Molecular Marker Assisted Development and Stability Analysis of Low Phytate, High Inorganic Phosphorus Soybean [Glycine max (L.) Merrill]

Suzannah Joy Wiggins
smelling@utk.edu

Recommended Citation
https://trace.tennessee.edu/utk_gradthes/1221
To the Graduate Council:

I am submitting herewith a thesis written by Suzannah Joy Wiggins entitled "Molecular Marker Assisted Development and Stability Analysis of Low Phytate, High Inorganic Phosphorus Soybean [Glycine max (L.) Merrill]." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Plant Sciences.

Vincent R. Pantalone, Major Professor

We have read this thesis and recommend its acceptance:

Dennis R. West, Dean A. Kopsell, Carl E. Sams

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)
Molecular Marker Assisted Development and Stability Analysis of Low Phytate, High Inorganic Phosphorus Soybean [Glycine max (L.) Merrill]

A Thesis presented for the
Master of Science Degree
The University of Tennessee, Knoxville

Suzannah Joy Wiggins
May 2012
DEDICATION

This thesis is dedicated to my husband, Benjamin Wiggins, my best friend and my teammate throughout life and especially graduate school; my parents Marlin and Bekah Mellinger, whose constant reminders of love and to “never sell yourself short” have helped me rise to every challenge; Drs. Bruce Greene and Janice Branson, who encouraged me to take on this challenge; and to the entire Mellinger and Wiggins families, who keep my life interesting and enjoyable all while teaching me something new every day.
ACKNOWLEDGEMENT

I want to thank the faculty, staff, and fellow students at the University of Tennessee College of Agriculture Sciences and Natural Resources for assisting me with my education and research and my friends and family for their encouragement during this process. I cannot begin to fully express my appreciation to my major professor, Dr. Vince Pantalone, for his willingness to take on an additional graduate student and provide the guidance and mentorship necessary to prepare me for earning this degree. I am additionally grateful for my committee members, Dr. Dennis West, Dr. Dean Kopsell, and Dr. Carl Sams for the time they invested in me and their advice to help me reach my project objectives. I am also greatly indebted to the members of the soybean breeding research team both past and present including Dr. Catherine Nyinyi, Deborah Landau-Ellis, Elizabeth Myer, Ben Fallen, Chris Smallwood, Ben Wiggins, Nicole Renworth-Tacey, and Greg Allen for their willingness to help with lab and field work and to share their valuable insights. Additional thanks go to the entire staff at the East Tennessee Research and Education Center, Highland Rim Research and Education Center, and the Research and Education Center at Milan, TN for assistance with field experiments. I am also so very thankful for my friends Will Phillips, Matthew Wiggins, Kelly Barnett and Sarah Atkinson who reminded me that we are all in this together and kept life fun during some of the most stressful times. And finally, my thanks go to the United Soybean Board and the Consortium for Plant Biotechnology Research for their financial support throughout my research.
ABSTRACT

Phytate [myo-inositol 1,2,3,4,5,6-hexakisphosphate] is a mixed cation salt of phytic acid which binds to other minerals, in many field crops including soybean [Glycine max (L.) Merr.]. Livestock with monogastric digestive systems lack the ability to break bonds between phytate and these minerals, causing phytate phosphorus (P) to be excreted in the waste and contributing to possible P and other mineral deficiencies. Discovery of single sequence repeat (SSR) and single nucleotide polymorphism (SNP) markers associated with low phytate QTL (cqPha-001 and cqPha-002) have aided in development of low phytate soybean lines. The objectives of this study were to 1) utilize SSRs and SNPs to identify the heterozygous and recessive allelic forms of low phytate loci during trait introgression into a high yielding commercial cultivar; 2) evaluate differences in genotypic values for agronomic and seed quality traits, and seed inorganic phosphorous (Pi) concentration between the recurrent parent 5601T and BC4, BC3, and BC2 derived progeny lines; and 3) evaluate those traits and environmental stability of inorganic P in the recurrent parent 5601T and the BC4 derived line TN09-239 across ten southern U.S. environments. Successful trait introgression was confirmed in the BC4 line TN09-239 via SNP assay. In comparison to the recurrent parent 5601T (222.9 ng µL⁻¹) the BC4 derived line TN09-239 (1675.9 ng µL⁻¹) contained significantly higher Pi (P<0.0001), which is inversely proportionate to phytate concentration, was taller (87.4 cm for 5601T and 119.7 cm for TN09-239) (P<0.0001) and yielded less (3500 kg ha⁻¹ for 5601T and 3151 kg ha⁻¹ for TN09-239) (P<0.05). Differences in Pi concentration, height, and yield (P<0.0001, P<0.05, and P<0.05 respectively) were also observed across the ten southern U.S. environments. A significant genotype by environment interaction was evident (P<0.01) and in a linear regression, 5601T was found to have a slope of zero (P>0.05), while TN09-239 had a slope significantly different from zero as well as from one (P<0.0001 and P<0.01). These data indicate that introgression of the
low phytate trait with a high yielding line adapted to the southern U.S. is possible, but stability of the trait across environments will be of concern in future research.
### LIST OF TABLES

Table 3.1: Primer, probe, and reagent volumes used for PCR of Gm19 soybean phytate SNP... 35
Table 3.2: Primer, probe, and reagent volumes used for PCR of Gm03 soybean phytate SNP... 35
Table 3.3: Nucleotide sequence for Gm19 and Gm03 phytate SNP forward primers obtained from Integrated DNA Technologies (Coralville, IA). ...................................................... 35
Table 3.4: Nucleotide sequence for Gm19 and Gm03 phytate SNP reverse primers obtained from Integrated DNA Technologies (Coralville, IA). ............................................................... 35
Table 3.5: Nucleotide sequence for Gm19 and Gm03 phytate SNP simple probes obtained from TIB MOLBIOL (Berlin, Germany). ................................................................................. 35
Table 3.6: SNP Chi Square Goodness of Fit Test for BC$_2$F$_2$ Progeny resulting from initial cross CX1834-1-2 x 5601T to create a low phytate soybean line (indicated by plants in class “aabb”). ............................................................................................................................. 40
Table 3.7: Results from SSR Assay on BC$_2$F$_1$ derived lines (recurrent parent 5601T by low phytate mutant line CX1834-1-2) to determine low phytate trait capture. ......................... 41
Table 3.8: SNP Assay results at Gm19 and Gm03 soybean phytate loci for DNA of BC$_2$, BC$_3$, and BC$_4$ derived lines TN07-604, TN07-602, and TN09-239 and their recurrent parent 5601T. ............................................................................................................................... 42
Table 4.1: Least Square Means estimations for agronomic traits for 5601T and the BC$_4$ line (TN09-239) developed for low phytate trait introgression grown in three replications at each at ETREC, HRREC, and RECM LN averaged over two years (2010 and 2011). .... 57
Table 4.2: Least Square Means estimations for seed quality traits for 5601T and its BC$_4$ line (TN09-239) developed for low phytate trait introgression grown in three replications at each at ETREC, HRREC, and RECM LN averaged over two years (2010 and 2011). .... 58
Table 4.3: Least Square Means estimations for fatty acid compositions for 5601T and its BC$_4$ line (TN09-239) developed for low phytate trait introgression grown in three replications at each at ETREC, HRREC, and RECM LN averaged over two years (2010 and 2011). 60
Table 4.4: Least Square Means estimations for amino acid compositions for 5601T and its BC$_4$ line (TN09-239) developed for low phytate trait introgression grown in three replications at each at ETREC, HRREC, and RECM LN averaged over two years (2010 and 2011). 61
Table 5.1: Agronomic trait means for 5601T and TN09-239 averaged over ten environments of the 2010 Uniform Preliminary V Soybean Test for the Southern Region. ......................... 77
Table 5.2: Least square means estimations for fatty acid composition of 5601T and TN09-239 over ten environments in the 2010 USDA Uniform Preliminary V Soybean Test, Southern Region................................................................. 79

Table 5.3: Least Square Means estimations for amino acid compositions for 5601T and its BC₄ line (TN09-239) grown over ten environments in the 2010 USDA Uniform Preliminary V Soybean Test, Southern Region. ....................................................................................... 80

Table 5.4: Summary of phytate stability parameters of 5601T and its BC₄ derived low phytate line TN09-239 grown in Rohwer, AR, Pine Tree, AR, McCune, KS, Pittsburg, KS, Portageville, MO, Stoneville, MS, Kinston, NC, Plymouth, NC, Jackson, TN, and Warsaw, VA in 2010...................................................................................................... 82

Table A1.1: Development of BC₂ line TN07-604 from 5601T and CX1834-1-2. ......................... 94

Table A1.2: Development of BC₃ line TN07-602 from 5601T and CX1834-1-2. ................. 95

Table A1.3: Development of BC₄ line TN09-239 from 5601T and CX1834-1-2 .................. 96

Table A4.1: Least squares means for all traits examined across all Tennessee locations (ETREC, HRREC, and RECMLN) over both 2010 and 2011 growing season. .............................. 97
LIST OF FIGURES

Figure 3.1: Phytate molecule bound to mineral cations, rendering the bonds unbreakable by monogastric digestive systems................................................................. 26

Figure 3.2: Standards plate from colorimetric assay developed by Chen et al., 1956 and adapted by Scaboo et al., 2009. Assay used to assist in phenotypic selection when breeding for low phytate soybean. Results given in ng/uL of inorganic phosphorus (Pi), which is inversely proportionate to phytate. Darker colors representing higher Pi levels (lower phytate) are more desirable than lighter colors representing lower Pi (higher phytate).... 27

Figure 3.3: Melting curve genotyping SNP assay expected melting temperature (approximately 54 to 55 °C) of soybean sample that has homozygous recessive genotype (low phytate) at Gm19................................................................. 36

Figure 3.4: Melting curve genotyping SNP assay expected melting temperature (approximately 60 to 61°C) of soybean sample that has homozygous dominant genotype (normal phytate) at Gm19................................................................. 36

Figure 3.5: Melting curve genotyping SNP assay expected melting temperature (approximately 62 to 63 °C) of soybean sample that has homozygous recessive genotype (low phytate) at Gm03................................................................. 37

Figure 3.6: Melting curve genotyping SNP assay expected melting temperature (approximately 57 to 58 °C) of soybean sample that has homozygous dominant genotype (normal phytate) at Gm03................................................................. 37

Figure 3.7: Melting curve genotyping SNP assay double melting peaks (approximately 54 to 55°C and 60 to 61°C) indicates heterozygous genotype for soybean sample at Gm19... 38

Figure 3.8: Melting curve genotyping SNP assay double melting peaks (approximately 57 to 58°C and 62 to 63 °C) indicate heterozygous genotype for soybean sample at Gm03.... 38

Figure 4.1: Inorganic phosphorus content for recurrent parent 5601T and BC4 line TN09-239 during 2010, 2011, and two year average over three replications each at three locations (ETREC, HHREC, and RECMLN). ................................................................. 62

Figure 5.1: Inorganic phosphorus means for 5601T and TN09-239 grown in 2010 across locations within designated regions in the Southern U.S. Locations within each region are as follows: East: Kinston, NC, Plymouth, NC, and Warsaw, VA; South: Jackson, TN; Delta: Pine Tree, AR, Rohwer, AR, Portageville, MO, Stoneville, MS; West: McCune, KS, Pittsburg, KS................................................................. 83

Figure 5.2: Genotype by environmental index linear regression against inorganic phosphorus for 5601T and its BC4 low phytate line TN09-239. Slope for TN09-239 (2.033) is significantly different than zero (P<0.0001) and did not exhibit stability for the low phytate trait when grown across ten locations in the southern U.S. in 2010. .................. 83
PART 1

Introduction
Soybean \([Glycine \text{ Max} \ (L.)]\) is a crop that traces its origins back to ancient China. While it is difficult to pinpoint when soybean domestication first began, Hymowitz and Newell (1981) indicate the first written record of soybean appeared in the \textit{Book of Odes}, which was written during the 11\textsuperscript{th} to 7\textsuperscript{th} centuries B.C. Soybean was so important to the survival of Chinese civilization that it was considered one of their five sacred crops (McMahon et al., 2002). While the common use of soybean grew throughout China, Japan, Indonesia and other East Asian countries during the early centuries A.D., it was not brought to Europe until the 18\textsuperscript{th} century and to the Americas until the 19\textsuperscript{th} century (Hymowitz and Newell, 1981).

In its center of origin (Asia), soybean is used as high protein food in most diets, in addition to use as livestock feed, which is the main use in the U.S. In Asia, tofu and tempeh are used as main dishes (rather than meat), soymilk is used as commonly as cow’s milk is used in the West, and miso is bouillon for soup (Hymowitz and Newell, 1981). However, in the Western Hemisphere, soybean is most often used as an oil crop and as a high-protein livestock feed component. It is also used as an ingredient in other foods and manufactured products such as paints, inks, dyes, biodiesel and lubricants (Aquaah, 2005).

A member of the Leguminosae family, soybean was first used in the U.S. as a forage crop and then as a way of replenishing soil nitrogen (N) for corn \((Zea \text{ mays})\) production. The use of soybean as an oilseed began in Europe in the early 20\textsuperscript{th} century. The onset of World War II caused the U.S. to become more self-reliant. Previously, the U.S. imported as much as 40\% of its edible fats and oils. The war spurred more farmers to grow soybean for both oil production and livestock feedstuffs. Through the rest of the 20\textsuperscript{th} century, more uses for soybean were discovered
and research of the crop increased, causing more acreage of farmland to be used for growing this crop (Carpenter et al., 2002).

The popularity and widespread cultivation of soybean in the Western Hemisphere can be attributed to its adaptability. The ability of soybean to grow in soils that are also suitable for growing corn, its level of drought tolerance, and its variety of non-food uses are all favorable qualities of this crop (Aquaah, 2005). Likewise, further development of productive and disease-resistant cultivars by sources such as private companies, U.S. Department of Agriculture (USDA), and land grant universities has increased the demand and popularity of soybean (McMahon et al., 2002). Another favorable attribute of soybean is its ability to fix N via *Bradyrhizobium japonica* bacteria. In the typical agronomic setting, soybean is self-pollinated, with a complete flower (Aquaah, 2005). In crop improvement settings, it is necessary and often most economical to manually cross-pollinate soybean in order to create the initial hybrids for population development of selfed pure-line progenies to achieve genetic gain. However, other methods such as using and gene insertion through agrobacterium transformation used to attain desired characteristics in the crop.

In U.S. legume crop oil production, soybean produces the second highest amount of oil per seed at 20%, while peanut (*Arachis hypogaea*) produces the most oil per seed at 48% but peanut is produced on far less land than soybean. This can explain why 29% of the U.S. vegetable oil market is provided by soybean. Soybean has the highest protein content (40%) among all legumes and cereals. Other notable attributes of soybean oil include its polyunsaturated fatty
acids, lack of cholesterol, biodiesel properties, and usefulness in food products as both oil and as the emulsifier, lecithin (Carpenter et al., 2002).

Soybean holds eight of the essential amino acids that are not produced naturally by the human body, but are needed for nutrition. Soybean meal is a common ingredient of diets of swine (*Sus domesticus*), poultry (*Gallus gallus domesticus*), pets and fish (Hymowitz and Newell, 1981). Additionally, the crop can be used in paints, textiles, pharmaceuticals, and as an anti-foaming agent. U.S. soybean production is centered in the Midwestern states, with Iowa, Illinois, Minnesota, and Nebraska being the top producing states. Acquaah (2005) states that 75% of all U.S. soybean acres are grown in this North Central or Corn Belt area. Other top regions for soybean production are the Mississippi River valley (which accounts for up to 10%) and the eastern seaboard.

According to the American Soybean Association’s Soy Stats 2011, soybean makes up 58% of the world’s oil seed production. The U.S. produces 35% of the world’s soybean crop, making it the top producing country in the world. The increasing popularity of no-till growing systems has also helped to expand soybean production in the U.S. and worldwide (Carpenter et al., 2002). Of the total animal consumption of U.S. soybean, 49% is by poultry, and 25% is by swine (“Soy Stats 2011”). Its acceptable cost and high protein composition of soybean have made soymeal a popular livestock feed component. However, soybean contains a phosphate, called phytate, which cannot be completely broken down by swine and poultry due to their monogastric digestive systems (Powers et al., 2006; Gillman et al., 2009; Raboy et al., 2000).
Since phytate cannot be fully broken down by non-ruminant livestock, it is passed in their waste. Runoff from livestock manure from concentrated animal production areas or from fields fertilized with manure is already attributed to agricultural non-point sources (NPS) of pollution in waterways, and phytate in the manure contributes to the problem (Gillman et al., 2009; Maupin et al., 2011; Sharpley et al., 1994). The use of livestock manure for fertilizer is popular among many farmers; however, applications are often made solely based on N content. This can cause excess application of phosphorus (P) from the non-digested phytate, which increases the likelihood of extra P in agricultural runoff (Sharpley et al., 1994). The extra P quickens the onset of eutrophication in fresh water bodies and estuaries. Eutrophication results in the growth of noxious aquatic weeds and lower dissolved oxygen content in the water (Daniel et al., 1994; Sharpley et al., 1994; Sharpley et al., 2000). While P is not the only cause of eutrophication, P inputs to fresh water and estuaries are often more easily managed than N and Ca (Sharpley et al., 1994).

Phytate chelates or binds to other minerals such as Ca$^{2+}$ and Zn$^{2+}$. When this bond occurs, neither the phytate nor the mineral it is bound to can be broken down during digestion in a non-ruminant animal (Gillman et al., 2009; Oltmans et al., 2005; Raboy et al., 2000). More specifically, phytate is a mixed cation salt of phytic acid (myo-inositol-1,2,3,4,5,6-hexakisphosphate) (Maupin et al., 2011; Gillman et al., 2009). This phytate is found in many grain and legume seeds, often in the grains’ endosperms and cotyledons of legumes (Reddy et al., 1989). Phytate and inorganic P in soybean have an inverse relationship. This relationship allows us to measure phytate by performing an assay for inorganic P levels (Wilcox et al., 2000; Chen et al., 1956; Raboy, 2002; Raboy et al., 1984). Research has begun to breed soybean lines with lower phytate
content in order to improve the metabolic growth in poultry and swine and to reduce P loads to surface water.

References


www.unitedsoybean.org.


PART 2

Literature Review
Controlling excess nutrient inputs to freshwater bodies is a chief concern to environmentalists in government and the general public. Inorganic phosphorus, specifically, is the top nutrient currently being targeted because it is more easily managed than N inputs (Daniel et al., 1994; Sharpley et al., 2000). The EPA found that agricultural NPS have a major effect on the inability to meet the Clean Water Act goals by contaminating water quality in freshwater bodies. The EPA also found NPS to be the main cause of eutrophication in freshwater lakes and streams. Livestock manures contain on average, 4.3% N, 1.4% P, and 2.2% K (Daniel et al., 1994). Often, animal manure is spread as fertilizer with N being the only nutrient of interest monitored. This means that the amount of manure required to provide adequate N is spread without concern for the P content or its runoff potential. Surface water eutrophication can be attributed to P from runoff and leaching from soils that have had animal manure applied – specifically poultry or swine manure (Walker et al., 2006; Sharpley et al., 2000). Animal manure can also be of concern in areas of intense non-ruminant livestock production, such as Iowa or North Carolina where hogs are raised. It is easiest to control excess P transport at its source (Daniel et al., 1994).

Not all P found in soybean is detrimental or otherwise unusable P. In soybean, between 60 and 85% of total P content is phytate, while the remaining 15 to 40% is made up of inorganic P and other forms of the mineral (Gillman et al., 2009; Hitz et al., 2002; Raboy, 2002). Such a high presence of phytate means that it can make up between 1% and 3% of the total seed (Gillman et al., 2009; Raboy, 2002). Moreover, as much as 70% of P in the diet of swine is unavailable to the animal and therefore excreted (Powers et al., 2006). The phytate found in animal waste can cause water pollution (Raboy, 2002), which is a significant problem in the U.S., England, and other
countries. Excess P in livestock excrement has caused concern that it may be one of the main origins of surface water quality deterioration (Raboy, 2002).

The reason phytate cannot be broken down has to do with chemical reactions taking place in the intestine of non-ruminant animals. Phytate is often present as a phytin salt, or phytic acid which has a negative charge that remains strong at various pH levels and is often bound to K\(^+\) or Mg\(^{2+}\) ions (Cichy and Raboy 2009). The chemical reactions bind the phytin salt with other nutrient cations such as Ca\(^{2+}\), Zn\(^{2+}\), and Fe\(^{2+}\), which can cause deficiencies of those mineral nutrients in swine and poultry, and the phytin salt compound is excreted in the manure (Raboy, 2002; Walker et al., 2006). According to Gillman et al. (2009), soybean is not the only plant with phytate. In fact many plants have P in the form of phytate. In addition to existing as a phytin salt, phytate can also be called phytic acid, which is produced in soybean, corn, wheat (*Triticum aestivum*), and other crops as a way of storing P in the seed (Hitz et al., 2002). Once the phytic acid and mineral are bound together, non-ruminant animals are unable to digest either (Oltmans et al., 2005). In many cases, the bonded phytate and mineral cause the respective deficiencies despite the seed containing these minerals. While adding inorganic P and minerals to feed rations is a common way for livestock producers to combat deficiencies, it does not solve the issue of phytate presence in manure and it adds an extra expense to the operation (Sharpley et al., 2000).

There have only been three methods identified for changing phytate availability in livestock. Currently the most common method for changing phytate levels is adding the enzyme phytase to livestock feeds. Another option being explored by Bilyeu et al. (2008) is to genetically engineer plants to express phytase in the seed to reduce phytic acid and still maintain the same total P
level. The third method being employed to change P availability is to breed plants to have lower phytate levels (Scaboo et al., 2009; Walker et al., 2006). According to Reddy et al. (1989), the enzyme phytase hydrolyzes myoinositol hexakisphosphate into inorganic orthophosphoate and other forms of phosphoric esters. When phytase is added to livestock feed, it breaks many of the bonds between the phytate and other minerals, allowing more P to be absorbed and digested by the livestock (Gillman et al., 2009). However, using phytase adds another expense to feeding livestock (Maupin et al., 2011; Scaboo et al., 2009). If phytate levels were reduced in soybean, livestock producers may lower feeding cost (Oltmans et al., 2004).

While the use of phytase is considered effective, it does not mitigate all phytate in livestock diets (Powers et al., 2006). It should be noted that while phytase reduces the amount of phytate excreted in livestock manure, it does not increase total seed P; it simply changes the form of P, and thus the bioavailability of inorganic P (Pi) (Powers et al., 2006). The total amount of P in the seed cannot be changed, just transformed from phytate to Pi or other forms. This combined with various studies prove that phytate and Pi have an inverse relationship (Powers et al., 2006; Gillman et al., 2009; Chen et al., 1956). The inverse relationship between phytate and Pi allows for the quantification of phytate levels by simply measuring for Pi (Gillman et al., 2009).

A study by Powers et al. (2006) examined the use of phytase and low phytate soybean meal in swine diets and its effect of phytate excretion. Total digestibility of P in swine from low phytate soybean meal diets was higher than normal phytate soybean meal diets. Likewise, swine fed a low phytate soybean meal diet that included phytase had even greater total P digestibility. Adding the phytase to livestock diets not only increases the bioavailability of P, but also the
metal ions complexed with phytate. When fed low phytate diets, waste P levels in livestock manure proportionally decreased as well. Likewise, in human studies, consumption of tortillas made with low phytate corn increased Fe absorption (Wilcox et al., 2000).

This leads toward the goal of lowering phytate in livestock feedstuff through breeding for low phytate soybean lines. Currently, there are several genetic sources for the low-phytate trait in soybean. They are the \textit{Gm-lpha-TW-1} allele, \textit{Gm-lpa-ZC-2} allele, the MIPS allele, and the combination of the pha1 and pha2 alleles (Cichy and Raboy 2009, Bilyeu et al. 2008, Anderson and Fehr, 2008). MIPS stands for Myoinositol-1-phosphate synthase, which causes Glc-6P to be changed into myoinositol-1-phosphate which then becomes phytic acid through phosphorylation (Bilyeu et al. 2008). Hitz et al. (2002) found that there was a missense mutation on GmMIPS1, otherwise known as the MIPS structural gene. Lines developed with this mutation were found to contain 50\% less phytic acid in the seed (Bilyeu et al. 2008, Hitz et al. 2002). Likewise, lines containing the \textit{Gm-lpha-TW-1} allele or the \textit{Gm-lpa-ZC-2} allele exhibited up to a 66\% and 50\% reduction in phytic acid, respectively (Cichy and Raboy 2009, Yuan et al. 2007).

This publication will focus on the pha1 and pha2 alleles. In order to find a low phytate line, Wilcox et al. (2000) used mutagenesis to create the M153 and M766 lines from the CX1515-4 line. Since the M153 mutant had the highest Pi (and thereby the lowest phytate), it was chosen to produce various low phytate lines including CX1834; that line was also developed by J.R. Wilcox of the USDA-ARS at Purdue University (Walker et al., 2006; Gillman et al., 2009). Crop mutants containing the low phytate gene were discovered in corn, barley (\textit{Hordeum vulgare}), and rice (\textit{Oryza sativa}) before they were developed in soybean (Raboy, 2002).
Originally Wilcox et al. (2000) believed that low phytate was controlled at just one locus. However, several other studies have implied that there are actually two recessive alleles at two different loci that control low phytic acid (Gillman et al., 2009; Oltmans et al., 2005; Raboy, 2002; Anderson and Fehr, 2008; Walker et al., 2006). They are often referred to as either \textit{lpa1} and \textit{lpa2} (typically by molecular geneticists) or \textit{pha1} and \textit{pha2} (typically by plant breeders). Both loci must be double homozygous recessive (\textit{pha1 pha1} and \textit{pha2 pha2}) to be considered low phytate. The \textit{lpa} trait inhibits the seed’s ability to turn P into phytic acid. Myo-inositol phosphates can also increase when phytate is reduced; however, these phosphates will contain only five or less esters per molecule rather than six P esters per molecule of phytic acid (Reddy et al., 1989). A study by Wilcox et al. (2000) found that many HPLC tests showed that Pi content increased inversely as phytate decreased in low phytate soybean lines while total P levels stayed the same.

According to Walker et al. (2006), there are two linkage groups that contain loci which control phytate levels in soybean seed. The locus near Satt237 on chromosome 3 [linkage group (LG) N] is responsible for up to 41% of phytate expression variation in the seed, which again, is proportionately inverse to phytate content. The other linkage group, chromosome 19 (LG L) contains a locus near Satt 527 which has accounted for 11% of phytate expression variation in soybean. Furthermore, the interaction between these two linkage groups when the \textit{lpa} mutations are present in both, contributes to another 8 to 11% of expression variation of phytate. The \textit{lpa} mutation within LG N is referred to as \textit{pha1}, and the mutation within LG L is referred to as \textit{pha2} (Walker et al., 2006; Gillman et al., 2009). The confirmed QTL at those loci are referred to as cqPha-001 and cqPha-002 (Scaboo et al., 2009). According to Oltmans et al. (2004), in order for
a specific soybean plant to be considered low phytate, the alleles at both chromosome 19 (LG L) and chromosome 3 (LG N) must be homozygous recessive for the plant to express the low-phytate trait. This information about these quantitative trait loci was discovered using simple sequence repeat markers (SSRs) (Gillman et al., 2009; Scaboo et al., 2009; Walker et al., 2006).

The use of SSRs is popular and those types of markers were widely used for identifying quantitative trait loci (QTL) in crops. However, the use of single nucleotide polymorphisms (SNPs) is growing in popularity due to its common occurrence in plants as well as the availability of many high-throughput analysis methods. A single base difference, found in homologous pieces of DNA, is what specifically qualifies as a SNP. However, the term “SNP” can also include single base pair insertions and deletions or INDELS within those DNA fragments (Hyten et al., 2010, Choi et al. 2007). The discovery of many new SNP markers has allowed researchers to update and expand the Universal Soy Linkage Panel (USLP 1.0) or a map of SNP QTL in soybean. SNP markers are also a preferred method of QTL detection due to their lower cost of analysis per data point and the rapidity of data collection (Hyten et al., 2010).

Some low phytate lines have been developed and tested not only for low phytate traits, but also to compare agronomic traits to standard check cultivars. Gillman et al. (2009) indicated that it is possible to decrease phytic acid in soybean without extreme cost to other desirable traits. Raboy (2002) concurred, sharing that early studies have shown that these low phytate plants do not have germination or yield problems. However, Oltmans et al. (2005) as well as Anderson and Fehr (2008) cited reduced seedling emergence in low phytate lines – some lines contained the \textit{pha1} and \textit{pha2} alleles and other lines carrying the MIPS gene. In some cases, seedling emergence in
low phytate lines was only 45% while emergence in non-low phytate lines was 68% (Oltmans et al., 2005).

There are differing thoughts on what could be causing the reduced seedling emergence: the possibility of phytate being required for seed germination and seed stock source location are the two most often attributed to the germination problem. Although phytate is present in the seed, Hitz et al. (2002) believed that it was not required for seed germination or viability, also believing that it is possible to reduce phytate in soybean without lowering soybean stand. A study by Spear and Fehr (2007) found that 50% of the lines derived from backcrosses between a normal cultivar (B019) and the low-phytate line CX1834-1-6 had field emergence that was not significantly different from the normal recurrent parent, and that 94% of the lines were not significantly different from the normal phytate parent in yield. These results suggest a potential for the development of low-phytate lines that do not have emergence problems. When that study was continued by Anderson and Fehr (2008), they discovered that the low phytate seed produced in Puerto Rico during the previous study again had a reduced emergence rate when compared to the normal phytate parent. This change in emergence was attributed to the seed decay from infection of seed storage fungi. Those fungi were present in the CX1834 seed produced in Puerto Rico in both studies, while seed from the normal cultivar and a check cultivar (IA-2005) that were also grown in Puerto Rico did not become infected with the seed storage fungi (Spear and Fehr, 2007; Anderson and Fehr 2008). Anderson and Fehr (2008) concluded that it is possible to produce low phytate lines with emergence that does not differ from normal phytate lines provided that the seed comes from seed sources that are free of these fungi.
Regardless of the cause of reduced seedling emergence in low phytate lines, it has been generally agreed that other agronomic traits within soybean are not diminished within a low phytate line (Maupin et al., 2011; Scaboo et al., 2009; Spear and Fehr, 2007). Even in a study by Oltmans et al. (2005) that showed lower yields, it was not determined whether low phytate lines would remain low yielding if seedling emergence was normal.

The purpose of this study is to:

i) Identify the heterozygous and recessive allelic forms of the low phytate loci during trait introgression into a high yielding commercial soybean cultivar.

ii) Test the null hypotheses that there are no significant differences between the recurrent parent and BC$_2$, BC$_3$, and BC$_4$ low phytate lines in genotypic values for seed yield, other agronomic traits, and seed quality traits when grown in Tennessee environments.

iii) Test the null hypotheses that there are no significant differences between the recurrent parent and the BC$_4$ low phytate progeny line in genotypic values for seed yield, other agronomic traits, seed quality traits, and soybean cyst nematode (SCN) resistance when grown across a broad geographic region of the southern U.S.
iv) Test the null hypothesis that there is no difference in the environmental stability of Pi concentration between the recurrent parent and the BC₄ derived progeny line across a broad geographic region of the southern U.S.

References


PART 3

Identifying Low Phytate Alleles via Single Sequence Repeats and Single Nucleotide Polymorphisms for Trait Introgression during Backcrossing in Soybean
Abstract:

Developing soybean [Glycine max (L.) Merr.] with reduced phytate adds value to soybean meal for livestock producers and assists with preventing fresh-water eutrophication caused by agricultural run-off. Single sequence repeat (SSR) markers known to be linked to quantitative trait loci (QTL) for the low phytate trait (Satt561 on Gm19 and Satt237 on Gm03) and a phenotypic assay were used to make backcross selections to create a low phytate, high yielding line adapted to the southern U.S. The backcross lines originated from a cross with the recurrent parent, 5601T and low phytate line CX1834-1-2. SSR and single nucleotide polymorphism (SNP) markers for confirmed phytate QTL cqPha-001 and cqPha-002 were used to verify successful introgression of the low phytate trait. Low phytate trait introgression was successful for the BC₄ line. The use of SNP markers for making future cross selections and confirming trait introgression was shown to be a more successful and cost-effective method than the use of SSRs for this trait. A chi-square goodness of fit test showed that there was a significant difference between the observed and expected genotypic ratios in a population of 252 F₂ plants. The F₂ generation did not exhibit the expected genotypic frequency (χ² = 25.79, P<0.05) in some of the progeny classes, but the sample size may not have been large enough to fully exhibit the proper inheritance ratio for all nine genotypic classes.
Introduction

Until the latter half of the 20th century, crop improvement was limited to classical breeding methods; however, marker-assisted selection (MAS) has become a useful tool to aid in accurate and concise development of lines carrying desired traits. Molecular markers previously used to identify quantitative trait loci (QTLs) in soybean included restricted fragment length polymorphisms (RFLPs) and simple sequence repeats (SSRs). Single-nucleotide polymorphisms (SNPs) are more commonly utilized today (Choi et al., 2007). MAS can lead to better and faster gains in plant breeding when compared to screening phenotypes for some traits. It can take place in early generations, can be done early in plant development, and in turn, allow for two or more selection events each year since post-harvest analysis may not be necessary – this can allow for selections to be made before harvest instead of waiting until harvested samples reach the lab and can be assayed. MAS can sometimes be less expensive and reduce required amounts of labor and time compared to phenotypic selection (Ha et al., 2007).

SSR marker polymorphisms and heritability in plants were first discovered in soybean (Hyten et al., 2010a). Today there are at least 1051 SSR markers mapped in the soybean genome (Hyten et al., 2010a). In a study done by Choi et al. (2007), SSR markers were found to be located very close to genes (between 0 and 0.5 cM). Successes in the use of MAS require markers that are known to be linked to desired traits and the availability of economical and efficient marker assay screening (Ha et al., 2007). Despite such a large number of SSR markers, there were 112 5-10cM gaps in the SSR map of the soybean genome (Choi et al., 2007).
SNPs are quickly becoming the MAS method of choice due to their reduced cost per data point and variety of high throughput assays (Ha et al., 2007; Hyten et al., 2010b). SNPs are more common than SSRs; and can improve the likelihood of successfully identifying QTL and other genetic relationships in organisms (Choi et al., 2007; Hyten et al., 2010a). SNPs are single base differences between homologous DNA fragments. This also includes insertions and deletions in the DNA fragments, referred to as INDELs. SNPs are the most commonly occurring molecular marker in soybean but compared to other crops, are actually low in occurrence. This was attributed to a domestication bottleneck that lowered variation by 50% as well as low variation of sequences in the ancestor G. soja (Choi et al., 2007; Hyten et al., 2010a, Hyten et al., 2010b). In a study by Choi et al. (2007), 291 out of the 1141 SNPs mapped in their study were mapped in 72 of the 112 5-10cM gaps in the SSR map of the soybean genome. Hyten concluded that SNPs are most likely the ideal marker for mapping the soybean genome because they are the most commonly occurring marker available, with a frequency of 1 SNP per 1000 bp of continuous sequence. Currently there are 7,108 mapped SNPs in the soybean genome (Hyten et al., 2010b).

Oltmans et al. (2004) and Walker et al. (2006) both found that the low phytate characteristic in soybean is controlled by recessive alleles at two independent loci. Those loci are referred to as pha1 [located on Gm03 (LG N)] and pha2 [located on Gm19 (LG L)], and the desired genotype for low phytate is $pha1 \; pha1 \; pha2 \; pha2$. Soybean lines with the low phytate trait are desirable for livestock producers because phytate cannot be completely broken down by swine ($Sus \; domesticus$) and poultry ($Gallus \; gallus \; domesticus$) due to their monogastric digestive systems (Powers et al., 2006; Gillman et al., 2009; Raboy et al., 2000). Phytate is often present as a phytin salt, or a phytate bound to $K^+$ or $Mg^{2+}$ ions. Chemical reactions bind the phytin salt with
other nutrient cations (Figure 3.1) such as Ca$^{2+}$, Zn$^{2+}$, and Fe$^{2+}$, which can cause deficiencies of those nutrients in swine and poultry, and the phytin salt compound is excreted in the manure (Raboy, 2002; Walker et al., 2006). Agricultural run-off containing excess P can contribute to the eutrophication of fresh-water bodies (Gillman et al. 2009). When Wilcox et al. (2000) developed low phytate soybean mutants to develop low phytate lines it was thought that the low phytate phenotype in soybean was controlled by an allele mutation at just one locus. This is the case in maize [*Zea mays (L.)*], barley (*Hordeum vulgare*), and rice (*Oryza sativa*), where the pha mutations (known as *lpa* mutations in these crops) were found to be controlled by a single gene. Wilcox et al. (2000) were testing for the low phytate phenotype via inorganic phosphorus assay (Figure 3.2) developed by Chen et al. (1956), where the darkest blue samples indicated the low phytate trait, and all lighter blue samples were considered to not have the trait. There was no indicator of a second locus that could be affecting the expression of the low phytate trait. When Walker et al. (2006) attempted to map the single pha locus using F$_2$ populations, but the observed percentage of progeny with the low phytate trait was much lower than the expected 25%, and there was a much higher percentage of individuals observed with intermediate levels of phytate. At this point Walker et al. (2006) suspected the requirement of another allele for complete expression of the low phytate phenotype. Via a bulked segregant analysis (BSA) approach and a genome-scan methodology, it was found that loci on Gm03 (LG N) and Gm19 (LG L) both contributed to phytate levels in soybean. The SSR markers associated with these linkage groups were Satt237 and Satt561, respectively (Walker et al., 2006). When the frequency of low-phytate progeny were still found to be lower than expected in the original population, it was determined that the low phytate allele must be homozygous recessive at both loci to achieve full expression of the low phytate trait (Walker et al. 2006). The research by Walker et al. (2006) complimented
Figure 3.1: Phytate molecule bound to mineral cations, rendering the bonds unbreakable by monogastric digestive systems.

previous research by Oltmans et al. (2004) where duplicate dominant epistasis at both loci was observed, so it was concluded that the alleles at both must be homozygous recessive for the low phytate trait. Scaboo et al. (2009) further affirmed that the low phytate trait is controlled by two independent QTL by selecting progeny containing both SSR markers Satt237 and Satt561 while developing low phytate lines and observing higher concentrations of inorganic phosphorus than in lines containing just one SSR marker or the other. However, both Satt237 and Satt561 markers collectively only explained approximately 63% of the variation in phytate in the seed (Walker et al., 2006).
Figure 3.2: Standards plate from colorimetric assay developed by Chen et al., 1956 and adapted by Scaboo et al., 2009. Assay used to assist in phenotypic selection when breeding for low phytate soybean. Results given in ng/uL of inorganic phosphorus (Pi), which is inversely proportionate to phytate. Darker colors representing higher Pi levels (lower phytate) are more desirable than lighter colors representing lower Pi (higher phytate).

Research by Scaboo et al. (2009) confirmed that the QTL on Gm03 (LG N) was linked to Satt237 and the one on Gm19 (LG L) was linked to Satt561, concomitant with the concentration of seed inorganic phosphorus. These confirmed QTL were designated cqPha-001 and cqPha-002, respectively. It was found that the *pha1* allele, located on Gm03 (LG N) accounted for 41% of low phytate expression variation compared to the 11% for *pha2* (Gillman et al. 2009).
While the use of SSRs has been useful in plant breeding and mapping of the soybean genome, it does have some limitations. When Scaboo et al. (2009) made selections for the low phytate trait using only SSR markers; the desired trait was only expressed in 50% of the recombinant inbred lines that were homozygous recessive at both Satt237 and Satt561. Although not all lines containing both SSR markers in the recessive allelic form exhibited the low phytate trait, all lines containing the low phytate trait in the study contained both desired SSR markers. This suggests genetic recombination occurred between the SSR marker locus and the QTL locus in a portion of the RIL. A budget analysis of phenotypic selections per 96 individuals cost $6.33 in that study and necessitated about 23 h of labor, while the price of genotyping those same individuals via SSR markers cost about $250 and required 48 h of labor (Scaboo et al. 2009). The cost of SSR marker assays has been attributed to the assay being gel based (Ha et al., 2007). The 112 gaps of 5 cM or more on the linkage map has also been viewed as a limitation when using only SSRs in MAS (Choi et al., 2007).

The discovery and use of SNPs have made it possible to map new QTL, some that correspond with SSRs and some that are located within 72 of the 112 gaps on the SSR linkage map (Choi et al., 2007). SNPs are becoming the marker of choice due to their common distribution throughout the genome, high frequency of occurrence, and the existence of assays that are both high-throughput and have automated genotyping abilities. This does not make SSRs completely obsolete. It has been suggested that SSRs serve as a good source of information to be used for SNP discovery because SSRs are already on the linkage map and may sometimes be close to the QTL for the trait being investigated, then more densely mapped SNPs in that regions could more precisely locate the QTL (Ha et al., 2007; Zhu et al., 2003).
After failing to show soybean low phytate genotypes were due to the MIPS gene mutation as they expected (Chappell et al., 2006), a study by Gillman et al. (2009) was conducted to test whether homologs of the known maize \textit{lpa} mutation exist in soybean. Homologs were found and their DNA sequence was searched to discover SNPs were present in the coding region. This suggested that if a primer sequence could successfully be designed, then the SNP assay would detect the mutant allele directly as part of the gene, avoiding recombination and limited success by previous SSR assays. Following that, Gillman et al. (2009) were first to use a SNP assay to assist in MAS for low phytate soybean lines and they demonstrated 100% effectiveness of SNP selections. The objective of this study was to utilize both SSRs and SNPs to identify the heterozygous and recessive allelic forms of low phytate cqPha-001 and cqPha-002 loci during trait introgression into a high yielding commercial cultivar.

\textbf{Materials and Methods:}

\textit{Development of the Backcross Lines}

The pedigrees of the backcross 2 (BC$_2$) line, TN07-604, backcross 3 (BC$_3$) line, TN07-602, and backcross 4 (BC$_4$) line, TN09-239 can be traced back to the original cross pollination reported by Scaboo et al. (2009) in 2000 between the 5601T cultivar developed by Pantalone et al. (2003) and the low phytate line Cx1834-1-2 developed by Wilcox (2000). This cross produced the line 56Cx-284. In the summer of 2004, 5601T was crossed with 56Cx-284 which resulted in the approximate genetic equivalent of BC$_1$F$_1$ seed. SSR markers for this material were evaluated by the Boerma lab at the University of Georgia. It was found that the BC$_1$F$_1$ material was heterozygous for the low phytate allele at three loci: Satt237 (LG N), Satt561 (LG L) and the closely linked Satt527 (LG L). The close proximity of Satt527 to Satt561 makes its use
redundant, and was not included in our study. The BC$_{1}$F$_{1}$ seed was planted in December 2004 at the USDA Tropical Agricultural Research Station (USDA-TARS) winter nursery in Puerto Rico. At flowering during the 2004-2005 winter nursery, 5601T was crossed with the BC$_{1}$F$_{1}$, and produced BC$_{2}$F$_{1}$ seed, which were planted in Knoxville during summer 2005. Leaves of individual BC$_{2}$F$_{1}$ plants were tested for SSR markers at the University of Tennessee. Once again, the presence of heterozygous alleles at both loci Satt237 and Satt561 were detected.

BC$_{2}$F$_{2}$ seed was harvested from a single plant and grown in the 2005-2006 winter nursery at USDA-TARS in Puerto Rico. A single BC$_{2}$F$_{2}$ harvested plant provided BC$_{2}$F$_{2:3}$ seed which was grown at the Plant Science Unit of East Tennessee Research and Education Center (ETREC) in Knoxville during the summer of 2006 to produce BC$_{2}$F$_{2:4}$ seed after bulk row harvest. During the summer of 2007, the BC$_{2}$F$_{2:4}$ seed was designated TN07-604, and entered into a preliminary MG V test with three replications at one location.

Meanwhile, 5601T was also crossed with a BC$_{2}$F$_{1}$ plant in order to create a BC$_{3}$ line. The BC$_{3}$F$_{1}$ seed was grown at USDA-TARS Puerto Rico in the 2005-2006 winter nursery. In the summer 2006, the BC$_{3}$F$_{2}$ plants were grown in Knoxville and bulk threshed at harvest. The resulting BC$_{3}$F$_{3}$ seed designated as TN07-602 was entered into 2007 preliminary MG V with three replications at one location.

Additionally, in summer 2006, the BC$_{3}$F$_{1}$ seed was planted in the conventional crossing block at ETREC. All plants in row 90 were tested for SSR markers at the University of Tennessee. Plant 23 was found to be heterozygous at both loci Satt237 and Satt561. 5601T was crossed with plant
90-23 to create BC₄F₁ seed. The BC₄F₁ seed was grown at USDA-TARS in Puerto Rico and
grown in the 2006-2007 winter nursery. BC₄F₂ seed was harvested in Puerto Rico in April 2007,
planted in Knoxville in summer 2007, and harvested as single BC₄F₂ plants. In summer 2008, the
BC₄F₂:3 seed was planted in Knoxville and the plant rows were bulk threshed individually. The
BC₄F₂:4 seed of one of the plant rows was designated TN09-239 and entered into a 2009
preliminary MG V field trial with three replications at one location (See Appendix 1 for charts
detailing the development of the BC₂, BC₃, and BC₄ lines.)

In 2009, 5601T was crossed with TN09-239 to create BC₅F₁ seeds at ETREC. The BC₅F₁ seed
was grown in winter nursery at USDA-TARS in Puerto Rico in 2009-2010 and screened for SSR
markers at the University of Tennessee. Ten rows of BC₅F₂ population were grown in Knoxville
during summer 2010 and SNP markers were analyzed on a sample of individual plants.

The recurrent parent (5601T), BC₂ (TN07-604), BC₃ (TN07-602) and BC₄ (TN09-239) derived
lines were planted in a replicated randomized complete block design (RCBD) in 2010 and 2011
with three replications at three locations: The East Tennessee Research and Education Center
(ETREC), Highland Rim Research and Education Center (HHREC) and the Research and
Education Center at Milan (RECMLN). At ETREC, the plots were planted in four, 3.66 m long
rows spaced 76.2 cm apart and were under irrigation for both years. At HHREC the plots were
planted in two, 7.62 m long rows (end trimmed to 4.88 m) spaced 76.2 cm apart for 2010 and the
location was not irrigated. In 2011 HRREC the plots were planted in four rows end trimmed to
3.66 m long and spaced 76.2 cm apart and were under irrigation. At RECMLN in 2010 the plots
were planted in two 9.14 m rows spaced 76.2 cm apart and the location was not irrigated. In
2011 the RECMLN plots were planted in four 3.66 m long rows spaced 76.2 cm apart and were under irrigation.

Procedure for Inorganic P (Pi) Assay

Eight seeds per replication, per location, per genotype were individually crushed using a steel cylinder and hammer. Up to a 0.5 mL subsample of each crushed sample was placed in its own labeled 1.5 mL microcentrifuge tube. One mL of Pi extraction buffer (12.5% trichloroacetic acid (TCA) and 25 mM MgCl$_2$) was added to each microcentrifuge tube and tubes were vortexed thoroughly. Samples were allowed to soak overnight (minimum requirement was 6 h and maximum was 18 h) at 4°C to soften bean tissue. The tubes were vortexed again during this incubation period. Before continuing the assay, the samples were vortexed, and the tissue was allowed to settle for 10-30 min. Tubes were centrifuged at 4000 rpm for 4 min. A 90 µL aliquot of dH$_2$O was pipetted into each well of a flat bottomed Corning Costar plate (Corning Inc., Corning, NY), and 10 µL of the sample extracts was added in their corresponding well. Next, 100 µL of Chen’s reagent (1 volume 6 N H$_2$SO$_4$, 1 volume 2.5% Ammonium molybdate, 1 volume 10% ascorbic acid, 2 volumes ddH$_2$O) was added to each well and allowed to incubate for exactly 1 h. At the end of incubation, Pi concentrations were measured using a Bio-Tek Powerwave XS microplate spectrophotometer set to measure reflectance at an 882nm wavelength.

SNP Analysis

SNP analysis was performed using the Simple Probe Assay and Melting Curve Genotyping Protocol, including primer and probe sequences developed at the Boerma lab at the University of
Georgia using information reported by Gillman et al. (2009). Fresh leaf samples were pressed onto FTA Classic Plant Cards (Whatman Intl. Ltd., Maidstone, Kent, UK). A single punch from each sample was taken using a 1.2 mm micro punch and transferred into a 96 well plate designed for use in the Roche Lightcycler 480 (Roche Applied Sciences, Indianapolis, IN). There were two plates each containing the same samples – one plate was used for SNP detection on Gm19 (LG L) and one was used for SNP detection on Gm03 (LG N). After the samples were transferred to their respective wells, 120 µL of Whatman FTA Purification Reagent was pipetted onto each sample and allowed to sit for 5 min. At the conclusion of 5 min, the reagent was pipetted out of the wells, but the punched samples remained in their wells. Once again, 120 µL of Whatman FTA Purification Reagent was pipetted onto each sample and allowed to sit for 5 min before being pipetted off (punches remained in their respective wells). The following step was done twice: 120 µL of Integrated DNA Technologies Inc. (Coralville, IA) 1xTE Solution was pipetted on each sample, allowed to sit for 5 min, and pipetted off, leaving the punch samples behind. The plate was then placed under a non-plastic guard to keep the sample from being disturbed while the remaining solution was allowed to evaporate overnight.

Forward (Table 3.3) and reverse (Table 3.4) primers and simple probes (Table 3.5) were used for polymerase chain reaction (PCR). The primer stock solutions were obtained from Integrated DNA Technologies (Coralville, IA) and the simple probe stock solutions were obtained from TIB MOLBIOL (Berlin, Germany) and reduced to 5 µM solutions. The Roche LightCycler 480 Genotyping Master Kit (Roche, Indianapolis) was also used in this procedure. Its contents include LightCycler 480 Genotyping Master Mix, PCR-grade H2O, and MgCl2 solution. For each
plate, the following solutions and volumes were pipetted onto each sample well as indicated in Tables 3.1 and 3.2.

Each plate was then centrifuged for 1 min at 1000 rpm before being put into the LightCycler. For the chromosome 19 (LG L) plate, the LightCycler ran 50 PCR cycles at 60 °C for 15 sec. Samples with low phytate characteristics were expected to have a common melting temperature of approximately 54 to 55 °C (Figure 3.3), while samples with normal phytate characteristics were expected to have a common melting temperature of 60 to 61 °C (Figure 3.4). For the chromosome 3 (LG N) plate, the LightCycler ran 50 PCR cycles at 55 °C for 15 sec. Results were read using Roche LightCycler 480 software version 1.5. Samples with low phytate characteristics were expected to yield melting peaks at 62 to 63 °C (Figure 3.5) and samples with normal phytate characteristics were expected to yield melting peaks at 57 to 58°C (Figure 3.6). Samples that exhibited two melting curves at either linkage group were classified as having heterozygous alleles (Figures 3.7 and 3.8).

**Data Analyses**

A chi square goodness of fit test was performed to determine if the BC₃F₂ population produced the expected genotypic inheritance pattern.

**Results and Discussion**

A chi-squared goodness of fit tested the null hypothesis that there was no significant difference between the observed and expected genotypic ratio in an F₂ population (in this case, the BC₃F₂ population). The data for this test is indicated in Table 3.6. We rejected the null hypothesis for
## Table 3.1: Primer, probe, and reagent volumes used for PCR of Gm19 soybean phytate SNP.

<table>
<thead>
<tr>
<th>Solution</th>
<th>Original Concentration</th>
<th>Volume (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forward Primer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF63MRP-F2</td>
<td>5 µM</td>
<td>3</td>
</tr>
<tr>
<td><strong>Reverse Primer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF63MRP-R2</td>
<td>5 µM</td>
<td>1.5</td>
</tr>
<tr>
<td><strong>Simple Probe</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF63MRP-SP[T]</td>
<td>5 µM</td>
<td>0.6</td>
</tr>
<tr>
<td>Genotype Master Mix</td>
<td>5X</td>
<td>2.0</td>
</tr>
<tr>
<td>MgCl2</td>
<td>25 mM</td>
<td>1.5</td>
</tr>
<tr>
<td>Water</td>
<td></td>
<td>6.4</td>
</tr>
</tbody>
</table>

## Table 3.2: Primer, probe, and reagent volumes used for PCR of Gm03 soybean phytate SNP.

<table>
<thead>
<tr>
<th>Solution</th>
<th>Original Concentration</th>
<th>Volume (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forward Primer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF165MRP-F5</td>
<td>5 µM</td>
<td>3</td>
</tr>
<tr>
<td><strong>Reverse Primer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF165MRP-R5</td>
<td>5 µM</td>
<td>1.5</td>
</tr>
<tr>
<td><strong>Simple Probe</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF165MRP-SP[G]-2</td>
<td>5 µM</td>
<td>0.6</td>
</tr>
<tr>
<td>Genotype Master Mix</td>
<td>5X</td>
<td>2.0</td>
</tr>
<tr>
<td>MgCl2</td>
<td>25 mM</td>
<td>1.5</td>
</tr>
<tr>
<td>Water</td>
<td></td>
<td>6.4</td>
</tr>
</tbody>
</table>

## Table 3.3: Nucleotide sequence for Gm19 and Gm03 phytate SNP forward primers obtained from Integrated DNA Technologies (Coralville, IA).

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Synthesized Concentration</th>
<th>Nucleotide Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>LG-L (Gm19)</td>
<td>SF165MRP F5 100 uM</td>
<td>GCC TGA ATT TAA ATG CAC GTC</td>
</tr>
<tr>
<td>LG-N (Gm03)</td>
<td>SF63MRP F2 100 uM</td>
<td>CCT GGA GGC ATC TGT TAT GAC</td>
</tr>
</tbody>
</table>

## Table 3.4: Nucleotide sequence for Gm19 and Gm03 phytate SNP reverse primers obtained from Integrated DNA Technologies (Coralville, IA).

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Synthesized Concentration</th>
<th>Nucleotide Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>LG-L (Gm19)</td>
<td>SF165MRP R5 100 uM</td>
<td>ACA AAT GTG AAG CTG AGG TT</td>
</tr>
<tr>
<td>LG-N (Gm03)</td>
<td>SF63MRP R2 100 uM</td>
<td>GCT GCT GCC ATG TAT GAA AGA</td>
</tr>
</tbody>
</table>

## Table 3.5: Nucleotide sequence for Gm19 and Gm03 phytate SNP simple probes obtained from TIB MOLBIOL (Berlin, Germany).

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Synthesized Concentration</th>
<th>Nucleotide Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>LG-L (Gm19)</td>
<td>SF165MRP-SP[G]-2 20 uM</td>
<td>5'-TTG GCT GTA CTG ATA XI AAT TCT CTC AAT AG - PH</td>
</tr>
<tr>
<td>LG-N (Gm03)</td>
<td>SF63MRP-SP[T] 20 uM</td>
<td>5'-CAA GCT GTT TC XI TTT CAC GAT CGT T - PH</td>
</tr>
</tbody>
</table>
Figure 3.3: Melting curve genotyping SNP assay expected melting temperature (approximately 54 to 55 °C) of soybean sample that has homozygous recessive genotype (low phytate) at Gm19.

Figure 3.4: Melting curve genotyping SNP assay expected melting temperature (approximately 60 to 61 °C) of soybean sample that has homozygous dominant genotype (normal phytate) at Gm19.
Figure 3.5: Melting curve genotyping SNP assay expected melting temperature (approximately 62 to 63 °C) of soybean sample that has homozygous recessive genotype (low phytate) at Gm03.

Figure 3.6: Melting curve genotyping SNP assay expected melting temperature (approximately 57 to 58 °C) of soybean sample that has homozygous dominant genotype (normal phytate) at Gm03.
Figure 3.7: Melting curve genotyping SNP assay double melting peaks (approximately 54 to 55°C and 60 to 61°C) indicates heterozygous genotype for soybean sample at Gm19.

Figure 3.8: Melting curve genotyping SNP assay double melting peaks (approximately 57 to 58°C and 62 to 63 °C) indicate heterozygous genotype for soybean sample at Gm03.
the BC$_5$F$_2$ population ($\chi^2 = 25.79$, P<0.05). The data show that there is a significant difference between the observed and expected data, and that the F$_2$ generation did not inherit the expected genotypic frequency in some of the progeny classes. A likely explanation may be that our sample size of 252 F$_2$ plants was not large enough to fully exhibit the expected inheritance ratio among all nine of the recombinant progeny classes. Each genotype class could be viewed as a targeted selection goal. This would be especially appropriate for the double homozygous recessive class that confers low phytate. In that context, the chi-square test would be a single degree of freedom test of 1/16 aabb versus 15/16 all other classes. Most of the genotype classes, including the low phytate class fit such one degree of freedom chi-squared tests, further suggesting that sample size limitation may have resulted in overall lack of fit of 1:2:1:2:4:2:1:2:1 ratio of nine genotypic classes.

There are a variety of methods that can be used to aid plant breeders in selection decisions during the development of cultivars with desired traits. Breeders are often achieving greater success when they combine certain MAS techniques with classical breeding methods. The use of SSR and SNP markers helped to confirm the low phytate trait capture in the BC$_4$ line TN09-239 and selected BC$_5$ lines.

All F$_1$ plants of the BC$_1$, BC$_2$, BC$_3$, and BC$_4$ generations were found to be heterozygous at both loci for the low phytate trait. The inconclusiveness of the SSR assay for individual BC$_5$F$_1$ (Table 3.7) for plants 127, 132, 137, 139, 143 and 145 (which showed no detectable SSR data at one or two loci) and plant 144 where the Gm03 genotype call was not correct, led us to run the SNP assay on the BC$_5$F$_2$ population (Table 3.6) as well as the established BC$_2$, BC$_3$, and BC$_4$ lines.
Table 3.6: SNP Chi Square Goodness of Fit Test for BC$_5$F$_2$ Progeny resulting from initial cross CX1834-1-2 x 5601T to create a low phytate soybean line (indicated by plants in class “aabb”).

<table>
<thead>
<tr>
<th>CLASSES</th>
<th>RATIO EXPECTED</th>
<th>SNP GENOTYPES</th>
<th>OBSERVED</th>
<th>EXPECTED</th>
<th>(O-E)$^2$/E</th>
</tr>
</thead>
<tbody>
<tr>
<td>AABB</td>
<td>1</td>
<td>T T</td>
<td>8</td>
<td>15.75</td>
<td>3.813492</td>
</tr>
<tr>
<td>AABb</td>
<td>2</td>
<td>T H</td>
<td>25</td>
<td>31.5</td>
<td>1.34127</td>
</tr>
<tr>
<td>AA.bb</td>
<td>1</td>
<td>T CX</td>
<td>9</td>
<td>15.75</td>
<td>2.892857</td>
</tr>
<tr>
<td>AaBB</td>
<td>2</td>
<td>H T</td>
<td>39</td>
<td>31.5</td>
<td>1.785714</td>
</tr>
<tr>
<td>AaBb</td>
<td>4</td>
<td>H H</td>
<td>65</td>
<td>63</td>
<td>0.063492</td>
</tr>
<tr>
<td>AaBB</td>
<td>2</td>
<td>H CX</td>
<td>25</td>
<td>31.5</td>
<td>1.34127</td>
</tr>
<tr>
<td>aaBB</td>
<td>1</td>
<td>CX T</td>
<td>26</td>
<td>15.75</td>
<td>6.670635</td>
</tr>
<tr>
<td>aaBb</td>
<td>2</td>
<td>CX H</td>
<td>45</td>
<td>31.5</td>
<td>5.785714</td>
</tr>
<tr>
<td>aabb</td>
<td>1</td>
<td>CX CX</td>
<td>10</td>
<td>15.75</td>
<td>2.099206</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td></td>
<td>252</td>
<td>252</td>
<td>25.79</td>
</tr>
</tbody>
</table>

H = Heterozygous for low phytate SNP locus  
CX = Homozygous for low phytate SNP locus  
T = Homozygous for normal phytate SNP locus  
8 df, $\chi^2$=15.51, test fails: significant deviations from expectations were observed  
A = Pha1, a = pha1  
B = Pha2, b = pha2

(Table 3.8). As suspected, the BC$_2$ derived TN07-604 line was not homozygous recessive at both loci, but instead was homozygous recessive on Gm19 (LG L) and homozygous dominant on Gm03 (LG N). This is certainly not the targeted genotype. The BC$_2$F$_2$ single plant given rise to the BC$_2$F$_{2.3}$ progeny line must have been homozygous recessive only at the Gm19 locus. The BC$_3$F$_2$ line TN07-602 also was not the targeted genotype because its genotype call was heterozygous on Gm19 (LG L) and homozygous recessive on Gm03 (LG N). The colorimetric assay as described by Scaboo et al. (2009) also showed that the BC$_3$ derived line appeared lighter in color than the BC$_2$ and BC$_4$ lines. It is thought that an error may have been previously made in early BC$_3$ line selection caused by using a sample of three seeds mixed together rather than a single seed per sample, giving a darker color and higher spectrophotometer reading presumed to
Table 3.7: Results from SSR Assay on BC₃F₁ derived lines (recurrent parent 5601T by low phytate mutant line CX1834-1-2) to determine low phytate trait capture.

<table>
<thead>
<tr>
<th>2009 CROSS</th>
<th>PURPOSE</th>
<th>FEMALE</th>
<th>MALE</th>
<th>PR Hill</th>
<th>Satt237</th>
<th>Satt561</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>create 5601T BC5 pha</td>
<td>5601T</td>
<td>TN09-239</td>
<td>VP10 - 125</td>
<td>HET</td>
<td>HET</td>
</tr>
<tr>
<td>15</td>
<td>create 5601T BC5 pha</td>
<td>5601T</td>
<td>TN09-239</td>
<td>VP10 - 127</td>
<td>T</td>
<td>N</td>
</tr>
<tr>
<td>15</td>
<td>create 5601T BC5 pha</td>
<td>5601T</td>
<td>TN09-239</td>
<td>VP10 - 132</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>15</td>
<td>create 5601T BC5 pha</td>
<td>5601T</td>
<td>TN09-239</td>
<td>VP10 - 137</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>15</td>
<td>create 5601T BC5 pha</td>
<td>5601T</td>
<td>TN09-239</td>
<td>VP10 - 139</td>
<td>N</td>
<td>HET</td>
</tr>
<tr>
<td>15</td>
<td>create 5601T BC5 pha</td>
<td>5601T</td>
<td>TN09-239</td>
<td>VP10 - 142</td>
<td>HET</td>
<td>HET</td>
</tr>
<tr>
<td>15</td>
<td>create 5601T BC5 pha</td>
<td>5601T</td>
<td>TN09-239</td>
<td>VP10 - 143</td>
<td>N</td>
<td>HET</td>
</tr>
<tr>
<td>15</td>
<td>create 5601T BC5 pha</td>
<td>5601T</td>
<td>TN09-239</td>
<td>VP10 - 144</td>
<td>CX</td>
<td>HET</td>
</tr>
<tr>
<td>15</td>
<td>create 5601T BC5 pha</td>
<td>5601T</td>
<td>TN09-239</td>
<td>VP10 - 145</td>
<td>N</td>
<td>HET</td>
</tr>
<tr>
<td>15</td>
<td>create 5601T BC5 pha</td>
<td>5601T</td>
<td>TN09-239</td>
<td>VP10 - 147</td>
<td>HET</td>
<td>HET</td>
</tr>
</tbody>
</table>

HET = Heterozygous for low phytate trait at locus
CX = Homozygous for low phytate trait at locus
T = Homozygous for normal phytate trait at locus
N = No data
Pha= low phytate
PRHill = winter nursery hill that was seed source for the BC5F1 lines

be double homozygous recessive low phytate. During early selection, the Pi assay described by Scaboo et al. 2009 was carried out but instead of using a single seed per sample (replicated eight times), three seeds were cracked together and a mixture of those seeds were used as a sample. After this potential error was recognized, all four lines were tested in the Pi assay with just one seed per sample. When we retroactively examined the BC₃ line’s pedigree in detail, as described above, we noticed that each BC₃F₂ row of plants was bulk harvested rather than by single plant harvest to produce the BC₃F₂:₃. Normally, and F₂ population is not bulk threshed due to the
Table 3.8: SNP Assay results at Gm19 and Gm03 soybean phytate loci for DNA of BC$_2$, BC$_3$, and BC$_4$ derived lines TN07-604, TN07-602, and TN09-239 and their recurrent parent 5601T.

<table>
<thead>
<tr>
<th></th>
<th>LG L</th>
<th>LG N</th>
</tr>
</thead>
<tbody>
<tr>
<td>TN07-604</td>
<td>CX</td>
<td>T</td>
</tr>
<tr>
<td>BC2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TN07-602</td>
<td>H</td>
<td>CX</td>
</tr>
<tr>
<td>BC3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TN09-239</td>
<td>CX</td>
<td>CX</td>
</tr>
<tr>
<td>BC4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

H = Heterozygous for low phytate trait at locus
CX = Homozygous for low phytate trait at locus
T = Homozygous for normal phytate trait at locus

variability present in that population. As a consequence, TN07-602 is a bulk collection of selfed progeny of the nine genotypic classes derived from its double heterozygous BC$_3$F$_1$, hence the allele call “H” at pha2 in the BC$_3$ derived line represents the presence of both types of alleles from a collection of homozygous dominant, heterozygous, and homozygous recessive plants at the pha2 locus that collectively constitutes line TN07-602. The BC$_3$F$_2$ line TN09-239 exhibited homozygous recessive alleles on both Gm19 (LG L) and Gm03 (LG N). The BC$_3$F$_2$ population exhibited all nine genotypic classes, with ten different plants with the double homozygous recessive alleles, inferring successful trait introgression proceeded with an expected BC$_5$ average of 98.45% of the remaining genome from the recurrent parent. For a plant breeder, this level of backcross recovery is typically sufficient for release of an improved cultivar following successful field testing of the new line.

The use of the SNP assay in MAS as well as to confirm trait introgression into a backcross line is a useful technique for soybean breeding. An extra benefit is its reduced cost and labor requirements. For example, when backcrossing a double homozygous recessive trait through
multiple generations of backcrossing, a common approach for every backcross generation is to self each BC$_x$F$_1$ plant in order to determine by progeny testing which pollen source actually carried one copy of the recessive form at each locus. This requires the breeder to make a series of excess crossing attempts from many unknown BC$_x$F$_1$ genotypes and retroactively discard all the newly created BC$_y$F$_1$ plants that arose from the BC$_x$F$_1$ plants whose progeny test failed to produce double homozygous recessive progeny. Alternately, the breeder can delay crossing an extra generation every cycle of backcrossing and use the progeny test detected BC$_x$F$_2$ plant found to be homozygous recessive as pollen donor. Compared to the estimate of Scaboo et al. (2009) for SSR genotyping, which estimated ~48 person h per 96 samples and a cost of US $3.60 for DNA extraction per sample and US $2.50 per SSR sample, SNP genotyping is much more economical, costing only $1.75 (including DNA extraction) and ~10 person h. When breeding for traits that have confirmed QTL as well as SNP markers, use of SNPs will most likely become the most popular method of genotyping among the current technologies available.

References


PART 4

Development of a Low Phytate Line adapted to Tennessee and the Southern Region
Abstract:

High yielding, low phytate soybean \([Glycine \textit{max} \text{ (L.) Merr.}]\) lines adapted to the southern U.S. are of value to Tennessee soybean and livestock producers. Between 60 and 85% of P in soybean is in the form of phytate, which cannot be digested by livestock with monogastric digestive systems and is excreted in their waste. This myoinositol hexakisphosphate also binds other minerals and can cause deficiencies in poultry (\textit{Gallus gallus domesticus}) and swine (\textit{Sus domesticus}). Producers often add the enzyme phytase to hydrolyze the phytate, but low phytate soybean lines could alleviate mineral deficiencies, allow for better digestion of P and lessen the environmental effects caused by excess P in agricultural runoff. A low phytate line was developed from backcrossing low phytate mutant line CX1834-1-2 to high yielding recurrent parent 5601T. Low phytate trait introgression was confirmed at both QTL cqPha-001 and cqPha-002 in the BC\textsubscript{4} derived line TN09-239. The purpose of this study was to compare agronomic and seed quality traits in TN09-239 to the recurrent parent. A field study was conducted for two years at three locations. TN09-239 was significantly higher for inorganic P (Pi) (P<0.0001), with TN09-239 containing 1657.9 ng μL\textsuperscript{-1} and significantly lower for yield (P<0.05), with TN09-239 yielding 3151 kg ha\textsuperscript{-1} than the recurrent parent yield (3500 kg ha\textsuperscript{-1}) and Pi (222.9 ng μL\textsuperscript{-1}). The BC\textsubscript{2} line TN07-604 and BC\textsubscript{3} line TN07-602 also exhibited differences in yield and Pi (P<0.01 and P<0.0001, respectively) with TN07-604 yielding 424 kg ha\textsuperscript{-1} less than the recurrent parent and having 373.2 9 ng μL\textsuperscript{-1} more Pi than the recurrent parent. TN07-602 yielded 347 kg ha\textsuperscript{-1} less than the recurrent parent and contained 83.3 ng μL\textsuperscript{-1} more Pi than 5601T. Although the low phytate trait and much of the recurrent parent genome have been captured, further backcrossing is necessary to recover the seed yield of 5601T and other traits desirable for southern U.S. producers.
Introduction

The high protein composition of soybean has made it a very popular livestock feed component. Of total livestock consumption of soybean \(\text{Glycine max}\) (L.) Merr. in the U.S., 48% is by poultry \((\text{Gallus gallus domesticus})\) and 26% is by swine \((\text{Sus domesticus})\) (Soystats 2011). However, soybean contains a phosphate, called phytate, which cannot be completely broken down by swine or poultry (Powers et al., 2006; Gillman et al., 2009; Raboy et al., 2000). Phytate is often present as a phytin salt, or a phytate bound to \(K^+\) or \(Mg^{2+}\) ions. Chemical reactions bind the phytin salt with other nutrient cations such as \(Ca^{2+}\), \(Zn^{2+}\), and \(Fe^{2+}\), which can cause deficiencies of those nutrients in swine and poultry while the phytin salt compound is excreted (Raboy, 2002; Walker et al., 2006). According to Gillman et al. (2009), soybean \(\text{Glycine max}\) (L.) Merrill is not the only plant with phytate. In fact many plants have P in the form of myo-inositol-1,2,3,4,5,6-hexa-kisphosphate. In addition to existing as a phytin salt, phytate is also called phytic acid, which is produced in soybean, as a way of storing P in the seed (Hitz et al., 2002). Once the phytic acid and mineral are bound together, non-ruminant animals are unable to digest either (Oltmans et al., 2005). In many cases, the bonded phytate and mineral cause the respective deficiencies despite the seed containing these minerals. While adding inorganic P (Pi) to feed rations is a common way for livestock producers to combat deficiencies, it does not solve the issue of phytate presence in manure and it adds an extra expense to the operation (Sharpley et al., 2000).

Since phytate cannot be fully broken down by non-ruminant livestock, it is passed in their waste. Runoff from livestock manure from concentrated animal production areas or from fields fertilized with manure, contributes to agricultural non-point sources (NPS) of pollution in
waterways, and phytate in the manure contributes to the problem (Gillman et al., 2009; Maupin et al., 2011; Sharpley et al., 1994). The extra P quickens the onset of eutrophication in fresh water bodies. Eutrophication results in the growth of noxious aquatic weeds and lower dissolved oxygen content in the water, killing aquatic life (Daniel et al., 2011; Sharpley et al., 1994; Sharpley et al., 2000).

Not all P found in soybean is detrimental or otherwise unusable P. In soybean, between 60 and 85% of total P content is phytate, while the remaining 15 to 40% is made up of inorganic P and other forms of the mineral (Gillman et al., 2009; Hitz et al., 2002; Raboy, 2002). Such a high presence of phytate can make up between 1% to 3% of the total seed weight (Gillman et al., 2009; Raboy, 2002). Alternately, as much as 70% of P in the diet of swine is unavailable to the animal and therefore excreted (Powers et al., 2006). According to Reddy et al. (1989), the enzyme phytase hydrolyzes myoinositol hexakisphosphate into inorganic orthophosphate and other forms of phosphoric esters. When phytase is added to livestock feed, it breaks many of the bonds between the phytate and other minerals, allowing more P to be absorbed and digested by the livestock although it is an added expense to producers (Gillman et al., 2009). Alternately, in order to find a low phytate soybean line, Wilcox et al. (2000) used mutagenesis to create the M153 and M766 lines from the CX1515-4 line. Since the M153 mutant had the highest inorganic P (and thereby the lowest phytate), it was chosen to produce various low phytate lines including CX1834-1-2, which was used as the low phytate male donor parent to create backcross progeny lines.
The purpose of this study was to test the null hypotheses that there are no significant differences in genotypic values for seed yield, plant height, plant lodging, plant maturity, seed size, seed protein concentration, seed amino acid composition, seed oil concentration, fatty acid composition of the oil, and seed inorganic phosphorous concentration between the recurrent parent 5601T and BC4, BC3, and BC2 derived progeny lines (TN09-239, TN07-602, and TN07-604, respectively).

Materials and Methods:

Development of the Backcross Lines

The pedigrees of the backcross 2 (BC2) line, TN07-604, backcross 3 (BC3) line, TN07-602, and backcross 4 (BC4) line, TN09-239 can be traced back to the original cross pollination reported by Scaboo et al. (2009) in 2000 between the 5601T cultivar developed by Pantalone et al. (2003) and the low phytate line Cx1834-1-2 developed by Wilcox (2000). This cross produced the line 56Cx-284. In the summer of 2004, 5601T was crossed with 56Cx-284 which resulted in the approximate genetic equivalent of BC1F1 seed. SSR markers for this material were evaluated by the Boerma lab at the University of Georgia. It was found that the BC1F1 material was heterozygous for the low phytate allele at three loci: Satt237 (LG N), Satt561 (LG L) and the closely linked Satt527 (LG L). The close proximity of Satt527 to Satt561 makes its use redundant, and was not included in our study. The BC1F1 seed was planted in December 2004 at the USDA Tropical Agricultural Research Station (USDA-TARS) winter nursery in Puerto Rico. At flowering during the 2004-2005 winter nursery, 5601T was crossed with the BC1F1, which produced BC2F1 seed, which were planted in Knoxville during summer 2005. Leaves of
individual BC$_2$F$_1$ plants were tested for SSR markers at the University of Tennessee. Once again, the presence of heterozygous alleles at both loci Satt237 and Satt561 were detected.

BC$_2$F$_2$ seed was harvested from a single plant and grown in 2005-2006 winter nursery at USDA-TARS in Puerto Rico. A single BC$_2$F$_2$ harvested plant provided BC$_2$F$_{2:3}$ seed which was grown at the Plant Science Unit of East Tennessee Research and Education Center (ETREC) in Knoxville during the summer of 2006 to produce BC$_2$F$_{2:4}$ seed after bulk row harvest. During the summer of 2007, the BC$_2$F$_{2:4}$ seed was designated TN07-604, and entered into a preliminary MG V test with three replications at one location.

Meanwhile, 5601T was also crossed with a BC$_2$F$_1$ plant in order to create a BC$_3$ line. The BC$_3$F$_1$ seed was grown at USDA-TARS Puerto Rico in the 2005-2006 winter nursery. In the summer 2006, the BC$_3$F$_2$ plants were grown in Knoxville and bulk threshed at harvest. The resulting BC$_3$F$_3$ seed designated as TN07-602 was entered into 2007 preliminary MG V with three replications at one location.

Additionally, in summer 2006, the BC$_3$F$_1$ seed was planted in the conventional crossing block at the East Tennessee Research and Education (ETREC) in Knoxville, TN. All plants in row 90 were tested for SSR markers at the University of Tennessee. Plant 23 was found to be heterozygous at both loci Satt237 and Satt561. 5601T was crossed with plant 90-23 to create BC$_4$F$_1$ seed. The BC$_4$F$_1$ seed was grown at USDA-TARS in Puerto Rico and grown in the 2006-2007 winter nursery. BC$_4$F$_2$ seed was harvested in Puerto Rico in April 2007 and planted in Knoxville in summer 2007 and harvested as single BC$_4$F$_2$ plants. In summer 2008, the BC$_4$F$_{2:3}$
seed was planted in Knoxville and bulk threshed. The BC₄F₂:₄ seed was designated TN09-239 and entered into a 2009 preliminary MG V field trial with three replications at one location (See Appendix 1 for charts detailing the development of the BC₂, BC₃, and BC₄ lines.)

In 2009, 5601T was crossed with TN09-239 to create BC₅F₁ seeds at ETREC. The BC₅F₁ seed was grown in winter nursery at USDA-TARS in Puerto Rico in 2009-2010 and screened for SSR markers at the University of Tennessee. Ten rows of BC₅F₂ population were grown in Knoxville during summer 2010 and SNP markers were analyzed on a sample of individual plants.

The recurrent parent (5601T), BC₂ (TN07-604), BC₃ (TN07-602) and BC₄ (TN09-239) derived lines were planted in a replicated randomized complete block design (RCBD) in 2010 and 2011 with three replications at three locations: The East Tennessee Research and Education Center (ETREC), Highland Rim Research and Education Center (HHREC) and the Research and Education Center at Milan (RECMLN). At ETREC, the plots were planted in four, 3.66 m long rows spaced 76.2 cm apart and were under irrigation for both years. At HHREC the plots were planted in two, 7.62 m long rows (end trimmed to 4.88 m) spaced 76.2 cm apart for 2010 and the location was not irrigated. In 2011 the HRREC plots were planted in four rows end trimmed to, 3.66 m long and spaced 76.2 cm apart and were under irrigation. At RECMLN in 2010 the plots were planted in two, 9.14 m rows spaced 76.2 cm apart and the location was not irrigated. In 2011 the RECMLN plots were planted in four, 3.66 m long rows spaced 76.2 cm apart and were under irrigation.
**Phenotypic Traits Observed:**

Flower color (white) was observed when more than 95% of all plants in the replicate were in full bloom (R2 stage) (Fehr and Caviness, 1977). Pubescence color (gray) and days to maturity were noted when 95% of the pods showed their mature color. Lodging score was assigned on a one to five scale at time of maturity, with one being an erect row and five being a row that is prostrate to the ground. Plot average plant height was also recorded at maturity, and seed yield (kg ha\(^{-1}\)) was measured at harvest and adjusted to a 13% moisture basis.

**Sample preparation for NIR Analyses:**

A sample of approximately 20 g of soybean seed from each replicate was ground in a water-cooled Knifetec 1095 Sample Mill (FOSS Tecator, S-26321, Hogana, Sweden) for 20 sec. This gave the whole ground soybean a uniform consistency and particle size. In order to run the near infrared instrument (NIR 6500, Foss North America), it was turned on and the lamp allowed to warm up for two h. Diagnostics were run to test instrument operability, wavelength accuracy, and NIR repeatability. The ground soybean samples were run through the NIR to obtain estimates of amino acid composition and protein and oil concentrations. The software used was Winisi II 2.85.3. The NIR remained on until all samples were scanned, and the diagnostics were run each day. While the NIR was in use, the lab humidity was maintained at 40% by a dehumidifier and the room temperature remained at 20 °C.

**Gas Chromatograph Analyses for Fatty Acids:**

For each replicate, five seeds were bashed with a steel cylinder and hammer so that the seed coat was cracked. The sample was placed into a test tube and 3 mL of an 8:5:2 (v/v/v) ratio of
Chloroform, Hexane and Methanol solvent was added. The samples were then capped and allowed to sit overnight. After extraction was completed, 100 µL of the extract from each sample was transferred into its respective vial. Seventy-five µL of methylation reagent [8:4:2 (v/v/v) ratio of Petroleum Ether, Ethyl Ether, and Sodium Methoxide/Methanol] and 0.75 mL of hexane was added to each vial before it was capped with a crimper. The fatty acid compositions were determined by a Hewlett Packard HP 6890 series gas chromatograph (Palo Alto, CA) system set using a model 7683 auto sampler and a model 7673 flame ionization detector, and an immobilized 30 mm x 0.53 mm inner diameter, Alltech AT-Silar capillary column with 0.5 µm fused stationery phase. Analysis was conducted under the following operating conditions: carrier, Helium (20 ml/min); 20:1(v/v) split injection; injection temperature 250 °C; detector temperature 275 °C, and column temperature 230 °C. The RM-1 standard (appropriate for measuring soybean oil) was used in order to calibrate and determine the relative fatty acid content of the experimental samples.

Procedure for Inorganic P (Pi) Assay

Eight seeds per replication, per location, per genotype were individually crushed using a steel cylinder and hammer. Up to a 0.5 mL subsample of each crushed sample was placed in its own labeled 1.5 mL microcentrifuge tube. One mL of Pi extraction buffer (12.5% trichloroacetic acid (TCA) and 25 mM MgCl₂) was added to each microcentrifuge tube and tubes were vortexed thoroughly. Samples were allowed to soak overnight (minimum requirement was 6 h and maximum was 18 h) at 4°C to soften bean tissue. The tubes were vortexed again during this incubation period. Before continuing the assay, the samples were again vortexed and the tissue was allowed to settle for 10 to 30 min. Tubes were centrifuged at 4000 rpm for four min. A 90
µL aliquot of dH₂O was pipetted into each well of a flat bottomed Corning Costar plate (Corning Inc., Corning, NY), and then 10 µL of the sample extracts was added in their corresponding well. Next, 100 µL of Chen’s reagent (1 volume 6 N H₂SO₄, 1 volume 2.5% Ammonium molybdate, 1 volume 10% ascorbic acid, 2 volumes ddH₂O) was added to each well and allowed to incubate for exactly 1 h. After incubation, inorganic P concentrations were measured using a Bio-Tek Powerwave XS microplate spectrophotometer set to measure reflectance at an 882nm wavelength.

Data Analyses

A randomized complete block design was used with three replications in each of three locations for two years as described above. The random factors included location, replication within location, and genotype by environment interaction, with the fixed term being the genotype. In order to test for significant genotypic and environmental effects, the MIXED procedure was run in SAS (SAS ver. 9.2, Cary, NC). Genotypic differences examined included yield, 16:0, 18:0, 18:1, 18:2, and 18:3 fatty acid content, composition of amino acids, protein and oil concentrations, and inorganic P concentration. The following model was assumed for the combined ANOVA:

\[ Y_{ijk} = \mu + L_i + R(L)_{k(i)} + G_j + (GL)_{ji} + e_{ijk} \]

Where \( Y_{ijk} \) is the specific measured value of genotype \( j \) in rep \( k \) within location \( i \); \( \mu \) is the mean overall; \( L_i \) is effect of location \( i \); \( (GL)_{ji} \) is the interaction between genotype \( j \) and location \( i \), and \( Y_{ijk}'s \) error term is \( e_{ijk} \).

The Pearson Correlation Coefficient was used to calculate the phenotypic correlation among seed quality traits.
Results and Discussion

Improving desired agronomic traits is very important to producers and maintaining seed quality traits such as fatty acid and amino acid composition in newly developed soybean lines is very important to processors. The descriptive statistics for agronomic traits in this study are shown in table 4.1. Least square means estimations and the differences of least square means between lines were obtained overall as well as by environment; only overall means averaged over locations and years are reported. Observations discussed here will focus on differences between the recurrent parent (5601T) and the BC$_4$ line (TN09-239) unless otherwise noted because DNA analyses showed that only TN09-239 inherited the targeted goal of both low phytate recessive genes. Differences in maturity, plant height, yield and seed weight were observed (P<0.001, P<0.0001, P<0.05, and P<0.0001 respectively, data not shown). The contrast between 5601T and the means of all backcross generation lines (TN09-604, TN07-602, TN09-239) was also significant for maturity, height, yield, and seed weight, (P<0.05, P<0.001, P<0.01 and P<0.001, respectively) but not significant for lodging. TN09-239 matured 1.7 days earlier (P<0.05) was 32 cm taller (P<0.0001), yielded 348 kg ha$^{-1}$ less (P<0.05) and had a seed weight of 1.8 g 100$^{-1}$ seed greater (P<0.0001) than 5601T (Table 4.1). This suggests that further backcrossing for the 5601T trait introgression project may be desirable to ensure the targeted low phytate line remains a maturity group V determinate, high yielding line which is desired by southern producers.

Significant differences were found among the four genotypes for inorganic phosphorus (P<0.0001) and for seed oil concentration (P<0.05) but not for seed protein concentration (data not shown). Likewise, the contrast between 5601T and all backcross generation lines (TN07-604, TN07-602, and TN09-239) was significant for inorganic phosphorus (P<0.0001), oil (P=0.05),
Table 4.1: Least Square Means estimations for agronomic traits for 5601T and the BC4 line (TN09-239) developed for low phytate trait introgression grown in three replications at each at ETREC, HRREC, and RECMLN averaged over two years (2010 and 2011).

<table>
<thead>
<tr>
<th>Trait Units</th>
<th>5601T Estimate</th>
<th>TN09-239 Estimate</th>
<th>Difference</th>
<th>SED</th>
<th>t-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maturity (days)</td>
<td>128.3</td>
<td>126.7</td>
<td>1.7</td>
<td>0.6</td>
<td>2.67</td>
<td>0.0203</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>87.4</td>
<td>119.7</td>
<td>-32.3</td>
<td>4.8</td>
<td>-6.79</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Yield (kg ha(^{-1}))</td>
<td>3500</td>
<td>3151</td>
<td>348</td>
<td>123</td>
<td>2.83</td>
<td>0.0126</td>
</tr>
<tr>
<td>Seed Weight (g 100(^{-1}))</td>
<td>13.0</td>
<td>14.9</td>
<td>-1.9</td>
<td>0.2</td>
<td>-7.47</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

SED=standard error of the difference

but not protein (data not shown). TN09-239 produced 1453 ng μL\(^{-1}\) more inorganic phosphorus and 4.2 g kg\(^{-1}\) more oil (P<0.01) than 5601T (Table 4.2). As expected, inorganic P for the BC\(_2\) (TN07-604) (596.1 ng μL\(^{-1}\)) and BC\(_4\) TN09-239 (1675.9 ng μL\(^{-1}\)) lines greatly increased relative to 5601T (222.9 ng μL\(^{-1}\)). However the inorganic P expressed in the BC\(_3\) (TN07-602) (306.2 ng μL\(^{-1}\)) line was not significantly higher than 5601T and was significantly lower than the BC\(_2\) line (P<0.05). This is especially puzzling because Single Nucleotide Polymorphism (SNP) markers run on the BC\(_2\) line showed that it was homozygous recessive for the low phytate genotype on Gm19 (LG L) but homozygous dominant for the normal phytate genotype on Gm03 (LG N).

SNP markers run on the BC\(_3\) line showed that the genotype for Gm19 (LG L) was heterozygous while the genotype on Gm03 (LG N) was homozygous recessive for the low phytate trait. Since Gm03 (LG N) is known to control up to 41% of the low phytate genotype expression and Gm19 (LG L) only controls up to 11% (Walker et al., 2006), it would be expected that the BC\(_3\) line would have more inorganic P than the BC\(_2\) line, given the results from the SNP assay. While it is known that SSR markers can falsely indicate the presence of a desired genotype, there have not been any reports of similar occurrences with SNP markers for this specific trait.
Table 4.2: Least Square Means estimations for seed quality traits for 5601T and its BC₄ line (TN09-239) developed for low phytate trait introgression grown in three replications at each at ETREC, HRREC, and RECMLN averaged over two years (2010 and 2011).

<table>
<thead>
<tr>
<th>Trait Units</th>
<th>5601T Estimate</th>
<th>TN09-239 Estimate</th>
<th>Difference</th>
<th>SED</th>
<th>t-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inorganic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphorus (ng µL⁻¹)</td>
<td>222.9</td>
<td>1675.9</td>
<td>-1453.0</td>
<td>105.5</td>
<td>-13.78</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Protein (g kg⁻¹)</td>
<td>409.44</td>
<td>408.18</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Oil (g kg⁻¹)</td>
<td>205.7</td>
<td>209.9</td>
<td>-4.3</td>
<td>1.3</td>
<td>-3.27</td>
<td>0.0052</td>
</tr>
</tbody>
</table>

SED=standard error of the difference
N/A = no significant difference at P<0.05 level was found among genotypes

There was a significant increase in oil observed between the BC₄ line (209.9 g kg⁻¹) and recurrent parent (205.7 g kg⁻¹); however, there was no significant change in protein between the two lines. Scaboo et al. (2009) reported a weak negative correlation between protein and Pi, however, it was in a population of 187 RILs, whereas correlations within this study would only be among four genotypes, so we cannot accurately predict what protein levels will be observed in further backcrossing this line. Significant differences were found among the four genotype means for all of the fatty acids except oleic (palmitic, P<0.001; stearic, P<0.01; linoleic, P<0.05; and linolenic, P<0.001) and the contrast between 5601T and the mean of all backcross generation lines (TN07-604, TN07-602, and TN09-239) was significant (data not shown) for all five fatty acids (palmitic, P<0.01; stearic, P<0.05; oleic, P<0.05; linoleic, P<0.01; and linolenic, P<0.05). TN09-239 produced 8.0 g kg⁻¹ more palmitic (P<0.001), 3.8 g kg⁻¹ more stearic (P<0.001), 11.6 g kg⁻¹ more oleic (P<0.05), 14.2 g kg⁻¹ less linoleic (P<0.01), and 9.3 g kg⁻¹ less linolenic (P<0.001) than 5601T (Table 4.3). Soybean lines with low palmitic acid content are more desirable due to growing health concerns over saturated-fat content (Cherrak et al., 2003). Lines with low palmitic as well as stearic content, are more likely to meet U.S. labeling requirements for low saturate oils since both of those fatty acids are saturated (Hulke et al., 2004). Low
linolenic and high oleic soybean lines are desirable for improved oxidative stability, aroma, and flavor (Cherrak et al., 2003; Fallen et al., 2012). Breeding for desirable fatty acid levels as well as the low phytate trait appears to be difficult. Fatty acid modifier QTL for all five fatty acids were detected on Gm19 (LG L) (Hyten et al. 2004a). The modifier QTL for palmitic, stearic, oleic, linoleic, and linolenic acids respectively (Hyten et al., 2004a) were located within 26.53, 4.37, 19.93, 11.93, and 11.97 cM of the Gm19 (LG L) pha2 (located at 62.57 on the USLP Consensus Map 4.0).

The only changes observed in amino acids among the four genotypes (Table 4.4) were for arginine (P<0.001), aspartic (P<0.0001), cysteine (P<0.005), histidine (P<0.05), phenylalanine (P<0.01), and tryptophan (P<0.0001) (estimates of all other amino acids can be found in appendix table A4.1). Of these differences, histidine, phenylalanine and tryptophan would be of the most concern because they are three of the nine essential amino acids needed by animals (Panthee et al., 2006; Rao, and Shewry 2009). Cysteine is also sometimes considered an essential amino acid because it must be synthesized from methionine (Rao and Shewry 2009). Although arginine, aspartic and (usually) cysteine, are considered non-essential amino acids, they do provide nitrogen to be used in the synthesis of protein and their compositions should not be completely ignored (Panthee et al., 2006). Likewise, changes between 5601T and TN09-239 were also observed in arginine and aspartic (P<0.0001), cysteine (P<0.01), phenylalanine (P<0.01) and tryptophan (P<0.0001). When comparing the amino acids exhibiting differences among the four genotypes with the amino acids exhibiting differences between 5601T and TN09-239, only histidine did not show a change between 5601T and TN09-239.
Table 4.3: Least Square Means estimations for fatty acid compositions for 5601T and its BC₄ line (TN09-239) developed for low phytate trait introgression grown in three replications at each at ETREC, HRREC, and RECMLN averaged over two years (2010 and 2011).

<table>
<thead>
<tr>
<th>Trait Units</th>
<th>5601T Estimate</th>
<th>TN09-239 Estimate</th>
<th>Difference</th>
<th>SED</th>
<th>t-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palmitic (g kg⁻¹)</td>
<td>125.6</td>
<td>133.6</td>
<td>-8.0</td>
<td>1.7</td>
<td>-4.7</td>
<td>0.0003</td>
</tr>
<tr>
<td>Stearic (g kg⁻¹)</td>
<td>37.7</td>
<td>41.5</td>
<td>-3.8</td>
<td>0.9</td>
<td>-4.4</td>
<td>0.0005</td>
</tr>
<tr>
<td>Oleic (g kg⁻¹)</td>
<td>23.1</td>
<td>24.2</td>
<td>-11.6</td>
<td>4.6</td>
<td>-2.5</td>
<td>0.0234</td>
</tr>
<tr>
<td>Linoleic (g kg⁻¹)</td>
<td>520.4</td>
<td>506.3</td>
<td>14.2</td>
<td>4.0</td>
<td>3.5</td>
<td>0.0031</td>
</tr>
<tr>
<td>Linolenic (g kg⁻¹)</td>
<td>85.9</td>
<td>76.6</td>
<td>9.3</td>
<td>1.8</td>
<td>5.1</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

SED=standard error of the difference

There were some differences for agronomic traits (yield, maturity, height, and seed size, Table 4.1) and for seed quality traits, including Pi, oil, fatty acid and amino traits (Tables 4.2 and 4.3); however, if they are closely monitored in further backcrosses, a line could be selected that is not adversely affected by these changes and developed into a desirable cultivar. One final consideration that remains to be seen is whether a determinate low phytate 5601T line can be developed. Hyten et al. (2004b) noted that the QTL for soybean stem termination (Dt1) is located on Gm19 (LG L) which happens to be near the pha002 region. In 2010 field studies, we noticed segregation for the Dt1 locus in the BC₄ line TN09-239 suggesting that breaking the apparent linkage of 26.5 cM between the Dt1 locus and the low phytate pha002 QTL on Gm19 (LG L) may be challenging and would require further research. The high yield of the recurrent parent has not been recovered yet in the backcross lines. This may be an artifact of the unexpected segregation of indeterminate and determinate plants expressed in the BC₄ line TN09-239, which showed overall plot measurements greater for height and lodging. However, a BC₅ line is expected to exhibit 98.4% of the recurrent parent’s traits (rather than the 96.9% exhibited in the BC₄) and a BC₆ line is expected to exhibit 99.2% of the recurrent parent’s traits. Therefore if the Dt1/pha002 genetic linkage could be broken, we speculate that, it may still be possible to create a
Table 4.4: Least Square Means estimations for amino acid compositions for 5601T and its BC$_4$ line (TN09-239) developed for low phytate trait introgression grown in three replications at each at ETREC, HRREC, and RECMLN averaged over two years (2010 and 2011).

<table>
<thead>
<tr>
<th>Trait (g kg$^{-1}$)</th>
<th>5601T Estimate</th>
<th>TN09-239 Estimate</th>
<th>Difference</th>
<th>SED</th>
<th>t-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arginine</td>
<td>74.2</td>
<td>75.8</td>
<td>-1.6</td>
<td>0.2</td>
<td>-6.6</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Aspartic</td>
<td>112.9</td>
<td>110.8</td>
<td>2.1</td>
<td>0.3</td>
<td>6.5</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Cysteine</td>
<td>15.6</td>
<td>14.9</td>
<td>0.7</td>
<td>0.2</td>
<td>4.0</td>
<td>0.0012</td>
</tr>
<tr>
<td>Histidine</td>
<td>31.0</td>
<td>29.8</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>50.1</td>
<td>49.4</td>
<td>0.7</td>
<td>0.2</td>
<td>4.0</td>
<td>0.0012</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>10.8</td>
<td>11.9</td>
<td>-1.1</td>
<td>0.2</td>
<td>-7.2</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

SED=standard error of the difference
NS=no significant difference (P>0.05)

Nevertheless low phytate *trait introgression* via backcrossing into a high yielding cultivar was successful in the BC$_4$ line TN09-239 as evidenced by its significantly higher inorganic phosphorus levels expressed in the seed in both 2010 and 2011 (Figure 4.1). However, due to the unexpected reduced yield, it would be prudent to continue further backcross or forward cross development of a low phytate line targeting high yield.
Figure 4.1: Inorganic phosphorus content for recurrent parent 5601T and BC₄ line TN09-239 during 2010, 2011, and two year average over three replications each at three locations (ETREC, HHREC, and RECMLN).
References


PART 5

Agronomic Performance and Seed Inorganic Phosphorus Stability of Low Phytate Soybean Line TN09-239 over Ten Southern U.S. Environments
Abstract:
Phytate (myo-inositol-1,2,3,4,5,6-hexa-kisphosphate) in soybean [Glycine max (L.) Merr.] cannot be absorbed by livestock with monogastric digestive systems and thus is often excreted in their waste, contributing to agricultural runoff pollution and fresh water eutrophication. The use of the enzyme phytase is often employed to break the phytin salt bonds and allow for phosphorus (P) absorption, but is an added cost for animal producers. The high yielding line 5601T was continuously backcrossed with low phytate mutant line CX1834-1-2 to develop the BC$_4$ derived line TN09-239. Both 5601T and TN09-239 were entered into the USDA 2010 Uniform Preliminary V Soybean Test for the Southern Region to assess potential differences in agronomic and seed quality traits, fatty and amino acid content, inorganic phosphorus (an indicator of low phytate), and SCN races 2 and 3 susceptibility. Significant differences were observed between 5601T and TN09-239 for lodging, plant height, yield, and seed weight (P<0.05). No significant differences were observed for oil or protein, however significant differences were observed for fatty acid compositions, palmitic, stearic, oleic, linoleic, and linoleic acids (P<0.01, P<0.01, P<0.05, P<0.05, and P<0.001, respectively) were found. Significant differences in amino acid compositions of crude protein were found for arginine, cysteine, leucine, and tryptophan (P<0.05, P<0.01, P<0.05, and P<0.001, respectively). 5601T and TN09-239 were also evaluated for stability across southern environments. A significant genotype by environment interaction was evident (P<0.01), and in a linear regression, 5601T was found to have a slope of zero (P>0.05) while TN09-239 had a slope significantly different from zero (P<0.0001). While much of the recurrent parent genome has been captured in TN09-239, expression of the low phytate is not uniform over all southern.
Introduction

Providing enough food for the world’s population is an ever present challenge for agricultural producers world-wide. With the world population exceeding seven billion people and only approximately 1.5 million square kilometers (roughly 1%) of the earth’s land used for agriculture (The World Fact Book, 2012), producers have been pushing the limits of agricultural practices and science to meet the needs of a growing world population. While innovations and scientific breakthroughs have helped to feed billions and increase the quality of life throughout much of the world, there have also been setbacks caused by agriculture, such as ill-effects, from herbicides, pesticides and erosion.

One such concern is pollution of fresh-water bodies due to agricultural runoff. Runoff of livestock manure from concentrated animal production areas or from fields fertilized with manure, has been identified as agricultural non-point source (NPS) pollution in waterways (Gillman et al., 2009; Sharpley et al., 1994). When excess nutrients from the manure (mainly N and P) make their way into fresh water bodies and estuaries, they quicken the onset of eutrophication. Eutrophication results in the growth of algae and noxious aquatic weeds and lowers dissolved oxygen content in the water. The reduced dissolved oxygen content can cause fish kills and other damage within the ecosystem (Daniel et al., 1994; Sharpley et al., 1994; Sharpley et al. 2000; Maupin et al. 2011). Many times, farmers that fertilize their fields with manure base their application rates solely on the percent N in the manure. The application rate may be correct for meeting the needs of the field crop and minimizing excess N runoff, but the P can be in excess of crop nutrient requirements and increase the likelihood of excess P in agricultural runoff (Sharpley et al., 1994).
Often, the excess P in agricultural runoff is in the form of a phosphate called phytate, which cannot be completely broken down by swine (*Sus domesticus*) and poultry (*Gallus gallus domesticus*) due to their monogastric digestive systems (Powers et al., 2006; Gillman et al., 2009; Raboy et al., 2000). Of the total consumption by animals of U.S. soybean, 49% is by poultry and 25% is by swine (Soystats, 2011). The high protein composition of soybean has made it a very popular livestock feed source, however phytate is often present as a phytin salt, or a phytate bound to K\(^+\) or Mg\(^{2+}\) ions. Chemical reactions bind the phytin salt with other nutrient cations such as Ca\(^{2+}\), Zn\(^{2+}\), and Fe, which can cause deficiencies of those nutrients in swine and poultry and the phytin salt compound is excreted in the manure (Raboy, 2002; Walker et al., 2006). Many plants have P in the form of myo-inositol-1,2,3,4,5,6-hexa-kisphosphate (Gillman et al., 2009). In addition to existing as a phytin salt, phytate can also be called phytic acid, which is produced in soybean [*Glycine max* (L.) Merrill] and corn (*Zea mays*), as a way of storing P in the seed (Hitz et al., 2002). Adding inorganic P (Pi) to feed rations is a common way for livestock producers to combat P deficiencies, but it does not solve the issue of phytate presence in manure and the cost of Pi feed amendments adds an extra expense to the operation (Sharpley et al., 2000).

Not all P found in soybean is detrimental or otherwise unusable P. In soybean, between 60 and 85% of total P content is phytate, while the remaining 15 to 40% is made up of inorganic P and other forms of the mineral (Gillman et al., 2009; Hitz et al., 2002; Raboy, 2002). Such a high presence of phytate means that it can make up between 1% to 3% of the total seed (Gillman et al., 2009; Raboy, 2002). Up to 70% of P in the diet of swine is unavailable to the animal and therefore excreted (Powers et al., 2006). According to Reddy et al. (1989), the enzyme phytase
hydrolyzes myoinositol hexakisphosphate into inorganic orthophosphate and other forms of phosphoric esters. When phytase is added to livestock feed, it breaks many of the bonds between the phytate and other minerals, allowing more Pi to be absorbed and digested by the livestock, although the phytase enzyme is an added expense to producers (Gillman et al., 2009).

Alternatively, in order to find a low phytate soybean line, Wilcox et al. (2000) used mutagenesis to create the M153 and M766 lines from the CX1515-4 line. Since the M153 mutant had the highest inorganic P (and thereby the lowest phytate), it was chosen to produce various low phytate lines including CX1834-1-2, which was used as the low phytate male donor parent to create backcross progeny lines.

The purpose of this study was to test the null hypotheses that there are no significant differences in genotypic values for seed yield, plant height, plant lodging, plant maturity, seed size, seed protein concentration, seed amino acid composition, seed oil concentration, fatty acid composition of the oil, resistance to soybean cyst nematode (SCN) races 2 and 3 and seed inorganic phosphorous concentration between the recurrent parent 5601T and its low phytate BC4, derived progeny line TN09-239. Another goal of this research was to evaluate environmental stability of the low phytate trait across ten southern U.S. environments.

Materials and Methods

Development of the Backcross Line

The pedigree of backcross 4 (BC4) line, TN09-239 can be traced back to the original cross pollination reported by Scaboo et al. (2009) in 2000 between the 5601T cultivar developed by
Pantalone et al. (2003) and the low phytate line Cx1834-1-2 developed by Wilcox (2000). This cross produced the line 56Cx-284. In the summer of 2004, 5601T was crossed with 56Cx-284 which resulted in the approximate genetic equivalent of BC$_1$F$_1$ seed. SSR markers for this material were evaluated by the Boerma lab at the University of Georgia. It was found that some of the BC$_1$F$_1$ plants were heterozygous for the low phytate allele at three loci: Satt237 Gm03 (LG N), Satt561 Gm19 (LG L) and the closely linked Satt527 Gm19 (LG L). The close proximity of Satt527 to Satt561 makes its use redundant, and that marker was not included in our study. The BC$_1$F$_1$ seed was planted at the USDA Tropical Agricultural Research Station (USDA-TARS) winter nursery in Puerto Rico. During flowering at the 2004-2005 TARS winter nursery, 5601T was crossed with the BC$_1$F$_1$, which produced BC$_2$F$_1$ seed, which were planted in Knoxville, TN during summer 2005. Leaf tissue DNA from this material was tested for SSR markers at the University of Tennessee. Once again, the presence of heterozygous alleles at loci Satt237 and Satt561 loci were detected in selected BC$_2$F$_1$ plants, while the remaining plants that were homozygous dominant at either or both loci were discarded.

Then in summer 2005, 5601T was crossed with a selected BC$_2$F$_1$ plant in order to create BC$_3$F$_1$ seeds. The BC$_3$F$_1$ seeds were grown at USDA-TARS Puerto Rico in the 2005-2006 Winter Nursery. In the summer 2006, the BC$_3$F$_2$ seed was planted in the conventional crossing block at the East Tennessee Research and Education (ETREC) in Knoxville, TN. All plants in row 90 were tested for SSR markers at the University of Tennessee. Plant 90-23 was found to be heterozygous at both loci Satt237 and Satt561. 5601T was crossed with plant 90-23 in 2006 to create BC$_4$F$_1$ seed. The BC$_4$F$_1$ seed was grown at USDA-TARS in Puerto Rico at the 2006-2007 Winter Nursery. BC$_4$F$_2$ seed was harvested and planted in Knoxville in summer 2007 and
harvested as single BC₄F₂-plants. In summer 2008, the harvested BC₄F₂;₃ seed was planted in individual rows in Knoxville and each row was bulk threshed separately in the fall. The BC₄F₂;₄ seed was designated TN09-239 and entered into a 2009 preliminary MG V field test with three replications at ETREC (See Appendix 1 for Figure A1.3 detailing the development of the BC₄ line TN09-239.

The recurrent parent (5601T) and the BC₄ derived line TN09-239 were planted in a randomized complete block design (RCBD) in 2010 with two replications at ten locations: Pine Tree, AR, Rohwer, AR, McCune, KS, Pittsburg, KS, Portageville, MO, Stoneville, MS, Kinston, NC, Plymouth, NC, Jackson, TN, and Warsaw, VA as entries in the 2010 USDA Preliminary Group V Regional Test (Gillen and Shelton, 2011). The plots were planted in four, 3.66 m long rows spaced 76.2 cm apart and were under irrigation.

*Phenotypic Traits Observed:*

Flower color (white) was observed when more than 95% of all plants in the plot were in full bloom (R2 stage) (Fehr et al., 1977) Pubescence color (gray) and days to maturity were noted when 95% of the pods showed their mature color. Lodging score was assigned on a one to five scale at time of maturity, with one being a plot with erect plants and five being a plot whose plants were prostrate to the ground. Plot average plant height was also recorded at maturity, and seed yield was measured at harvest and adjusted to a uniform moisture basis.
Sample preparation for NIR Analyses:

A sample of approximately 20 g of soybean seed from each replicate was ground in a water-cooled Knifetec 1095 Sample Mill (FOSS Tecator, S-26321, Hogana, Sweden) for 20 sec. This gave the ground soybean sample a uniform consistency and particle size. In order to run the near infrared instrument (NIR 6500, Foss North America), it was turned on and the lamp allowed to warm up for two h. Diagnostics were run to test instrument operability, wavelength accuracy, and NIR repeatability. The ground soybean samples were run through the NIR to obtain estimates of amino acid composition and protein and oil concentrations. The software used was Winisi II 2.85.3. The NIR remained on until all samples were scanned, and the diagnostics were run each day. While the NIR was in use, the lab humidity was maintained at 40% by a dehumidifier and the room temperature remained at 20 °C.

Gas Chromatography Analyses for Fatty Acids:

For each replicate, five seeds were bashed with a hammer impacting a steel cylinder retained within a steel bearing so that the seed coat was cracked. The sample was placed into a test tube and 3 mL of an 8:5:2 (v/v/v) ratio of chloroform, hexane and methanol solvent was added. The samples were then capped and allowed to sit overnight. After extraction was completed, 100 μL of the extract from each sample was transferred to its respective vial. Seventy-five μL of methylation reagent [8:4:2 (v/v/v) ratio of Petroleum Ether, Ethyl Ether, and Sodium Methoxide/Methanol] and 0.75 mL of hexane was added to each vial before it was capped with a crimper. The fatty acid compositions were determined by a Hewlett Packard HP 6890 series gas chromatograph (Palo Alto, CA) system set using a model 7683 auto sampler and a model 7673 flame ionization detector, and an immobilized 30 mm x 0.53 mm inner diameter, Alltech AT-
Silar capillary column with 0.5 µm fused stationery phase. Analysis was conducted under the following operating conditions: carrier, Helium (20 ml/min); 20:1(v/v) split injection; injection temperature 250 °C; detector temperature 275 °C, and column temperature 230 °C. The RM-1 standard (appropriate for measuring soybean oil) was used in order to calibrate and determine the relative fatty acid content of the experimental samples.

*Procedure for Inorganic P (Pi) Assay*

Eight seeds per replication, per location, per genotype were individually crushed using a steel cylinder and hammer. Up to a 0.5 mL subsample of each crushed sample was placed in individually labeled 1.5 mL microcentrifuge tubes. One mL of Pi extraction buffer (12.5% trichloroacetic acid (TCA) and 25 mM MgCl$_2$) was added to each microcentrifuge tube and tubes were vortexed thoroughly. Samples were allowed to soak overnight (minimum requirement was six h and maximum was 18 h) at 4°C to soften bean tissue. The tubes were vortexed again during this incubation period. Before continuing the assay, the samples were again vortexed and the tissue was allowed to settle for 10 to 30 min. Tubes were centrifuged at 4000 rpm for 4 min. A 90 µL aliquot of dH$_2$O was pipetted into each well of a flat bottomed Corning Costar plate (Corning, Lowell, MA), and then 10 µL of the sample extracts were added in their corresponding well. Next, 100 µL of Chen’s reagent (Chen et al., 1956) (consisting of 1 volume 6 N H$_2$SO$_4$, 1 volume 2.5% Ammonium molybdate, 1 volume 10% ascorbic acid, 2 volumes ddH$_2$O) was added to each well and allowed to incubate for exactly 1 h. After incubation, inorganic P concentrations were measured using a Bio Tek Powerwave XS microplate spectrophotometer (BioTek Instruments, Winooski, VT) set to measure reflectance at an 882nm wavelength.
Data Analyses

Statistical analyses for seed yield, maturity, height, lodging, seed weight, and protein and oil concentration were conducted by location and overall locations using a mixed model ANOVA (PROC MIXED in SAS) by Gillen and Shelton (2011). The fixed effect was the line and replication was the random effect. The LSD (ɑ=0.05) and coefficient of variation (CV) were calculated from the yield output from PROC Mixed. For the remaining traits (seed fatty acids, seed amino acids and seed inorganic P), a seed sample of each replicate of 5601T and TN09-239 was obtained from each field testing collaborator and analyzed by the University of Tennessee as randomized complete block design with two replications in each of ten locations: Pine Tree, AR; Rohwer, AR; McCune, KS; Pittsburg, KS; Portageville, MO; Stoneville, MS; Kinston, NC; Plymouth, NC; Jackson, TN; and Warsaw, VA.

The random factors included location, replication within location, and genotype by environment interaction, with the fixed term being the genotype. In order to test for significant genotypic and environmental effects, the MIXED procedure was run in SAS (SAS ver. 9.2, Cary, NC). Genotypic differences examined included 16:0, 18:0, 18:1, 18:2, and 18:3 fatty acid content by gas chromatography, composition of amino acids, protein and oil concentrations by near infrared spectroscopy, and inorganic P concentration using plate spectrophotometry as described above. The following model was assumed for the combined ANOVA:

\[ Y_{ijk} = \mu + L_i + R(L)_{k(i)} + G_j + (GL)_{ji} + e_{ijk} \]

Where \( Y_{ijk} \) is the specific genotypic difference of genotype \( j \) in rep \( k \) within location \( i \); \( \mu \) is the mean overall; \( L_i \) is effect of location \( i \); \( (GL)_{ji} \) is the interaction between genotype \( j \) and location \( i \), and \( Y_{ijk} \)’s error term is \( e_{ijk} \).
Stability of Pi content was calculated using analysis of variance and least square means estimation, similar to the procedure described by Fallen et al. (2012). In order to estimate the stability of Pi content across locations, the REG procedure was used in SAS v.9.2 (Cary, NC). The stability parameters in this method were the regression coefficient \( b_i \) and deviation from regression \( sd_i \) (Eberhart and Russell, 1966). The model used for estimating the linear regression was:

\[
Y_{ij} = \mu + b_i I_j + \delta_{ij} +
\]

Where \( Y_{ij} \) is the mean of genotype \( i \) within location \( j \), the mean of genotype \( i \) over all locations is \( \mu \), \( b_i \) is the regression coefficient that measures genotype \( i \)'s response to varying locations, \( I_j \) is the environmental index (defined as the mean deviation of all genotypes at a given location from the overall mean), \( \delta_{ij} \) is the deviation from regression for genotype \( i \) in location \( j \). Additionally, coefficients of determination \( (R^2) \) for the genotypes were estimated to measure the percent total variance explained by the regression model.

**Results and Discussion**

Maintaining desired agronomic traits is very important to producers while maintaining or improving seed quality traits such as protein, oil, fatty acid, and amino acid concentration in newly developed and improved soybean lines is very important to processors and their customers. Performance data obtained from the 2010 Uniform Soybean Tests for the Southern Region (Gillen and Shelton, 2011) indicated that there were significant differences in lodging, plant height, yield, and seed size between 5601T and TN09-239. TN09-239 had a 0.9 higher lodging score, was 40.6 cm taller, had a 2.3 g/100 greater seed size, and yielded 470.3 kg/ha\(^{-1}\) less than 5601T but there was no significant difference for days to maturity (Table 5.1). This
Table 5.1: Agronomic trait means for 5601T and TN09-239 averaged over ten environments of the 2010 Uniform Preliminary V Soybean Test for the Southern Region.

<table>
<thead>
<tr>
<th></th>
<th>Agronomic</th>
<th>Seed Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Maturity† (days)</td>
<td>Lodging††</td>
</tr>
<tr>
<td>5601T</td>
<td>-2</td>
<td>1.4</td>
</tr>
<tr>
<td>TN09-239</td>
<td>-2</td>
<td>2.3</td>
</tr>
<tr>
<td>LSD (0.05)</td>
<td>-2†</td>
<td>0.4</td>
</tr>
</tbody>
</table>

† Actual maturity in days not provided by Uniform Soybean Test Southern States results. Values indicate the number of days difference from a designated check averaged over environments. †† Lodging, score 1-5 where 1=all plants erect and 5=all plants prostrate DW Protein and oil concentration are reported on a seed dry weight basis.

suggests that further backcrossing may be desirable to ensure that the low phytate line becomes more genetically similar to its high yielding recurrent parent 5601T, and expresses similar agronomic attributes.

There was no observed difference in oil or protein concentration between the BC₄ line TN09-239 and recurrent parent 5601T (Table 5.1). Scaboo et al. (2009) reported a significant but weak negative correlation between seed protein and Pi concentration, however that was in a population of 187 RILs from a cross of the unrelated parental lines 5601T and CX1834-1-2, whereas this study only examined two lines which were highly related through backcrossing. Scores indicating susceptibility to SCN races 2 and 3 were identical for 5601T and TN09-239, where each scored the highest possible value (5), averaged over seven replications conducted in greenhouse pots at the USDA-ARS, Jackson, TN (data not shown). The susceptibility of 5601T to SCN has been previously documented (Pantalone et al., 1993).
Significant differences were found in the contrasts between 5601T and TN09-239 across all locations for each of the fatty acids (palmitic, P<0.01; stearic, P<0.01; oleic, P<0.05; linoleic, P<0.05; and linolenic, P<0.001). TN09-239 produced 5 g kg$^{-1}$ more palmitic, 5 g kg$^{-1}$ more stearic, 28 g kg$^{-1}$ more oleic, 27 g kg$^{-1}$ less linoleic, and 14 g kg$^{-1}$ less linolenic than 5601T (Table 5.2). Soybean lines with low palmitic acid content are more desirable due to growing health concerns over saturated fat content (Cherrak et al., 2003). Lines with low palmitic as well as low stearic content are more likely to meet U.S. labeling requirements for low saturate oils (Hulke et al., 2004). Higher oleic soybean lines are desirable for improved oxidative stability, aroma, and flavor (Cherrak et al., 2003, Fallen 2012). Breeding for desirable fatty acid levels as well as the low phytate trait appears to be problematic only for targeted low saturated fat soybeans. Development of low linolenic and high oleic soybeans appears to be benefited. Fatty acid modifier QTL for all five fatty acids were detected on Gm19 (LG L) (Hyten et al. 2004a). The modifier QTL (Hyten et al., 2004a) were located within 26.53, 4.37, 19.93, 11.93, and 11.97 cM of the Gm19 (LG L) pha2 (located at 62.57 on the USLP Consensus Map 4.0) for palmitic, stearic, oleic, linoleic, and linolenic acids, respectively. The linkage between the low phytate QTL and palmitic content could be problematic. Hyten et al. (2004a) observed that growth habit QTL Dt1 (located ~77 cM on the USLP Consensus Map 4.0) and palmitic acid content are not independent, noting that the Williams cultivar in that study had an indeterminate growth habit.

The low phytate QTL is located closer to the Dt1 locus than it is to the palmitic locus. Stem termination growth habit was observed to be segregating in TN09-239. Thus its BC4F1 plant with the targeted double heterozygous genotype must have also been heterozygous at Dt1 because it produced progeny segregating for that trait. It may prove difficult to break the genetic
Table 5.2: Least square means estimations for fatty acid composition of 5601T and TN09-239 over ten environments in the 2010 USDA Uniform Preliminary V Soybean Test, Southern Region.

<table>
<thead>
<tr>
<th>Trait Units</th>
<th>5601T Estimate</th>
<th>TN09-239 Estimate</th>
<th>Difference</th>
<th>SED</th>
<th>t-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palmitic (g/kg&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>112</td>
<td>117</td>
<td>-5</td>
<td>0.1</td>
<td>-4.29</td>
<td>0.0026</td>
</tr>
<tr>
<td>Stearic (g/kg&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>33</td>
<td>38</td>
<td>-5</td>
<td>0.1</td>
<td>-3.89</td>
<td>0.0046</td>
</tr>
<tr>
<td>Oleic (g/kg&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>198</td>
<td>226</td>
<td>-28</td>
<td>1.1</td>
<td>-2.56</td>
<td>0.0335</td>
</tr>
<tr>
<td>Linoleic (g/kg&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>572</td>
<td>545</td>
<td>27</td>
<td>0.8</td>
<td>3.1</td>
<td>0.0140</td>
</tr>
<tr>
<td>Linolenic (g/kg&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>86</td>
<td>72</td>
<td>14</td>
<td>0.2</td>
<td>5.93</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

SED=standard error of the difference

linkage between indeterminate growth habit (Dt1) and the LG L phytate locus pha2 to provide producers with a low phytate, determinate line adapted to the southern U.S. Yet if we are able to break that Dt1 linkage, then we would also be likely to overcome the linkage with elevated palmitic from the fatty acid modifier QTL near pha-002.

The only changes observed in amino acids (Table 5.3) were for arginine (P<0.05), cysteine (P<0.05), leucine (P<0.05) and tryptophan (P<0.0001) Of these differences, leucine and tryptophan would be of the most concern because they are two of the nine essential amino acids needed by animals (Panthee et al., 2006; Rao and Shewry, 2009). Cysteine is also sometimes considered an essential amino acid because it must be synthesized from methionine (Rao, and Shewry 2009). Although arginine and (usually) cysteine are considered non-essential amino acids, they do provide N to be used in the synthesis of protein and their compositions should not be completely ignored (Panthee et al., 2006).
Table 5.3: Least Square Means estimations for amino acid compositions for 5601T and its BC₄ line (TN09-239) grown over ten environments in the 2010 USDA Uniform Preliminary V Soybean Test, Southern Region.

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>5601T Estimate (g kg⁻¹ CP)</th>
<th>TN09-239 Estimate (g kg⁻¹ CP)</th>
<th>Difference</th>
<th>SED</th>
<th>t-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>50.6</td>
<td>50.1</td>
<td>0.5</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Arginine</td>
<td>77.0</td>
<td>79.5</td>
<td>-2.5</td>
<td>1.0</td>
<td>-2.63</td>
<td>0.0251</td>
</tr>
<tr>
<td>Aspartic</td>
<td>115.5</td>
<td>113.6</td>
<td>0.9</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Cysteine</td>
<td>17.1</td>
<td>15.6</td>
<td>1.5</td>
<td>0.4</td>
<td>3.74</td>
<td>0.0039</td>
</tr>
<tr>
<td>Glutamine</td>
<td>168.6</td>
<td>167.7</td>
<td>1.0</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Glycine</td>
<td>61.7</td>
<td>60.9</td>
<td>0.8</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Histidine</td>
<td>33.8</td>
<td>32.8</td>
<td>0.9</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>47.6</td>
<td>48.0</td>
<td>-0.4</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Leucine</td>
<td>73.7</td>
<td>72.5</td>
<td>1.2</td>
<td>0.6</td>
<td>2.3</td>
<td>0.0481</td>
</tr>
<tr>
<td>Lysine</td>
<td>55.5</td>
<td>55.4</td>
<td>0.1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Methionine</td>
<td>16.3</td>
<td>16.2</td>
<td>0.1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>50.5</td>
<td>50.2</td>
<td>0.3</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Proline</td>
<td>55.4</td>
<td>56.2</td>
<td>-0.8</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Serine</td>
<td>60.0</td>
<td>60.8</td>
<td>-0.7</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Threonine</td>
<td>44.4</td>
<td>43.8</td>
<td>0.5</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>41.2</td>
<td>40.4</td>
<td>0.7</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>12.1</td>
<td>12.9</td>
<td>-0.9</td>
<td>0.2</td>
<td>-5.27</td>
<td>0.0004</td>
</tr>
<tr>
<td>Valine</td>
<td>57.8</td>
<td>58.3</td>
<td>-0.5</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

SED=standard error of the difference
N/A=not applicable because trait was not significant at the P=0.05 level.
Successfully developing a high yielding, low phytate soybean cultivar for the southern U.S. may also require that a minimum value of elevated Pi be expressed in seeds produced in environments where soybeans are grown in the south. Such a situation is found for NuSun sunflower, whose oil has significantly higher oleic acid than normal sunflower, but the values of elevated oleic acid are not consistent across the growing environments. Hence the industry adopted a minimum standard for a harvested sunflower crop to be designated as NuSun. The BC₄ derived line TN09-239, with confirmed low phytate alleles was entered into the 2010 Uniform Preliminary V Soybean Test for the Southern Region to test its seed Pi levels. As expected, when averaged across replications and ten locations, there was a highly significant difference for seed Pi between 5601T (170 ng µL⁻¹) and TN09-239 (1959 ng µL⁻¹) (P<0.0001), with TN09-239 having 1789 ng µL⁻¹ more Pi than 5601T. However when Pi concentration was compared between individual environments, it was found that the expression of the low phytate trait varied over environment, but even at the environment with the lowest phytate levels (Pittsburg, KS), the value for TN09-239 (1556 ng µL⁻¹) was significantly higher than that for 5601T (154 ng µL⁻¹).

5601T and TN09-239 were significantly different for Pi (P<0.0001). The raw Pi data was arranged into regions defined by the USDA (Gillen and Shelton 2011) as West: Pittsburg, KS and McCune, KS; East: Kinston, NC, and Plymouth, NC; South: Jackson, TN; and Delta: Rohwer, AR, Pine Tree, AR, Portageville, MO, and Stoneville, MS and genotypic means for Pi within each location were used to observe potential stability issues across the southern U.S. There was variation of as much as 658 ng µL⁻¹ of Pi across regions for TN09-239 while very little variation was observed for 5601T (38 ng µL⁻¹), as shown in figure 5.1. The data indicated that there was a significant genotype by environment interaction (P<0.01). Regression was used
to test if the Pi concentration for each line fit a linear equation, whose stability parameters are shown in Table 5.4. While each line was tested for the expected slope of zero, only 5601T fit that slope (P>0.05) and TN09-239 had a slope that was significantly different from zero (P<0.0001) (Table 5.4). TN09-239 also showed a slope significantly different from one (P<0.01). The environmental index and means for each line and the equation describing their linear trend line are shown (Figure 5.2). The slope of TN09-239 (2.03 ng µL$^{-1}$ unit change in Pi for each unit change in environmental index) indicates that this line is not stable for the low phytate trait across southern U.S. environments, yet the mean value of Pi for TN09-239 is typically an order of magnitude greater than that of 5601T in every environment (Figure 5.2). Further research may be necessary to determine how much variation for phytate is genetically controlled by the cqPha-001 and cqPha-002 loci and why a significant portion (37%) remains unexplained by the duplicate dominate epistatic QTL (Gillman et al., 2009) and how much is affected by planting environments, as well as whether the low phytate trait can be stably expressed across


<table>
<thead>
<tr>
<th>Genotype</th>
<th>Inorganic Phosphorus Content (ng µL$^{-1}$)</th>
<th>Stability Parameters</th>
<th>Slope Different than Zero</th>
<th>Slope Different than One</th>
</tr>
</thead>
<tbody>
<tr>
<td>5601T</td>
<td>170</td>
<td>CV = 0.0456</td>
<td>0.3657</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>TN09-239</td>
<td>1959</td>
<td>CV = 0.7562</td>
<td>&lt;0.0001</td>
<td>0.0013</td>
</tr>
</tbody>
</table>

CV = coefficient of variability is calculated by 100*std. deviation/mean and measures the amount of variability in the data relative to the mean; bi = regression coefficient that measures the response of line i to varying environments; sdi = a genotype’s deviation from the regression line; R$^2$ = coefficient of determination, determined for each genotype to measure the percentage of the total variance explained by the regression model; †CV is calculated for individual genotype mean.
Figure 5.1: Inorganic phosphorus means for 5601T and TN09-239 grown in 2010 across locations within designated regions in the Southern U.S. Locations within each region are as follows: East: Kinston, NC, Plymouth, NC, and Warsaw, VA; South: Jackson, TN; Delta: Pine Tree, AR, Rohwer, AR, Portageville, MO, Stoneville, MS; West: McCune, KS, Pittsburg, KS.

Figure 5.2: Genotype by environmental index linear regression against inorganic phosphorus for 5601T and its BC₄ low phytate line TN09-239. Slope for TN09-239 (2.033) is significantly different than zero (P<0.0001) and did not exhibit stability for the low phytate trait when grown across ten locations in the southern U.S. in 2010.
environments in a newly developed soybean line.

Our study used highly related lines (BC$_4$-derived versus recurrent parent) and it was evident that the recessive forms of cqPha-001 and cqPha-002 when combined in a single genotype expressed nearly a ten-fold increase in Pi in every environment tested. We neglected to collect soil samples at each location. A future study that analyzed soil Pi time of planting and harvest may further indicated environmental effects on genetic control of seed phytate. Moreover, an important genetic linkage (Dt1 and pha-002) will need to be broken to produce a determinant low phytate backcross line to more accurately compare trait expression with 5601T. We have initiated the BC$_5$ generation to target that goal.

Although low phytate trait introgression via backcrossing into a high yielding cultivar was generally successful in the BC$_4$ line, its unexpected reduced yield, excessive plant height and lack of stability for the low phytate trait expression indicated that it would be prudent to continue further backcrosses or forward crosses to develop a superior low phytate line targeting high yield. Although low phytate trait introgression via backcrossing into a high yielding cultivar was generally successful in the BC$_4$ line, its unexpected reduced yield, excessive plant height and lack of stability for the low phytate trait expression indicated that it would be prudent to continue further backcrosses or forward crosses to develop a superior low phytate line targeting high yield for the southern US.
References


coloration of a mutation that confers a decreased raffinosaccharide and phytic acid

Hulke, Brent S., Walter R. Fehr, and Grace A Welke. 2004. Agronomic and seed characteristics
of soybean with reduced phytate and palmitate. Crop Sci. 44: 2027-2031.

Hyten, David L., Vincent R. Pantalone, Arnold M. Saxton, Michael E. Schmidt, and Carl E.
Sams. 2004a. Molecular mapping and identification of soybean fatty acid modifier


seed traits of soybean lines with low-phytate phosphorus. Crop Sci. 45: 593-598.

Sci. 43:1123.


PART 6

Conclusion and Future Research
Conclusions

In this study, successful low phytate introgression into a commercial cultivar was confirmed, however not all desired traits from the recurrent parent were fully recovered. It was observed that the use of SNPs in MAS is useful for confirming trait introgression as well as allowing for more precise selections by indicating the presence (or lack of presence) of desired genotypes in progeny rows. When compared to phenotypic and SSR assays that test for the low phytate trait, SNP genotyping is less expensive and less time consuming. While the use of SNPs is likely to become the most popular method of genotyping among currently available technologies, the use of SSRs are still necessary to develop a more comprehensive map of confirmed SNP markers and it is likely that in the future, more precise, useful and cost effective technologies will be developed to allow breeders to map the soybean genome.

When comparing agronomic and seed quality traits and Pi concentration of the backcross lines to the recurrent parent 5601T, changes were observed for several traits. However the traits of most concern are yield, height, Pi concentration, and fatty acid composition. The desired increases in Pi concentration were observed in the backcross lines however yield was reduced. A BC5 derived line is expected to recover 98.4% of the recurrent parent’s genome, however with yield being controlled by so many QTL, it is not certain that further backcrosses will recover the high yield observed in the recurrent parent.

It also remains to be seen is whether a determinate low phytate line with desirable low saturated fat levels can be developed. Since QTL for all five fatty acids are known to exist on Gm19 within 60 cM of the pha002 locus, there may be some difficulties in breeding for such a line.
While the low phytate trait is generally of concern to livestock producers and low saturate fats is a concern for human consumption, breeding for a line with both qualities may soon be of value to soybean producers since soybean oil and meal are separated before being processed into livestock feed components, human food components, or other products. Likewise, Hyten et al. (2004b) noted that the QTL for soybean stem termination (Dt1) is located on Gm19 at position 89.1 cM on USLP 3.0 (~77 cM on USLP 4.0). In this study, we noticed breaking the apparent linkage of 26.53 cM between the Dt1 locus and the low phytate pha002 QTL on Gm19 would require further research.

The final conclusion of this study is that while the low phytate trait in the BC4 derived line TN09-239 was present and expressed in all plots planted across the ten southern US environments, variation was observed in the low phytate trait expression across those environments.

**Future Research**

Currently the pha1 and pha2 alleles are known to account for only up to 63% of phytate trait expression. Further discovery of QTL accounting for the remaining 37% of trait expression would be useful for breeders to be able to breed for low phytate soybean lines. However, this study did not examine the degree of phytate reduction needed to improve livestock health and reduce environmental impacts. Nor was non-phytate and non-inorganic P monitored for changes during the development of the low phytate line. In future research, it would be prudent to be sure that current low phytate lines in the research realm have not already achieved the desired phytate reduction. Additionally, if research was conducted with a goal of gaining a better understanding
of the relationship of all forms of phosphorus within soybean seed, more trait linkages could be
discovered, perhaps leading to a better understanding of the soybean genome and uncovering
potential genetic stumbling blocks when breeding for low phytate or other desired traits. In
addition to assaying for relationships between all forms of phosphorus in the seed, examining the
effect of soil P on P content in the seed could be helpful in identifying whether or not soil P has
any effect on phytate levels in soybean seed.

If BC₅ and BC₆ lines are going to be further developed, plants with determinate growth habits
should be screened for low phytate alleles. If such plants are found, they should be advanced and
evaluated for other desirable traits as well as growth habit and low phytate traits. It may become
necessary to repeat the development of a low phytate line through backcrossing or cross a
backcross line with an unrelated cultivar with desirable traits, if desired traits cannot be captured
in the BC₅ and BC₆ lines. If this were the case, the use of SNP genotyping will aid in more
accurate selections, have reduced lab costs, and possibly allow for selections to happen sooner
since the SNP assay can be done before harvest.
Table A1.1: Development of BC<sub>2</sub> line TN07-604 from 5601T and CX1834-1-2.

**Development of TN07-604 (BC<sub>2</sub> Entry)**

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Cross</th>
<th>Creates</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer 2004</td>
<td>Knoxville</td>
<td>5601T x 56Cx-284</td>
<td>BC&lt;sub&gt;1&lt;/sub&gt;F&lt;sub&gt;1&lt;/sub&gt; seeds</td>
<td>Assume genetic equivalent of creating BC&lt;sub&gt;1&lt;/sub&gt;F&lt;sub&gt;1&lt;/sub&gt; because 56Cx-284 is the offspring of 5601T x CX1834-1-2</td>
</tr>
<tr>
<td>Winter 2004-2005</td>
<td>Puerto Rico</td>
<td>5601T x TN05-PR-074</td>
<td>BC&lt;sub&gt;2&lt;/sub&gt;F&lt;sub&gt;1&lt;/sub&gt; seeds</td>
<td>BC&lt;sub&gt;1&lt;/sub&gt;F&lt;sub&gt;1&lt;/sub&gt; plant designated TN05-PR-074, tested for SSR markers by the University of Georgia and found to be heterozygous at targeted loci, Satt 237, Satt561, Satt 527</td>
</tr>
<tr>
<td>Summer 2005</td>
<td>Knoxville</td>
<td>None</td>
<td>BC&lt;sub&gt;2&lt;/sub&gt;F&lt;sub&gt;2&lt;/sub&gt; seeds</td>
<td>BC&lt;sub&gt;2&lt;/sub&gt;F&lt;sub&gt;1&lt;/sub&gt; plants designated row 452-9, grown in field F8. Tested all plants in row 452-9 for SSR markers, by University of Tennessee and found to be heterozygous at both loci, Satt 237 and Satt 561 for low phytate alleles. The single plant 452-9-04 was chosen to collect BC2F2 seed.</td>
</tr>
<tr>
<td>Winter 2005-2006</td>
<td>Puerto Rico</td>
<td>None</td>
<td>BC&lt;sub&gt;2&lt;/sub&gt;F&lt;sub&gt;2:3&lt;/sub&gt; seed</td>
<td>BC&lt;sub&gt;2&lt;/sub&gt;F&lt;sub&gt;2&lt;/sub&gt; seed planted in Winter Nursery produced fifty hills, Hill # 419 (designated as TN06PR-419) was threshed to provide BC&lt;sub&gt;2&lt;/sub&gt;F&lt;sub&gt;2:3&lt;/sub&gt; seed</td>
</tr>
<tr>
<td>Summer 2006</td>
<td>Knoxville</td>
<td>None</td>
<td>BC&lt;sub&gt;2&lt;/sub&gt;F&lt;sub&gt;2:4&lt;/sub&gt;</td>
<td>TN06PR-419 planted in rows 30,121-30,122</td>
</tr>
<tr>
<td>Summer 2007</td>
<td>Knoxville</td>
<td>None</td>
<td>BC&lt;sub&gt;2&lt;/sub&gt;F&lt;sub&gt;2:4&lt;/sub&gt; seed Source TN07-604</td>
<td>BC&lt;sub&gt;2&lt;/sub&gt;F&lt;sub&gt;2:4&lt;/sub&gt; seed entered into 2007 PYT5CON test, designated TN07-604, used as BC&lt;sub&gt;2&lt;/sub&gt; entry for 2010 and 2011 Low Phytate Tests</td>
</tr>
</tbody>
</table>
Table A1.2: Development of BC\textsubscript{3} line TN07-602 from 5601T and CX1834-1-2.

**Development of TN07-602 (BC\textsubscript{3} Entry)**

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Cross</th>
<th>Creates</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer 2004</td>
<td>Knoxville</td>
<td>5601T x 56Cx-284</td>
<td>BC\textsubscript{1}F\textsubscript{1}</td>
<td>Assume genetic equivalent of creating BC\textsubscript{1}F\textsubscript{1} because 56Cx-284 is the offspring of 5601T x CX1834-1-2</td>
</tr>
<tr>
<td>Winter 2004-2005</td>
<td>Puerto Rico</td>
<td>5601T x TN05-PR-069</td>
<td>BC\textsubscript{2}F\textsubscript{1}</td>
<td>BC\textsubscript{1}F\textsubscript{1} seed designated TN05-PR-069, tested for SSR markers by the University of Georgia and found to be heterozygous at all three loci, Satt 237 (Linkage Group N), Satt 561, and Satt 527 (Linkage Group L)</td>
</tr>
<tr>
<td>Summer 2005</td>
<td>Knoxville</td>
<td>5601T x 452-4-03</td>
<td>BC\textsubscript{3}F\textsubscript{1}</td>
<td>BC\textsubscript{2}F\textsubscript{1} seed designated as and grown in row 452-4, in field F8. Tested all plants in row 452-4 for SSR markers, by University of Tennessee and found to be heterozygous at both loci, Satt 237 and Satt 561 for low phytate alleles. The single plant 452-4-03 was chosen to make backcross.</td>
</tr>
<tr>
<td>Winter 2005-2006</td>
<td>Puerto Rico</td>
<td>None</td>
<td>BC\textsubscript{3}F\textsubscript{2}</td>
<td>Ten hills result from planting of BC\textsubscript{3}F\textsubscript{1} seed, designated as hills TN06PR-274 - TN06PR-283, Hill TN06PR-275 threshed to provide BC3F2 seed</td>
</tr>
<tr>
<td>Summer 2006</td>
<td>Knoxville</td>
<td>None</td>
<td>BC\textsubscript{3}F\textsubscript{3}</td>
<td>BC\textsubscript{3}F\textsubscript{2} plants designated 20,698-20,699 and bulk threshed at harvest</td>
</tr>
<tr>
<td>Summer 2007</td>
<td>Knoxville</td>
<td>None</td>
<td>BC\textsubscript{3} Seed Source TN07-602</td>
<td>BC\textsubscript{3}F\textsubscript{3} seed entered into 2007 PYT5CON test, designated TN07-602, used as BC\textsubscript{3} entry for 2010 and 2011 Low Phytate Tests</td>
</tr>
</tbody>
</table>
### Table A1.3: Development of BC4 line TN09-239 from 5601T and CX1834-1-2.

#### Development of TN09-239 (BC\textsubscript{4} Entry)

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Cross</th>
<th>Creates</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer 2004</td>
<td>Knoxville</td>
<td>5601T x 56Cx-284</td>
<td>BC\textsubscript{1}F\textsubscript{1}</td>
<td>Assume genetic equivalent of creating BC\textsubscript{1}F\textsubscript{1} because 56Cx-284 is the offspring of 5601T x CX1834-1-2</td>
</tr>
<tr>
<td>Winter 2004-2005</td>
<td>Puerto Rico</td>
<td>5601T x TN05-PR-069</td>
<td>BC\textsubscript{2}F\textsubscript{1}</td>
<td>BC\textsubscript{2}F\textsubscript{1} seed designated TN05-PR-069, tested for SSR markers by the University of Georgia and found to be heterozygous at all three loci, Satt 237, Satt561, Satt 527</td>
</tr>
<tr>
<td>Summer 2005</td>
<td>Knoxville</td>
<td>5601T x 452-4-01</td>
<td>BC\textsubscript{3}F\textsubscript{1}</td>
<td>BC\textsubscript{3}F\textsubscript{1} seed designated 452-4, grown in field F8. Tested all plants in row 452-4 for SSR markers, by University of Tennessee and found to be heterozygous at both loci, Satt 237 and Satt 561 for low phytate alleles. The single plant 452-4-01 was chosen to make backcross.</td>
</tr>
<tr>
<td>Summer 2006</td>
<td>Knoxville</td>
<td>5601T x 90-23</td>
<td>BC\textsubscript{4}F\textsubscript{1}</td>
<td>BC\textsubscript{4}F\textsubscript{1} seed planted in row 90 of conventional crossing block. All plants in row 90 were tested for SSR markers by University of Tennessee and plant 23 was found to be heterozygous at both loci, Satt 237 and Satt 561 for low phytate alleles and therefore chosen for backcross.</td>
</tr>
<tr>
<td>Winter 2006-2007</td>
<td>Puerto Rico</td>
<td>None</td>
<td>BC\textsubscript{4}F\textsubscript{2}</td>
<td>BC\textsubscript{4}F\textsubscript{1} seed planted in hills VP07-065 to VP07-74, hill VP07-072 threshed to provide BC\textsubscript{4}F\textsubscript{2} seed.</td>
</tr>
<tr>
<td>Summer 2007</td>
<td>Knoxville</td>
<td>None</td>
<td>BC\textsubscript{4}F\textsubscript{3}</td>
<td>BC\textsubscript{4}F\textsubscript{2} seed planted in rows 20,083-20,289. Seed from plant #50 was harvested for BC\textsubscript{4}F\textsubscript{3} seed.</td>
</tr>
<tr>
<td>Summer 2008</td>
<td>Knoxville</td>
<td>None</td>
<td>BC\textsubscript{4}F\textsubscript{4}</td>
<td>BC\textsubscript{4}F\textsubscript{3} seed planted in row 30,356 and bulk harvested.</td>
</tr>
<tr>
<td>Summer 2009</td>
<td>Various</td>
<td>None</td>
<td>BC\textsubscript{4} Seed Source TN09-239</td>
<td>Entered in PYT5CON test. Became seed stock for TN09-239 which is the BC\textsubscript{4} entry in 2010 and 2011 Low Phytate Tests as well as the BC\textsubscript{4} entry in the Tennessee Multiple Environments Test.</td>
</tr>
</tbody>
</table>
Table A4.1: Least squares means for all traits examined across all Tennessee locations (ETREC, HRREC, and RECMLN) over both 2010 and 2011 growing season.

<table>
<thead>
<tr>
<th>TRAIT (units)</th>
<th>Recurrent Parent 5601T</th>
<th>BC2 TN07-604</th>
<th>BC3 TN07-602</th>
<th>BC4 TN09-239</th>
<th>SED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maturity (days)</td>
<td>128.3</td>
<td>125.3</td>
<td>128.9</td>
<td>126.7</td>
<td>0.62</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>87.4</td>
<td>98.9</td>
<td>91.4</td>
<td>119.7</td>
<td>0.17</td>
</tr>
<tr>
<td>Lodging</td>
<td>2.6</td>
<td>2.9</td>
<td>2.6</td>
<td>2.9</td>
<td>1.87</td>
</tr>
<tr>
<td>Yield (g kg(^{-1}))</td>
<td>3499.6</td>
<td>3076.2</td>
<td>3153.7</td>
<td>3151.2</td>
<td>123.07</td>
</tr>
<tr>
<td>Seed Weight (g 100(^{-1}))</td>
<td>13.1</td>
<td>13.9</td>
<td>13.3</td>
<td>14.9</td>
<td>0.24</td>
</tr>
<tr>
<td>Pi (ng µL(^{-1}))</td>
<td>222.9</td>
<td>596.1</td>
<td>306.2</td>
<td>1675.9</td>
<td>105.46</td>
</tr>
<tr>
<td>Protein (g kg(^{-1}))</td>
<td>409.4</td>
<td>406.2</td>
<td>403.1</td>
<td>408.2</td>
<td>2.26</td>
</tr>
<tr>
<td>Oil (g kg(^{-1}))</td>
<td>205.7</td>
<td>205.5</td>
<td>207.1</td>
<td>209.9</td>
<td>1.30</td>
</tr>
<tr>
<td>Palmitic (g kg(^{-1}))</td>
<td>125.6</td>
<td>128.6</td>
<td>125.2</td>
<td>133.6</td>
<td>1.70</td>
</tr>
<tr>
<td>Stearic (g kg(^{-1}))</td>
<td>37.7</td>
<td>38.8</td>
<td>39.0</td>
<td>41.5</td>
<td>0.87</td>
</tr>
<tr>
<td>Oleic (g kg(^{-1}))</td>
<td>230.5</td>
<td>236.4</td>
<td>241.4</td>
<td>242.1</td>
<td>4.60</td>
</tr>
<tr>
<td>Linoleic (g kg(^{-1}))</td>
<td>520.4</td>
<td>512.4</td>
<td>509.1</td>
<td>506.3</td>
<td>4.02</td>
</tr>
<tr>
<td>Linolenic (g kg(^{-1}))</td>
<td>85.9</td>
<td>84.1</td>
<td>85.4</td>
<td>76.6</td>
<td>1.84</td>
</tr>
<tr>
<td>Alanine (g kg(^{-1}))</td>
<td>48.0</td>
<td>47.3</td>
<td>47.5</td>
<td>47.7</td>
<td>0.42</td>
</tr>
<tr>
<td>Arganine (g kg(^{-1}))</td>
<td>74.2</td>
<td>75.0</td>
<td>74.8</td>
<td>75.8</td>
<td>0.24</td>
</tr>
<tr>
<td>Aspartic (g kg(^{-1}))</td>
<td>112.9</td>
<td>112.4</td>
<td>112.5</td>
<td>110.8</td>
<td>0.32</td>
</tr>
<tr>
<td>Cysteine (g kg(^{-1}))</td>
<td>15.6</td>
<td>15.6</td>
<td>15.6</td>
<td>14.9</td>
<td>0.17</td>
</tr>
<tr>
<td>Glutamine (g kg(^{-1}))</td>
<td>166.4</td>
<td>169.4</td>
<td>170.1</td>
<td>166.8</td>
<td>1.98</td>
</tr>
<tr>
<td>Glycine (g kg(^{-1}))</td>
<td>55.7</td>
<td>53.6</td>
<td>54.0</td>
<td>55.3</td>
<td>1.16</td>
</tr>
<tr>
<td>Histidine (g kg(^{-1}))</td>
<td>31.0</td>
<td>29.6</td>
<td>29.6</td>
<td>29.8</td>
<td>0.60</td>
</tr>
<tr>
<td>Isoleucine (g kg(^{-1}))</td>
<td>46.9</td>
<td>47.0</td>
<td>47.1</td>
<td>47.1</td>
<td>0.08</td>
</tr>
<tr>
<td>Leucine (g kg(^{-1}))</td>
<td>72.8</td>
<td>73.5</td>
<td>73.8</td>
<td>71.8</td>
<td>0.57</td>
</tr>
<tr>
<td>Lysine (g kg(^{-1}))</td>
<td>59.6</td>
<td>60.7</td>
<td>61.3</td>
<td>59.0</td>
<td>0.93</td>
</tr>
<tr>
<td>Methionine (g kg(^{-1}))</td>
<td>15.8</td>
<td>15.2</td>
<td>15.3</td>
<td>15.5</td>
<td>0.09</td>
</tr>
<tr>
<td>Phenylalanine (g kg(^{-1}))</td>
<td>50.1</td>
<td>50.0</td>
<td>50.2</td>
<td>49.4</td>
<td>0.17</td>
</tr>
<tr>
<td>Proline (g kg(^{-1}))</td>
<td>52.6</td>
<td>52.4</td>
<td>52.1</td>
<td>52.6</td>
<td>0.36</td>
</tr>
<tr>
<td>Serine (g kg(^{-1}))</td>
<td>54.3</td>
<td>53.8</td>
<td>53.4</td>
<td>55.0</td>
<td>0.80</td>
</tr>
<tr>
<td>Threonine (g kg(^{-1}))</td>
<td>42.7</td>
<td>42.1</td>
<td>42.1</td>
<td>42.6</td>
<td>0.38</td>
</tr>
<tr>
<td>Tryptophan (g kg(^{-1}))</td>
<td>10.8</td>
<td>10.9</td>
<td>11.2</td>
<td>11.9</td>
<td>0.15</td>
</tr>
<tr>
<td>Tyrosine (g kg(^{-1}))</td>
<td>40.1</td>
<td>39.6</td>
<td>39.6</td>
<td>39.7</td>
<td>0.37</td>
</tr>
<tr>
<td>Valine (g kg(^{-1}))</td>
<td>55.9</td>
<td>54.4</td>
<td>54.7</td>
<td>55.8</td>
<td>0.59</td>
</tr>
</tbody>
</table>

SED = standard error of the difference
VITA

Suzannah J. Wiggins was born September 18, 1987 in York, Pennsylvania. She was raised on the banks of the Susquehanna River outside Wrightsville, Pennsylvania and graduated from Eastern York High School in 2005. She attended Tennessee Technological University in Cookeville, TN, graduating with a B.S. in Agriculture with a concentration in Education.

Suzannah enrolled at the University of Tennessee in July 2010 as a graduate research assistant and plans to receive her Master of Science degree in May 2012. She has accepted a position with Pioneer Hi-Bred International as an Agronomist Trainee in Renwick, Iowa and plans to move on to a full time Agronomist position within the company.