the phosphorus in roots and tubers suggesting that phytic acid might have a role in seed germination. The proportion of phytic acid in the seeds of grain crops ranges from 0.4 to 10.7% of the total weight of the seed (Oatway et al., 2001). Only trace levels of lower inositol phosphates and other inorganic phosphorus (<5%), and organic phosphorus in forms of DNA, RNA, free nucleotides, phospholipids, sugar phosphates (10 to 20%) are found in agricultural seed (Larson et al., 2000).

Phytate is influenced by genetics, environmental fluctuations, location, irrigation conditions, soil type, year, and fertilizer application (Oatway et al., 2001). Phytic acid is stored in grains and legumes in a range of 60-97% of the total phosphorus (Oatway et al., 2001). Phytic acid concentrations vary depending on the food type and the processing of the grain. Phytic acid contents can be reduced by over 50% by processing of the grain.

In maize kernels, 90% of the phytic acid is accumulated in the embryo and 10% in the aleurone layer. In rice, barley, and wheat 90% of the phytic acid is found in the aleurone layer and 10% in the embryo (Shi et al., 2003). Phosphorus from phytate serves several functions including second messenger ligand, DNA double strand break repair, RNA export, ATP metabolism, phosphorus and mineral storage, as well as a role in the physiological response of guard cells to ABA (Raboy, 2001).

### **Phytic Acid: Soil Interaction**

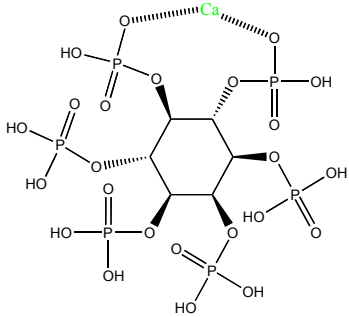
Phytic acid is a major source of phosphorus in agricultural systems. The total amount of phosphorus in soil can be very high but the majority of the phosphorus is not available to plants because of soil-phosphorus interactions. Up to 80% of phosphorus in fertilizer becomes unavailable to plants due to soil interactions (Grotz and Guerinot, 2002) and over 50% of all fertilizer phosphorus is converted into phytate which plants normally do not take up from the soil (Raboy, 2001; Schachtman et al., 1998). Manure with a high concentration of phosphorus can be applied to fields and plants can still be phosphorus deficient if the majority of the manure's phosphorus is in the bound form.

## **Phytic Acid: Nutrient Interactions**

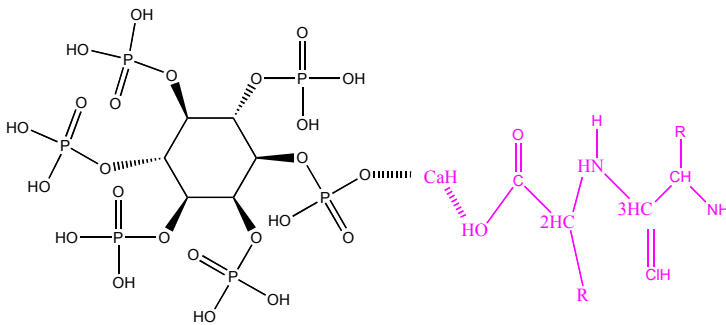
Since phytic acid has twelve replaceable protons it is a chelating agent. It is a negatively charged molecule which binds metal ions making them unavailable for digestion by monogastric animals. The majority of phytate is a mixture of phytate and mineral cations forming phytin salts. When phytic acid binds to minerals, proteins, and starch; the interaction will alter the solubility, functionality, digestion, and absorption of these components to some degree (Figure 2) (Oatway et al., 2001). Two of the nutrients that most limit plant growth (iron and phosphorus) are often abundant in soil but not available for uptake because phytic acid bonds must be broken to release both iron and phosphorus into the soil (Grotz and Guerinot, 2002). These chelating functions of phytic acid have serious nutrient effects for both human and poultry diets.

**Figure 2: Phytic Acid Interactions with (A) mineral ions, (B) proteins, and (C) starch.**

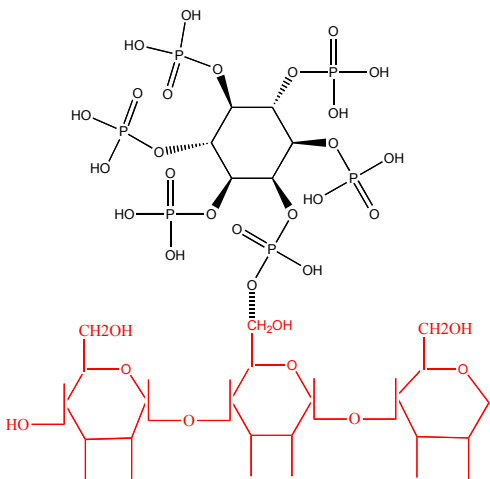
A.



B.



C.



## **Phytic Acid: Phytase**

Phytase (*myo*-inositol hexakisphosphate phosphohydrolase) is an enzyme that hydrolyses phytate releasing phosphorus and minerals from the inositol ring. During germination, phytase enzymes break down the phytin salts enabling the plant to use phosphorus and minerals for growth (Raboy, 2001). Inositol phosphates (IP1-IP5) and *myo*-inositol phosphates formed by breaking the inositol ring can be used throughout germination for the plant's phosphorus needs (Oatway et al., 2001; Hegeman et al., 2001).

Phytase is poorly understood biochemically and its exact role in phytic acid biosynthesis and other biosynthetic pathways is not completely understood (Viveros et al., 2000). Two types of phytase, 3-phytase and 6-phytase, play a role in the breakdown of phytic acid. Three-phytase is observed in microorganisms and 6-phytase is in seeds of higher plants. Both phytases are pH dependent and most active at slightly acidic levels (pH 5.1) (Oatway et al., 2001).

Phytase in cereal grains are present primarily in the aleurone layer of the grain and are not activated until germination. Therefore it cannot break down the phytic acid in animal feed since phytase is inactivated in dried seed.

### ***Myo*-inositol biosynthesis and MIPS gene**

*Myo*-inositol is a precursor to many compounds whose function in plants is phosphorus storage, signal transduction, stress protection, hormonal homeostasis, and cell wall biosynthesis (Hegeman et al., 2001). The *myo*-inositol biosynthesis pathway has important roles in plant metabolism including the synthesis of phytic acid. A complete elimination of the *myo*-inositol pathway would stop the formation of phytic



acid, but it would also disrupt many other biochemical pathways needed for germination.

The first step in *myo*-Inositol biosynthesis is the conversion of D-glucose-6-phosphorus to 1L-*myo*-inositol-phosphorus (Loewus and Murthy, 2000) by the isomerase D-*myo*-inositol-3-phosphate (MIPS). The source of inositol ring to convert glucose-6-phosphate to inositol 3 is the enzyme *myo*-inositol(3)phosphate synthase gene (MIPS) (Hegeman et al., 2001). There is evidence that the inositol backbone for phytic acid might be derived from MIPS activity (Raboy et al., 2000).

The complete DNA sequence of the MIPS gene as well as the cDNA sequence has been determine in a number of sources such as *Saccharomyces cerevisiae*, *Spirodela polyrrhiza*, *Arabidopsis thaliana*, *Citrus paradisi*, *Nicotiana tabacum*, *Glycine max (L.) Merr.*, *Zea mays*, and *Hordeum vulgare*. All of these sequences show regions of high conservation at the nucleotide level (Hegeman et al., 2001). Some plant species have the MIPS region comprised of gene families. A gene family is a group of paralogous genes. Maize has seven MIPS genes (Larson and Raboy, 1999), Arabidopsis two regions, and soybean has at least four regions (Hegeman et al., 2001).

A MIPS cDNA (GmMIPS1) sequence of soybean was isolated. GmMIPS1 is 1729 base pairs in length with a 1533 open reading frame. The open reading frame encodes a protein of 510 amino acids with the predicted molecular mass of 56.5 kD. The cDNA sequence showed a high homology to other published MIPS sequences (Hegeman et al., 2001).

## *Nutritional Issues Concerning Phytate*

### **Phytic Acid: Human Issues**

Micronutrient malnutrition affects more than a third of the world's population and iron deficiency affects over 3.5 billion people (Mendoza, 2002). The primary cause of nutrient deficiency is the poor bioavailability of nutrients from plant-based diets consumed in low income countries. Phytate binds to proteins and minerals affecting their digestibility. Anionic phosphate groups of phytate bind to the cationic groups of proteins, especially the basic amino acids. Phytate also binds trypsin and chymotrypsin. Once protein is bound to phytate the ability of animals to digest the protein is reduced (Biehl and Baker, 1997). Phytate also binds minerals such as potassium, magnesium, calcium, iron and zinc; the bound minerals cannot be digested by humans unless the phytic acid structure is broken. The ability of phytic acid to bind minerals and protein reduces the ability of grains to provide a balanced diet based solely on these grains.

In third world countries phytate consumption can cause mineral deficiency in humans (Raboy, 2001). The reduction of bioavailability of minerals and phosphorus in the diet of humans and livestock it is considered to be an anti-nutritional component of grain crops (Shi et al., 2003). In 1990, the World Health Organization, United Nations Children's Fund, and the World Summit for Children advocated the elimination of micronutrient malnutrition by the year 2000 (Gibson and Hitz, 2001). To eliminate micronutrient malnutrition either a more diverse diet must be made available to third world countries or phytic acid present in the diet must be reduced to enable more proteins and minerals to be utilized. Dietary phytate can have both

positive and negative effects in terms of human health depending on the diet of the individual consumer.

### **Phytic Acid: Zinc**

The amount of zinc needed in the diet depends on how balanced the daily diet is, if there has been major blood loss, or stresses imposed by puberty or pregnancy. The major determinant of the amount of zinc absorbed from plant material is phytic acid. A concentration of phytic acid of 15 molar to one molar of zinc has serious negative effects on zinc absorption (Lopez et al., 2002). Zinc becomes bound to the inositol ring structure when there are five or more phosphate groups attached to the ring and a high concentration of phytic acid and low concentration of zinc means that the majority of the zinc will be bound (Gibson and Hitz, 2001).

One of the common approaches used to improve zinc absorption is fermentation. Organic acids (acetic, citric, lactic, and malic acids) in the fermentation process form soluble connections to zinc; reducing the zinc: phytic acid structure (Lopez et al., 2002). Zinc becomes bound to the fermentation acids rather than to phytic acid.

### **Phytic Acid: Iron**

Iron absorption is affected by the concentration of phytic acid in bread (Lopez et al., 2002). Iron chelated to inositol rings with three or more phosphate groups is a major issue in diets because 30% of the world's population is iron deficient (Gibson and Hitz, 2001).

The ability of phytic acid to chelate iron makes it an excellent antioxidant, which makes it a good food preservative. Phytate, though not on the “generally recognized as safe” (GRAS) list prepared by the Food and Drug Administration of the United States in 1995, is used as a food additive outside of the United States. Phytic acid is added to meats, canned seafood, fruits, vegetables, cheese, miso, soy sauce, juices, alcoholic beverages, and other food to prevent discoloration, increase nutritional quality and prolong shelf life. In 1997, sodium phytate was listed as GRAS and is used as a preservative for baked goods in the U.S. (Oatway et al., 2001).

### **Phytic Acid: Calcium**

The relationship between phytic acid and calcium has not been completely determined. Some researchers have found that phytic acid inhibits calcium absorption but others have not found the same negative relationship (Lönnerdal et al., 1989; Miyazawa et al., 1996). A high concentration of calcium can hinder phytic acid breakdown. Calcium is needed in the diet but, too much can cause kidney calcifications. Phytic acid-rich foods can help maintain an adequate calcium level in the urine thereby reducing calcifications (Lopez et al., 2002). By balancing the amount of calcium absorption, phytic acid can help maintain a healthy diet.

### **Phytic Acid: Diet**

Phosphorus in the phytate form is not available for human consumption until the phosphate groups are removed from the inositol molecule (Oatway et al., 2001). High concentration of phytate in commercial soy protein isolates may affect the bioavailability of proteins and trace minerals such as, zinc and iron, because of the

chelating mechanism (Brooks and Charles, 1984). Functional properties of soy proteins in commercial food products can be adversely affected by phytase (Brooks and Charles, 1984). Since humans do not have phytase naturally in their system, humans depend on food processing to degrade phytic acid. Cooking or processing can degrade phytic acid; milling of wheat and polishing of rice grains decreases phytic acid by up to 90%; soaking legume seeds reduces phytic acid by 20%; and germination of the seed reduces phytic acid by 50% (Hurrell et al., 2002). Studies have found some food processes will increase the insolubility of the phytic acid complex further decreasing mineral availability. For example, extrusion cooking of starches rich in phytate inactivates phytase enzymes (Oatway et al., 2001).

Phytic acid should not be completely taken out of the diet; phytic acid's ability to bind minerals can actually help in a balanced diet to ensure that one does not get diseases associated with excess minerals. Phytic acid has been described as hypocholerolemic, an antioxidative, anticarcinogenic and hypolipidemic (Oomah et al., 1996). Phytic acid in the diet could have the same positive effect as a high-fiber diet in a situation where the population is able to attain a well-balanced diet, but in low-income areas the positive aspects of phytic acid in the diet is overshadowed by the negative aspects.

### **Poultry Issues**

Deficiencies in phosphorus limit growth and bone development in poultry (Waldroup et al., 2000). Micronutrients are drastically reduced in livestock feed as a consequence of phytic acid chelating metal ions such as iron, zinc, calcium, potassium, sodium, chlorine, and magnesium (Erdman, 1981; Wilcox et al., 2000).

Poultry, like all nonruminant animals, have a limited amount of phytase to breakdown phytic acid. Poultry farmers need to combat the problem of limited phosphorus in the poultry diet in order to produce healthy birds. Since most of the phosphorus present in soybeans is phytate phosphorus; inorganic phosphorus has historically been added to poultry feed to supplement their diet (Huff et al., 1998).

### **Pollution Concerns**

After the second world war (WWII) excess phosphorus in manure became a serious problem in agricultural systems. Before WWII, farming communities only supported the local economy, therefore, the phosphorus sink and source were in the same locale and manure nutrients were recycled. The excess phosphorus in the manure of animals was a small enough that manure could be used on local fields without the excess becoming a detriment to the water supply. After WWII agricultural systems became more specialized, separating the grain and animal producing areas (separating the sink and source of phosphorus). Animal products have become the main component in the agrarian economy and most of the soybean industry focuses on feed for animals. Since 1995 livestock has increased 10 to 30% whereas farms have decreased 40 to 70% (Sharpley et al., 2001). The concentration of phosphorus entering the soil has increased dramatically leading to concerns about farmlands as a major source of pollution.

Eutrophication is a condition where waters are rich in mineral and organic nutrients promoting a proliferation of plant life which reduces dissolved oxygen which forces the extinction of some marine organisms. Excess nutrients in water runoff can potentially cause algal blooms, which produce toxins harmful to both the

environmental system and humans (Preusch et al., 2002). Massive blooms of *cyanobacteria* cause fish kills, unpalatable water, and the formation of carcinogens when water is chlorinated (Sharpley et al., 2001).

In 1996, the Environmental Protection Agency (EPA) of the U.S. identified eutrophication as the most widely spread impairment to water sources in the US with agriculture being a major contributor of phosphorus (www.EPA.gov). Nitrogen and phosphorus are nutrient-limiting for algae growth in the Chesapeake Bay watershed (Preusch et al., 2002). The poultry industry in Maryland has placed pressure on the Chesapeake Bay watershed by the quantities of fresh poultry litter (FPL) produced near the water-system. In 1999, over 680,000 Mg of FPL were applied to farmland surrounding the Delmarva region (Delaware, Maryland, and Virginia) and 425,666 Mg of FPL in the Virginia farmland region (Preusch et al., 2002). The Chesapeake Bay region must now try to solve water quality issues arising from the disposal of poultry wastes in a way that is cost effective for the industry.

### **Solutions for the Poultry Industry for Phytate**

There are two current ways to deal with the low availability of phosphorus to monogastric animals and the excess phosphorus in the waste of monogastrics: 1) the addition of microbial phytase to the diets to increase available phosphorus; and 2) genetic alteration of the grain used in the poultry industry (Li et al., 2000).

In the 1960s scientists at the International Minerals and Chemicals Corporation determined the amount of supplemental inorganic phosphorus in monogastric animals could be greatly reduced by adding phytase enzyme to the feed (Huff et al., 1998). Adding extra phytase enzyme to a broiler's diet increases the

availability of phosphorus to poultry (Nelson et al., 1971; Waldroup et al., 2000). Commercial development of phytase enzyme for the animal industry has resulted in a decrease of phosphorus in animal manure (Huff et al., 1998). Phytase supplements can break down up to 50% of feed phytate reducing the amount of phosphorus waste in the manure (Raboy, 2001).

Phytase from *Asperigillus niger* was isolated and used in poultry feed to help reduce the amount of phosphorus wasted in poultry production. The use of phytase in poultry feed has improved the absorption of phosphorus by poultry (Lan et al., 2002). Using phytase in broiler hens feed allows 56.8 to 59.1% more phosphorus from the seed to be digested (Lan et al., 2002). As phytic acid in the seed is broken down by phytase the availability of minerals such as calcium, iron, and zinc will also increase (Lan et al., 2002). Minerals released from the bond with phytic acid can be used in seed germination by the plant or absorbed in the digestive tract by animals. Lei et al. (1993) showed that supplementing swine diets with phytase also reduced fecal phosphorus.

There are three approaches to genetically alter grain crops for increased micronutrient availability: 1) increase concentration of micronutrients in plants; 2) increase concentration of promoter compounds; and 3) decrease anti-nutrients. Increasing micronutrients in plants has been attempted in rice. Three constructs were expressed in the endosperm of rice. *Phaseolus* phytoferritin was expressed to increase the concentration of the micronutrient of iron in the plant. A cysteine-rich metallothionein-like protein is in the pathway to create iron. Overexpression of this promoter increased iron concentrations in rice. *Aspergillus fumigatus* phytase is



inserted into the rice genome in order to decrease phytate, an anti-nutrient. By decreasing phytate, the concentration of iron that can be absorbed in human consumption is increased (Mendoza, 2002). This engineered rice has not been tested yet for human consumption, but it has the potential to alleviate iron deficiencies in areas where rice is the main diet staple.

The second approach is to increase the concentration of promoter compounds. It is hoped that by increasing the amount of proteins there will be an increase in mineral utilization. Increasing the concentration of cysteine and lysine has been seen to have a positive effect on the absorption of zinc (Mendoza, 2002). Breeding is currently being done to increase promoter compounds, but it is not known if a higher concentration of proteins rich in cysteine and lysine will negatively affect other functions of the plant.

The third approach is the development of cultivars with lower levels of micronutrient chelating agents such as phytate. Breeders have started producing plants with decreased concentrations of phytic acid. Since phytic acid is one of the main inhibitors of iron and zinc it was hypothesized that a reduction in the level of phytic acid would increase mineral absorption.

### **Low Phytate Acid Mutations**

#### **Generation of Mutations**

Ethylmethane sulfonate (EMS) induces point mutations in DNA. It is an alkylating agent, which adds an alkyl group to the O-side chain of guanine and thymine. Alkylated guanine pairs to thymine instead of cytosine when the DNA structure is altered by EMS (Figure 3). EMS mutates each cell of the seed embryo

independently of other cells. Seeds are typically soaked in EMS for 10 to 20 hours, washed, and grown. The M1 generation of plants contains chimeras and only mutations in their reproductive tissue are inherited in the developing seed (M2 generation) (McCallum et al., 2000). The M2 generation of seeds are grown and screened for the desired mutation.

Wilcox et al. (2000) generated a soybean mutation line by soaking seeds of CX1515-4 for 24 hours in 18mM solution of EMS. Seeds were then rinsed with water and planted. M3 seeds were tested for HIP phenotype and two progenies were classified putative LPA mutations, M153 and M766 (Figure 4). Both of these two independent, heritable, and nonlethal mutants are phenotypically similar to *lpa1-1* type.

Figure 3: Ethylmethane sulfonate is an alkylating agent that induces a point mutation.

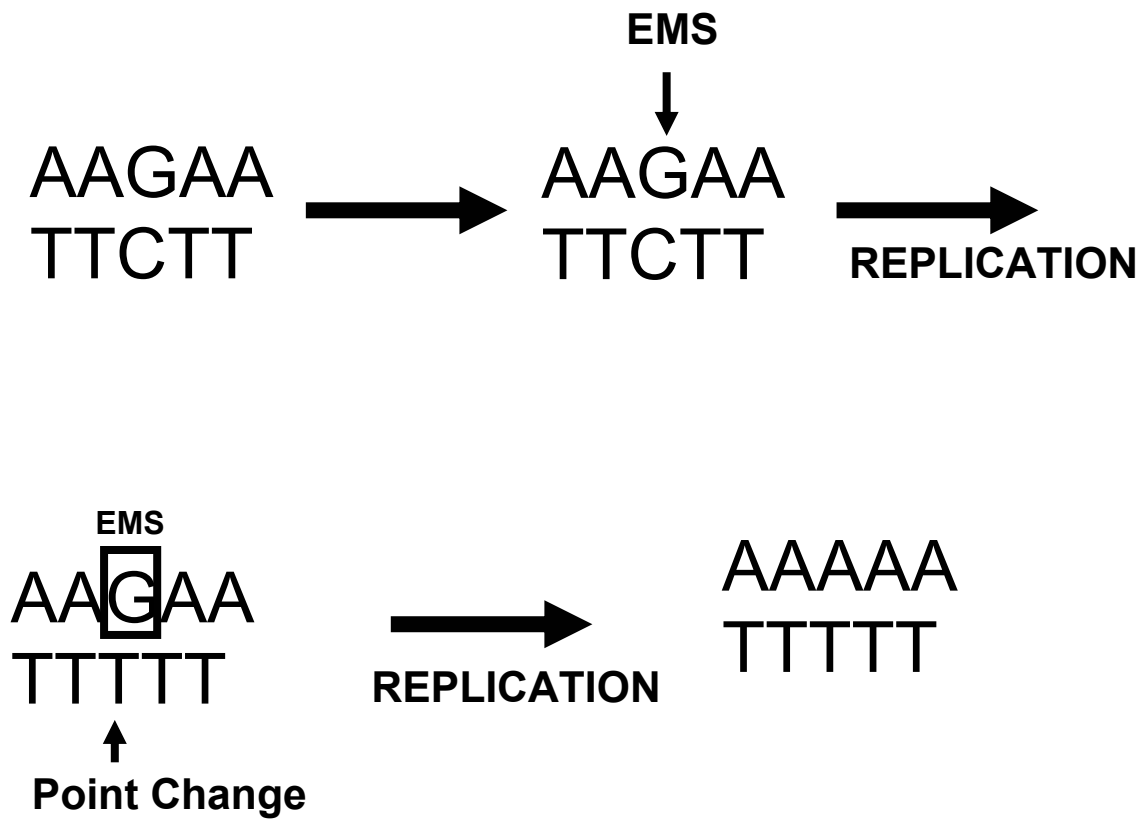
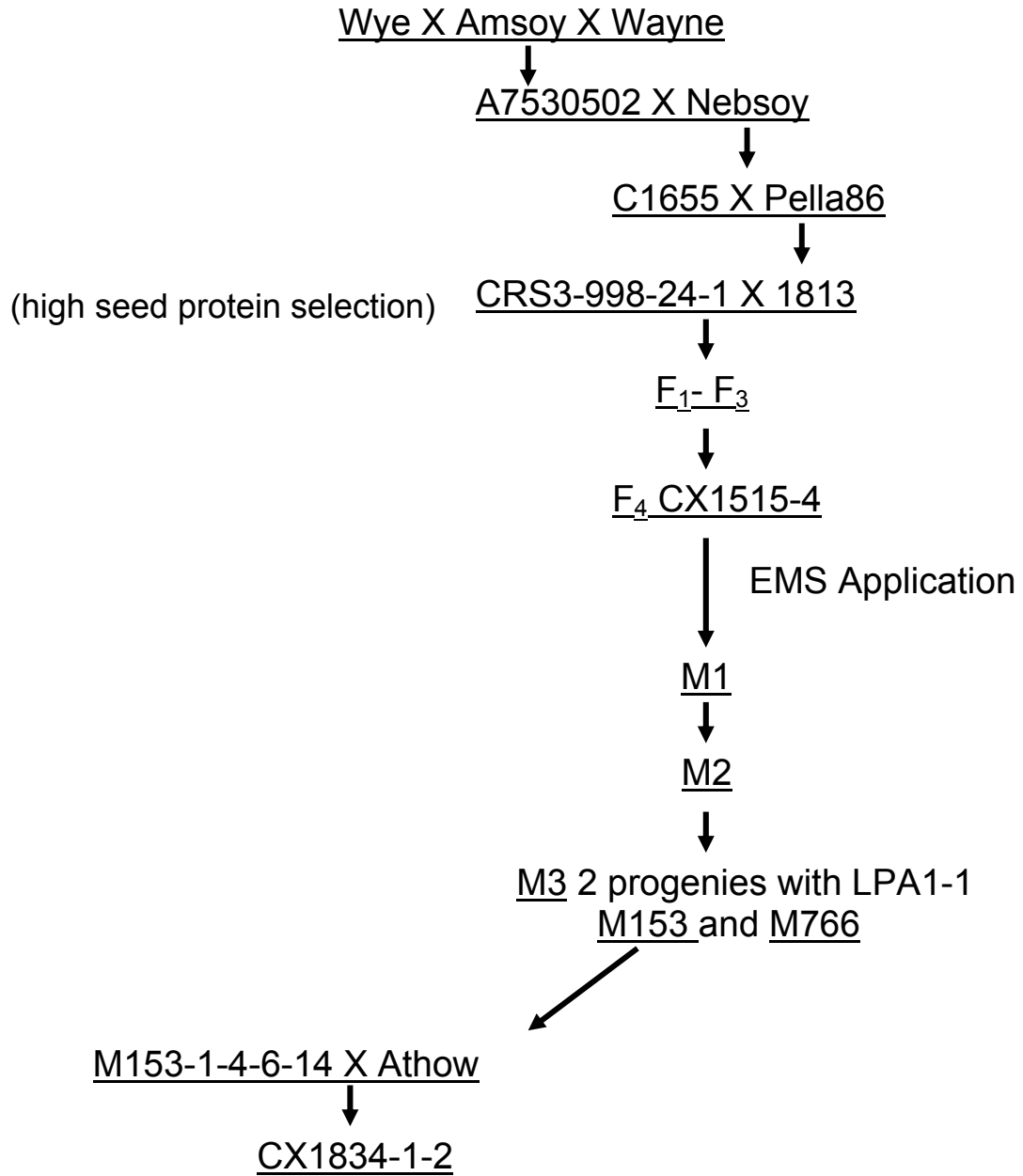


Figure 4: Generation of M153 and M766, both LPA putative mutants (Wilcox et al., 2000).



## **Germination of Mutations**

Is germination compromised by lower concentrations of phytic acid?

Soybeans have higher phytate stores than are needed for normal seed functions in germination. Without compromising normal germination, phytic acid can be reduced by two-thirds in soybean seed (Hegeman et al., 2001). This indicates that low phytic acid mutants can be grown and used as feed.

Using phosphate starvation during seed development it was shown that soybean seeds with reduced phytic acid still germinate; indicating high levels of phytate in soybean seeds is not necessary for seedling growth (Raboy and Dickinson, 1987). Phytic acid is the main storage of phosphorus in seeds and serves to prevent phosphorus and mineral leaching. However, the high levels of phytic acid are not necessary for germination of domesticated crops (Larson et al., 2000).

## **Mutant Genotypes**

LPA mutant genotypes of corn (*Zea mays*), rice (*Oryza sativa*), barley (*Hordeum vulgare*), and soybean (*Glycine max*) have been isolated. These mutants have genetically reduced amounts of phytic acid of at least 70% (reductions range from 50% to 95%) in the seed. Reduction of phytic acid in these mutants corresponds to an equal increase of inorganic phosphorus. Alleles with reductions of greater than 90% cause severe loss of plant growth functions (Raboy, 2001).

There are two types of loss-of-function mutations; low phytic acid 1-1 is the first recessive allele of the *lpa1* locus and low phytic acid 2-1 is the first recessive allele of the *lpa2* locus. Both of the mutant types produce normal levels of total phosphorus. A hypothesis has been proposed that the difference in the types of

phosphorus that increase in the mutant seeds are based on the mutation's location in the phytate synthesis pathway. LPA1 mutations are from a change in the gene encoding substrate supply early in the pathway whereas LPA2 mutations occur in late in the inositol phosphate pathway. Both mutations reduce phytic acid content in the seed, but each mutant accumulates different inositol phosphorus types because the mutants affect different parts of the pathway.

### **LPA1 Mutants**

In seed of *lpa1-1* mutants the phytate concentration is reduced by two-thirds compared to the wild-type and has an equivalent increase in inorganic phosphorus. The *lpa1-1* phenotype has low phytic acid, but it does not accumulate inositol polyphosphates (IP<sub>3</sub> to IP<sub>5</sub>) (Raboy et al., 2000). The maize *lpa1-1* mutation has reduced phytic acid that is matched by a molar equivalent increase in phosphorus (Dorsch et al., 2003). It has been suggested that the low phytate phenotype described by Wilcox et al. (2001) is the result of two mutations (Oltmans et al., 2004).

### **LPA2 Mutants**

LPA 2 mutants are apparently located later in the phytic acid biochemical pathway after *myo*-inositol formation. In the maize *lpa2-1* mutant, the reduction of phytic acid is coupled with an increase in both phosphorus and inositol phosphate with five or fewer phosphorus esters (IP<sub>3</sub> to IP<sub>5</sub>) (Dorsch et al., 2003). In the typical wild-type seeds, inositol phosphates with six or more phosphorus esters are present more frequently than five or fewer phosphorus esters (Raboy, 2001).

Shi et al. (2003) used a *Mutator* insertion to knockout ZmIpk, a maize inositol phosphate kinase gene. After ZmIpk was silenced, seeds showed reduced phytic acid and increased *myo*-inositol phosphates (IP<sub>3</sub> to IP<sub>5</sub>) similar to the low phytic acid mutant lpa2-2 observed in maize. Shi et al. (2003) cloned and sequenced the ZmIpk gene from the lpa2-2 maize genotype and observed that the allele had a nucleotide mutation causing premature termination of the ZmIpk open reading frame. The lpa2-2 mutation in maize might be from a mutation in the ZmIpk gene, which correlates with the theory that lpa2-2 mutations affect the phytic acid biosynthesis pathway after formation of *myo*-inositol.

### **High Inorganic Phosphorus (HIP) Phenotype**

The high Inorganic Phosphorus (HIP) phenotype is associated with a homozygous LPA mutation. Normal seeds have low levels of inorganic phosphorus, typically less than 0.5 mg inorganic phosphorus per gram. A simple color assay has been developed using a Chen reagent (Chen et al., 1956) that detects the different concentrations of inorganic phosphorus allowing selection of high inorganic phosphorus genotypes in plant breeding programs (Raboy, 2002).

### **LPA mutants: solutions for feed nutrition and pollution problems?**

Human diets could be enhanced by using LPA seeds in a limited nutrition diet. Mendoza (2002) studied the effect of low phytic acid mutants on mineral absorption in humans. The absorption in humans of iron, zinc, and calcium from their diet was tested to determine if there was a difference between wild-type maize and low phytic acid maize. There was an increase in the absorption of all three minerals when using

the low phytic acid maize. Maize low phytic acid mutants could be incorporated into the diet to increase mineral absorption.

LPA mutants also have been shown to have a positive effect on dietary problems in swine and poultry. LPA seeds could both alleviate excess phosphorus in manure and also incorporate nutrients from feed in the animal diets. Bird weight was shown to increase by 16% when compared to a diet that included only normal seed. Calcium absorption increased in some poultry experiments (blood calcium from 29 to 36% and bone calcium from 11 to 13%) (Mendoza, 2002). Phosphorus absorption was also enhanced by the use of low phytic acid mutants (blood phosphorus from 28 to 36% and bone phosphorus of 10%) (Mendoza, 2002). Fecal waste phosphorus was reduced from 9% to 40%; this reduction helps alleviate the excess of phosphorus found in the environment.

#### **LPA mutations associated with the MIPS gene**

One copy of the maize MIPS gene mapped to the same location as the *lpa-1* trait on chromosome 1S (Larson and Raboy, 1999). A mutation affecting MIPS expression could be the cause of the low phytate accumulation in the seeds (Hegeman et al., 2001).

Hitz et al. (2002) characterized a mutation in soybean that decreased both phytic acid and raffinose production. The mutation was linked to a single base change in the myo-inositol 1-phosphate synthase gene which caused a decrease in the viability of the gene but not the complete loss of function.

The association between the MIPS gene and the LPA phenotype is not always present in other cereals. In barley and rice the LPA phenotype is not associated with



mutations in known MIPS regions. One hypothesis is that barley and rice contain cryptic copies of genes encoding MIPS activity. Also, it could be that the LPA phenotype is a mutation in regulatory genes affecting MIPS expression and not the actual MIPS gene itself (Hegeman et al., 2001).

### **Single Nucleotide Polymorphism**

A locus is deemed a single nucleotide polymorphism (SNP) when it exists in at least two variants and the allele frequency of the most common variant is less than 99% (Landegren et al., 1998). There are four reasons for the use of SNPs as markers in genetic analysis: 1) SNPs are prevalent in the sequence and therefore large sets of markers can be found near most loci of interest; 2) some SNPs in the gene sequence can directly affect protein structure and expression levels, and manipulation of the SNP can alter genetic mechanisms of disease; 3) SNPs are stable and inherited unlike many repeated sequences used as markers; and 4) SNPs can be used to analyze sequences with large throughputs (Landegren et al., 1998). There are various strategies that are in use for SNP determination. One of the more popular is mismatch distinction by polymerases and ligases. Polymerization reactions have very stringent requirements on correct base pairing at the 3' end of hybridizing primers. This idea can be used to amplify one allele of a SNP (Landegren et al., 1998). Once a SNP is determined it could be used as a marker in breeding.

### **Ligase Chain Reaction (LCR)**

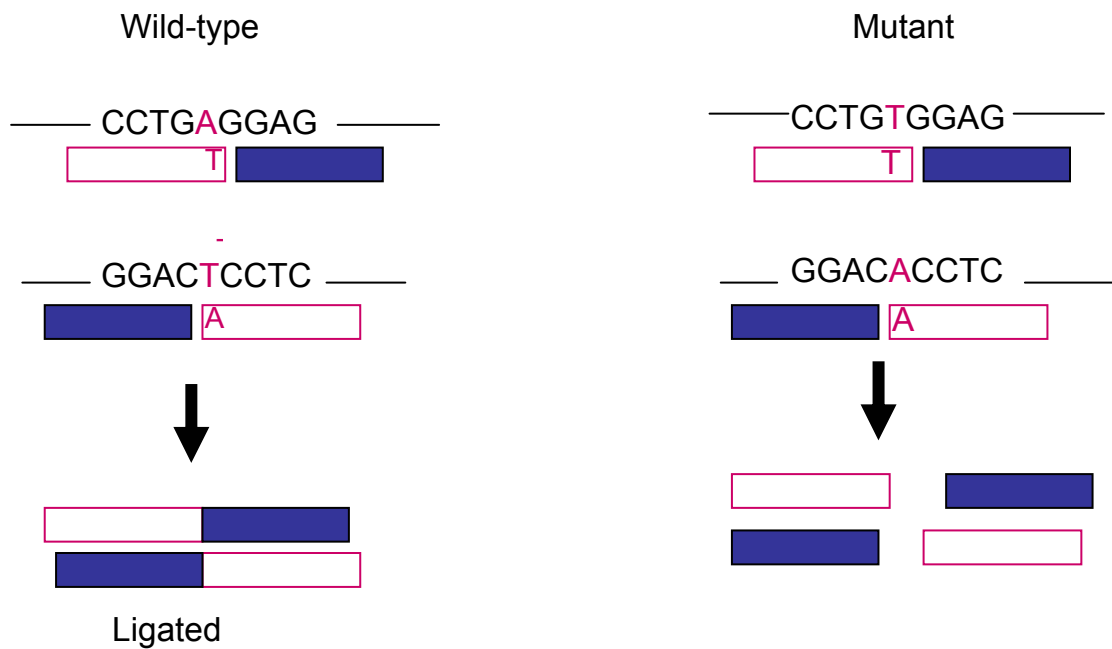
The ligase-mediated gene detection technique can be used to detect a difference of a single base pair between two genotypes. This procedure has been

used to detect mutations that cause sickle cell anemia and is used as a procedure to screen for other diseases (Weiss, 1991). Ligase chain reaction (LCR) uses two adjacent oligonucleotides which when ligated will form a complementary strand to the target DNA. Oligonucleotides are designed to completely cover the target DNA sequence. The 3' end of one oligonucleotide is adjacent to the 5' end of the other oligonucleotide. The oligonucleotides are perfectly base-paired with the target DNA and are 20 to 25 base pairs in length (Barany, 1991). These oligonucleotides are incubated with the DNA sample and ligated. If there is a base pair difference between the DNA and the oligonucleotides then the ligation will not be complete (Karthigesu et al., 1995).

LCR is a cyclic two-step reaction (Figure 5). Double-stranded DNA is unwound to become single-stranded during a melting step. Then in a cooling step two adjacent, complementary oligonucleotides anneal to the single-stranded target and ligate together. This product is the template for the next ligation reaction cycle. If the sequence is present without a SNP mutation then the oligonucleotides align to the DNA perfectly. Ligase interprets the two oligonucleotides as a complementary copy of the target DNA and welds the two oligonucleotides together creating permanent, covalent bonds that create a full length copy a DNA complementary sequence. After numerous cycles many copies of the complementary sequence are made and can be viewed visually on an agarose gel. If there is a mutation in the sequence, then the oligonucleotides might anneal to the DNA, but the SNP creates a mismatch between the oligonucleotide and the target DNA and therefore ligase does not create a complementary copy. This mismatch allows the wild-type to be visualized on an

agarose gel as a positive band whereas the mutant sequence will have no band (a negative result) (Weiss, 1991).

**Figure 5: Using LCR for SNP detection creates a positive/negative product on an agarose gel.**



### Genomic Mapping

Genetic mapping provides an indirect estimate of the distance between two regions of a genome. Markers are variations in the genome that can be observed between two genotypes. Genetic mapping is a useful tool if markers are linked to a gene of interest. Linked means that the markers are inherited together; indicating they are both close together on the same chromosome. If a marker and a gene of interest are linked then that marker can be used to test genetically if the genotype has

a variation that is important. In other words, if there is a mutation of interest and one has a marker that is linked to this mutation then other genotypes can be screened to determine if they too have this mutation. The goal in marker design is to have a marker that is completely linked to the gene of interest. The marker must be on the same chromosome and close to the gene of interest in order for both the marker and the gene of interest to be inherited together. If a marker is on the same chromosome, but not close to the gene of interest then the linkage between the two is known as partial linkage. In partial linkage the marker and gene of interest might be inherited together or they might segregate independently from each other. If the marker only has partial linkage then it can not be used to test for the gene of interest.

## **Materials and Methods**

### **DNA Population**

A low phytate mutant (LPA) soybean genotype (CX1834-1-2) from Purdue University was crossed with an adopted elite line (MD96-5722). Phenotypic marker traits (such as flower color) were used to confirm that the F<sub>1</sub> plants were the result of successful crosses. The F<sub>2</sub> generation of this cross was grown at the Wye Research and Education Center near Queenstown, Maryland. F<sub>2</sub> leaves and F<sub>3</sub> seeds were harvested individually from 115 plants.

### **HIP Phenotype Assay**

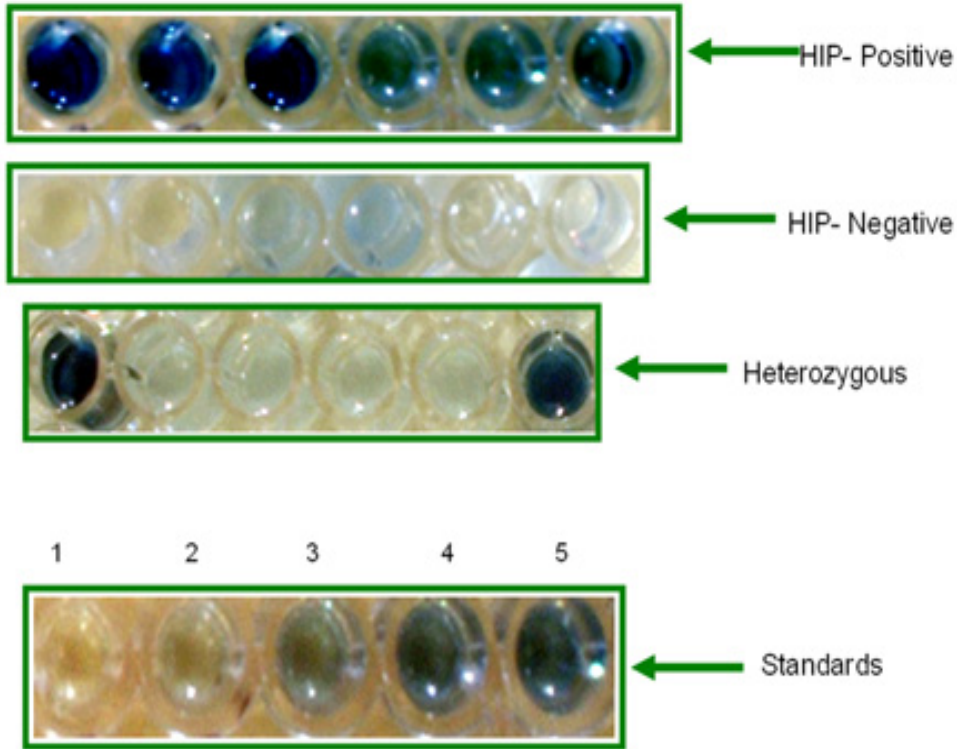
Six F<sub>3</sub> seeds from each of the 115 F<sub>2</sub> plants were individually crushed with a hammer. Each crushed seed was extracted overnight in 2.5mL of 12.5% (v/w) TCA:25 mM MgCl<sub>2</sub> at 4° C with gentle shaking. Extracts were allowed to settle for 30 minutes at room temperature the next morning. Aliquots of each single-seed extract were assayed for inorganic phosphorus by a modified version of the Chen et al. (1956) protocol. A 10 µl aliquot of each single-seed extract, 90 µL of DD H<sub>2</sub>O, and 100 µL of the colorimetric reagent were placed in a microtitre plate well. The colorimetric reagent consisted of one volume of 3M H<sub>2</sub>SO<sub>4</sub>, one volume of 2 mM ammonium molybdate, one volume of 10%(v/v) ascorbic acid and two volumes of DD H<sub>2</sub>O. Assays were incubated at room temperature for 2 hours. The assays of the extraction mixture were visually scored for presence or absence of HIP. In general a dark blue color indicates a presence of HIP, whereas, a clear is a negative for the HIP

phenotype. The five phosphorus standards used in the assay were made by diluting 1mM  $K_2HPO_4$  to 0.0  $\mu\text{g P}$ ; 0.15  $\mu\text{g P}$ ; 0.46  $\mu\text{g P}$ ; 0.93  $\mu\text{g P}$ , and 1.39  $\mu\text{g P}$ .

The two parental genotypes were assayed and the results were used to determine the classification of each of the  $F_2$  genotypes (Guttieri et al., 2004). The inorganic phosphorus standard closest to the wild-type genotype inorganic phosphorus concentration was used to label HIP negative seeds (clear color assay). The inorganic phosphorus standard closest to the LPA genotype inorganic phosphorus concentration was used to label HIP positive seeds (dark color assay). The concentrations between the parental genotypes were classified as intermediate. The range of intermediate concentrations was divided in half based on concentration of inorganic phosphorus. The lower half of the range was classified as HIP negative intermediates (light blue assay) and the top half of the range was classified as HIP positive intermediates (blue assay) (Figure 6).

$F_2$  genotypes were classified into the three HIP phenotypes based on five classifications (Guttieri et al., 2004). 1) HIP negative: all the seeds tested HIP negative 2) HIP heterozygous: at least one seed was classified HIP negative and one seed was classified HIP positive. 3) HIP heterozygous: at least one seed was classified as HIP positive and the rest of the seeds were classified as HIP negative intermediates. 4) HIP positive: none of the seeds tested HIP negative and all were HIP positive intermediates. 5) HIP positive: all the seeds tested HIP positive.

**Figure 6: High Inorganic Phosphorus Phenotype (HIP) assays were performed on six seeds for each genotype.**



### **Oil and Protein Concentration**

The concentration of both oil and protein from dry-weight analysis of each parental and F<sub>2</sub> genotypes were measured using the Infratech model 1255 feed and food analyzer. Two standards, one with high oil concentration and one with high protein concentration, were used to ensure that the infrared machine was adjusted properly.

### **DNA Extraction**

Genomic DNA from the two parental genotypes and the F<sub>2</sub> population was extracted by a phenol/chloroform method. Leaves were collected from plants grown at the Wye Research and Education Center approximately six weeks after planting.

One hundred to 150 micrograms of leaves of each genotype were ground using a micropestle attached to a cordless drill. One milliliter of extraction buffer, composed of 50mM Tris (pH 8.0), 10mM EDTA (pH 8.0), 100mM LiCl, 2% SDS, and 10 $\mu$ g/ml proteinase K (Sigma Aldrich) was heated to 50°C and added to the ground leaves. The solution of extraction buffer and leaves were mixed for 15 minutes using an orbit rotator. The mixture was then centrifuged for 15 minutes at 8000 rpm. The supernatant was spilt into two tubes and RNase A (10 $\mu$ g/ $\mu$ l) was added. The mixture was incubated at 37°C for 2 hours. Five hundred microliters of 99% phenol was added and the solution was inverted 10 times. The solution was then centrifuged for 2 minutes at 10,000 g. The supernatant was placed into a new tube and 500 $\mu$ l of chloroform:isoamyl alcohol (24:1) was added. The solution was again mixed and spun at 10,000 g. The supernatant was collected and 250 $\mu$ l phenol and 250 $\mu$ l chloroform:isoamyl alcohol (24:1) was added. This mixture was again mixed and spun. The supernatant was collected into a fresh, sterile tube. DNA was precipitated by adding 0.25 volumes of 10M ammonium acetate and two volumes of 100% ethanol. The solution was kept at -20°C for at least 8 hours to precipitate the DNA. The solution was centrifuged for 15 minutes at 12,000 g. The resulting pellet was washed with 500 $\mu$ l of 70% ethanol. The DNA pellet was air dried until ethanol the residue had evaporated. Thirty microliters of 10mM Tris (pH 8.0) was added to dissolve the pellet. DNA concentration was measured by a Spectronic Genesys 2 spectrophotometer or a Biomate 3 thermospectronic spectrophotometer and diluted to a final concentration of 50ng/ $\mu$ l. The DNA was stored at -20°C until needed.



### **MIPS-Based Primer Design**

Based on the soybean MIPS cDNA sequence (GI 13936690) primers were designed to sequence the genomic sequence of the MIPS gene in the two parental genotypes. The primers were designed using the program Primer 3 ([http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi/](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi/)) covering the entire cDNA sequence.

### **MIPS Polymerase Chain Reaction (PCR) Amplification**

DNA from the parental genotypes was used with primers designed to amplify segments of the MIPS gene genomic sequence. The total volume for each PCR reaction was 50 $\mu$ L which included: 50 ng/ $\mu$ L DNA template, 1X PCR Buffer (10 mM KCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 20 mM MgSO<sub>4</sub>, 0.1% Triton X-100 (pH 8.8)), 1.75-2.0 mM MgCl<sub>2</sub>, 100-150  $\mu$ M dNTP mix (equal parts dATP, dCTP, dGTP, dTTP), 0.2  $\mu$ M of each primers; and 0.3  $\mu$ L of NEB Taq DNA polymerase (5 units/ $\mu$ L) or 0.1  $\mu$ L of lab generated Taq DNA polymerase. PCRs were run in a PTC 100 MJ Research (Watertown, MA) machine. An initial denaturation at 93 °C for 3 minutes, then 25 cycles of denaturation 93 °C for 30 seconds; 53°C-65°C annealing for 45 seconds; and 72°C for 45 seconds for extension, and a 10 minute 72°C extension.

PCR products were visualized on a 1.7% agarose gel stained with ethidium bromide. Bands between 300bp and 1.5kb were excised from the agarose gel and cleaned for cloning using the protocol of QIAprep Gel Extraction kit.

### **Nested PCR**

Nested PCR was also used to produce fragments for analysis from regions where multiple bands were too numerous to extract singly. A PCR was run with a single pair of primers and the results were checked on an agarose gel. Then this PCR was used as the template for a second PCR using primer pairs which are nested inside the first. In this way instead of getting numerous multiple bands, only a single band was produced. These reactions were also visualized using a 1.7% agarose gel. All reactions that gave a 500bp to 1.5kbp band size were kept for further analysis.

### **Cloning of MIPS**

The Topo-TA cloning kit (Invitrogen) was used to clone PCR products that were previously isolated. The PCR products are ligated into a pCR 2.1-TOPO vector. One to four microliters of fresh PCR product were added to 1  $\mu$ L salt solution and 1  $\mu$ L of the TOPO vector. This was incubated for 15 minutes at room temperature. Two microliters of the reaction was added to One Shot chemically competent *E. coli* and this mixture was incubated on ice for 15 minutes. The cells are then heat shocked by placing in a 42° C water bath for 30 seconds. Two hundred and fifty micro liters of SOC medium were added to the reaction and shaken at 37° C, 200rpm for 1 hour. The cells were then plated on LB medium, which contained kanamycin and x-gal. These plates were incubated at 37° C for at least 24 hours until colonies of cells were observed. White cells and whitish-blue cells were taken from the plates and cultured overnight. The white cells were placed in a solution of 5mL LB broth with 15  $\mu$ L kanamycin (200mg/mL) and grown overnight at 37° C with shaking at 250 rpm.

Plasmid DNA was isolated using the genelute plasmid mini-prep kit (Sigma-Aldrich). The cells were pelleted from the overnight culture by centrifugation (Sorvall RC 26 Plus) at 12000 g for 5 minutes. The cells were resuspended, lysed, neutralized, and washed using the protocol provided with the genelute kit. Plasmid DNA was analyzed by restriction digestion to determine the size of the cloned DNA fragment.

A restriction digest was performed using 10  $\mu$ L plasmid DNA, 1  $\mu$ L *ECORI* enzyme, and 1X *ECORI* buffer in a total reaction solution of 20  $\mu$ L. The reaction solution was incubated at 37° C for 2 hours. The reaction solution was run on a 1.7% agarose gel for 1.5 hours at 90 volts. Plasmid DNA, the same size as the cloned PCR product was selected. For each PCR product, two colonies were examined.

Plasmid DNA (150-200 ng/ $\mu$ l) was then sent to the University of Maryland College Park DNA sequencing facility in the Center for Biosystems Research (CBR) for sequence analysis. Sequence analysis was determined using an Applied Biosystems DNA sequencer (model 3100). Each sequence fragment was analyzed twice with a forward primer and a reverse primer (a total of four sequences).

#### **Alignment of MIPS sequence**

Fragments were processed by CBR sequencing facility and screened using the NCBI Vector Screen program to eliminate any vector sequence contamination. The fragment was run through BLAST to see if it had a high match to the published soybean cDNA MIPS gene sequence (Figure 7). The two forward and two reverse sequences were aligned and scanned for mismatches. If three of the four sequences had the same nucleotide then it was assumed that the fourth was a sequencing error.

If the sequences did not agree on one nucleotide, a code for possible nucleotides was inserted. For example, if two of the sequences had the nucleotide adenosine and the other two had tyrosine then Y, unspecified pyrimidine nucleoside, was used for the combined sequence. The sequence from each parental genotype was compared to one another and mismatches between the parental were highlighted for further study.

Sequencher program (Gene Codes Corporation, Michigan) was used to combine fragments into contigs. A contig is a group of fragments that represent overlapping regions of the sequence. These contigs were further analyzed for similarity to published MIPS sequences.

### **Ligase Chain Reaction (LCR)**

The ligation reaction was a 25  $\mu$ L reaction volume containing 50 pmol of each oligonucleotide designed to flank the mismatch nucleotide, 25 ng DNA template, 1X Ampligase reaction buffer (20 mM Tris-HCL (pH 7.6), 25 mM potassium acetate, 10 mM magnesium acetate, 10 mM DTT, 600  $\mu$ M nicotinamide adenine dinucleotide (NAD), 0-1% Triton X-100), and 50 units Ampligase (Stratagene, California). These reactions were heated at one cycle of 92° C for 4 minutes and 60° C to 70° C for 3 minutes; then 30 cycles of: denaturing at 92° C for 1 minute, and ligation at 60° C to 70° C for 1 minute.

**Figure 7: MIPS soybean cDNA sequence from NCBI Blast website (Hegeman et al., 2001).**

GTGAAAAATAATGTTTCATCGAGAATTTTAAGGTTGAGTGTCTAATGTGA  
AGTACACCGAGACTGAGATTTCAGTCCGTGTACAACACTACGAAACCACCGAA  
CTTGTTACAGAGAACAGGAATGGCACCTATCAGTGGATTGTCAAACCCAA  
ATCTGTCAAATACGAATTTAAAACCAACATCCATGTTCTAAATTAGGGG  
TAATGCTTGTGGGTTGGGGTGGAAACAACGGCTCAACCCTCACCGGTGGT  
GTTATTGCTAACCGAGAGGGCATTTCATGGGCTACAAGGACAAGATTCA  
ACAAGCCAATTACTTTGGCTCCCTACCCAAGCCTCAGCTATCCGAGTTG  
GGTCCTTCCAGGGAGAGGAAATCTATGCCCCATTCAAGAGCCTGCTTCCA  
ATGGTTAACCCCTGACGACATTGTGTTTGGGGGATGGGATATCAGCAACAT  
GAACCTGGCTGATGCCATGGCCAGGGCAAAGGTGTTTGACATCGATTTGC  
AGAAGCAGTGGAGGCCTTACATGGAATCCATGCTTCCACTCCCCGGAATC  
TATGACCCGGATTTTCATTGCTGCCAACCAAGAGGAGCGTGCCAACAACGT  
CATCAAGGGCACAAAGCAAGAGCAAGTTCAACAAATCATCAAAGACATC  
AAGGCGTTTAAGGAAGCCACCAAAGTGGACAAGGTGGTTGTACTGTGGA  
CTGCCAACACAGAGAGGTACAGTAATTTGGTTGTGGGCCTTAATGACACC  
TGGAGAATCTCTTGGCTGCTGTGGACAGAAATGAGGCTGAGATTTCTCCT  
TCCACCTTGTATGCCATTGCTTGTGTTATGGAAAATGTTCTTTCATTAAT  
GGAAGCCCTCAGAACACTTTTGTACCAGGGCTGATTGATCTTGCCATCGC  
GAGGAACACTTTGATTGGTGGAGATGATTTCAAGAGTGGTCAGACCAAAA  
TGAAATCTGTGTTGGTTGATTTCTTGTGGGGGCTGGTATCAAGCCAACAT  
CTATAGTCAGTTACAACCATCTGGGAAACAATGATGGTATGAATCTTTCG  
GCTCCACAAACTTTCCGTTCCAAGGAAATCTCCAAGAGCAACGTTGTTGA  
TGATATGGTCAACAGCAATGCCATCCTCTATGAGCCTGGTGAACATCCAG  
ACCATGTTGTTGTTATTAAGTATGTGCCTTACGTAGGGGACAGCAAGAGA  
GCCATGGATGAGTACACTTCAGAGATATTCATGGGTGGAAAGAGCACCAT  
TGTTTTGCACAACACATGCGAGGATTCCCTCTTAGCTGCTCCTATTATCTT  
GGACTTGGTCCTTCTTGCTGAGCTCAGCACTAGAATCGAGTTTAAAGCTG  
AAAATGAGGGAAAATTCCACTCATTCCACCCAGTTGCTACCATACTCAGC  
TACCTACCAAGGCTCCTCTGGTTCCACCGGGTACACCGGGTGTGAATGC  
ATTGTCAAAGCAGCGTGCAATGCTGGAAAACATAATGAGGGCTTGTGTTG  
GATTGGCCCCAGAGAATAACATGATTCTCGAGTACAAGTGAAGCATGGGA  
CCGAAGAATAATATAGTTGGGGTAGCCTAGCTGAATGTTTTATGTTAATA  
ATATGTTTGCTTATAATTTTGCAAGTGTAATTGAATGCATCAGCTTCATTA  
ATGCTTTAGAGCGGGGCATATTCTGTTTACTAGGAACATGAATGAATGTA  
GTATAATTTT GTGTAAAAAAAAAAAAAAAAAAAAAAAAA

## Results and Discussion

### High Inorganic Phenotype Assay

The high inorganic phosphorus (HIP) phenotype assay is a quick way to determine if a genotype can be considered to possess a low phytic acid genotype (Raboy et al., 2000). A genotype with low phytic acid stores phosphorus as inorganic phosphorus, which gives a positive HIP phenotype result. Inorganic phosphorus extracted from each genotype was measured in the F<sub>3</sub> seeds of a cross between the wild-type and the LPA mutant. A ratio of 64 wild-types: 34 intermediate types; 17 LPA types was observed (Table 1). This ratio did not fit the 1:2:1 ratio anticipated for a single gene segregation ( $\chi^2= 58.96$ , P= 0.001, 2 df.). This indicated that more than one gene controlling the low phytic acid mutations in the LPA soybean genotype, but the results do not fit typical ratios associated with multiple genes.

**Table 1: Phenotypic classification of F<sub>2</sub> genotypes by the HIP soybean assay conducted on F<sub>3</sub> seeds\*.**

<b>HIP Phenotype<sup>†</sup></b>	<b>HIP -/-</b>	<b>HIP +/-</b>	<b>HIP +/+</b>
<b>Observed</b>	<b>64</b>	<b>34</b>	<b>17</b>
<b>Expected</b>	<b>28.5</b>	<b>57</b>	<b>28.5</b>

\*  $\chi^2= 58.96$  P < 0.001 , 2 df.

<sup>†</sup> HIP -/- = HIP Negative  
HIP +/- = HIP Intermediate  
HIP +/+ = HIP Positive

There were two main areas for experimental errors: the procedural execution, and the scoring constraints. Using a hammer to break the seeds for inorganic

phosphorus extraction produced pieces of seed that were relatively large and not uniform. Pilu (2003) and Larson et al. (2000) both used the HIP phenotype assay for primary determination of mutants. Larson et al. (2000) ground rice seeds using a hammer, whereas, Pilu (2003) extracted uniformly by grinding maize seeds individually with mortar and pestle. The extraction of inorganic phosphorus could have been effected by the size of the seed pieces used. The extraction buffer will interact with more surface area if the fragments are small, potentially affecting the amount of inorganic phosphorus extracted.

The definition of a positive HIP phenotype may bring error into the HIP phenotype assay method. Larson et al. (2000) indicated that for the HIP assay seeds testing higher than the third standard ( $0.49\mu\text{g}$ ) were HIP positive and lower than the third standard were HIP negative. Raboy et al. (2000) classified HIP positive seeds as those that had a dark blue color and HIP negative as a clear. Larson et al. (2000), and Raboy et al. (2000) both stated that wild-type seeds had an inorganic phosphorus concentration of less than  $0.15\mu\text{g}/\mu\text{l}$  and low phytic acid seeds have inorganic phosphorus concentrations higher than  $0.7\mu\text{g}/\mu\text{l}$ .

In the studies of Raboy et al. (2000) and Larson et al. (2000) the HIP phenotype was controlled by one gene. The HIP phenotype assays had clear differences between the three phenotypic classes. In studies with more than one gene, intermediate types can affect the category into which a particular line is placed. Guttieri et al. (2004) developed a classification system to categorize the intermediates into the three HIP phenotypes. In this study, only six seeds per genotype were assayed. If more than one gene is involved in the LPA mutation some genotypes

labeled HIP negative might have been mislabeled and are HIP heterozygous. The mislabeling could have occurred because more than six seeds would have to be analyzed to ensure that all genotypic variations of the F<sub>3</sub> seeds were observed. The mislabeling of HIP negative could be the reason that a typical ratio could not be determined for the HIP types.

Oltmans et al. (2004) used a mutant line provided by Wilcox et al. (2000) as one of the parental genotypes in a cross between wild-type and mutant. In 2001 she crossed CX1834-1-6 (LPA mutant) to A00-711013 at the Agricultural Engineering and Agronomy Research Center in Iowa. CX1834-1-6 used in the Oltmans et al. (2004) study is from the cross of Anthow and M153 performed by Wilcox et al. (2000), which is the same lineage as the LPA mutant used in this study. Oltmans et al. (2004) progeny test resulted in 127 phenotypic of normal phytate types, 73 phenotypic of intermediate types, and 10 phenotypic of low phytate types. These results, similar to current study, did not fit a 1:2:1 ratio expected for a single gene affecting HIP.

Oltmans et al. (2004) speculated that the low phytate mutation could be controlled by recessive alleles at two independent loci exhibiting duplicate dominant epistasis, but there were insufficient F<sub>3</sub> seeds to test the two-gene model. Reciprocal crosses of the wild-type and the mutant were made and in the F<sub>2</sub> generation 197 seeds had normal phytate and 13 low phytate (15:1 ratio); indicating recessive alleles at two independent loci both showing dominant epistasis (Oltmans et al., 2004).

Though Oltmans et al. (2004) and this study used a LPA mutant genotype from Wilcox's et al. (2000) program, the ratios of the three HIP phenotypes differed.



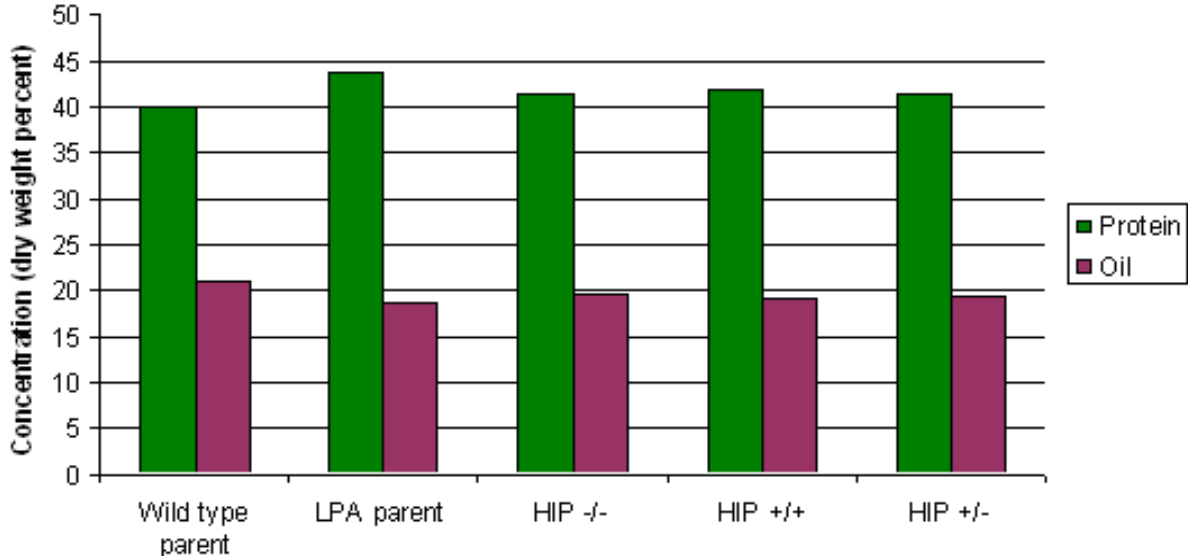
The choice of wild-type parent might have affected the results of the HIP assay. Oltmans et al. (2004) also changed the methodology used to extract inorganic phosphorus, reducing the time the ground pieces of seeds incubated in the extraction buffer. Different soaking times of the seeds in TCA: magnesium chloride could have affected the amount of inorganic phosphorus extracted from each seed. No tests have been conducted to determine if a reduction in extraction time will affect the amount of inorganic phosphorus eluted from a seed. If more than one gene is involved in the mutation in soybeans, it could affect how the HIP phenotypes are classified. Both studies did find that more than one gene might be controlling the LPA trait in soybean.

To determine if the phenotypes of HIP were correctly classified, a ferric phosphate assay can be performed. Pilu (2003) and Larson (2000) both used this assay to further evaluate LPA mutants. This test was used to ensure that the genotypes which were classified as HIP positive were also low in phytic acid. A ferric phosphate assay would allow the classification of the intermediates and might alter the ratio of phenotypic classes.

### **Oil and Protein Concentration**

Oil content and protein content were evaluated in a population segregating for the HIP phenotype. An ANOVA was done using SAS (SAS Institute, Inc 1996) and no statistical difference among the three classes of HIP phenotypes: HIP negative, HIP intermediate, and HIP positive in oil and protein concentration was observed (P=0.5 oil; P=0.866 protein) (Figure 8).

**Figure 8: Oil and Protein Concentrations\* of the two parental genotypes and the three HIP types<sup>†</sup>.**



\* Differences between the three phenotypes are not statistically significant for either oil or protein (P=0.5 oil; P=0.866 protein).

<sup>†</sup> HIP -/- = Negative HIP

HIP +/+ = Positive HIP

HIP +/- = Indeterminate HIP

Soybeans are grown for two reasons: protein and oil. Soybeans supply a quarter of the world's edible oil and two-thirds of protein meal (Li and Burton, 2002).

Soybeans are used as feed for poultry because of the seeds high protein content (Oatway et al., 2001), if protein content in a genotype is reduced that genotype will not be used. The main goal of any breeding program, therefore, is to increase these two components or at least not sacrifice the ability of the plant to produce high

quantities for the sake of another characteristic. Low phytic acid plants could be used in a breeding program without reduction of protein or oil concentrations.

### **PCR and Nested PCR for Cloning**

To determine genetic differences in the MIPS gene between both parental genotypes, the gene was cloned from each genotype and the sequences were compared. A published cDNA sequence (Hegeman et al., 2001) was used to design primers for the genomic sequencing of the MIPS gene. This cDNA sequence was isolated using reverse transcriptase polymerase chain reaction (RT-PCR) with RNA from developing soybean seeds. Comparing the sequence to published sequences, it had high identity to other published MIPS sequences; and was most similar to the tobacco MIPS sequence.

The MIPS cDNA sequence (GI 13936690) from soybean was used to generate PCR products using the Primer3 program. PCR products were generated to help design a marker from a MIPS sequence that is related to the HIP phenotype. A total of 56 primer pairs were created from the original published cDNA sequence (Table 2). These primers produced three types of results using the wild-type parental DNA and the mutant parental DNA as templates: 1) single bands; 2) multiple bands; and 3) smears (Figure 9).

**Table 2: Primers designed based on the published MIPS sequence using Primer3 program for use in PCR-based cloning.**

<b>Primer Pair</b>	<b>Comments*</b>
1 TAAGGTTGAGTGTCTAATG TTCCATTAGGACACTCAACC	Cloned 1.2 kb
4 TGGAGGCCTTACATGGAATC GAGCCTTGGTGAGGTAGCTG	Polymorphic
5 CAGAATATGCCCCGCTTCG GCTTCATTAATGCTTTAGAG	Cloned 700bp
13 CTTGTGGGGGCTGGTAGC TAAAGCATTAATGAAGCTGATG	Cloned 500bp
13-Nested Primer Pair ATTCCAACCTCACAGGATTACC CTCGCCCCGTATAAGACAAA	
15 TAAACAGAATATGCCCCGC GCGGGGCATATTCTGTTTAG	Cloned 400bp
20 TAAGGTTGAGTGTCTAATG TACTCATTAGGACACTCAACC	Cloned 1Kb
21 TAAGGTTGAGTGTCTAATG CATTAGGACACTCAACCTTA	Cloned <sup>†</sup> 800bp
22 GAATCCTCGCATGTGTTGTGGC CCATTAGGACACTCAACCTTA	Cloned 350bp

<b>Primer Pair</b>	<b>Comments*</b>
23 TATTCCTTTTGTAAATTCATTC GAATCCTCGCATGTGTTGTG	Cloned <sup>†</sup> 700bp, 500bp
23-Nested Primer Pair TAAGGTTGAGTGTCTAATG GACTCAACCTAACATTAG	
24 GAATCCTCGCATGTGTTGTGG GCCACAACACATGCGAGGATTC	Cloned <sup>†</sup> 1Kb
25 TTNGAGCGGATAACANTTTCC GACATTAGGACTCAACTTA	Cloned <sup>†</sup> 750bp
26 CTCTAAAGCATTAATGAAGCT CCAAACACAATGTCGTCAGGG	Cloned 450bp
27 CAGCATAATGAAGCTGATGCAAT CAAACACAATGTCGTCAGGG	Cloned <sup>†</sup> 650bp
28 AGCTCAGCACTAGAATCGAGT GCTTCATTAATGCTTTAGAG	Cloned <sup>†</sup> 650bp
29 CTTGTGGGGGCTGGTATCAAG TGGTCCNTCTTGCTGANCTCAC	Cloned <sup>†</sup> 600bp
30 GTCCTTCACCGACCGATCCAAAC GAGCGGGGCATATTCTGTTTA	Cloned <sup>†</sup> 750bp
31 CCCTGACGACATTGTGTTTGTC CATCAGCTTCATTATGCTG	Cloned <sup>†</sup> 450bp
32 TGCAGAATTCGCCCTTCCCTG GCTTCATTAATGCTTTAGAG	Cloned <sup>†</sup> 400bp

<b>Primer Pair</b>	<b>Comments*</b>
33 CTTGGAGATTCCTTGGAGAAC CTAGTANTTATCAACAACATTTT	Cloned <sup>†</sup> 650bp
34 CGTAGGGGACAGCAAGAGAGG GAAATTACAACCTTTATAGGCC	Cloned <sup>†</sup> 650bp
35 CGTAGGGGACAGCAAGAGAGGC GAACCTTGCCTTCTTTCATGCT	Cloned <sup>†</sup> 650bp
36 CTGCCAACCAAGAGGAGCGTGC GGATGATTCCTTGTGGGGGT	Cloned <sup>†</sup> 650bp
37 ACGTAGGGGACAGCAAGAG CTTCATTAATGCTTTAGAGAAG	Cloned 300bp
38 GAATCCTCGCATGTGTTGTGG CATTAGGACACTCAACCTTA	Cloned 500bp
39 CTCTAAAGCATTAAATGAAGCTG TCAAACACAATGTCGTCAGGG	Cloned 900bp
40 CTCTAAAGCATTAAATGAAGCTG CAAACACAATGTCGTCAGGG	Cloned <sup>†</sup> 700bp
41 CTCTAAAGCATTAAATGAAGCT TCAAACACAATGTCGTCAGGG	Cloned <sup>†</sup> 950bp
42 TAAACAGAATATGCCCCGCTAC GTCTCTTGCTGTCCCCTACGT	Cloned <sup>†</sup> 450bp

<b>Primer Pair</b>	<b>Comments*</b>
43 GGATTCTCTTCTTTATTCCTTTG CTGGGTGGAATGAGTGGAAT	Cloned <sup>†</sup> 650bp
43-Nested Primer Pair AAACAGAATATGCCCCGCTTCG GCTTCATTAATGCTTTAGAGA	
44 GCCCTTCCCTGACGACATTGTG GCTTCATTAATGCTTTAGAG	Cloned <sup>†</sup> 800bp
45 GGATTCTCTTCTTTATTCCTTTG CTTGGAGATTCCTTGGA	Cloned <sup>†</sup> 600bp, 1Kb
45-Nested Primer Pair AAGAGGAGCGTGGGTCCC GCCACACTGGCAGGTCGG	
46 GGATTCTCTTCTTTATTCCTTTG CTTCCAGGGAGAGGAAATC	Cloned <sup>†</sup> 650bp, 1Kb
46-Nested Primer Pair CCCTGACGACATTGTGTTTG CTCCAATTCTGGACCCA	
47 GAATGTGTCATGGAGGGTGT CTCTCTTGCTGTCCCCTAGC	Cloned 300bp
47-Nested Primer Pair GGAGATTCCTTGGAACAT GCTAGGGACAGCAAGAGAG	
48 CTCTAAAGCATTAATGAAGC GATAATAGGAGCAGCTAGAGG	Cloned 750bp
49 GGATTCTCTTCTTTATTCCTTTG CTCTCTTGGTGTCCCCTAGC	Cloned 200bp

<b>Primer Pair</b>	<b>Comments*</b>
50 ACGTAGGGGACAGCAAGAGA TAACAGAATATGCCCCGCT	Cloned <sup>†</sup> 600bp, 900bp
50-Nested Primer Pair CTGCCAACCAAGAGGAGCGTG GTGGGGGGNCTGGGAATCA	
51 GAATCCTCGCATGTGTTGTGGC GCCATTAGGACACTCAACCT	Cloned <sup>†</sup> 400bp
52 GGATTCTCTTCTTTATTCCTTTTG CTCTAAAGCATTAAATGAAGCTGATG	Cloned <sup>†</sup> 850bp, 500bp
52-Nested Primer Pair TAAGGTTGAGTGTCTAATG GACCTTATAAGTGAAATGGAGA	
53 CCTTCCAGGGAGAGGAAATC CTCTAAAGCATTAAATGAAGCTGATG	Cloned <sup>†</sup> 650bp
53-Nested Primer Pair CCGACCTTAAGCGGGAATTTG GTAATTACGAAATCTCTTCCCG	
54 CCTTCCAGGGAGAGGAAATC CTCTAAAGCATTAAATGAAGCTGATG	Cloned <sup>†</sup> 400bp
54-Nested Primer Pair ATTTGTCTTATACGGGGCG GTCAGAGAACGACAGGGGATGC	
55 CCTTCCAGGGAGAGGAAATC CTCTAAAGCATTAAATGAAGCTGATG	Cloned 500bp
55-Nested Primer Pair GCCAACCAAGAGGAGCGTGC TTTATGCTAATTCATTA	



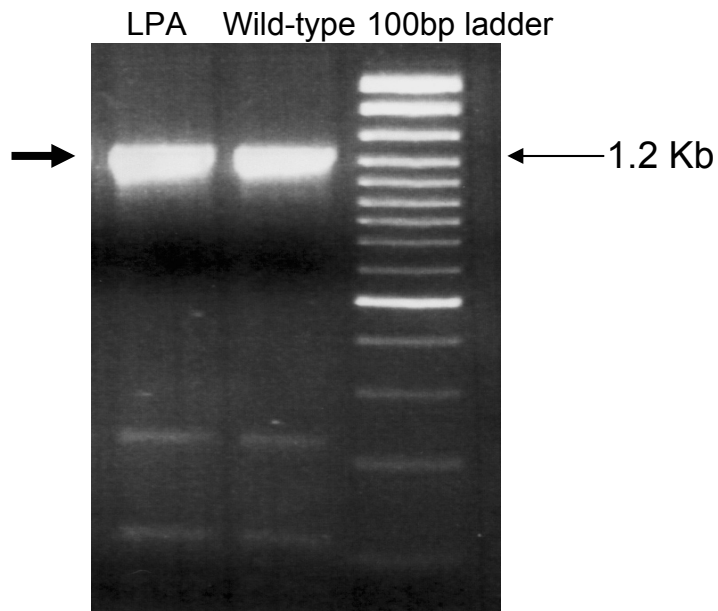
<b>Primer Pair</b>	<b>Comments*</b>
56 TAAACAGAATATGCCCCGCTTC CAGCTTCATTAATGCTTTAGAG	Cloned <sup>†</sup> 700bp

\* Cloned= TOPO-TA cloning of the PCR fragment  
Bp= size of the band(s) cloned

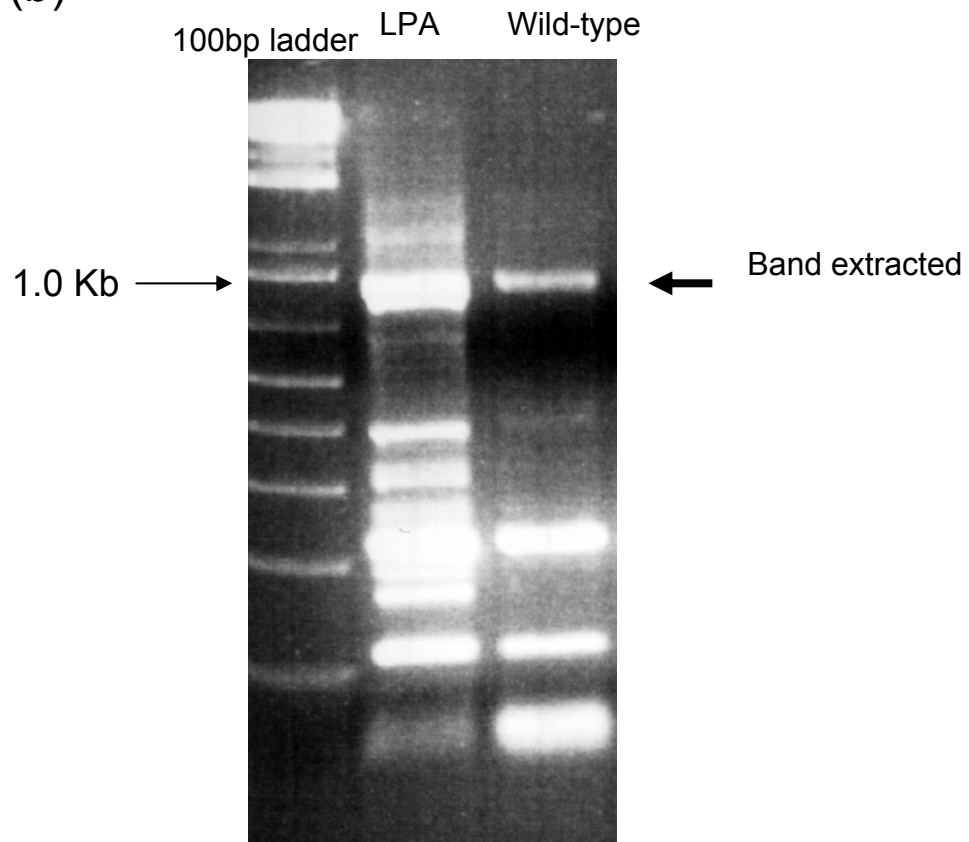
<sup>†</sup> Band(s) were extracted from an agarose gel and then cloned

**Figure 9: Three types of readable results were produced on an agarose gel with the primer pairs designed. (a) Single: the PCR product is a single fragment (b) Smear with an extractable band (c) Smear used in a nested PCR to produce an extractable single band**

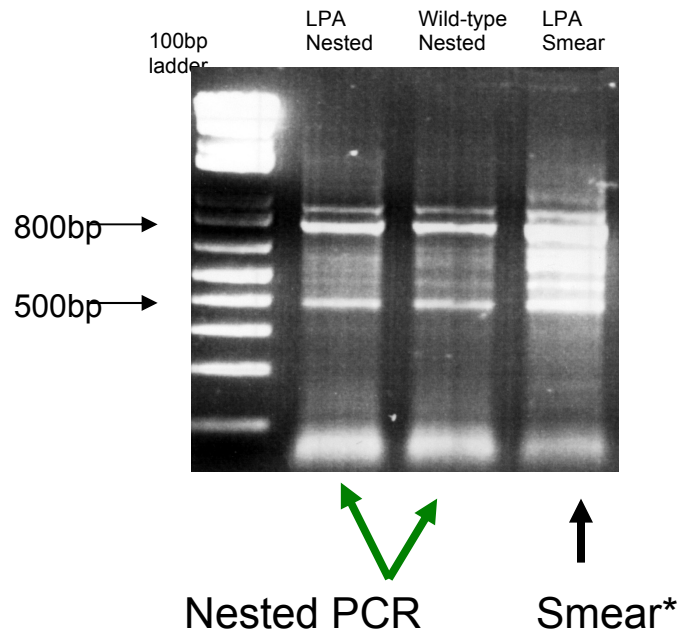
(a)



(b)



(c)



\*The smear PCR product is used as the template for the nested PCR. This nested PCR produced two bands (800bp and 500bp) that were extracted from the gel.

Twenty-one of the primer pairs used in a PCR reaction produced single fragments which were visualized on an agarose gel. These reactions were directly cloned into a vector and sequenced.

Nine primer pairs' reactions produced multiple fragments. The individual fragment bands produced on the agarose gel were extracted. Multiple fragments were produced when primers annealed to more than one part of a genome. Which fragments were used to clone and sequence? The distance between primers could be estimated based on the MIPS cDNA sequence. The genomic sequence produced was the size of the cDNA sequence or greater than the cDNA sequence because of included introns. The excised fragments which met that size requirement were cloned and sequenced.

Twelve PCRs produced reactions which were visualized on agarose gels as smears. A smear occurs when numerous fragments are produced by a PCR reaction to such a degree that it is not possible to distinguish individual fragments. This PCR reaction was used as a template in a second reaction with another set of primers (nested PCR). The nested PCRs which produced single fragments were cloned.

Primer pair 4 produced a 700bp fragment in the wild-type parental genotype but not in the mutant parental genotype. Primer 4 was tested with F<sub>2</sub> genotypes to determine if there was a linkage between the HIP phenotype and the polymorphism.

### **Cloning and sequence alignment**

Larson et al. (2000) showed a link between the lpa1-1 mutation and the MIPS gene in rice, barley, and maize. The MIPS gene in all three of these species was

mapped to chromosomes that contain orthologous regions, and the *lpa1-1* mutation in each of these species were mapped to MIPS regions.

Pilu (2003) characterized a low phytic acid mutant (*lpa241*) in maize and determined that the mutation originally assigned to *lpa2-1* type mutation is more similar to *lpa1-1* type mutations. Pilu (2003) used RT-PCR to compare MIPS expression between wild-type and mutant seedlings; a lower level of expression was observed in the mutant. Pilu (2003) sequenced MIPS cDNA from the mutant seedling; a 3' 925bp sequence was produced with 100% homology to the published MIPS sequence (AF056326). The 5' prime end of the sequence, however, could not be amplified using primers designed in the 5'UTR region. This suggested a mutation of the MIPS gene on the 5' prime end which could affect transcript processing of the MIPS gene (Pilu 2003).

A change in the MIPS gene might be the cause of the *lpa1-1* mutation in soybean and could be the result of a SNP between the two parental genotypes. If there is a link between the MIPS gene and *lpa1-1* mutation as Larson et al. (2000) and Pilu (2003) suggest, evaluating the MIPS gene for genetic differences between the parents would help determine if the *lpa1-1* mutation in CX1834-1-2 is a mutational change of the MIPS gene.

Thirty seven PCR products were cloned in order to compare the genetic differences of the MIPS sequences between the mutant and wild-type parental genotypes. Each fragment was run through the BLAST program (Altschul et al., 1990) to ensure a low e-value with the published cDNA MIPS sequence (Hegeman et

al., 2001). If the fragments did not show similarity to the published sequence more plasmids from that clone were sequenced and analyzed.

Each fragment was sequenced four times, two colonies both sequenced in the reverse direction and forward direction. One hundred twenty-six fragments were sequenced and 58 fragments had similarity to the published cDNA MIPS gene sequence. The fragments sequenced from each parental genotype were compared. Any base differences between the two were highlighted.

Sequences from each parental genotype were assembled into contigs. A contig is a group of fragments that represent overlapping regions of similar sequences. The Sequencher program (Gene Codes Corporation, Michigan) trimmed the ends of the fifty-eight fragments that had similarity to the published sequence. These fragments were then processed with a 80% match percentage and minimum overlap of 20bp constraint that produced fifteen contigs.

The MIPS gene is part of a multigene family (Hegeman et al., 2001) so a contig can be produced which has similarity to the gene but is not an actual fragment of the specific MIPS gene of interest. For instance, many contigs had sequence similarities to the beginning of the published cDNA sequence and to the end of that sequence. These contigs were less than a kilobase in length; it is doubtful that one contig spanned the length of the gene. It was more likely the primers used to produce the contig annealed to other locations on the genome similar to the MIPS gene. Two contigs showed similarity to the published cDNA sequence that made sense based on the size of the contig and where the contig were homologous to the published sequence (Figure 10). For each of these sequences, the contig would have perfect

homology to a part of the cDNA sequence that was considered one of the exons, then there was a part on the contig that did not match the published sequence (indicative of an intron) and then it was perfectly aligned to the next exon on the published sequence (Figure 10).



**Figure 10: Sequenced fragment\* of the MIPS gene composed of the two contigs which were homologous to the Hegeman et al. (2001) published cDNA sequence.**

CTGCCAACCA AGAGGAGCGT GCCAACACG TCCATCAAGG GCACAAAGCA  
 AGAGCAAGTT CAACAAATCA TCAAAGACAT CAAGTATGCC CTACTIONCAAT  
 AATAATCTAC AAAATCATCA TGATATGTGT GATCAGTATG TATTTTTTTT  
 TCAGCTTCTA ACTGTTGACA TGATTATGAA CTCAGGGCGT TTAAGGAAGC  
 CACCAAAGTG GACAAGGTGG TTGTACTIONTG GACTGCCAAC ACAGAGARGT  
 CAGTAATTTG GTTGTGGGCC TTAATGACAC CATGGAGAAT CTCTTGGCTG  
 CTGT<sup>A</sup>GGACAG AAATGAGGCT GAGATTTCTC CTTCCACCTT GTATGCCATT  
 GCTTGTGTTA TGGAAAATGT TCCTTTCATT AATGGAAGCC CTCAGAACAC  
 TTTTGTACCA GGTTGGATGC TTGCCNTTTA TGTTCCCCTC CACCAACYTT  
 TAAATTGKTT GNATTTATGC TAAT<sup>A</sup>TCATTT AATTAACCTT GGGCTTGKGG  
 NTTGNCAACA GGSCTGGATT GGATCTTTGN CCATCGCGAG GGAA<sup>C</sup>AACNTT  
 TGGATTGGTG NGGAAGATGA CTTTCAAGAG GKGGYCNAAA CCAAATGAA  
 AATCTGGNGK TGGGNWKGAN TTNCT<sup>TG</sup>CTTG <sup>T</sup>GGGGKNCTGG <sup>A</sup>GTATCAAGGT  
 ACATTGATTT TATACTAATG TCATATCATT GGTGTTATTT TTACTAGCAC  
 TATCTACCTT TGCATGCTTT GTTTAAAATT TATGCATTGT GCTAATTTTA  
 TATGGCAAAA ATGGAGAAAC TAGTTGATTT TTTCAAATGT TCTGTTGATG  
 ATTGGCAGCC AACATCTATA GTCAGTTACA ACCATCTGGA AACAAATGATG  
 GTATGAATCT TTCGGCTCCA CAACTTTCC GTTCCAAGGA AATCTCCAAG  
 AGCAACGTTG TTGATGATAT GGTC AACAGC AATGCCATCC TCTATGAGCC  
 TGGTGAACAT CCAGACCATG TTGTTGTTAT TAAGGTAAAT TTTGTTTCAC  
 CCATTTTTCT GTTTCTTTCT CTTGTCAGGG CTTTGATTAT TCTATCTGCT  
 TTGTTGCCTT TGCAGTATGT GCCTTACGTA GGGGACAGCA AGAGAGCCAT  
 GGGATGAGTA CACTTCAGAG ATATTCATGG GTGGAAAGAG CACCATTGGT  
 TTTGCACAAC ANATGCGAGG ATTCCCCCTC <sup>G</sup>TA<sup>A</sup>CTGCTCC TATTAYNTNG

GACTTGGTCC NTCTTGSTGA NCTTCASCNC TAGAATCGAG TTAAAGCTG  
 AAAATGAGGT CTGTATGCAT TGCTAAATAA TTTTCATTGCT <sup>T</sup>GGATTGTATT  
 AGTCATTGTA TTTTTTTTCA GCCTTCTGAG GTCCACTATG TTGGTGGTGA  
 TTGCTGCAGG GAAAAATTCC ACTCATTCCC MCCCAGTTGC TACCATCCTT  
 CAGCTACCTC ACCAAGGCTC CTCCTGGTGA GTTTAATCTA ATCTTATTCC  
 AATGATGATT GATGAGTACA TATCCTTATT TCTATATAGG AATTAAGTTG  
 AAAGAAACAT ATCCTCTGAA TTTTAGTATC GAGATAAAAT TGTTCOAATT  
 GTGAAGTTAA ATCATTGTTG TTCAGTGAAT TTCCCTTTTT TGCCTTATG  
 ATTTTTTTTT ATCAAAAGTT GGCAAGGTTA CTCACAACAG TACTCTCACA  
 GGTTCCACCG GGTACACCAG TGGTGAATGC ATTGTCAAAG CAGCGTGCAA  
 TGCTGGAAAA CATAATGAGG GCTTGTGTTG GATTGGCCCC AGAGAATAAC  
 ATGATTCTCG AGTACAAGTG AAGCATGGGA CCGAAGAATA ATATAGTTGG  
 GGTAGCCTAG CTGAATGTTT TATGTTAATA ATATGTTTGC TTATAATTTT  
 GCAAGTGTA TGAATGCAT CAGCTTCATT AATGCTTTAG AG

\* Exon segments were labeled by the dark blue font type. These parts of the sequence correspond directly to the published cDNA sequence (Hegeman et al., 2001). The areas of maroon are places where the wild-type parent differed from the mutant parent. In these areas the letter above the sequence indicates the base pair that was observed in the mutant parent.

Hegeman et al. (2001), using conserved primers from homologous regions of published MIPS sequences, partially sequenced a genomic fragment of the MIPS gene. This fragment was compared to the cDNA to determine the location of the exons; a total of nine were determined. The two contigs in this study covered seven regions of the nine exons from Hegeman et al. (2001).

There was difficulty in aligning fragments into a single contig because the fragments were not solely fragments of the one MIPS gene. Hegeman et al. (2001) used Southern blot analysis with a hybridization probe to show there are four or more

loci with MIPS homology. Two contigs which showed similarity to the published cDNA sequence were joined by running the Sequencher program (Gene Codes Corporation, Michigan) at a 70% match percentage. The combined contigs were of 1892bp in length and begin at the 621 base pair site of the published soybean cDNA sequence (Figure 11).

The MIPS fragment composed of the two aligned contigs was run with the BLAST program (Altschul et al., 1990), and three sequences matched with the contig: 1) *Phaseolus vulgaris* 1L-*myo*-inositol-1-phosphate synthase gene (GI 14582466) (Johnson et al., unpublished), 2) *Glycine max myo*-inositol-1-phosphate synthase mRNA (GI 14764465) (Hitz et al., 2002), and 3) *Glycine max myo*-inositol-3-phosphate synthase mRNA (GI 13936690) (Hegeman et al., 2001). The three published sequences and the contig were aligned using MultAlin (Figure 11) (Corpet, 1998). The consensus sequence indicated that there was a high degree of similarity between all four sequences and the contig sequence showed homology to published MIPS sequences.





1821 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920 1930 1940 1950  
Phaseolus contig | TGAGATTTCCTCCACCTGGTTGCCATTGCTTGTTGACGGAAATGTTCCATTCATACCGAGCCCTCAGAACACTTTTGCACGGTTGAACTTACCTTTTAGTTTCCCAAC---|  
Hitz | TGAGATTTCCTCCACCTGGTTGCCATTGCTTGTTGACGGAAATGTTCCATTCATACCGAGCCCTCAGAACACTTTTGCACGGTTGAACTTACCTTTTAGTTTCCCAAC---|  
Hegenan | TGAGATTTCCTCCACCTGGTTGCCATTGCTTGTTGACGGAAATGTTCCATTCATACCGAGCCCTCAGAACACTTTTGCACGGTTGAACTTACCTTTTAGTTTCCCAAC---|  
Consensus | TGAGATTTCCTCCACCTGGTTGCCATTGCTTGTTGACGGAAATGTTCCATTCATACCGAGCCCTCAGAACACTTTTGCACGGTTGAACTTACCTTTTAGTTTCCCAAC---|  
1951 1960 1970 1980 1990 2000 2010 2020 2030 2040 2050 2060 2070 2080 2090  
Phaseolus contig | TTAATTTGATTAATTTGATGCAAT---ATTAACTTTC---TTGCCATCAGGGTTG---AARACTGTTTGGATTGGAGG---AGATGACTT---CAAGA  
Hitz | TTAATTTGATTAATTTGATGCAAT---ATTAACTTTC---TTGCCATCAGGGTTG---AARACTGTTTGGATTGGAGG---AGATGACTT---CAAGA  
Hegenan | TTAATTTGATTAATTTGATGCAAT---ATTAACTTTC---TTGCCATCAGGGTTG---AARACTGTTTGGATTGGAGG---AGATGACTT---CAAGA  
Consensus | TTAATTTGATTAATTTGATGCAAT---ATTAACTTTC---TTGCCATCAGGGTTG---AARACTGTTTGGATTGGAGG---AGATGACTT---CAAGA  
.t.aattg..t..att.atgc.a.....attaac.t.tg.....ct.tg.ccaatcgcgagg.aahac...t.t.gatttggct.gg...t.ttaactag.ac.abcta.a.cc  
2081 2090 2100 2110 2120 2130 2140 2150 2160 2170 2180 2190 2200 2210  
Phaseolus contig | G-TGGTC-AGACCAAAATGAAATCC---GTGCTGG---TAGATTCCTTG666666---CTGG-ATCACAGGTATATCGTATTATACACTA-TGTTATATCATGGGTGTTTTCTTTACTAGTAAATCATATCC  
Hitz | G-KGGYCNAAACCAAAATGAAATCC---GTGCTGG---TAGATTCCTTG666666---CTGG-ATCACAGGTATATCGTATTATACACTA-TGTTATATCATGGGTGTTTTCTTTACTAGTAAATCATATCC  
Hegenan | G-TGGTC-AGACCAAAATGAAATCC---GTGCTGG---TAGATTCCTTG666666---CTGG-ATCACAGGTATATCGTATTATACACTA-TGTTATATCATGGGTGTTTTCTTTACTAGTAAATCATATCC  
Consensus | G-TGGTC-AGACCAAAATGAAATCC---GTGCTGG---TAGATTCCTTG666666---CTGG-ATCACAGGTATATCGTATTATACACTA-TGTTATATCATGGGTGTTTTCTTTACTAGTAAATCATATCC  
2211 2220 2230 2240 2250 2260 2270 2280 2290 2300 2310 2320 2330 2340  
Phaseolus contig | TTTCTATATTGGTTAGAAATTTAGTTGCTAAATTTATGTTGCACAAAGTAGAAATTAATTTGATTTTCAAGTGTCTCTGTTACCATTGACAGCCACATCATAGTTAGTTACAGCATC  
Hitz | TTTCTATATTGGTTAGAAATTTAGTTGCTAAATTTATGTTGCACAAAGTAGAAATTAATTTGATTTTCAAGTGTCTCTGTTACCATTGACAGCCACATCATAGTTAGTTAGTTACAGCATC  
Hegenan | TTTCTATATTGGTTAGAAATTTAGTTGCTAAATTTATGTTGCACAAAGTAGAAATTAATTTGATTTTCAAGTGTCTCTGTTACCATTGACAGCCACATCATAGTTAGTTAGTTACAGCATC  
Consensus | TTTCTATATTGGTTAGAAATTTAGTTGCTAAATTTATGTTGCACAAAGTAGAAATTAATTTGATTTTCAAGTGTCTCTGTTACCATTGACAGCCACATCATAGTTAGTTAGTTACAGCATC  
.t.t..ab..tt.gttta.aatttatg.atbgtctaatlta.t.tg.caaaa.g.ag.aa..a.ttgatb.t.tcaaa.tggtct..ggt..a..atg.cagccacacatc.atatgct.ebgtt.ttaactag.ac.abcta.a.cc  
2341 2350 2360 2370 2380 2390 2400 2410 2420 2430 2440 2450 2460 2470  
Phaseolus contig | TGCGAARCAATGATGGTATGAAATCTCTGCTCCACAAACCTTCGGTTCCAGGGAATCTCCAGAGTAGCTTGTGATGATGGTCCAGCGCAATGCCATCTCTATGAGCTGGTGAACATCCCGA  
Hitz | TGCGAARCAATGATGGTATGAAATCTCTGCTCCACAAACCTTCGGTTCCAGGGAATCTCCAGAGTAGCTTGTGATGATGGTCCAGCGCAATGCCATCTCTATGAGCTGGTGAACATCCCGA  
Hegenan | TGCGAARCAATGATGGTATGAAATCTCTGCTCCACAAACCTTCGGTTCCAGGGAATCTCCAGAGTAGCTTGTGATGATGGTCCAGCGCAATGCCATCTCTATGAGCTGGTGAACATCCCGA  
Consensus | TGCGAARCAATGATGGTATGAAATCTCTGCTCCACAAACCTTCGGTTCCAGGGAATCTCCAGAGTAGCTTGTGATGATGGTCCAGCGCAATGCCATCTCTATGAGCTGGTGAACATCCCGA  
2471 2480 2490 2500 2510 2520 2530 2540 2550 2560 2570 2580 2590 2600  
Phaseolus contig | CCATGTTGTTGTTATTAAGGTAAATTTGTTTCCACCTATTTTCTTTATTTCTTCTCTGTCAGGCTTTGATCTGTCGCTTTGTTGTTGCTTTGACAGTATGTCCTTACGTTCCGAGCAGCAGA  
Hitz | CCATGTTGTTGTTATTAAGGTAAATTTGTTTCCACCTATTTTCTTTATTTCTTCTCTGTCAGGCTTTGATCTGTCGCTTTGTTGTTGCTTTGACAGTATGTCCTTACGTTCCGAGCAGCAGA  
Hegenan | CCATGTTGTTGTTATTAAGGTAAATTTGTTTCCACCTATTTTCTTTATTTCTTCTCTGTCAGGCTTTGATCTGTCGCTTTGTTGTTGCTTTGACAGTATGTCCTTACGTTCCGAGCAGCAGA  
Consensus | CCATGTTGTTGTTATTAAGGTAAATTTGTTTCCACCTATTTTCTTTATTTCTTCTCTGTCAGGCTTTGATCTGTCGCTTTGTTGTTGCTTTGACAGTATGTCCTTACGTTCCGAGCAGCAGA  
2601 2610 2620 2630 2640 2650 2660 2670 2680 2690 2700 2710 2720 2730  
Phaseolus contig | GAGCCATGG--ATGAGTACACTTCAGAGATATTCATGGGTTGGAHAAAGCACCATTTG--TGCACAAATACATGCGAGGATTTCTCTCT--TGGCTCCCTATTAATTTAGACTTTGGTCTTCTTCTGATGCT  
Hitz | GAGCCATGG--ATGAGTACACTTCAGAGATATTCATGGGTTGGAHAAAGCACCATTTG--TGCACAAATACATGCGAGGATTTCTCTCT--TGGCTCCCTATTAATTTAGACTTTGGTCTTCTTCTGATGCT  
Hegenan | GAGCCATGG--ATGAGTACACTTCAGAGATATTCATGGGTTGGAHAAAGCACCATTTG--TGCACAAATACATGCGAGGATTTCTCTCT--TGGCTCCCTATTAATTTAGACTTTGGTCTTCTTCTGATGCT  
Consensus | GAGCCATGG--ATGAGTACACTTCAGAGATATTCATGGGTTGGAHAAAGCACCATTTG--TGCACAAATACATGCGAGGATTTCTCTCT--TGGCTCCCTATTAATTTAGACTTTGGTCTTCTTCTGATGCT



Base pairs in which the parental genotypes differed could be differences caused by the EMS mutation (especially the guanine to thymidine mutations) (Figure 10). The differences could also be mutations between the two parents which have happened by random chance and do not affect the gene.

To create low phytic mutants, Wilcox et al. (2000) used ethylmethane sulfonate (EMS). EMS indiscriminately changes the guanine base to the thymidine at random positions in the genome. Two of the nucleotide differences observed between the two parental genotypes are a change from the guanine nucleotide to the thymidine nucleotide (630bp; 1291bp) (Figure 10). The other base differences observed between the two parental genotypes are not the base changes usually observed with an EMS mutation (304bp; 474bp; 545bp; 627bp; 628bp; 641bp; 1193bp) (Figure 10). These differences could still be used as markers if the changes are heritable and are associated with the HIP phenotype (Czarnik et al., 2004). To evaluate if any of these differences can be used as a marker for the HIP phenotype it must be determined if the difference is associated with the HIP phenotype. A genetic test for a SNP must then be developed.

### **LCR Primer Design and Reaction**

To detect SNPs a test must be able to amplify the target sequences, distinguish polymorphic bases, have low background noise, and be easily applied (Barany, 1991). Ligase chain reaction (LCR) fits these requirements and was chosen to detect SNP differences between the parental genotypes.

LCR primers were designed for areas where the parental genotypes differed (Table 3). There were nine SNP differences between parental genotypes; five SNPs



were too close to each other or the sequence around the SNP was not unambiguously determined and therefore it could not be used in ligase primer design. The primer sets were named based on the location in the MIPS fragment in which there was a SNP distinguishing the two parental genotypes.

Four primers were designed and used with the LCR reaction kit (Stratagene, California) based on typical primer constraints (Table 3). The reaction product was run on a 2% agarose gel along with two controls (one positive and one negative) for 3 hours. The expected band was less than 100bp, but no bands were observed in any of the reactions. This procedure was redone for each primer pair increasing the ligation temperature by two degrees each time. At a ligation temperature of 66° C both parental genotypes showed a band, therefore there was no SNP detection between the two genotypes.

**Table 3: LCR primers designed for SNP detection.**

Set Name	Primer A	Primer B	Primer C	Primer D
304bp	CTTAGAG AACCGAC GACA*	CCTGTCTTCATC CGACTCTA	GTCGTCGGTTC TCTAAG	TAGAGTCGGAT GAAGACAGGT
474bp	AACNGGT AGCGCTC CCTTT	TGNAAACCTAAC CACNCCTTC	AAGGGAGCGCT ACCNGTT	GAAGGNGTGG TTAGGTTNCA A
1193bp	CTCCTAA GGGGGAG ATT	GACGAGGATAA TRNANCCT	ATCTCCCCCTT AGGAG	AGGNTNYATTA TCCTCGTCA
1291bp	GATTTATT AAAGTAA CGAC	CTAACATAATCA GTAACATAA	TCGTTACTTTA ATAAATC	TTATGTTACTG ATTATGTTAGG

\*Colored base pair indicates a SNP difference between the two parental

genotypes.

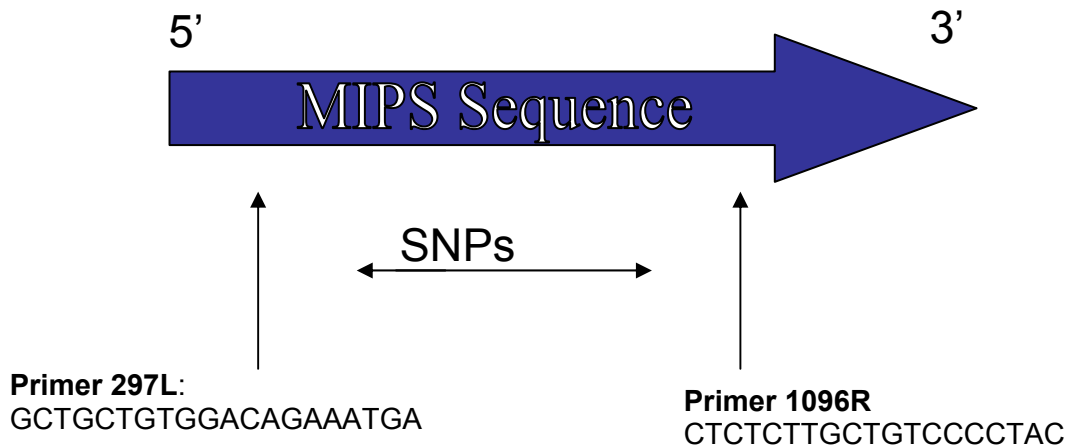
The LCR reactions did not show a polymorphism between the two parental genotypes, which conflicted with sequence data results. Negative results from LCR reaction could be due to the protocol. The LCR protocol used the entire genome as template for the LCR reaction as suggested by Stratagene's protocol. The LCR primers designed could be annealing to multiple regions of the genome due to the multiple regions of MIPS activity (Hegeman et al., 2001).

PCR-coupled LCR has been shown to improve the results of LCR products (Wiedmann et al., 1993). Instead of using the genome as a template, a PCR is performed to produce a target fragment. This target fragment is then used as a template in the LCR reaction. Ward (1998) amplified a small segment (350bp) of the *Lactococcus lactis* 16S rRNA gene sequence before using the fragment in a LCR protocol. Ward (1998) reduced the amount of mismatching that could occur when using LCR primers and was able to get positive results by using a fragment instead of a complete genome.

LCRs using a fragment of the soybean genome as a template were performed. Two primers (297R and 1096L) were designed flanking the entire contig sequence to amplify a PCR fragment that contained all the predicted SNPs. The primers (Figure 13) flanked the area on the contig between 297bp and 1096bp which included all the SNPs between the two parental genotypes. The PCR fragment was used as a template in the LCR reaction. Using this technique, oligonucleotide sets 304bp and 1291bp both produced bands in both genotypes, and oligonucleotide sets 445bp and 1182bp did not produce bands in either genotype. The LCR reaction did not produce results

consistent with the sequence data. If a test was developed to differentiate between the two parental genotypes, the SNP data could be used to test for linkage to the HIP phenotype, but the LCR reaction could not differentiate between the parental genotypes.

**Figure 12: A PCR fragment was designed from the sequenced MIPS fragment in which all SNPs between the two parental genotypes were included between the two primer pairs.**



Hitz et al. (2002) found a SNP caused a change from lysine to asparagine in a single, recessive LPA mutation in soybean. This change in the myo-inositol 1-phosphate synthase gene reduced phytic acid and raffinose stored in the seed. A cDNA clone from the mutant genotype and the wild-type genotype were the same except for a base change from G to T. This change affected an amino acid which was in a conserved region of the *myo*-inositol 1-phosphate synthase gene.

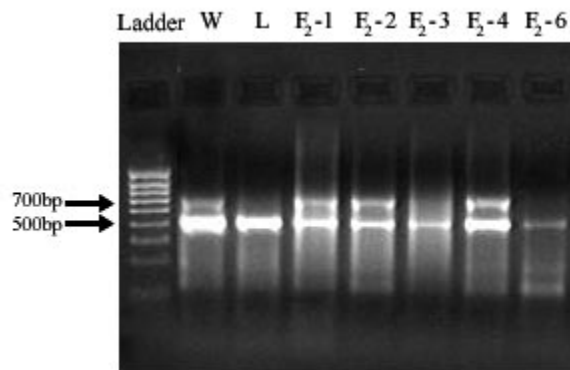
Hitz et al. (2002) designed two sets of primers. The first set had a G at the end of the 3' end, while the other set had the SNP change of T at the 3' end. Both

sets of primers were used with the wild-type genotype and the LPA mutant genotype. The PCR product had a bigger yield of the 600bp fragment if the primer set perfectly matched the template (Hitz et al., 2002). Tests, such as the one used by Hitz et al. (2002), could be used to differentiate the two parental genotypes.

### **Marker Design and F<sub>2</sub> Analysis**

The goal of this research was to design a marker linked to the HIP phenotype. Primer pair 4 showed a polymorphism between the two parental genotypes. A 700bp fragment was observed in the wild-type genotype, but not the LPA genotype. To use the polymorphism as a marker, it first had to be determined if the polymorphism was heritable. One hundred fourteen F<sub>2</sub> genotypes were used in PCR reactions with primer4 under the same conditions as the parental genotypes (Figure 13).

**Figure 13: Agarose gel electrophoresis showing polymorphism observed using primer 4\*.**



\*Ladder = 100bp ladder was used to measure the size of the DNA fragments produced.

W= Wild-type parent genotype

L= LPA parental genotype

F<sub>2</sub> = F<sub>2</sub> genotypes

The polymorphic difference observed was measured as a dominant trait (Table 4). In a dominant trait the heterozygous type and the homozygous wild-type that can not be differentiated; three-fourths of the population show the trait and one-fourth does not show the trait; a 3: 1 ratio. The F<sub>2</sub> genotypes tested at an observed 83:31 ratio which was an excellent fit to the expected 3:1 ( $\chi^2= 0.23$ , P=0.63). The polymorphism was a genetically heritable trait so it could be tested with the HIP phenotype to determine if there was a linkage between the two.

**Table 4: Presence/absence of 700bp fragment from a PCR reaction with primer pair 4\*.**

	700bp Fragment	
	Present	Absent
<b>Observed</b>	83	31
<b>Expected</b>	86.25	28.75

\*  $\chi^2= 0.235$  Significant P < 0.001, 1 df.

The HIP negative phenotype was linked to the presence of the band in the primer 4 PCR and the HIP positive phenotype was linked to the absence of a band in the primer 4 PCR (Table 5). Twelve HIP intermediate genotypes had an absence of a band while 22 possessed a band in the primer 4 PCR. HIP intermediate genotypes did not associate with the presence or absence of the band. Some intermediates possessed the band while others did not. The primer 4 polymorphism might be linked to only one of the genes controlling the LPA mutant which would explain why the

heterozygotes are segregating. If the genes controlling the LPA mutation segregated as a unit then one would expect to observe all the HIP intermediates to possess the 700bp band.

**Table 5: Polymorphism with primer 4, observed and expected HIP phenotype in the F2 population derived from a cross of a HIP mutant and a wild-type\*.**

	<b>700bp</b>	<b>HIP Type<sup>†</sup></b>	
<b>Present</b>	<b>83</b>	<b>HIP -/-</b>	<b>61</b>
		<b>HIP +/-</b>	<b>22</b>
		<b>HIP +/+</b>	<b>0</b>
<b>Absent</b>	<b>31</b>	<b>HIP -/-</b>	<b>3</b>
		<b>HIP +/-</b>	<b>12</b>
		<b>HIP +/+</b>	<b>16</b>

\* Only 16 of the 17 HIP positive genotypes had a PCR that produced results.

<sup>†</sup> HIP -/- = HIP Negative  
 HIP +/- = HIP Intermediate  
 HIP +/+ = HIP Positive

Is the polymorphism involved with the HIP trait? The polymorphic fragment is heritable and is linked to the HIP positive phenotype. None of the genotypes classified as HIP positive possess the primer 4 700bp band. Larson et al. (2000) has previously suggested a link between the MIPS regions of rice, barley, and maize and an *lpa1-1* mutation. This linkage is suggested by the polymorphic data and HIP phenotype data presented in this study. Since the soybean genome has at least four regions of MIPS activity (Hegeman et al., 2000) it is possible that the HIP phenotype and the polymorphism both are linked to the MIPS gene family but to different

regions of activity. The polymorphism could be linked to only one gene, and the heterozygotes are segregating because at least two genes are involved in the trait. If there are multiple genes involved in the LPA mutation and both have to be recessive in order to show a HIP positive phenotype then the genetic polymorphism would have segregation in the HIP intermediates.

## Conclusions

The introduction of a LPA mutation in soybean could be an important breeding goal. Phosphorus waste can be reduced by using LPA mutants (Mendoza, 2002). The LPA mutation could be important trait in a soybean breeding program. Eliminating excess phosphorus will reduce pollution in the water systems near farm lands. LPA grain provides farmers a protein supplement that needs no phosphorus supplement; reducing feed costs. It was determined that oil and protein content were not reduced making this mutation easily incorporated into a breeding program.

HIP phenotype assay is a quick method to indirectly estimate phytic acid content. The assay determines if a seed has high amounts of inorganic phosphorus content which is typical of LPA seeds. Conflicts in results stem from the determination of phenotype class based on inorganic phosphorus concentration. Multiple genes associated with the LPA mutation in soybean could have effected the labeling of the HIP phenotypes.

A MIPS gene fragment was sequenced from both parental genotypes to determine single nucleotide polymorphisms between the two. LCR primers were designed around four of the nine SNPs to design a genetic test for SNPs. The LCR did not work for any of the four primers. This does not mean that there is not a difference between the genotypes only that the LCR could not test for the SNPs between the parental genotypes.

A PCR run with primer 4 was polymorphic between the two parental genotypes. This difference was tested with the F<sub>2</sub> genotype and found to be heritable, and the absence of the band was associated with the HIP positive type. All but three



of the genotypes classified as HIP negative did possess the 700bp band. The HIP intermediates did not show an association to either the presence or absence of the band. One of the genes involved with the low phytic acid mutation is linked to the primer 4 marker.

Further study must be done to investigate the relationship between the MIPS gene and the HIP phenotype. A test must be designed that can be used to evaluate the SNPs between the parental genotypes. If this test can be developed, SNPs can be tested in the F<sub>2</sub> genotypes and a relationship between the SNP and the HIP phenotype could be tested.

## Literature Cited

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215:403-410.

American Soybean Association. 2000. General statistics. American Soybean Association Homepage. Available at <http://www.asa-europe.org>. Verified 28 July 2003.

Barany, F. 1991. Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proceedings of the National Academy of Sciences of the United States of America*. 88(1):189-193.

Biehl, R.R., and Baker, D.H. 1997. Microbial Phytase improves amino acid utilization in young chicks fed diets based on soybean meal but not diets based on peanut meal. *Poultry Science*. 76:355-360.

Brooks, J.R., and Charles, V.M. 1984. Phosphorus and phytate content of soybean protein components. *Journal of Agricultural Food Chemistry*. 32:672-674.

Chen, P.S., Toribara, T.Y., and Warner, H. 1956. Microdeterminations of phosphorus. *Analytical Chemistry*. 28:1756-1758.

Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Resources*. 16:10881-10890.

Czarnik, U., Zabołewicz, T., Galinski, M., Pareek, C.S., and Walawski, K. 2004. Silent point mutation polymorphism of the bovine *CD18* encoding gene. *Journal of Applied Genetics*. 45:73-76.

Dorsch, J.A., Cook, A., Young, K.A., Anderson, J.M., Bauman, A.T., Volkmann, C.J., Murthy, P.P.N., and Raboy, V. 2003. Seed phosphorus and inositol phosphate phenotype of barley low phytic acid genotypes. *Phytochemistry*. 62:691-706.

Environmental Protection Agency. Available at <http://www.epa.gov>. Verified 13 May 2004.

Erdman, J.W. 1981. Bioavailability of trace minerals from cereals and legumes. *Cereal Chemistry*. 58:21-26

Gibson, R.S., and Hitz, C. 2001. Dietary diversification/modification strategies to enhance micronutrient content and bioavailability of diets in developing countries. *British Journal of Nutrition*. 85(2): S159-S166.

- Grotz, N., and Guerinot, M.L. 2002. Limiting nutrients: an old problem with new solutions? *Current Opinion in Plant Biology*. 5:158-163.
- Guttieri, M., Bowen, D., Dorsch, J.A., Raboy, V., and Souza, E. 2004. Identification and characterization of a low phytic acid wheat. *Crop Science*. 44:418-424.
- Hegeman, C.E., Good, L.L., and Grabau, E.A. 2001. Expression of D-myo-inositol-3-phosphate synthase in soybean. Implications for phytic acid biosynthesis. *Plant Physiology*. 125:1941-1948.
- Hitz, W.D., Carlson, T.J., Kerr, P.S., and Sebastian, S.A. 2002. Biochemical and molecular characterization of a mutation that confers a decreased raffinose and phytic acid phenotype on soybean seeds. *Plant Physiology*. 128:650-660.
- Huff, W.E., Moore, P.A., Waldroup, P.W., Waldroup, A.L., Balog, J.M., Huff, G.R., Rath, N.C., Daniel, T.C., and Raboy, V. 1998. Effect of dietary phytase and high available phosphorus corn on broiler chicken performance. *Poultry Science*. 77:1899-1904.
- Hurrell, R.F., Reddy, M.B., Burri, J., and Cook, J.D. 2002. Phytate degradation determines the effect of industrial processing and home cooking on iron absorption from cereal-based foods. *British Journal of Nutrition*. 88:117-123.
- Johnson, M.D., Lackey, K.H., and Pope, P.M. Subcellular localization of 1L-myo-inositol-1-phosphate synthase. Unpublished direct submission 2000.  
[www.ncbi.nlm.nih.gov/blast/](http://www.ncbi.nlm.nih.gov/blast/)
- Karthigesu, V.D., Mendy, M., Fortuin, M., Whittle, H.C., Howard, C.R., and Allison, L.M.C. 1995. The ligase chain reaction distinguishes hepatitis B virus S-gene variants. *FEMS Microbiology Letters*. 131:127-132.
- Lan, G.Q., Abdullah, N., Jalaudin, S., and Ho, Y.W. 2002. Efficacy of supplementation of a phytase-producing bacterial culture on the performance and nutrient use of broiler chickens fed corn-soybean meal diets. *Poultry Science*. 81:1522-1532.
- Landegren, U., Nilsson, M., Kwok, P.Y. 1998. Reading bits of genetic information: methods for single-nucleotide polymorphism analysis. *Genome Research*. 8:769-776.
- Larson, S.R., and Raboy, V. 1999. Linkage mapping of maize and barley myo-inositol 1-phosphate synthase DNA sequences: correspondence with a low phytic acid mutation. *Theoretical Applied Genetics*. 99:27-36.
- Larson, S.R., Rutger, J.N., Young, K.A., and Raboy, V. 2000. Isolation and genetic mapping of a non-lethal rice (*Oryza sativa L.*) low phytic acid 1 mutation. *Crop Science*. 40:1397-1405.

- Lei, X.G., Ku, P.K., Miller, E.R., Yokoyama, M.T., and Ullray, D.E. 1993. Supplementing corn-soybean meal diets with microbial phytase maximizes phytate P utilization of weanling pigs. *Journal of Animal Science*. 71:3368-3375.
- Li, Y.C., Ledoux, D.R., Veum, T.L., Raboy, V., and Ertl, D.S. 2000. Effects of low phytic acid corn on phosphorus utilization, performance, and bone mineralization in broiler checks. *Poultry Science*. 79:1444-1450.
- Li, H., and Burton, J.W. 2002. Selecting increased seed density to increase indirectly soybean seed protein concentration. *Crop Science*. 42:393-398.
- Loewus, F.A., and Murthy, P.P.N. 2000. Review: myo-inositol metabolism in plants. *Plant Science* 150:1-19.
- Lönnerdal, B., Sandberg, A.S., Sandstöm, B., and Kunz, C. 1989. Inhibitory effects of phytic acid and other inositol phosphates on zinc and calcium absorption in suckling rats. *Journal of Nutrition*. 119:211-214.
- Lopez, H.W., Leenhardt, F., Coudray, C., and Remesy, C. 2002. Minerals and phytic acid interactions: is it a real problem for human nutrition? *International Journal of Food Science and Technology*. 37:727-739.
- Meis, S.J. 2003. Seed source effect on field emergence of soybean lines with reduced phytate and raffinose saccharides. *Crop Science*. 43:1336-1339.
- Mendoza, C. 2002. Effect of genetically modified low phytic acid plants on mineral absorption. *International Journal of Food Science and Technology*. 37:759-767.
- McCallum, C.M., Comai, L., Greene, E.A., and Henikoff, S. 2000. Tilling, a new technique for reverse genetics. *Plant Physiology*. 123:439-442.
- Miyazawa, E., Iwabuchi, A., and Yoshida, T. 1996. Phytate breakdown and apparent absorption of phosphorus, calcium, and magnesium in germfree and conventionalized rats. *Nutrition Research*. 16:603-613.
- Nelson, T.S., Shieh, T.R., Wodzinski, R.J., and Ware, J.H. 1971. Effect of supplemental phytase on the utilization of phytate phosphorus by chicks. *Journal of nutrition*. 101:1289-1294.
- Oatway, L., Vasathan, T., and Helm, J.H. 2001. Phytic acid. *Food Reviews International*. 17(4):419-431.
- Oltmans, S.E., Fehr, W.R., Welke, G.A., and Cianzio, S.R. 2004. Inheritance of low-phytate phosphorus in soybean. *Crop Science*. 44:433-435.

- Oomah, B.D., Kenaschuk, E.O., and Mazza, G. 1996. Phytic acid content of flaxseed as influenced by cultivar, growing season, and location. *Journal of Agriculture Food Chemistry*. 44:2663-2666.
- Pilu, R. 2003. Phenotypic, genetic and molecular characterization of a maize low phytic acid mutant (lpa241). *Theoretical and Applied Genetics*. 107:980-987.
- Preusch, P.L., Adler, P.R., Sikora, L.J., and Tworkaski, T.J. 2002. Nitrogen and Phosphorus availability in composted and uncomposted poultry litter. *Journal of Environmental Quality*. 31:2051-2057.
- Raboy, V. 2001. Seeds for a better future: 'low phytate' grains help to overcome malnutrition and reduce pollution. *Trends in Plant Science*. 6: 458-462.
- Raboy, V. 2002. Progress in breeding low phytate crops. *Journal of Nutrition*. 132:503S-505S.
- Raboy, V., and Dickinson D. 1987. The timing and rate of phytic acid accumulation in developing soybean seeds. *Plant Physiology*. 85:841-844.
- Raboy, V., Gerbasi, P.F., Young, K.A., Stoneberg, S.D., Pickett, S.G., Bauman, A.T., Murthy, P.N., Sheridan, W.F., and Ertl, D.S. September 2000. Origin and seed phenotype of Maize low phytic acid 1-1 and low phytic acid 2-1. *Plant Physiology*. 124:355-368.
- Schachtman, D.P., Reid, R.J., and Ayling, S.M. 1998. Phosphorus uptake by plants: From soil to cell. *Plant Physiology*. 116:447-453.
- Sharpley, A.N., Kleinman, P., and McDowell, R. 2001. Innovative management of Agricultural phosphorus to protect soil and water resources. *Communications Soil Science Plant Anal*. 32:1071-1100.
- Shi, J., Wang, H., Wu, Y., Hazenbroek, J., Meeley, R.B., and Ertl, D.S. 2003. The maize low-phytic acid mutant lpa2 is caused by mutation in an inositol phosphate kinase gene. *Plant Physiology*. 131:507-515.
- Viveros, A., Centeno, C., Brenes, A., Canales, R., and Lozano, A. 2000. Phytase and acid phosphatase activities in plant feedstuffs. *Journal of Agriculture Food Chemistry*. 48:4009-4013.
- Waldroup, P.W., Kersey, J.H., Saleh, E.A., Fritts, C.A., Yan, F., Stillborn, H.L., Crum, R.C., and Raboy, V. 2000. Nonphytate phosphorus requirement and phosphorus excretion of broiler chicks fed diets composed of normal or high available phosphate corn with and without microbial phytase. *Poultry Science*. 79:1451-1459.

Ward, L.J.H. 1998. Two methods for the genetic differentiation of *Lactococcus lactis* ssp. *Lactis* and *cremoris* based on differences in the 16S rRNA gene sequence. FEMS microbiology letters. 166:15-20.

Weiss, R. 1991. Hot prospect for new gene amplifier. Science. 254:1292-1293.

Wiedmann, M., Barany, F., and Batt, C.A. 1993. Detection of *Listeria monocytogenes* with a nonisotopic polymerase chain reaction-coupled ligase chain reaction assay. Applied and Environmental Microbiology. 59:2743-2745.

Wilcox, J.R., Premachandra, G.S., Young, K.A., and Raboy, V. 2000. Isolation of high seed inorganic P, low-phytate soybean mutants. Crop Science. 40:1601-1605.