# Mapping Quantitative Trait Loci for Soybean Quality Traits from Two Different Sources

Elizabeth Marie Clevinger

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
In
Crop and Soil Environmental Sciences

Dr. M.A. Saghai Maroof, Committee Chair
Dr. Glenn Buss, Committee Member
Dr. Cynthia Denbow, Committee Member

May 11, 2006
Blacksburg, Virginia

Keywords: soybean, oligosaccharides, stachyose, phytate phosphorous, quantitative trait loci

Mapping Quantitative Trait Loci for Soybean Quality Traits from Two Different Sources

Elizabeth M. Clevinger

## ABSTRACT

Soybeans are economically and agriculturally the most important legume in the world, providing protein and oil to the food and animal feed industries and base ingredients for hundreds of chemical products. Their value could be enhanced, however, if the oil and protein content remained high and the oligosaccharide and phytate contents were lowered to make soybeans more acceptable for human and animal consumption.

A soybean population of 55 families segregating for genes controlling quality traits was chosen for this study. Both parental lines have high sucrose and low stachyose. The former contains a high level of phytate while the latter is low phytate. The objective of this experiment was to determine whether or not both parents had the same gene(s) for low stachyose. An additional objective was to determine quantitative trait loci (QTL) controlling quality traits: sucrose, stachyose and phytate. An acetonitrile precipitation method and a modified colorimetric method were used to determine amounts of sugars and phytate, respectively. The phenotypic data for stachyose was analyzed and it was determined that two recessive genes control low stachyose content in this population. A map was constructed using 141 SSR markers and 15 molecular linkage groups (MLGs) were identified. After analyzing trait and marker data in QTL Cartographer, potential QTL were found on MLGs: B1, C2, D1b, F, M and N. Sucrose and stachyose QTL were identified on B1, C2, M and N. Phytate QTL were observed on B1, D1b, F and N. The markers identified for quality traits in this population may be useful in marker-assisted selection and the germplasm should be useful for the development of a cultivar.

## *Dedication*

My parents, Larry and Freeda Clevinger, who always believed in me

~

Josh Taylor for keeping me grounded and telling me that I would succeed at whatever I set my mind to and allowing me to succeed. I couldn't have done it without you.

~

Dominic Tucker for making me laugh every morning in lab

## TABLE OF CONTENTS

## LIST OF FIGURES AND TABLES

## Introduction

Soybeans first emerged as a domesticated plant around the eleventh century B.C.; which means it is one of the oldest cultivated crops in agriculture (Verma and Shoemaker, 1996). They are economically and agriculturally the most important legume in the world, providing protein and oil to the food and animal feed industries and base ingredients for hundreds of chemical products (Hedley, 2001). Soybeans were originally planted in the United States as a forage crop and then became important as an oilseed crop by the early twentieth century. Today it is one of the world's leading sources of seed oil and seed protein (Verma and Shoemaker, 1996). Soybeans represented 30.2 million hectares of crop production in the United States in 2000, with a value of over $12 billion. Soybeans accounted for 56% of the world oilseed production, with the United States accounting for 25% of the world production of soybeans. This makes soybeans an integral component of the United States' agriculture industry (Kerley and Allee, 2003).

Protein, oil, sugars, starch, organic acids and amino acids account for 75% of the dry weight in soybean (*Glycine max* L.) seeds (Yazdi-Samadi et al., 1977). Soybean cultivars that are typically grown in the United States, on a dry matter basis, contain about 21% oil, 40% protein and 11% soluble carbohydrates, the latter fraction being composed of fructose, sucrose, raffinose and stachyose (Openshaw and Hadley, 1977). Kuo et al. (1988) found that the carbohydrate, sucrose, is the predominant soluble saccharide in soybean seeds.

Soybeans are primarily grown for their oil and protein content. Their value could be enhanced, however, if the oil and protein content remained high and the

oligosaccharide content was lowered to make soybeans more acceptable for human consumption (Hymowitz et al., 1972). A high level of soluble oligosaccharides is important in the soyfood industry, for products such as, natto (Cober et al., 1997).

Phytic acid is a nearly ubiquitous component of plant seeds and is usually the most abundant form of phosphate in those seeds (Hitz et al., 2002). It typically represents from 65% to 85% of seed total phosphorous and acts as a phosphate storage compound. When consumed in feeds and foods, phytic acid will bind to nutritionally important minerals such as calcium and iron, which can contribute to mineral deficiencies (Raboy, 2002).

Diet rations fed to poultry and swine are typically supplemented with either phytase or inorganic phosphate to increase phosphorous availability (Lei and Porres, 2003; Adeola et al., 1995). Eutrification of surface water can result from runoff and leaching of excessive phosphorous in soils to which livestock or poultry manure has been applied (Ertl et al., 1998). This shows that the excretion of phytate by non-ruminant animals is therefore potentially detrimental to the environment and supplementation of soybean meal may exacerbate the problem. Reducing phytate and increasing inorganic phosphorous in soybean meal would increase the amount of phosphorous available to non-ruminants, which would in turn decrease the amount of supplemental inorganic phosphorous added to their ration and lower their fecal phosphorous (Cromwell et al., 2000; Cromwell, 2002).

There is a great interest in developing lines that are high sucrose, low stachyose and low phytic acid. The reason for this is that sucrose is digestible, while stachyose is not digestible by monogastric animals. A line that is high sucrose and low stachyose

would be better for both animal feed and human consumption due to the improved digestibility of the soybeans. Low phytic acid lines are important due to the environmental concerns of phosphorous runoff due to the addition of phosphorous to animal diets since phytic acid is largely unavailable to animals in feed. Low phytic acid lines would make phosphorous available and there would be no need for adding more phosphorous to the feed that would be unnecessary.

# LITERATURE REVIEW

## *Importance of Soybean:*

Soybean seed can be processed for protein and oil or utilized directly as a raw product in the manufacturing of several oriental soyfood items (Maughan et al., 2000). However, soybeans are primarily grown for their oil and protein content (Hymowitz et al., 1972). It is one of the more important sources of protein for both human consumption and as a fodder (Zhang et al., 2004). They also contain significant amounts of sucrose and the galactosyl oligosaccharides, raffinose and stachyose. These three oligosaccharides are of potential importance in the acceptance of soybeans as food, and are also of industrial significance since their total concentration can be greater than 10% of the seed dry weight (Kennedy et al., 1985).

The animal feed industry uses 77% of the soybean meal produced primarily as an amino acid and protein source in diets (Kerley and Allee, 2003). Protein meal from soybean is used as a component in animal feed but in reduced proportions due to the soluble sugars raffinose and stachyose. These soluble sugars may cause flatulence in dogs and digestive disturbances in baby pigs and chicks due to their inability to digest these saccharides. If these sugars were reduced then the amount of soy-protein in rations could be increased (Hartwig et al., 1997). The inability of monogastric animals to digest raffinose saccharides is due to the lack of the enzyme α-galactosidase, which is necessary to break down raffinose and stachyose into sucrose and galactose, which monogastric animals can readily digest (Neus et al., 2005). The poor digestion of raffinose

saccharides reduces the amount of metabolizable energy obtained from soybean meal (Meis et al., 2003). Soybean meal is used to meet the animal's requirement for limiting amino acids because soybean meal is usually the most cost-effective source of amino acids. Soybean meal is also one of the best protein sources for complementing the limiting amino acid profile of corn protein (Kerley and Allee, 2003).

Soybeans contain a variety of anticarcinogenic phytochemicals. This variety of phytochemicals has demonstrated anticancer activity, including protease inhibitors, phytate, phytosterols, saponins and isoflavones. Soybeans also contain bioactive proteins that exhibit anticancer activity, including lectins and lunasin. Lectins have the ability to agglutinate cells, which is a well-recognized physiologic effect that is dependent on their specific, high-affinity binding to particular carbohydrate moieties on the cell surface. It has been reported that soybean agglutinin is capable of inhibiting the growth of a transplanted tumor in rats (de Mejia et al., 2003). Lunasin is a major component of the Bowman-Birk protease inhibitor, a cancer-preventing component from soybeans. When the lunasin gene is injected into cancer cells, cell division is arrested and apoptosis is induced. It also inhibits the transformation of normal embryo fibroblast cells into cancerous cells by preventing chromatin acetylation and oncogene activation in cells with mutated tumor suppressor gene (de Mejia et al., 2003).

***Description of Molecular Markers:***

Some of the complications that are encountered when using visible phenotypes to select a superior plant in a breeding program may be overcome by using DNA markers that cosegregate with the gene(s) of interest. There are many different kinds of molecular

markers that may be used in plant diagnostics including: Random Amplified Polymorphic DNA (RAPD), Restriction Fragment Length Polymorphisms (RFLPs), Amplified Fragment Length Polymorphisms (AFLPs), Single Nucleotide Polymorphisms (SNPs) and Simple Sequence Repeats (SSRs) or microsatellites.

The RAPD procedure does not require species-specific primers and a single set of primers can be used with any genome. It also uses less genomic DNA per assay than other assays such as the RFLP procedure. The procedure only detects one allele at a locus, the one that gave rise to the amplification. This means that these are dominant markers and cannot be used to distinguish heterozygous loci in a population (Rafalski et al., 1991).

The RFLP procedure is used for the construction of genetic maps of agronomically important species and for the mapping of genetic traits but the method has become less popular now due to the availability of more rapid and less labor-intensive methods. An advantage to this method over RAPDs is that it uses codominant markers, which can detect heterozygous loci but it is still undesirable for plant breeding projects because it is not a high throughput method. This procedure requires species-specific probes, several micrograms of DNA and many steps including blotting and hybridization (Rafalski et al., 1991).

The AFLP technique is based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA. This method allows for the specific co-amplification of high numbers of restriction fragments and can be used on DNA of any origin or complexity. A benefit of this method is that DNA fingerprints can be produced

without any prior sequence knowledge. It essentially combines the RFLP technique with PCR (Vos et al., 1995).

SNPs are an abundant form of genome variation and account for around 90% of human DNA polymorphism. They are single base pair positions in genomic DNA at which different sequence alleles exist in normal individuals in some populations. The RFLP technique takes advantage of SNPs. Other procedures to detect SNPs involve target PCR amplification, which is time-consuming and costly. Other assays to detect SNPs are currently being investigated (Brookes, 1999). One such method is temperature gradient capillary electrophoresis (TGCE); this method unlike other SNP detection methods can be used even in the absence of prior knowledge of the sequence information (Hsia et al., 2005).

Microsatellites or as they are usually known, simple sequence repeats (SSRs), are another class of genetic markers. The repeats are ubiquitous in eukaryotic genomes and consist of repeats of 2-5 nucleotides in a continuous stretch (Schlotterer and Tautz, 1992; Valdes et al., 1993). They are a good choice for a genetic marker because they are: highly polymorphic, multiallelic, abundant, spread evenly throughout the eukaryotic genome, show Mendelian inheritance and are codominant. This marker system provides codominant markers, so that the heterozygote can be distinguished from either homozygote. The most polymorphic of the microsatellites are those that are perfect arrays and contain no interruptions. They may also be compound, made up of continuous or adjacent tandem arrays of different motifs (Goldstein and Schlotterer, 1999). In order for SSRs to be used as genetic markers, primers are synthesized using sequences that are

complementary to the flanking regions which are conserved DNA sequences in the genome (Maughan et al., 1995).

***Molecular Linkage Groups (MLGs) in Soybean:***

Genes located on the same chromosome show evidence of linkage to one another and in this way linkage groups may be established. In theory, the number of linkage groups should correspond to the haploid number of chromosomes (Klug and Cummings, 2002). For most characteristics the soybean behaves genetically like a diploid. It is then generally regarded to have 2n=40 as the diploid chromosome number. Genetic evidence of gene duplication also suggests that soybean is a polyploid and there is a strong tendency for polyploids to evolve into a diploid state through sequence diversification and chromosome rearrangement. It is therefore likely that the soybean genome exists in multiple copies (Verma and Shoemaker, 1996).

Soybeans have 20 molecular linkage groups and this goes along with the idea that the number of linkage groups should correspond to the haploid number of chromosomes. The linkage groups of soybeans are designated as follows: A1, A2, B1, B2, C1, C2, D1a, D1b, D2, E, F, G, H, I, J, K, L, M, N and O.

One of the genetic linkage maps available for soybean is an integrated map that combines data from five soybean mapping populations. The five populations used were 'Minsoy' x 'Noir 1', 'Minsoy' x 'Archer', 'Archer' x 'Noir 1', 'Clark' x 'Harosoy' and A81-356022 x PI468916. These five maps were combined into one integrated genetic map spanning 2,523.6 cM across 20 linkage groups that contained 1,849 markers which

included 1,015 SSRs, 709 RFLPs, 73 RAPDs, 24 classical morphological traits, six AFLPs, ten isozymes and 12 others (Cregan et al., 1999; Song et al., 2004).

***Description of Quantitative Trait Loci (QTL):***

The conflict between the Mendelian theory of particulate inheritance and the observation that most traits in nature exhibit continuous variation was eventually resolved by the concept that quantitative inheritance can result from the segregation of multiple genetic factors, modified by environmental effects (Lander and Botstein, 1989). Quantitative traits, in comparison to qualitative traits are characterized by a continuum of phenotypes and are controlled by the joint action of many genes. The plant's environment has a greater influence on a quantitative trait than on a qualitative trait. Yield is one such trait that is profoundly affected by environmental factors.

Based on the construction of a soybean genetic linkage map, quantitative trait locus (QTL) mapping has been reported for a number of agronomic traits in soybean. An example of soybean QTL analysis, can be found in Maughan et al. (2000), where seventeen marker loci associated with sucrose were found in a segregating $F_2$ population which contained 149 individuals. These loci mapped to seven different genomic regions that were significantly associated with sucrose content in soybean. One hundred seventy-seven genetic markers and four morphological markers were assayed for linkage relationships in an $F_2$ mapping population. The seventeen markers mapped to the following linkage groups of soybean: A1, A2, E, F, M, L and I (Maughan et al., 2000).

Another example comes from Zhang et al. (2004), who found 63 quantitative trait loci for nine agronomic traits including: days to flowering, days to maturity and plant

height. These 63 loci were mapped to 12 linkage groups. They found that most of the QTL were clustered and that some of the QTL even mapped to the same loci. This pleiotropism was common for most of the QTL and it was also observed that one of the QTL influenced up to five traits. The number of QTL for the different traits ranged from 4 to 10 (Zhang et al., 2004).

An additional example of QTL mapping in soybean, can be found in Primomo et al. (2005), where they identified QTL associated with isoflavone content of soybeans in a population of $F_{4:6}$ recombinant inbred lines that was produced from the cross 'AC756' x 'RCAT Angora'. Ninety-nine SSR markers were mapped and 17 of these showed highly significant associations with isoflavone content. These 17 QTL were located on 11 independent linkage groups and explained from 3.5 to 10.5% of the phenotypic variation for specific isoflavone content in this population. This low level of variation was indicative of the quantitative nature of isoflavone inheritance in soybean seeds. It was stated that the QTL identified in this study could be useful for developing soybean varieties with desirable isoflavone content in the seed through marker-assisted selection (Primomo et al., 2005).

***Description of Sucrose:***

Sucrose is a primary constituent of soybean seed (Hymowitz and Collins, 1974; Maughan et al., 2000). It is a nonreducing disaccharide of an α-D-glucopyranose and β-D-fructofuranose joined by a (1→2) linkage (Dey et al., 1985). It is the only disaccharide that is present in appreciable amounts in legume seeds (Hedley, 2001).

Sucrose is the primary photosynthate translocated into developing soybean embryos (Lowell and Kuo, 1989). Hymowitz et al. (1972) examined 60 different soybean lines representing virtually the entire range of oil and protein percentages expected in soybean seed. Sucrose on average represented 60% of the total sugar content of the 60 lines that were sampled. Sucrose was also found to be positively correlated with oil content and negatively associated with the protein content. Total sugar was found to be positively correlated with sucrose as well. Sucrose is also positively correlated with raffinose but negatively correlated with stachyose (Hymowitz et al., 1972).

There are three enzymes associated with sucrose synthesis in green plants; they are sucrose phosphate synthase, sucrose phosphatase and sucrose synthase. The enzymes occur in the cytoplasm of the cells and are not located within organelles. The substrates of sucrose phosphate synthase are UDP-glucose and fructose 6-P. This enzyme is important in the regulatory control of sucrose synthesis. Sucrose phosphatase has a high affinity for sucrose. Sucrose phosphatase is involved in the control of sucrose synthesis by virtue of the fact that it inhibits sucrose. Sucrose synthase synthesizes sucrose and is a ubiquitous enzyme in higher plants and occurs in almost all plant tissues. Another one of the important roles of sucrose synthase is the breakdown of sucrose. The enzyme is also one of the major soluble proteins in many growing plant tissues (Dey et al., 1985).

***Description of Raffinose:***

Raffinose [β-D-fructofuranosyl-*O*-α-D-galactopyranosyl- (1→6)-α-D-glucopyranoside] is a trisaccharide that is a α-galactosyl derivative of sucrose. It may be found in the cotyledons, seed coats, and hypocotyls (Bentsink et al., 2000).

Raffinose family oligosaccharides (RFOs) have multiple functions in plants. The most common RFOs are the trisaccharide raffinose, the tetrasaccharide stachyose and the pentasaccharide verbascose (Minorsky, 2003). They serve as transport carbohydrates in the phloem, as storage reserves and cryoprotectants in frost-hardy plant organs. They also accumulate in maturing seeds, where they may play a role in the acquisition of desiccation tolerance and storability (Karner et al., 2004). RFOs are components of the carbohydrate reserves of many seed types, ranking second only to sucrose in abundance as soluble carbohydrates. RFOs are found at especially high concentrations in legume seeds. Due to the absence of α-galactosidase and β-fructosidase activity in human and animal intestine mucosa, RFOs escape digestion and are metabolized by bacteria to hydrogen, carbon dioxide and methane. RFOs have thus been identified as the principal flatulence causing factors present in legumes and other seeds. This is the single most important reason as to why people do not eat more legumes. This problem also affects monogastric animals (Tewary and Muller, 1992).

Protein meal from soybean contains the soluble sugars raffinose and stachyose and because of these sugars, the quantity of soybean meal must be limited in rations to avoid flatulence in dogs and digestive disturbances in baby pigs and chicks. Soybean meal is also a common source of protein for livestock (Meis et al., 2003). It is also an excellent source of protein for humans (Guimaraes et al., 2001). Flatulence is caused when soybean protein accounts for more than 30% of the protein in rations. This means

that if the levels of these oligosaccharides were reduced then the amount of soy proteins in rations could be increased (Hartwig et al., 1997).

Soybean meal has very low metabolizable energy compared to its gross energy and this is due to the very poor digestibility of the carbohydrate fraction. The major reason for the low metabolizable energy is the oligosaccharides, raffinose and stachyose. In addition to their indigestibility, raffinose and stachyose can produce diarrhea that may increase digesta passage rate and decrease digestion and absorption of dietary nutrients (Parsons et al., 2000).

Studies in soybean, maize and brassica have suggested that soluble oligosaccharides might be involved in the protection of seeds against damage during seed dehydration and aging, and therefore in seed survival and storability. It has also been proposed that the soluble oligosaccharides may be involved in the protection of membranes, proteins and nucleic acids against damage that occurs during and upon the withdrawal of water in the drying of seeds (Bentsink et al., 2000).

***Description of Stachyose:***

Stachyose is a tetrasaccharide [*O*-α-D-galactopyranosyl- (1→6)-*O*-α-D-galactopyranosyl- (1→6)-α-D-glucopyranosyl-β-D-fructofuranoside]. It coexists with raffinose and other related oligosaccharides.

Stachyose has been recognized as an important transport carbohydrate in a large number of woody plants, cucurbits and legumes (Dey et al., 1985). Stachyose is also part of the raffinose family of oligosaccharides as stated above. Since this oligosaccharide is part of this family, it also has the same antinutritional properties that are associated with

this family such as flatulence. The same problems associated with raffinose may also be considered problems of stachyose since they are both RFOs. The biosynthesis of stachyose proceeds by the reversible addition of galactose units from galactinol (*O*-α-D-galactopyranosyl- (1→1)-L-*myo*-inositol) to sucrose. This chain elongation is caused by the action of stachyose synthase (Karner et al., 2004).

The implications of low stachyose are many. As stated before, RFOs have been identified as the principal flatulence causing factors present in legumes and other seeds. This is the single most important reason as to why people do not eat more legumes. This problem also affects monogastric animals (Tewary and Muller, 1992). Protein meal from soybean contains the soluble sugars raffinose and stachyose and because of these sugars, the quantity of soybean meal must be limited in rations to avoid flatulence in dogs and digestive disturbances in baby pigs and chicks (Meis et al., 2003). Soybean meal also has very low metabolizable energy compared to its gross energy and this is due to the very poor digestibility of the carbohydrate fraction. The major reason for the low metabolizable energy is the oligosaccharides, raffinose and stachyose. In addition to their indigestibility, raffinose and stachyose can produce diarrhea that may increase digesta passage rate and decrease digestion and absorption of dietary nutrients (Parsons et al., 2000). These reasons of flatulence, decreased absorption of dietary nutrients and low metabolizable energy due to normal levels of the RFOs, raffinose and stachyose, are why it is important to decrease the amount of stachyose in soybean.

***Description of Phytic Acid:***

*Myo*-inositol (1,2,3,4,5,6) hexa*kis*phosphate is ubiquitous in eukaryotic species, and is typically the most abundant inositol phosphate found in eukaryotic cells. It serves several cellular functions: as a second messenger ligand, in DNA double-strand break repair and in RNA export from the nucleus, in ATP metabolism and in phosphorous and mineral storage (Raboy, 2001). Phytic acid represents from 1% to several percentage of seed dry weight and typically is deposited in seeds as mixed phytate or phytin salts of potassium and magnesium. Phytic acid phosphorous typically represents from 65% to 85% of seed total P. It is presumed to act as a phosphate storage compound (Raboy, 2002).

Phytate accounts for a major portion of the stored reserves of phosphate and myo-inositol in legumes such as soybean. *Myo*-inositol is a precursor to compounds in plants that function not only in phosphorous storage, but also in signal transduction, stress protection, hormonal homeostasis and cell wall biosynthesis (Hegeman et al., 2001).

Five physiological roles for phytate have been suggested: 1) as a phosphorous store 2) as an energy store 3) as a source of cations 4) initiation of dormancy and 5) as a source of myo-inositol (a cell wall polysaccharide precursor) (Reddy et al., 1989).

Phytate is synthesized and deposited during seed development in discrete regions called globoids. The globoid crystals are located within the proteinaceous matrix of protein bodies. These are present in the cotyledons of legumes and not in the seed coats. Globoid crystals are structurally and chemically distinct areas within the protein bodies. The size and number will vary depending on the species. The globoids of soybeans

contain phytic acid and the minerals: potassium, magnesium and calcium (Reddy et al., 1989).

Mutant lines with greatly reduced levels of phytic acid have been described in maize (*Zea mays*) (Raboy et al., 2000), barley (*Hordeum vulgare*) (Larson et al., 1998), rice (*Oryza sativa*) (Larson et al., 2000) and soybean (*Glycine max*) (Wilcox et al., 2000; Sebastian et al., 2000). The fact that these lines were recovered and can be seed propagated suggests that a high level of stored phosphate is not a requirement for seed germination or seedling growth (Hitz et al., 2002).

When consumed in feeds and foods, phytic acid will bind to nutritionally important mineral cations that it encounters in the intestinal tract, such as calcium, iron and zinc. This is due to the fact that phytic acid is an effective chelator of positively charged cations. This phenomenon can contribute to human mineral deficiencies (Raboy, 2002).

The myo-inositol 1-phosphate synthase (MIPS) genes of maize (*Zea mays* L.) and barley (*Hordeum vulgare* L.) have been sequenced and genetically mapped by Larson and Raboy (1999). Seven loci with MIPS-homologous sequences were mapped in the maize genome. They were mapped to chromosome positions 1S, 4L, 5S, 6S, 8L, 9S and 9L. It is hypothesized that one of these MIPS loci will contain a functional MIPS gene. A single-copy barley MIPS gene was mapped to a locus on chromosome 4H (Larson and Raboy, 1999).

***Biosynthesis of Phytic Acid/Phytate:***

The first step in the synthesis of *myo*-inositol is the conversion of D-Glc-6-P to D-*myo*-inositol-3-phosphate by the isomerase D-*myo*-inositol-3-phosphate synthase (MIPS). This proposed phytic acid biosynthetic pathway also includes phosphorylation steps that are catalyzed by kinases that have not been well studied in plants (Hegeman et al., 2001).

As Raboy and Dickinson (1987) discovered, there are two distinct phases of phosphorous metabolism in developing soybean seeds. The first period being that of cell division and growth during the first three weeks and pollination which is devoted to synthesis of the phosphorous compounds needed for these processes and only a small amount of phytic acid is seen. After this period, phytic acid accumulation begins and the accumulation of other phosphorous compounds declines. A steady rate of phytic acid accumulation is maintained until very late in seed maturation (Raboy and Dickinson, 1987).

***Importance of Low Phytic Acid***

Phytate in soybean meal is largely unavailable to swine, poultry and non-ruminant animals because they do not contain the phytase enzyme in their digestive systems (Erdman, 1979). Phytate binds to nutritionally beneficial metals including zinc, calcium and magnesium, which reduce their availability (Raboy et al., 1984). Diet rations fed to poultry and swine are typically supplemented with either phytase or inorganic phosphate to increase phosphorous availability (Lei and Porres, 2003; Adeola et al., 1995). Eutrification of surface water can result from runoff and leaching of excessive phosphorous in soils to which livestock or poultry manure has been applied (Ertl et al., 1998). This shows that the excretion of phytate by non-ruminant animals is therefore potentially detrimental to the environment and supplementation of soybean meal may

exacerbate the problem. Reducing phytate and increasing inorganic phosphorous in soybean meal would increase the amount of phosphorous available to non-ruminants, which would in turn decrease the amount of supplemental inorganic phosphorous added to their ration and lower their fecal phosphorous (Cromwell et al., 2000; Cromwell, 2002).

***Soybean Lines with Unique Quality Traits:***

PI200508 and V99-5089 both have high sucrose and low stachyose contents. The former contains a high level of phytate while the latter is low phytate. High phytate lines are ones in which a very small amount of phosphorous, around 25%, is available to monogastric animals. The opposite is true of low phytate lines, which have around 75% of phosphorous available (Walker et al., 2006). PI200508 has a reduced amount of both raffinose and stachyose of around 1% DW, which results in an increased amount of sucrose (Neus et al., 2005), as does V99-5089 which has stachyose values less than 0.5% DW.

Previous work on V99-5089 at Virginia Tech has revealed unique characteristics of this line as shown in Huhn (2003). The stachyose content of $F_{2:3}$ families in the previous work by Huhn (2003), (V99-5089 x V99-5089-6180, V99-5089 x Essex, PI87013 x MFL-552), showed a more distinct bimodal segregation than either sucrose or raffinose content. The bimodal distribution suggested simple inheritance of stachyose content. The populations in the two V99-5089 crosses fit a 3 high:1 low monogenic ratio suggesting that a single diallelic gene is controlling stachyose content in these crosses. The population studied in this experiment, though, has both parents, V99-5089

and PI200508, with low stachyose content, both have stachyose contents of less than 1% DW.

***Objectives:***

The overall objective of this project was to determine whether or not PI200508 and V99-5089 have the same gene(s) for low stachyose/high sucrose. An additional objective was to identify quantitative trait loci (QTL) for soybean sugars (sucrose and stachyose) using molecular markers and sugar data collected from two consecutive years from the cross of the segregating population, V99-5089 x PI200508. Another facet of the project was to identify QTL for phytic acid in this same soybean population using the same molecular markers and phytic acid data collected from two consecutive years. Emergence tests were also conducted for one year to determine the effects of sucrose, stachyose and/or phytic acid to see if there was a correlation between any of these traits and emergence in this population.

# MATERIALS AND METHODS

## *Selection of Parents*

Seed of V99-5089 soybean grown in 2004 at Blacksburg, VA, was observed to have high sucrose (11.2%), and low raffinose (0.5%) and stachyose (0.4%) contents. It is also known to have low phytic acid. V99-5089 was derived from a cross of PI87013 and MFL-552. PI200508 is known to have low stachyose content as well; it is also known to have a high phytic acid content. PI200508 soybean seed grown in Blacksburg, VA, in 2004 was observed to have a sucrose value of 9.4%, a raffinose value of 1.3% and a stachyose value of 1%.

## *Genetic Materials and Field Work*

Genetic materials of this study included an $F_2$ population of 55 individuals resulting from crossing V99-5089 by PI200508. In 2004, analyses were done using seed generation $F_4$ and in 2005, analyses were done using seed generation $F_5$. The samples for analyses were bulk samples made up seed from of all plants left in a row after a single representative plant had been taken. Fifty-five families were planted at Blacksburg in 2004 and 2005. The parental lines, V99-5089, PI200508 and a standard check cultivar, Essex, were planted in the field along with the progeny rows. In the 2005 growing season, 75 $F_4$ seeds from each of the 55 families were planted in 5 foot rows with 2.5 feet row spacing in two replications in the same field to obtain emergence data. Emerged seedlings were counted on 5 July 2005 for these 55 $F_5$ families of V99-5089 x PI200508. The sources for these families were bulk samples from 2004.

There were three generations of seeds used for the sugar and phytate assays performed in this study. In 2004, there were $F_3$, $F_4$ and $F_5$ populations and then their corresponding generations for 2005. The $F_4$ plants (470-1 to 470-13, 471-1 to 471-10, 472-1 to 472-10, 473-1 to 473-22) were the main plants of this study and were made up 55 families.

There were 57 $F_2$ DNA samples that were used in the SSR marker analysis. They included two lines, 473-23 and 473-24 that were not continued in the field for the 55 families.

## Laboratory Procedures

### *Soluble Oligosaccharide Extraction and Analysis*

Soluble sugar extraction was based on a method developed by Cicek et al. (2006) at Virginia Polytechnic Institute and State University. The procedure involves extracting soybean meal in double distilled water and removing all components other than the soluble oligosaccharides via acetonitrile precipitation, evaporation and centrifugation (Huhn, 2003).

The seed generations for bulk row samples were as follows: $F_4$ plants in 2004 and the corresponding $F_5$ plants in 2005. It is important to note that each family came from a different $F_2$ plant and was continued from there. Bulk samples were made by combining all plants from a row for the 55 families. The bulk sample method was chosen because of time-constraint issues. Approximately 40-50 grams of seed were taken from those that had a large amount of seed available and 10-20 grams from those that only had a small

amount of seed. Large amounts of seed were ground using the C&H Laboratory Grinder Size 8 with a 1.0 mm screen and small amounts of seed were ground using the Cyclone Sample Mill with a 1.0 mm screen (Cicek, 2001).

Approximately 1 gram of a bulk ground sample of seed from each family was weighed and placed in a glass tube with 10 mL of double distilled water and vortexed. The tubes were then placed on a horizontal shaker and allowed to shake for 15 minutes. The samples then were transferred to plastic centrifuge tubes and centrifuged for 8 minutes at 1.5 rcf (4000 rpm). 5 mL of supernatant was then pipetted into a glass tube and 7 mL of 100% acetonitrile (HPLC grade) was added. The tube was then well shaken by hand and allowed to stand for one and half to two hours in order to precipitate soluble proteins. 1.25 mL of the supernatant was transferred to a 1.5 mL microcentrifuge tube and centrifuged for 17 minutes at 15.7 rcf (13,000 rpm). 1 mL of the supernatant was pipetted into a reactotherm bottle and allowed to evaporate with air at 98°C until dry. The reactotherm bottle was allowed to cool completely and then the sample was redissolved with 0.4 mL of 65% acetonitrile. The sample was then covered with plastic wrap to prevent evaporation and allowed to stand for 20 minutes. The sample was mixed well using the pipette and transferred into a 1.5 mL microcentrifuge tube. The sample was centrifuged for 6 minutes at 15.7 rcf (13,000 rpm) and then poured into a Fisher Scientific brand target DP vial for HPLC analysis and capped. For convenience, a maximum of 35 samples were extracted at a time, plus one control sample. The control sample used for all analyses was taken from a single lot of V99-5089 and used to normalize the data when a particular run was more than two standard deviation units away from the average of all other observations. This was done by averaging all the

observations of V99-5089 and bringing the significantly different values to that average by using the necessary conversion factor, which was then applied to all data from that run.

Moisture content was determined by the difference in weight of samples weighed at the time of extraction and after oven drying. Since all samples were allowed to equilibrate at the same temperature and moisture, samples (2 g) from four random ground samples were evaluated to obtain average moisture content for each run of samples.

An HPLC system utilizing an HP1047A refractive index detector operated by the March 2000 edition of Agilent Technologies ChemStation *Plus* equipped with an astec 250x4.6 mm $NH_2$ polymer column preceded by a 5 μ 10x4.6 mm guard column was used for quantification of soluble saccharides in the extracted samples. Three concentration standards were prepared from stock solutions to create a standard curve to be used for integration of the sample curves. Stock solutions were 5% solutions of sucrose, raffinose, and stachyose prepared by dissolving purified sugar powder, obtained from Sigma-Aldrich, in double-distilled water. Standard 1 contained 2.5 mg/mL sucrose, 1.25 mg/mL raffinose, and 1.25 mg/mL stachyose. Standard 2 was twice as concentrated as standard 1 and standard 3 was twice as concentrated as standard 2. Ten milliliters of stock solution were prepared at a time and stored in the refrigerator; new standards were prepared with each run of samples. Integration values (mg/mL) obtained were converted to percent dry weight (DW) by the following formula which takes into account the dilutions that occur throughout the extraction process as well as the initial sample weight and moisture:

(1) (2) (3) (4) (5)

[[ChemStation value (mg/mL) x .02 x 80 x 12 x 2] ÷ dry wt. (g) = % CHO

(1) multiply by .02 to convert mg/mL to mg of sugar injected on HPLC; 5 µl x .02=mL

(2) multiply by 80 because 10 µl of a 400 µl solution is injected on the HPLC

(3) multiply by 12 because 1 mL of the 12 mL solution is used for the remaining steps of the extraction procedure

(4) multiply by 2 because 5 mL of the initial 10 mL solution is used for extraction

(5) divide by 10 to convert mg sugar/g sample to % sugar on a dry weight basis (Huhn, 2003).

***Phytate Determination and Analysis***

Phytate was extracted using a modified colorimetric method developed by Gao et al. (unpublished). The procedure involves extracting soybean meal in hydrochloric acid, removing protein complexes with a neutral salt, then diluting the phytate extract to a detectable range, and reacting with modified Wade reagent. The details are as follows.

A 0.50 g sample of the same ground soybean seed used for the sugar assay, was weighed and placed into a 15 mL Falcon centrifuge tube. Ten mL of 0.65 M (2.4%) HCl was added to the sample and vortexed until the sample went into solution. This solution was shaken overnight (16 hours) in a reciprocal shaker. The tube was then centrifuged at maximum speed in a Sorvall RT6000B centrifuge at 10°C for 20 minutes and the upper phase was poured into another 15 mL Falcon tube containing 1 g of NaCl. The tube was vortexed until the NaCl was dissolved into the solution and allowed to precipitate for 30 minutes at 4°C; the sample was then centrifuged at maximum speed at 10°C for 20

minutes. One mL of supernatant was diluted 25 times with double distilled water in a 50 mL falcon tube. Three mL of the diluted sample was taken with an electronic pipette into a 15 mL Falcon tube and 1 mL of Wade reagent (0.03% $FeCl_3$-$6H_20$ + 0.3% sulfosalicylic acid) was added using a 1 mL fixed volume automatic pipette. The tube was vortexed and centrifuged at maximum speed at 10°C for 10 minutes.

Seven calibration standards were prepared from a stock of 25% NaCl and sodium phytate from Sigma (560 ppm phytic acid-P). The stock was prepared by dissolving 25 g of NaCl in 100 mL water. The phytic acid stock was made using 0.3047 g phytic acid dissolved in double distilled water to make 100 mL. The stock was then diluted 10 times to make a 56 ppm phytic acid-P working stock. The seven calibration standards were made by pipetting 0, 1, 2, 3, 5, 7 and 10 mL of the 56 ppm phytic acid-P working stock into seven different 50 mL volumetric flask and adding 1 mL of 25% NaCl to each flask and then adding double distilled water to 50 mL. The calibration standards 0 to 6 contain, 0 ppm, 1.12 ppm, 2.24 ppm, 3.36 ppm, 5.6 ppm, 7.84 ppm and 11.2 ppm phytic acid-P, respectively. These standards were mixed with the Wade reagent as, 3 mL standard to 1 mL Wade reagent.

Both samples and standards were read on a Beakman Coulter DU 640 spectrophotometer using a wavelength of 500 nm. Double distilled water was used to zero the machine after every ten samples (Gao et al., unpublished).

Moisture content was determined by the difference in weight of samples weighed at the time of extraction and after oven drying. Since all samples were allowed to equilibrate at the same temperature and moisture, samples (2 g) from four random ground

samples were evaluated to obtain average moisture content for each run of samples (Huhn, 2003).

The sample phytic acid-P (Y ppm) was calculated according to the standard curves decrease in absorbance at 500 nm against the phytate concentration.

The final phytic acid content of each sample:

= Y (ppm) x dilution factor x 10 ml / Dry weight (g)

= μg P/g

OR

= ($10^{-3}$ x Y (ppm) x dilution factor x 10 ml / Dry weight (g))/0.2816

= mg PA/g, where 0.2816 is the conversion coefficient between phytic acid-P and phytic acid.

***DNA Extraction***

Young leaves of field-grown $F_2$ plants were collected for DNA extraction in Blacksburg in July 2003. The tissue was transported on dry ice and stored at -80°C until extraction. It is important to note that there is DNA from 57 $F_2$ individuals and only 55 of the families were continued after 2003. The DNA from the 57 individuals is what was used for marker analysis in Mapmaker.

DNA from each $F_2$ plant and the parental lines was extracted using a modified version of the protocol described previously by Saghai Maroof et al. (1984). The tissue was ground in a mortar and pestle in liquid nitrogen with CTAB extraction buffer (50 mM Tris, 0.7 M NaCl, 10 mM EDTA, 1% hexadecyltrimethylammonium bromide, 0.1% 2-mercaptoethanol), and incubated at 65°C for 1 hour in a shaker bath. An equal volume

of chloroform was then added to the samples and allowed to shake for an hour. The samples were then centrifuged at 3.3 rcf (6000 rpm) for 15 minutes. The supernatant was transferred to a new tube and two-thirds of that volume in ice-cold isopropanol was added. A DNA pellet was precipitated by inverting the tube; the pellet was hooked with a glass rod and transferred to a glass tube containing 20 mL 76% EtOH/10 mM $NH_4Ac$. After standing overnight in the solution the DNA pellet was air-dried and transferred to a new microcentrifuge tube containing 10 mM $NH_40AC$/0.25 mM EDTA. After the DNA was dissolved in the solution a fluorometer reading was taken for each of the samples to measure the DNA concentration (Saghai Maroof et al., 1984).

***Simple Sequence Repeat (SSR) and Molecular Marker Analysis***

SSR procedures were as described by Yu et al. (1994). Briefly, a 10 µl PCR reaction contained 50 ng of genomic $F_2$ DNA, 0.1 mM of each primer, 10x reaction buffer, 3 mM $MgCl_2$, 200 mM each of dATP, dGTP, and dTTP, 5 mM dCTP, 1.0 U of AmpliTaq DNA polymerase (Perkin-Elmer Cetus, Norwalk, CT), and 0.05 µCi [$\alpha$-$^{32}$P] dCTP. The reaction was denatured at 94°C for 3 min, followed by 36 cycles at 94°C for 30 sec, 47°C for 30 sec, 68°C for 1 min. and with a final extension step of 68°C for 7 min. PCR products were denatured for 8 min at 94°C after loading buffer (98% deionized formamide, 10 mM EDTA (pH 8.0), 0.25% xylene cyanol FF, 0.25% bromophenol blue) was added. Five microliters of each sample was loaded on a polyacrylamide denaturing gel and separated at 1500-V constant power in 1x TBE (Tris-Borate-EDTA) running buffer, using a DNA sequencing unit (BioMax STS 45i). Gels

were immediately covered with plastic wrap and exposed to X-ray film for an hour to overnight depending on the intensity of the radioactivity (Yu et al., 1994).

A large set of SSR markers were tested using parental DNA and a total of 146 polymorphic markers representing 20 MLGs were identified and used for construction of a map based on the V99-5089 x PI200508 population.

***Map Construction and QTL analysis***

The program Mapmaker 3.1 (Lander, 1987) was used to create a map based on the V99-5089 x PI200508 population. The default settings of LOD 3.0 and maximum distance of 50 centimorgans (cM) were used to create a skeletal map. When these settings did not bring together groups that were known to go together, i.e. were less than 35 cM apart on the public soybean map (Cregan et al., 1999), then the default settings were lowered to LODs of 2.5 or 2.0, as needed. This lowering of LOD had the effect of lessening the stringency and allowing more markers to link together.

The program QTL Cartographer 2.5 (Wang et al., 2005) was used to map where the QTL were located for sucrose, stachyose and phytate in this population. This is a statistical program that maps quantitative traits using a map of molecular markers that is obtained from Mapmaker 3.1 (Lander et al., 1987). An LOD of 2.5 was set in this program to determine QTL that were significant for sucrose, stachyose or phytate in this population for either year.

***Statistical Methods***

Statistical analyses were done using Minitab 14 (Minitab Inc, 2003). One way analysis of variance (ANOVA) was performed for the traits: sucrose, stachyose and phytate against corresponding marker data for the 55 families to see which markers were significant at p=0.05. Tukey's honestly significant difference (HSD) test set at p=0.05, was used to test all possible two way comparisons among the three possible genotypes for each locus.

Chi-square analysis was done for each marker to determine goodness of fit to monogenic ratios. This was determined using a chi-square analysis provided by QTL Cartographer (Wang et al., 2005) and using a p-value set at 0.05.

R-square analysis was determined to see which markers were explaining a large amount of phenotypic variation for sucrose, stachyose and phytate content in this population. These values were obtained from QTL cartographer (Wang et al., 2005).

# RESULTS

## *Emergence Test*

Average stand counts are shown in Table 1. One thing to note is that the $F_5$ plants were grown in a field in duplicate replications so that an average could be determined. Analysis of variance was performed on emergence data and it was determined that there were significant differences among entries for stachyose with a p-value of 0.02 but it showed a weak correlation to emergence with a correlation coefficient of 0.32. Phytate also has a significant effect on emergence with a p-value of less than 0.01, the low correlation coefficient of 0.36 indicates lack of a strong association. Twenty-six out of the 55 lines had better than average emergence of 56.8%.

## *Map Construction*

Markers were chosen based upon several QTL for sugar traits on various linkage groups previously identified in other populations by Maughan et al. (2000), Cicek (2001) and Maroof et al. (unpublished). Additional markers were tested from the same linkage groups to see if they would show significant associations in this population as well and identify any quantitative trait loci for sucrose, stachyose or phytate. Markers from other linkage groups were chosen based upon their polymorphism in the population, so that a foundation of markers could be used to construct a skeletal map to represent all of the linkage groups. This was done since this was the first time that a map was being constructed for this population.

One hundred forty-one polymorphic SSR markers were mapped in the V99-5089 x PI200508 population using $F_2$ DNA of 57 individuals. These markers helped to construct a map covering all twenty MLGs of soybean (Figure 1). Using Mapmaker 3.1 (Lander et al., 1987) set at an LOD of 3.0, 2.5 or 2.0, it was determined that 15 single markers remained unlinked, which could be due to being too many centimorgans away from the mapped markers. Linkage groups A2bottom, D1b, Mtop and N were constructed using an LOD of 2.0. Linkage groups B1 and Fmiddle were done using an LOD of 2.5. All other linkage groups were constructed using an LOD set at 3.0. The different LODs of 2.0 and 2.5 were used to bring markers together that were not linking together at the default LOD of 3.0. These lower LODs are less stringent and allow markers to become associated with linkage groups that were not linking together at the higher LOD. These less stringent LODs also inflate the distances between markers since they are less stringent. Other markers were mapped together with another marker in a group of two, a doublet. Doublets are not put on linkage groups in Mapmaker, though it was attempted to bring these small linkage groups together with their respective MLG by mapping strategic markers but this did not always help and so it was decided to lower the LOD in Mapmaker to bring these groups together. The linkage groups were compared to the public soybean genetic map (Cregan et al., 1999) and were found to be in the same order.

***Distribution of Sugar Traits***

The average parental values for sucrose, stachyose and phytate for 2004 and 2005 may be seen in Table 2. Scatter plots comparing the values for the 55 families for

sucrose and stachyose in 2004 and 2005 may be seen in Figures 2 and 3. Frequency distributions of all three traits for 2004 and 2005 may also be seen in Figures 4-7.

The average stachyose value for the parent, V99-5089, was 0.4% DW for the year, 2004 (Table 2). In the $F_4$ population of 55 families, three had values lower than 0.4% DW (5.5%) and 52 progeny had values above 0.4% DW (94.5%). The average stachyose value for PI200508 was 1.0% DW. The $F_4$ population had 17 out of 55 families at or below 1% DW (30.9%) and 38 above (69.1%).

In 2005, the average stachyose value for the parent, V99-5089, was 0.26% DW for 2005. The population had 5 families that were at or below 0.26% DW (9.1%) and 50 above (90.9%). The average stachyose value for PI200508 was 0.81% DW for 2005. The population had 15 families that were at or below 0.81% DW (27.3%) and 40 families that were above (72.7%).

***Phytate Results***

V99-5089 had an average phytate phosphorous reading of 10.65 mg/g using values obtained in the summer of 2005. PI200508 had an average reading of 18 mg/g phytate phosphorous for 2004 and 2005 (Table 2). A range of 9.05 mg/g to 15.09 mg/g of phytate phosphorous was observed in the progeny of V99-5089 by PI200508 in 2004. The observed range in 2005 was 9.95 mg/g to 18.6 mg/g. A scatter plot comparing the values from 2004 and 2005 may be seen in Figure 8.

V99-5089 had a phytate range of 9.46 to 12.69 mg/g for the year 2005 and PI200508 had a range of 15.11 to 20.16 mg/g for 2004 and 2005. These ranges of values were used to determine the number of individuals in the segregating population that had

phytate values close to V99-5089, intermediate between V99-5089 and PI200508 and those that had values close to PI200508. Using a scale of 15 mg/g or more for high phytate or those that are similar to PI200508 since the lowest PI200508 value was 15.11 mg/g, values of 14.99 to 13.01 mg/g were considered to be intermediate phytate values and those with values of 13.00 mg/g or less were considered to be low phytate or similar to V99-5089 since it's highest phytate value was 12.69 mg/g. If these values are used as a guideline and averages of the 55 families from both years are used, then 5 families out of the 55 are considered high (9.1%). Forty-three out of the 55 families are considered to have intermediate values (78.2%). Only 7 out of the 55 families were considered to have low phytate values or those similar to V99-5089 (12.7%). A frequency distribution of the 55 families in 2004 and 2005 may be seen in Figures 9 and 10.

***Quantitative Trait Loci Mapping***

Significant QTL for sucrose were found on MLGs: B1, C2 and D1b. Stachyose QTL were found on MLGs: B1, C2 and M. Significant QTL for phytate were found on four MLGs including: B1, D1b, F and N. It should be noted that a set LOD of 2.5 was used to distinguish QTL for sucrose, stachyose and phytate in both years.

Figures 11 - 14 display the LOD likelihood curves for all markers that detected QTL at least in one year, for at least one sugar trait. Refer to Figure 1 to see marker placement for each linkage group. Figure 11 for MLG B1 has an LOD of 4.2 for 2005 sucrose between BE806308 and Sat270. It also has an LOD of 3.4 for 2004 stachyose between Satt251 and Satt197. Figure 12 for MLG C2top has two peaks; the first peak is

for 2004 stachyose with an LOD of 4.0 between Sat62 and Satt422 and the second is for 2005 stachyose with an LOD of 3.4 between Sat62 and Satt281. Figure 13 for MLG C2middle has 4 peaks. 2004 and 2005 stachyose both have an LOD of 8.0 and 7.2, respectively between Satt643 and Satt316; 2005 sucrose also has a smaller LOD of 3.6 between those same markers. 2004 sucrose has an LOD of 3.2 between Satt643 and Satt277. Figure 14 for MLG Mtop has two peaks; the first is at an LOD of 3.1 for 2005 stachyose and is located between Satt540 and Satt435. The second peak is at an LOD of 3.5 for 2004 stachyose and is also located between Satt540 and Satt435.

QTL Cartographer results for phytate for the years 2004 and 2005 can be seen in Figures 15 - 18. Figure 15 for B1 has LODs of 6 and 4.4 for 2004 and 2005, respectively, between Sct26 and Sat331. Figure 16 for MLG D1b has an LOD of 4.5 for 2004 phytate between Satt296 and AC13959.1. Figure 17 for MLG Fmiddle has an LOD of 7 for 2005 and an LOD of 6.4 for 2004 between Satt425 and Satt663. Figure 18 for MLG N has 4 peaks. 2005 phytate has an LOD of 5.6 between Sat33 and Satt549, an LOD of 7.0 between Satt237 and Sat91. 2004 phytate has an LOD of 3.2 between Sat33 and Satt549 and an LOD of 2.6 between Satt237 and Sat91.

## Statistical Analysis

### *One-way Analysis of Variance (ANOVA) and Quality Traits*

Analysis of variance found that sucrose and stachyose yielded 27 markers total that were significant at a p-value set at 0.05 for 2004. Twelve of these markers were from the sucrose trait and came from MLGs: B1, B2, C2, D1b, F, G, J, K and N. The last 15 markers came from stachyose and were located on MLGs: B1, C1, C2, F, K, M, N

and O. One-way ANOVA analysis was done for the 2005 sugar data and the results were as follows. Sucrose yielded 8 markers that were significant on MLGs: B1, C2, F, G, K, and N. There were 8 markers found to be significant for stachyose on MLGs: B1, B2, C1, C2, F, G and M. ANOVA analysis was also done for the two-year average of both sucrose and stachyose. This analysis found that for sucrose 11 markers out of the 20 markers found significant for either 2004 or 2005 were also significant using the 2-year average. Stachyose had 10 out of 23 markers that were significant using the 2-year average for stachyose. These values may be seen in Table 3.

One-way ANOVA analysis was also performed against all SSR markers tested and the phytate phosphorous data for both 2004 and 2005. The 2004 phytate data yielded 18 markers that were significant at a p-value set at 0.05 (Table 4). These markers were located on MLGs: A1, A2, B1, C2, D1b, E, H, I, J, K, M and N. The 2005 data yielded 13 markers that were significant on MLGs: A2, B1, C2, D1b, G, H, K, M and N.

***Chi-square analysis***

Most markers mapped to a linkage group were found to be segregating normally for an $F_2$ population, in a 1(AA):2(Aa):1(aa) manner. The following markers were found not to be segregating in a normal manner: Satt207 (MLG A2), BE820148 (MLG A2), BE806308 (MLG B1), Satt726 (MLG B2), Satt510 (MLG F), Satt496 (MLG I) and Satt518 (MLG K). This was determined using a chi-square analysis provided by QTL Cartographer (Wang et al., 2005) and determining that all were segregating normally if the $\chi^2$ value was below 5.991 with a p-value set at 0.05.

***R-square analysis***

R-square values were determined for QTL by using the output given by QTL Cartographer (Wang et al., 2005). These values may be seen in Table 5.

**Table 1 Emergence percentage of the parental lines, Essex and 55 families in 2005**

| Name | Emergence Percentage | Name | Emergence Percentage |
|---|---|---|---|
| V99-5089 | 65.4 | 473-6 | 55.7 |
| PI200508 | 83.4 | 473-7 | 62.9 |
| Essex | 74.0 | 473-8 | 52.1 |
| 470-1 | 57.9 | 473-9 | 57.2 |
| 470-2 | 77.9 | 473-10 | 57.2 |
| 470-3 | 60.5 | 473-11 | 48.6 |
| 470-4 | 59.7 | 473-12 | 44.3 |
| 470-5 | 68.6 | 473-13 | 38.6 |
| 470-6 | 78.6 | 473-14 | 47.9 |
| 470-7 | 62.9 | 473-15 | 50.0 |
| 470-8 | 51.4 | 473-16 | 50.0 |
| 470-9 | 47.9 | 473-17 | 50.0 |
| 470-10 | 65.8 | 473-18 | 56.3 |
| 470-11 | 66.2 | 473-19 | 47.9 |
| 470-12 | 62.9 | 473-20 | 39.3 |
| 470-13 | 69.3 | 473-21 | 50.0 |
| 471-1 | 62.1 | 473-22 | 52.6 |
| 471-2 | 50.0 | | |
| 471-3 | 69.3 | | |
| 471-4 | 80.0 | | |
| 471-5 | 55.8 | | |
| 471-6 | 62.1 | | |
| 471-7 | 58.6 | | |
| 471-8 | 51.5 | | |
| 471-9 | 52.9 | | |
| 471-10 | 65.3 | | |
| 472-1 | 50.0 | | |
| 472-2 | 35.1 | | |
| 472-3 | 57.2 | | |
| 472-4 | 47.9 | | |
| 472-5 | 50.0 | | |
| 472-6 | 47.2 | | |
| 472-7 | 65.7 | | |
| 472-8 | 65.7 | | |
| 472-9 | 57.8 | | |
| 472-10 | 48.1 | | |
| 473-1 | 60.7 | | |
| 473-2 | 55.2 | | |
| 473-3 | 58.6 | | |
| 473-4 | 42.2 | | |
| 473-5 | 56.5 | | |

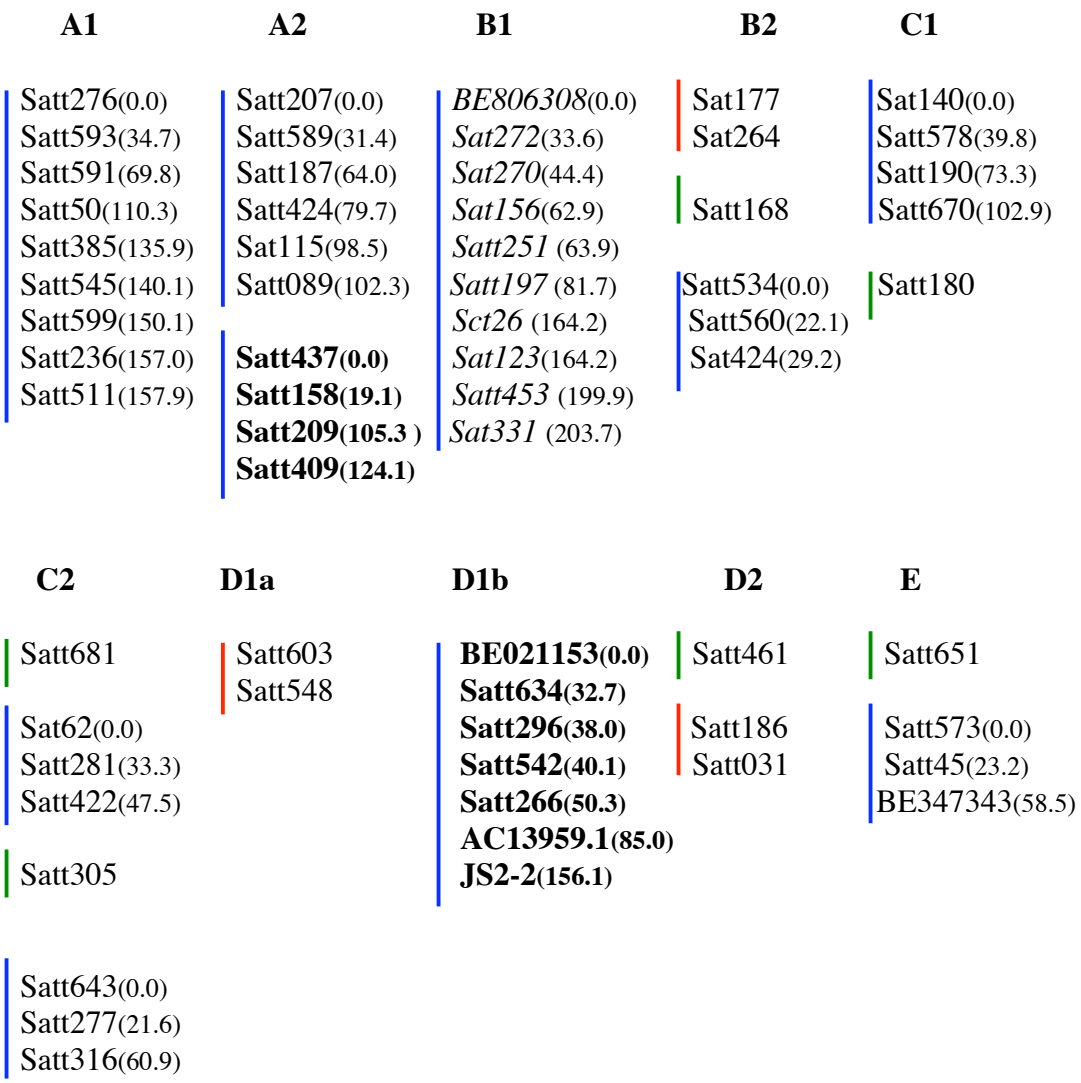


**Figure 1 : Soybean molecular linkage map based on $F_2$ DNA from the population of V99-5089 x PI200508. MLGs are labeled as their respective linkage group IDs unless it is fragmented then it is labeled as top, middle or bottom depending on the markers. For each MLG, markers are listed on the right side of the bar. Numbers in parentheses are distances from the top of each MLG in centimorgans. Those in bold were done at an LOD of 2.0.Those in italics were done at an LOD of 2.5; all others were done at a default of LOD 3.0. Interruptions between groups of a MLG are to signify groups that are not linking together. Red lines signify doublets, green single markers and blue are groups of 3 or more.**

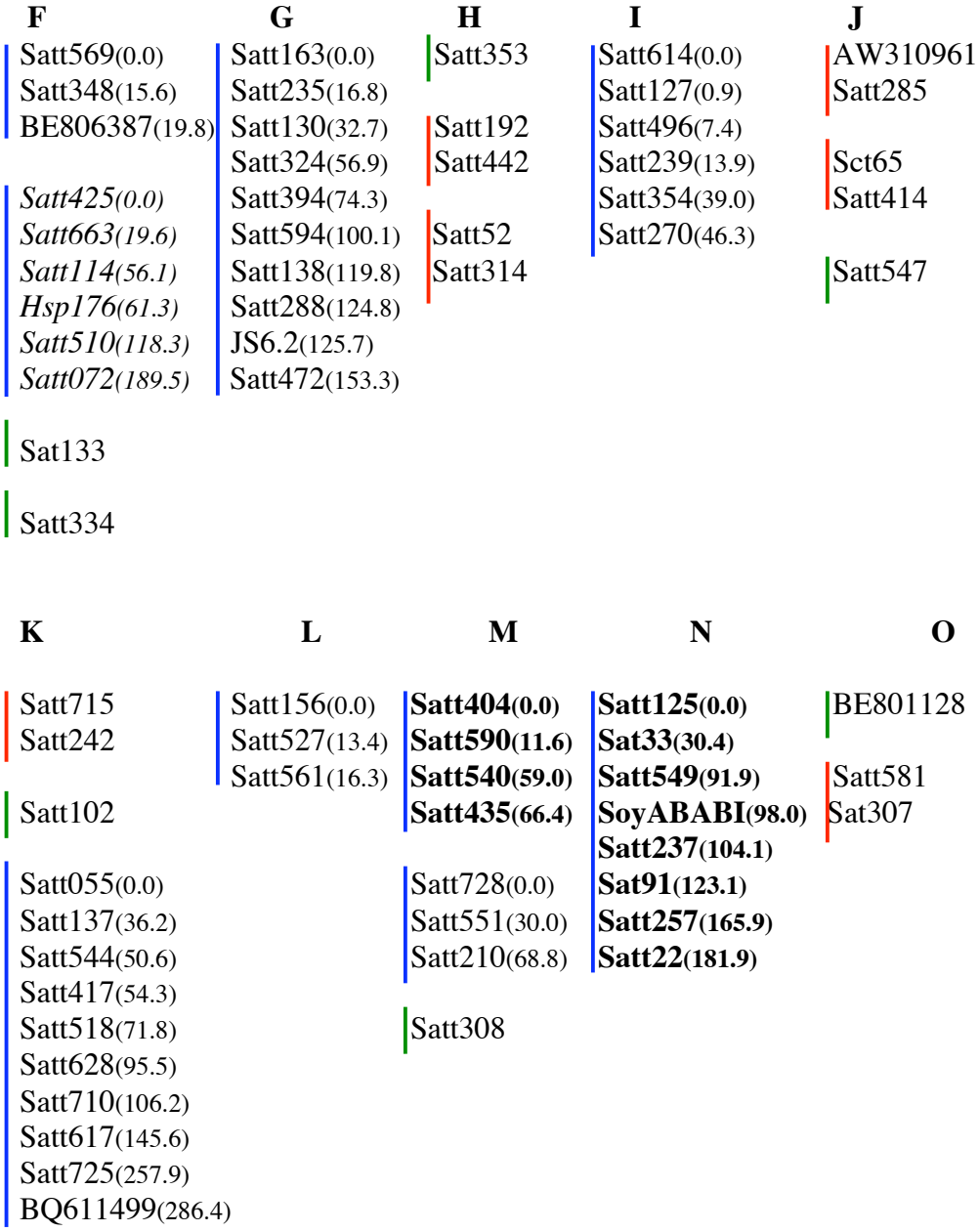


**Figure 1 continued: Soybean molecular linkage map based on $F_2$ DNA from the population of V99-5089 x PI200508. MLGs are labeled as their respective linkage group IDs unless it is fragmented then it is labeled as top, middle or bottom depending on the markers. For each MLG, markers are listed on the right side of the bar. Numbers in parentheses are distances from the top of each MLG in centimorgans. Those in bold were done at an LOD of 2.0.Those in italics were done at an LOD of 2.5; all others were done at a default of LOD 3.0. Interruptions between groups of a MLG are to signify groups that are not linking together. Red lines signify doublets, green single markers and blue are groups of 3 or more.**

**Table 2 Average Sugar and Phytate Trait Values and the corresponding standard deviations for V99-5089 and PI200508 in 2004 and 2005**

| Line | 2004 Sucrose (% DW) | 2005 Sucrose (% DW) | 2004 Stachyose (%DW) | 2005 Stachyose (%DW) | 2004 Phytate (mg/g) | 2005 Phytate (mg/g) |
|---|---|---|---|---|---|---|
| V99-5089 | 11.02+/- 0.77 | 12.56 +/- 2.39 | 0.40+/- 0.07 | 0.26+/- 0.07 | - | 10.65+/- 1.02 |
| PI200508 | 9.40 +/- 0.62 | 9.30 +/- 0.29 | 1.00+/- 0.09 | 0.81+/- 0.22 | - | 18+/- 2.29 |

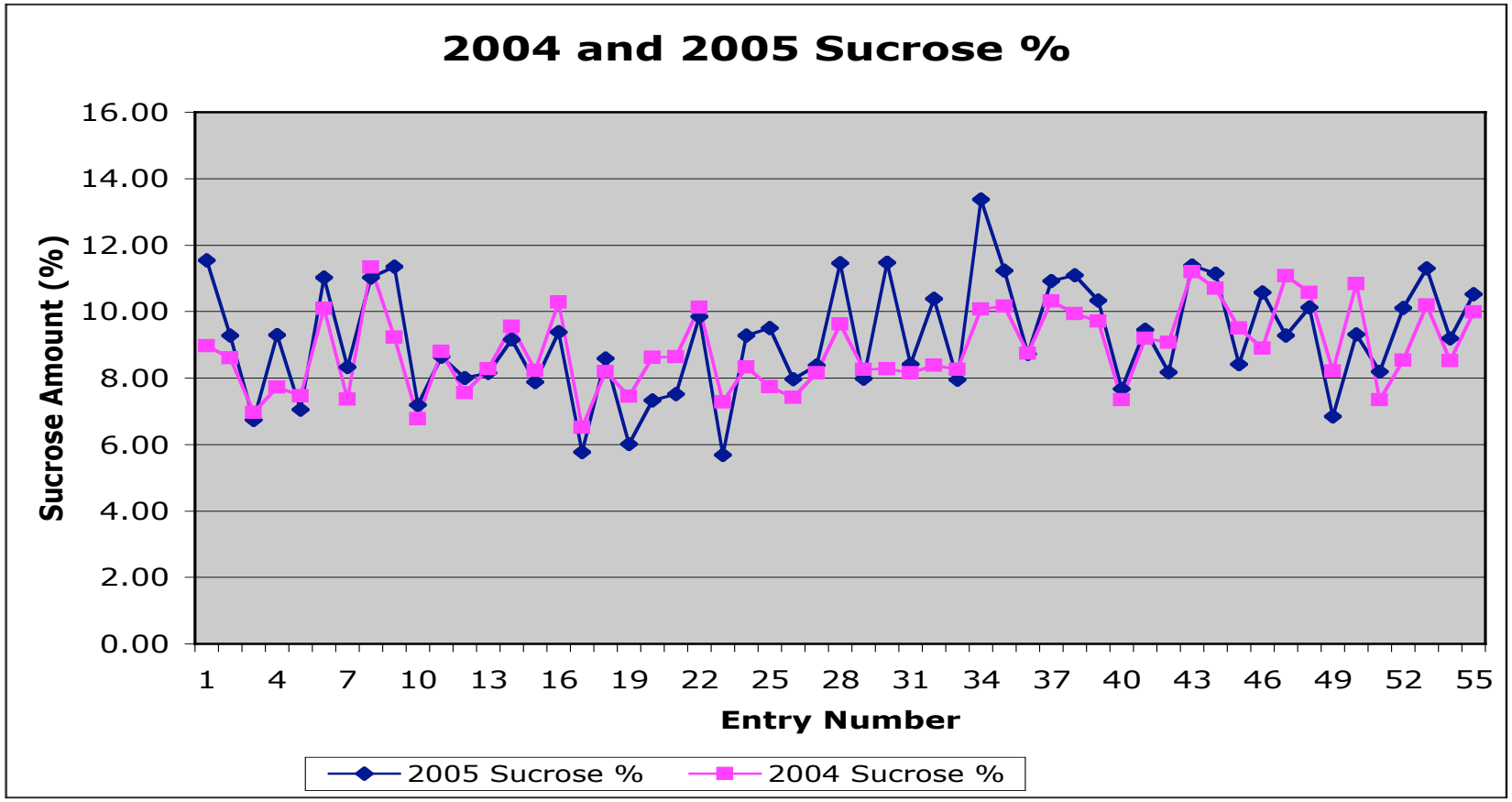


**Figure 2: A connected scatterplot of the 55 $F_5$ families from 2005 and 2004 and the respective sucrose percentage for that year.**

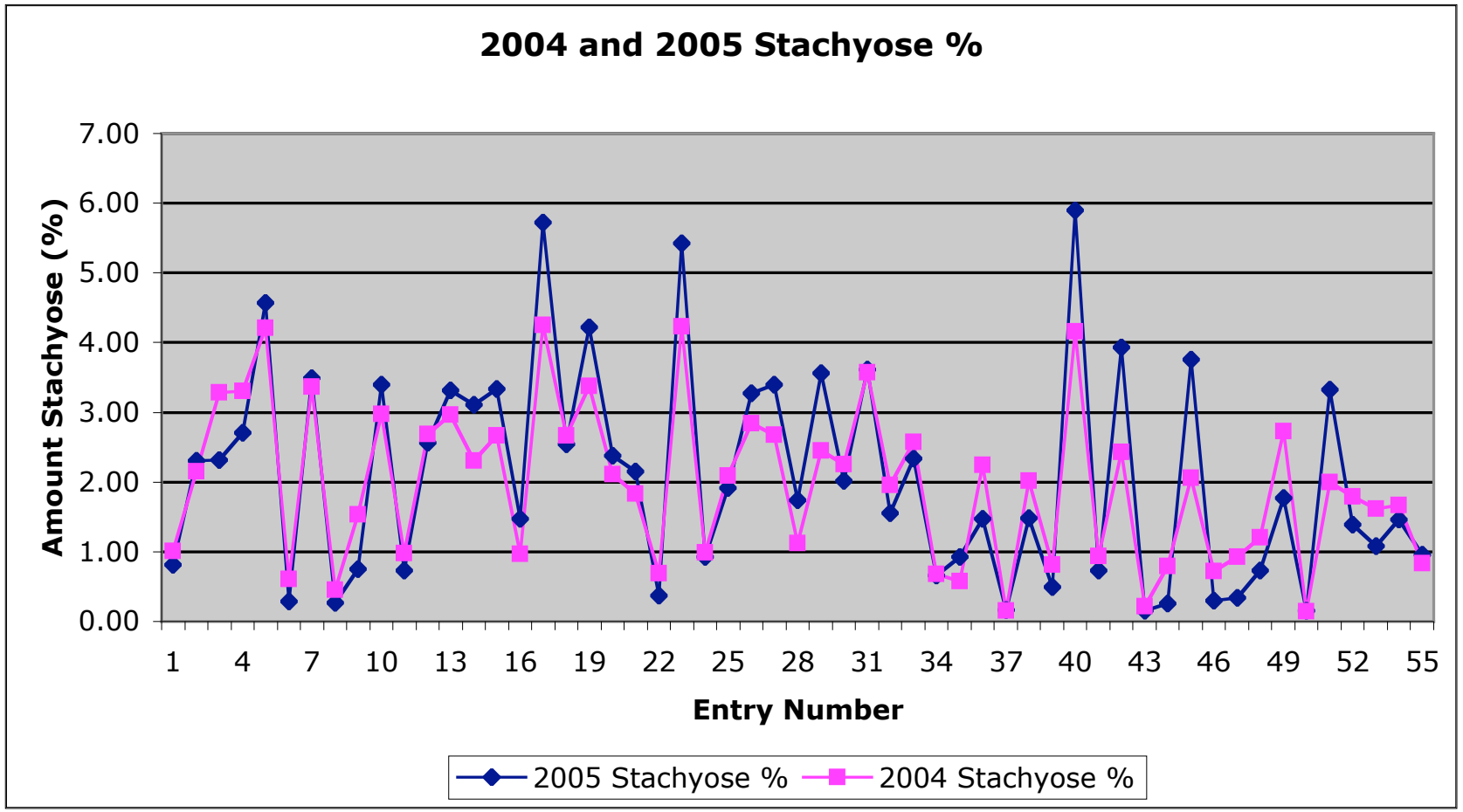


**Figure 3: A connected scatterplot of the 55 $F_5$ families from 2005 and 2004 and the respective stachyose percentage for that year.**

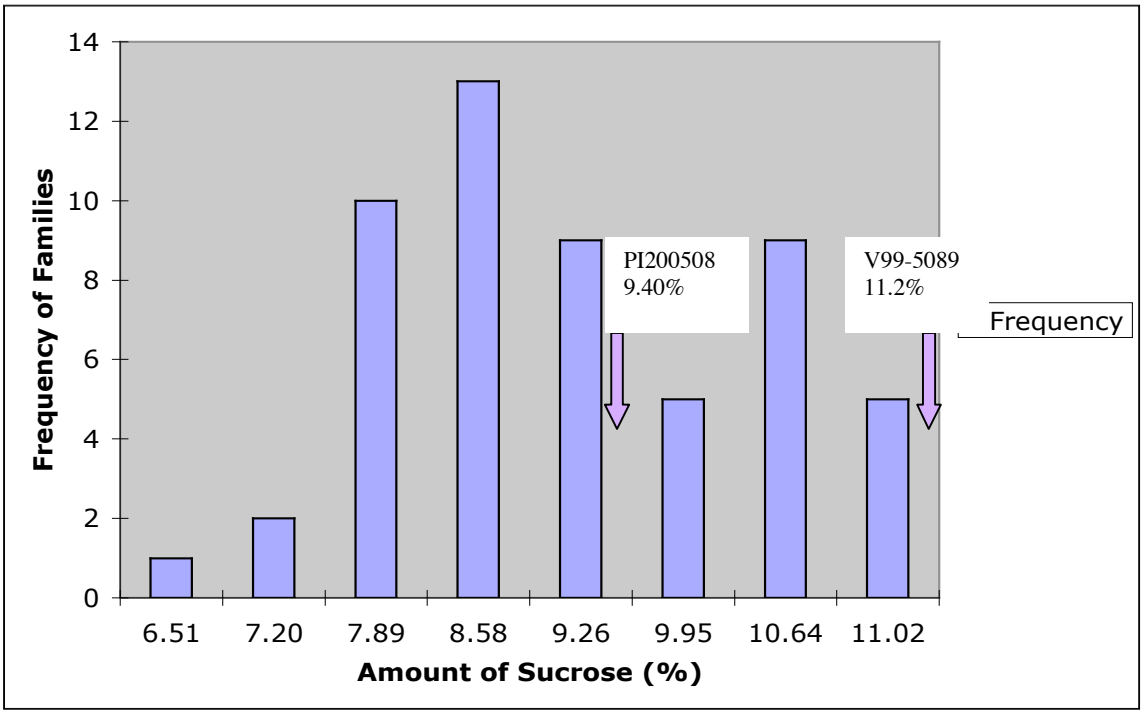


**Figure 4: Frequency distribution of sucrose in 2004 for the 55 $F_4$ families.**

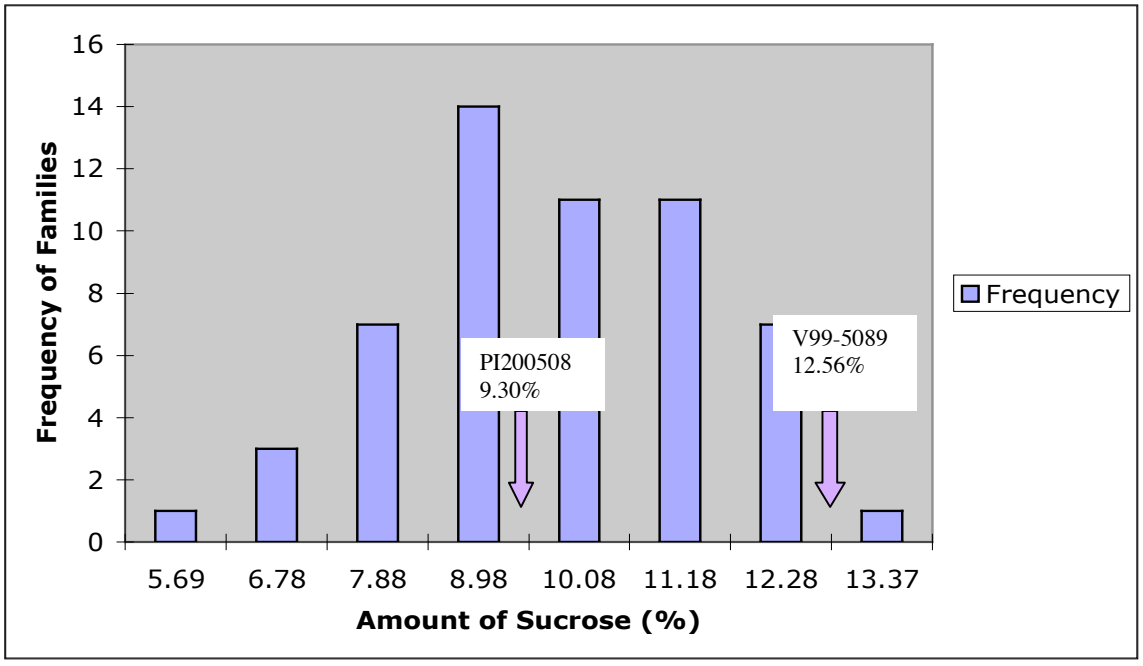


**Figure 5: Frequency distribution of sucrose in 2005 for the 55 $F_5$ families.**

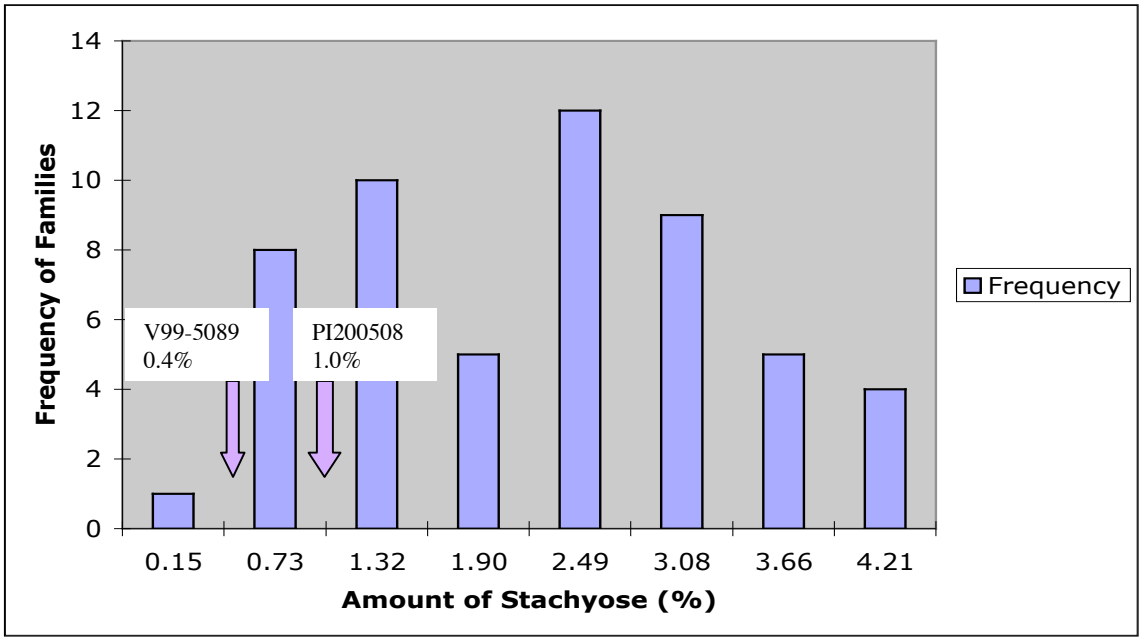


**Figure 6: Frequency distribution for stachyose in 2004 for the 55 $F_4$ families.**

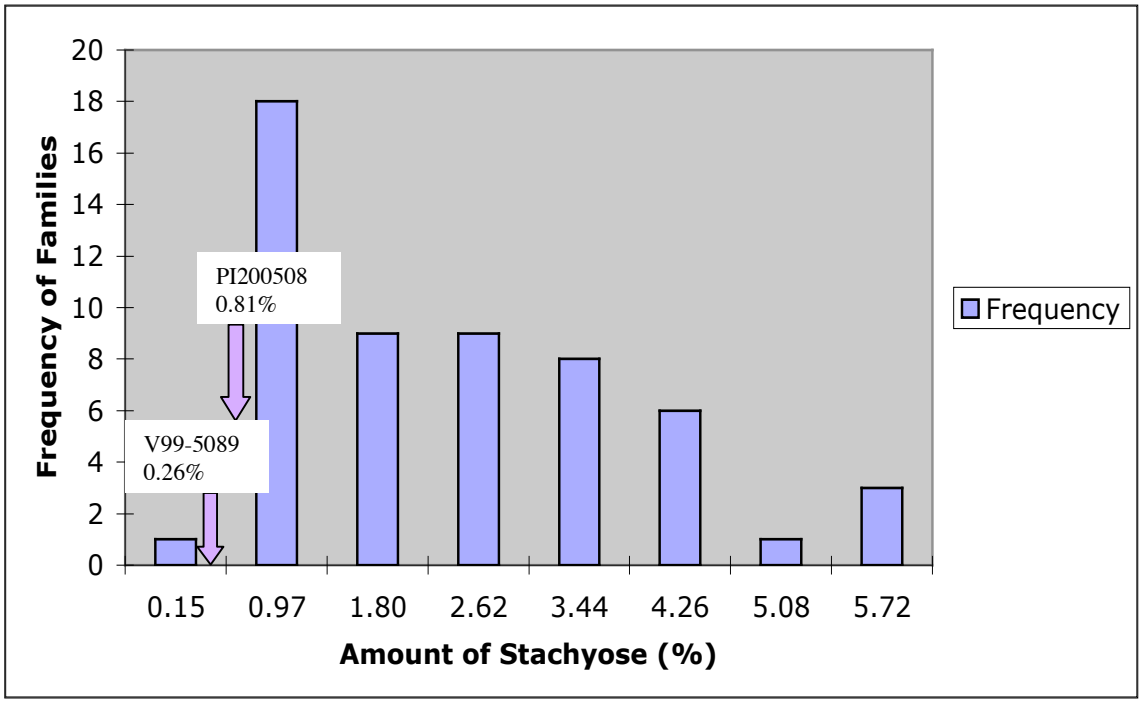


**Figure 7: Frequency distribution for stachyose in 2005 for the 55 $F_5$ families.**

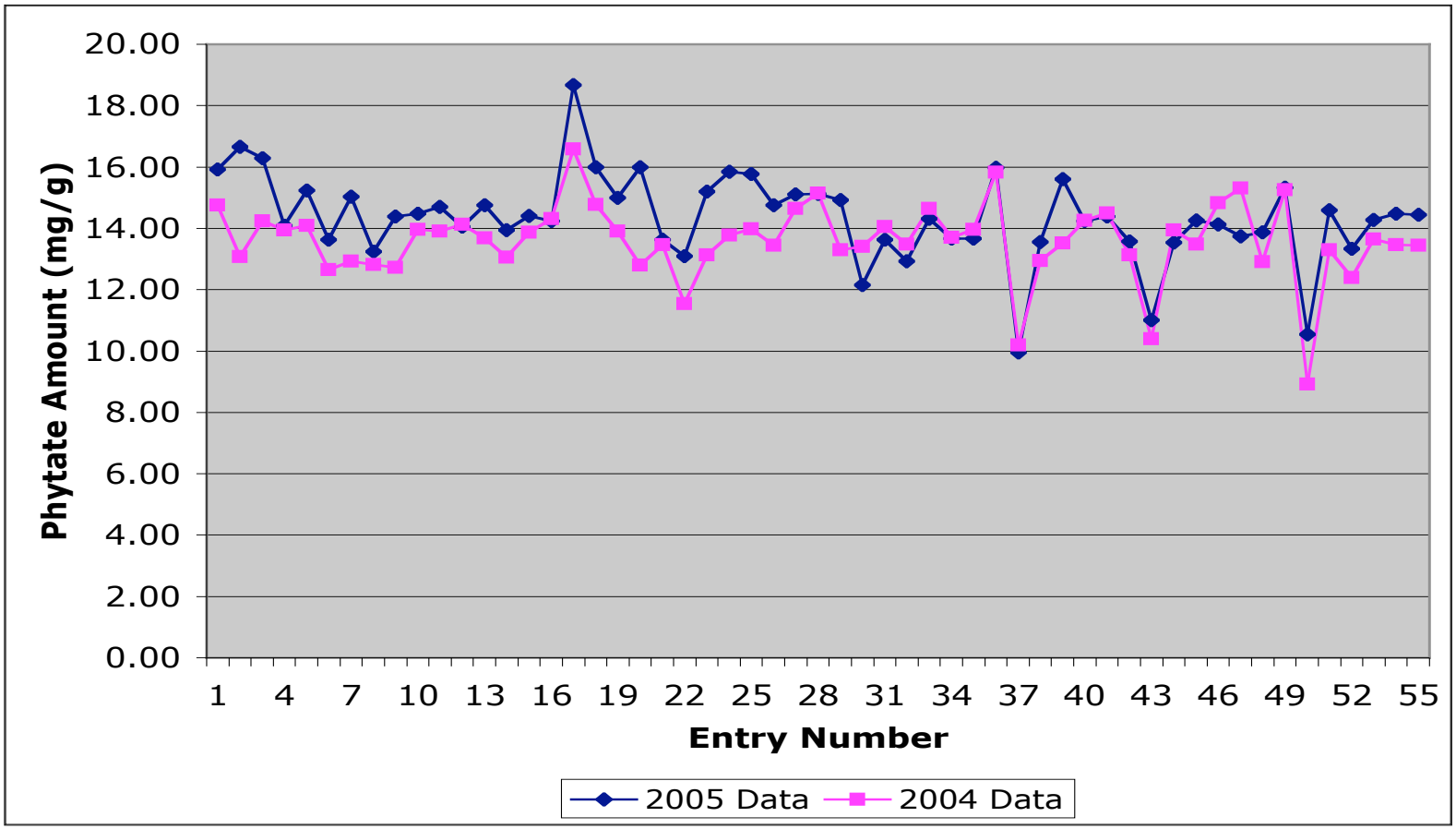


**Figure 8: Comparison of phytate amounts for the years 2004 and 2005 for the 55 families.**

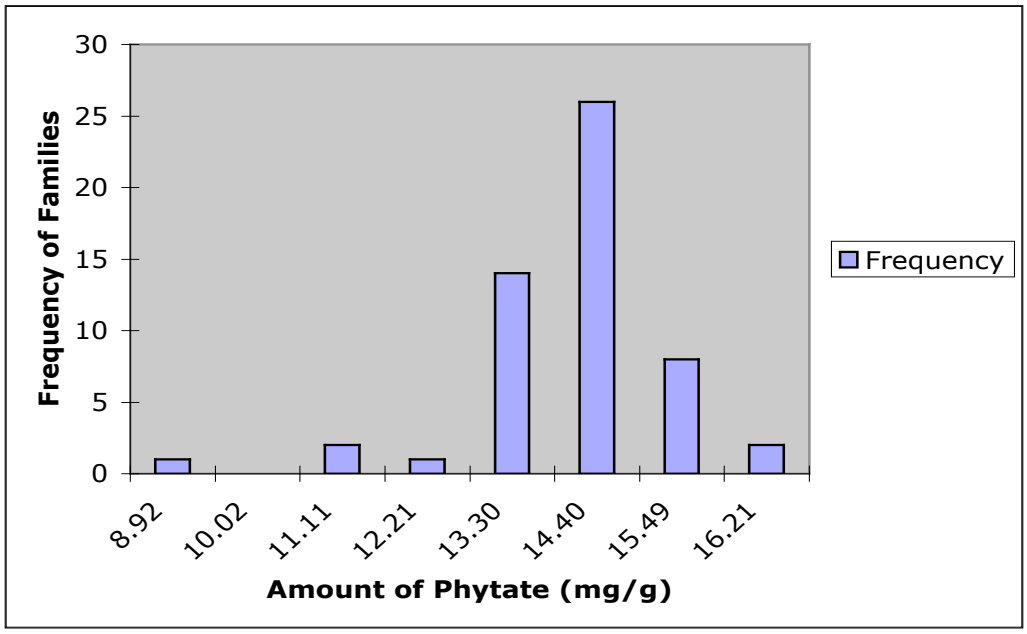


**Figure 9: Frequency distribution of phytate for the 55 families in 2004.**

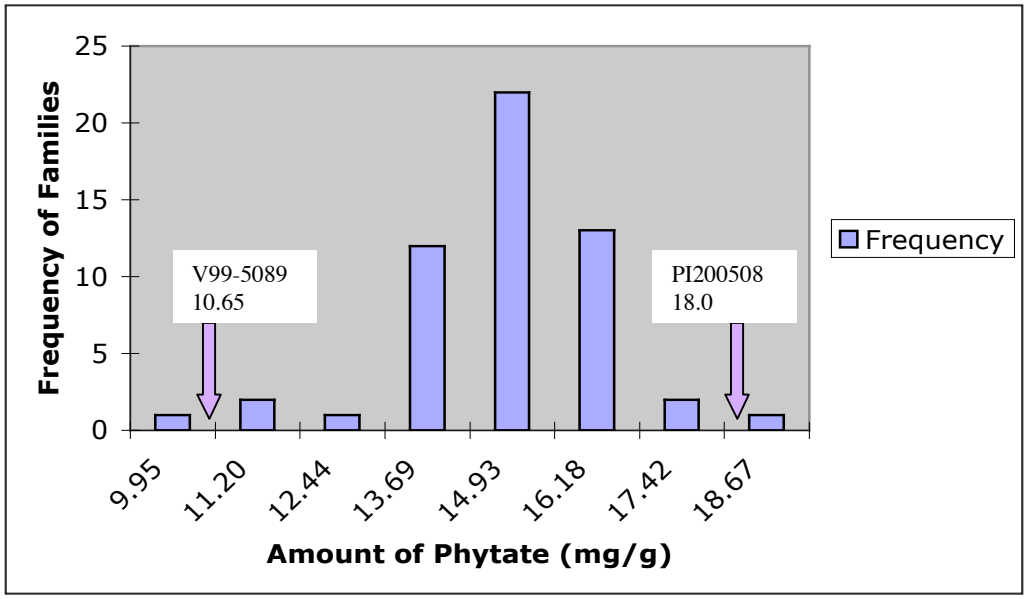


**Figure 10: Frequency distribution of phytate for the 55 families in 2005.**

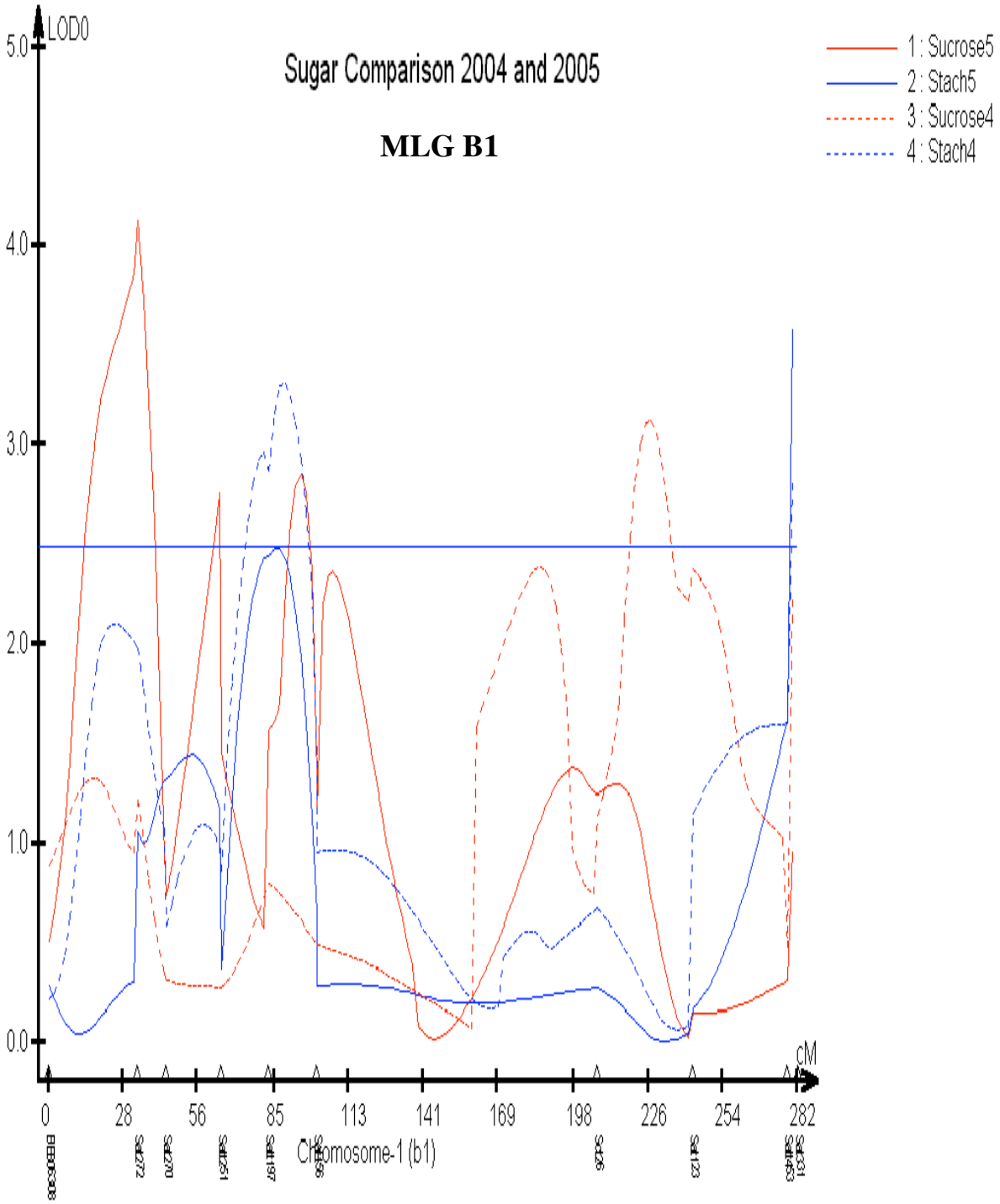


**Figure 11: QTL analysis of sucrose and stachyose for 2004 and 2005 on MLG B1. The line horizontal to the linkage group indicates the set LOD of 2.5.**

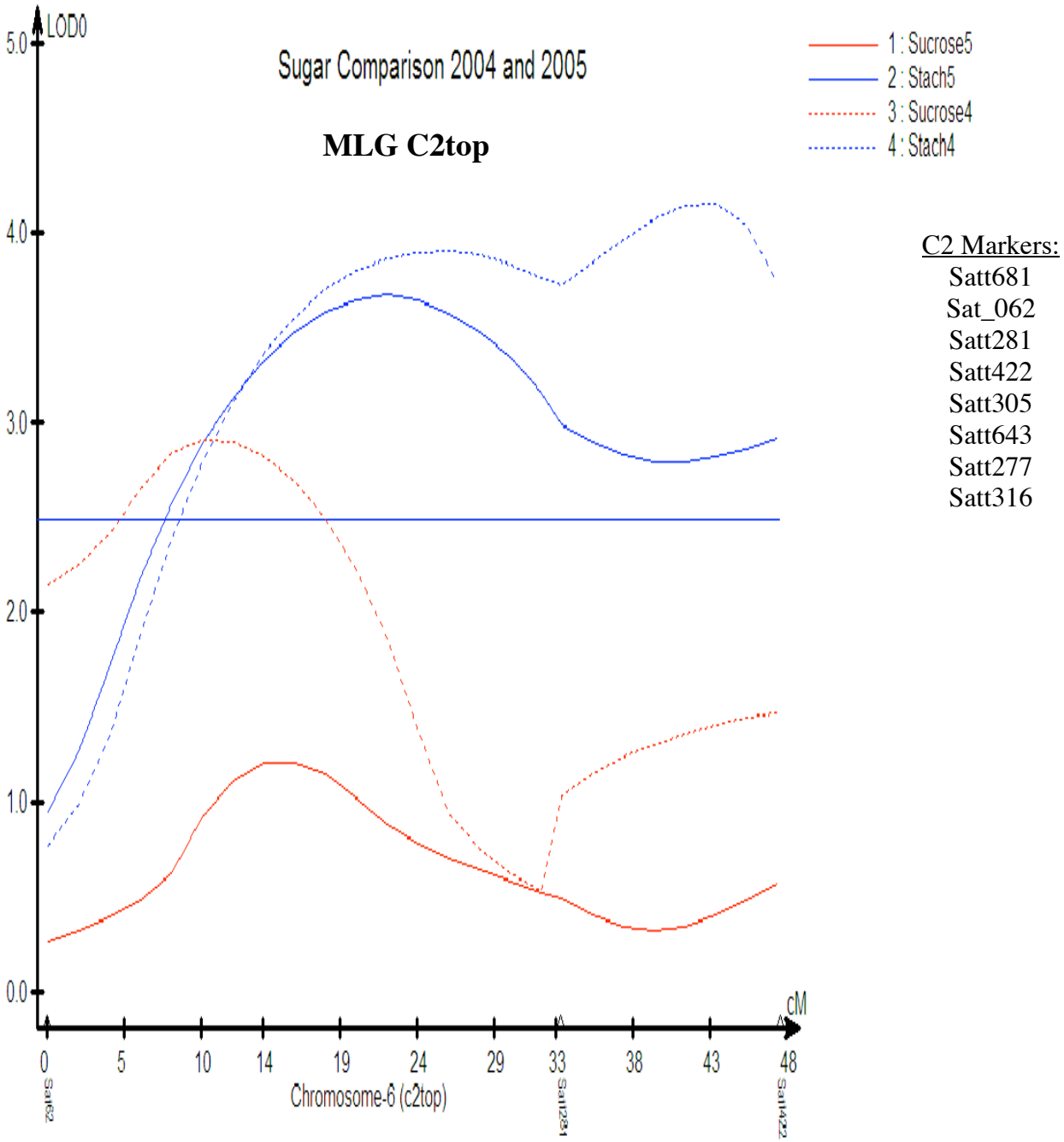


**Figure 12: QTL analysis of sucrose and stachyose for 2004 and 2005 on MLG C2top. The line horizontal to the linkage group indicates the set LOD of 2.5.**

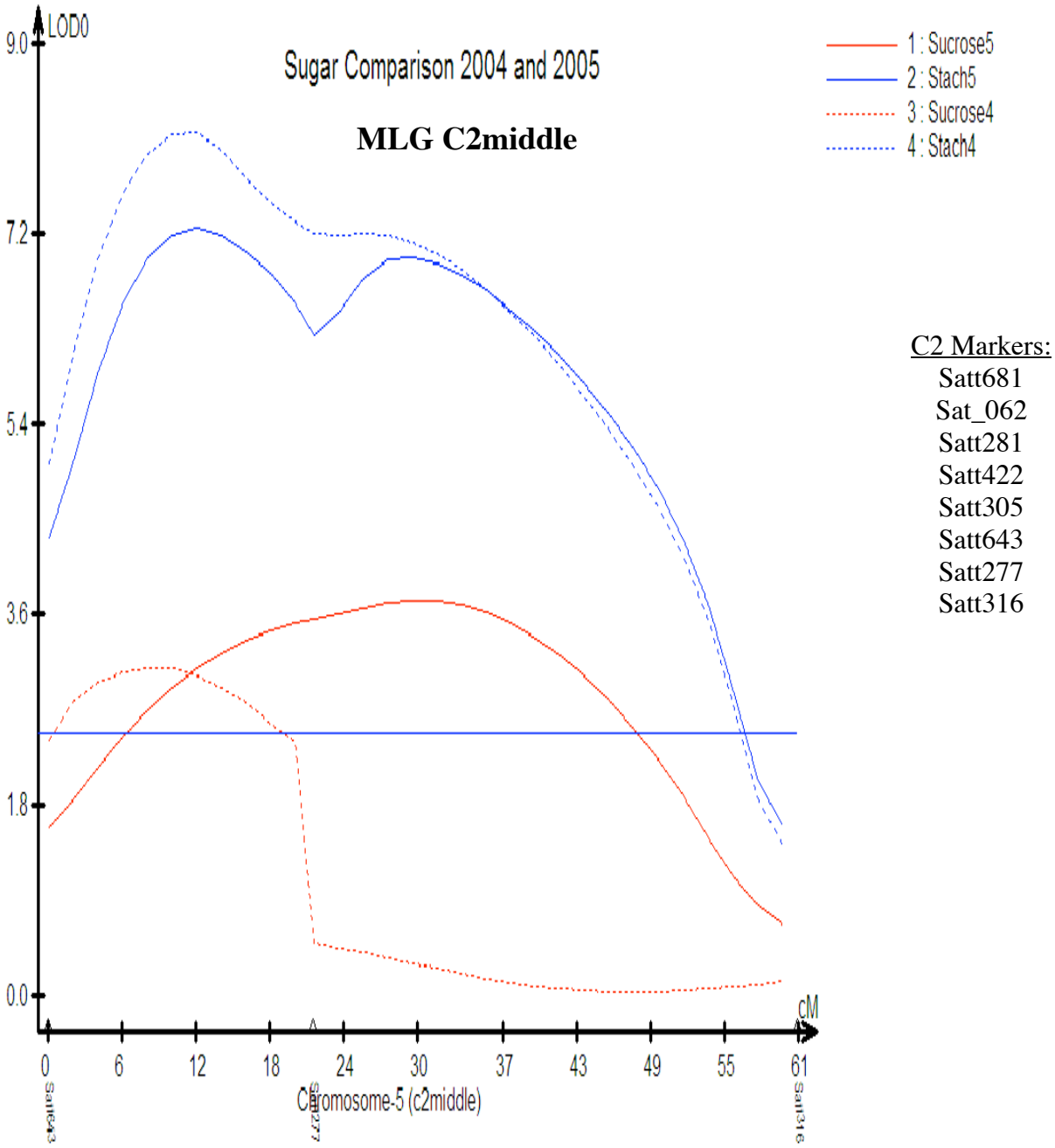


**Figure 13: QTL analysis of sucrose and stachyose for 2004 and 2005 on MLG C2middle. The line horizontal to the linkage group indicates the set LOD of 2.5.**

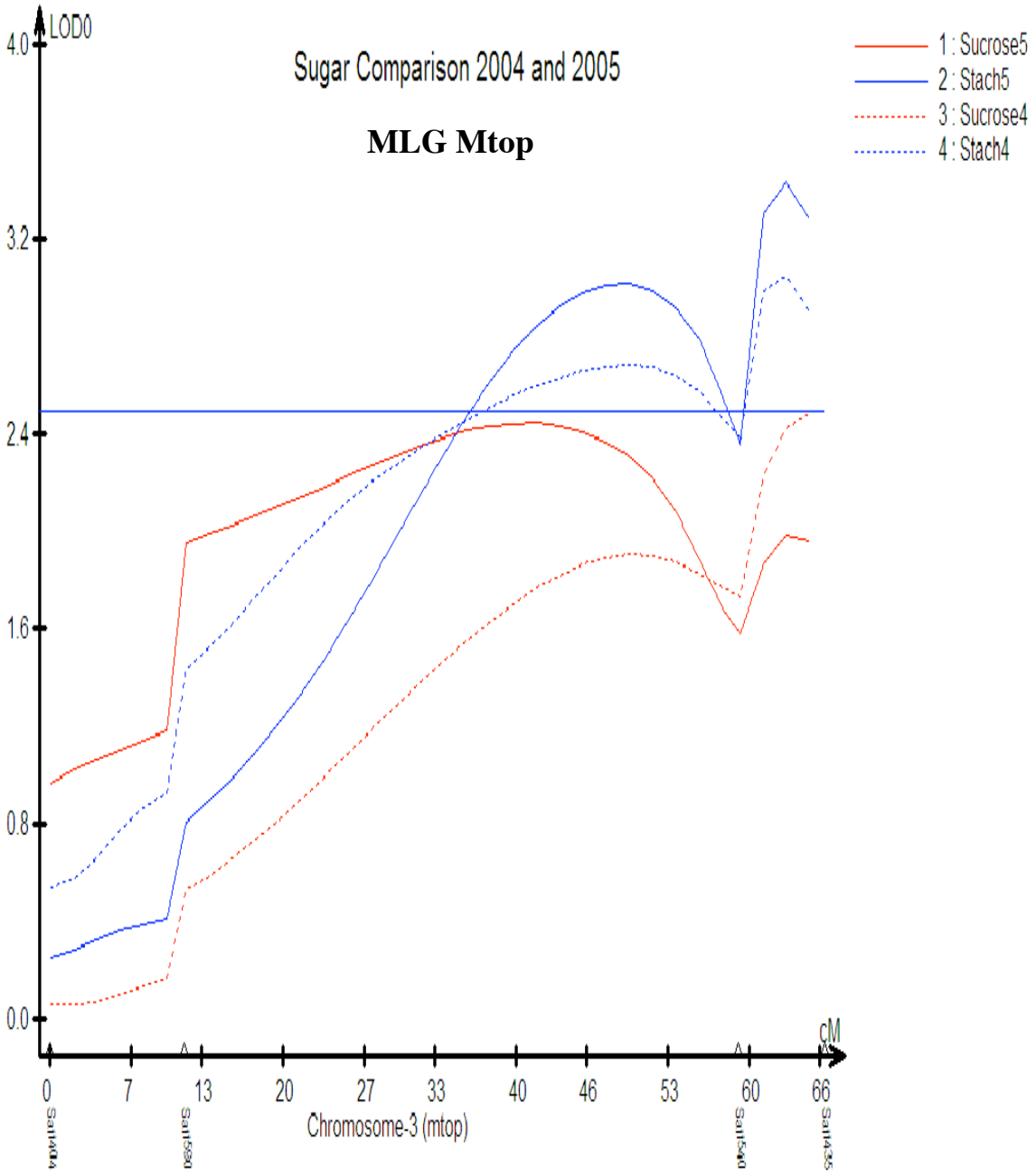


**Figure 14: QTL analysis of sucrose and stachyose for 2004 and 2005 on MLG Mtop. The line horizontal to the linkage group indicates the set LOD of 2.5.**

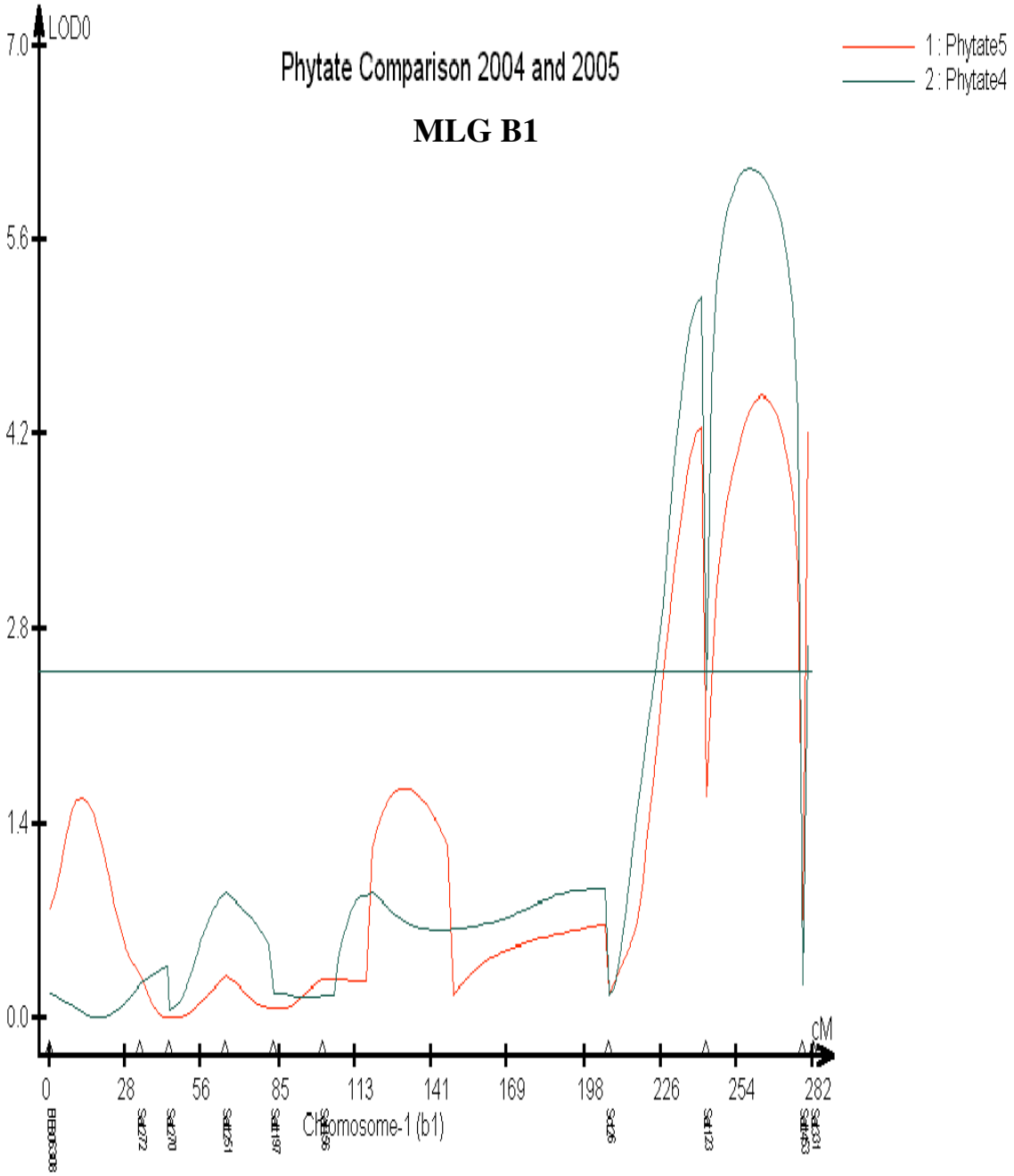


**Figure 15: QTL analysis of Phytate for 2004 and 2005 on MLG B1. The line horizontal to the linkage group indicates the set LOD of 2.5.**

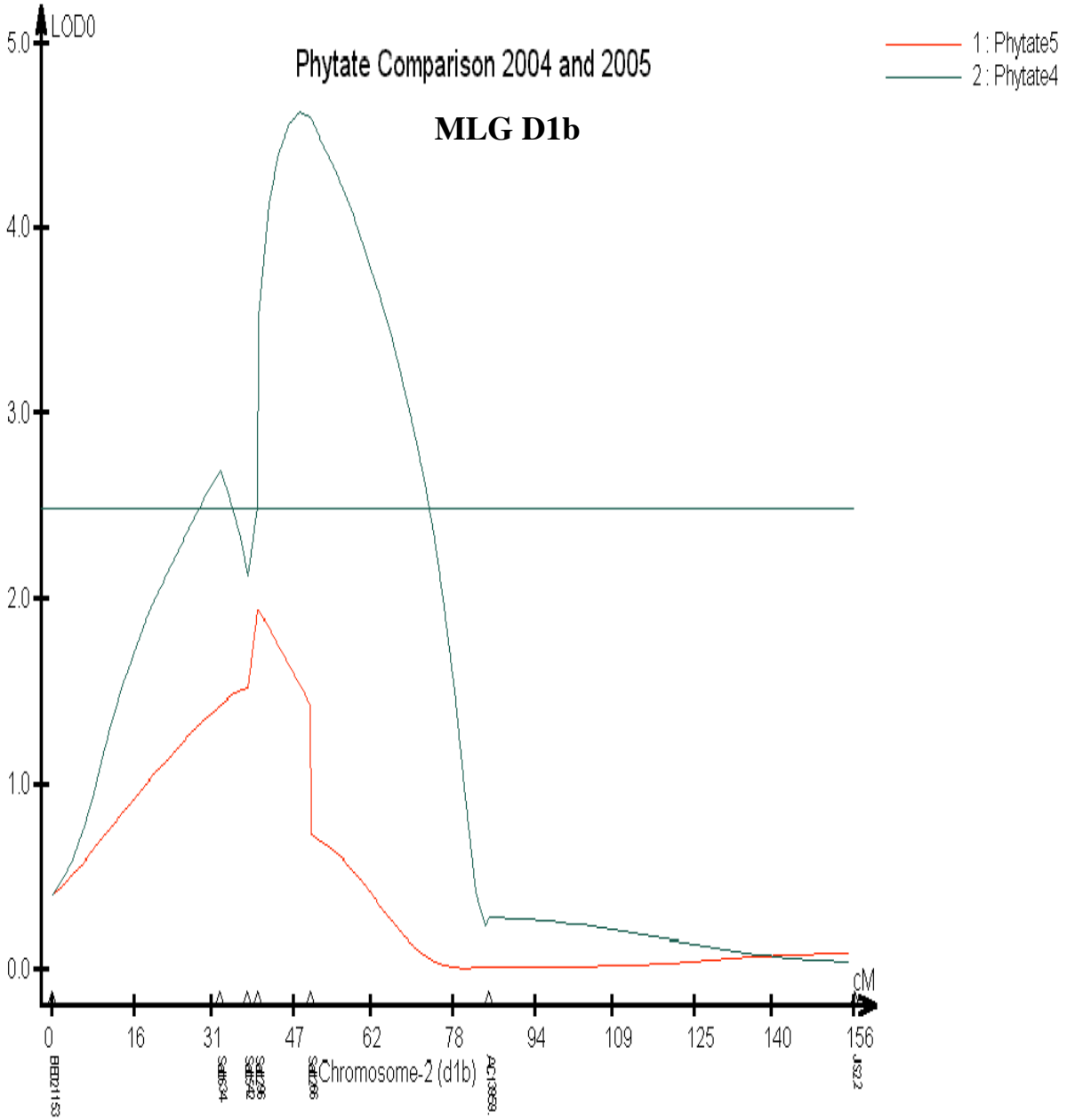


**Figure 16: QTL analysis of Phytate for 2004 and 2005 on MLG D1b. The line horizontal to the linkage group indicates the set LOD of 2.5.**

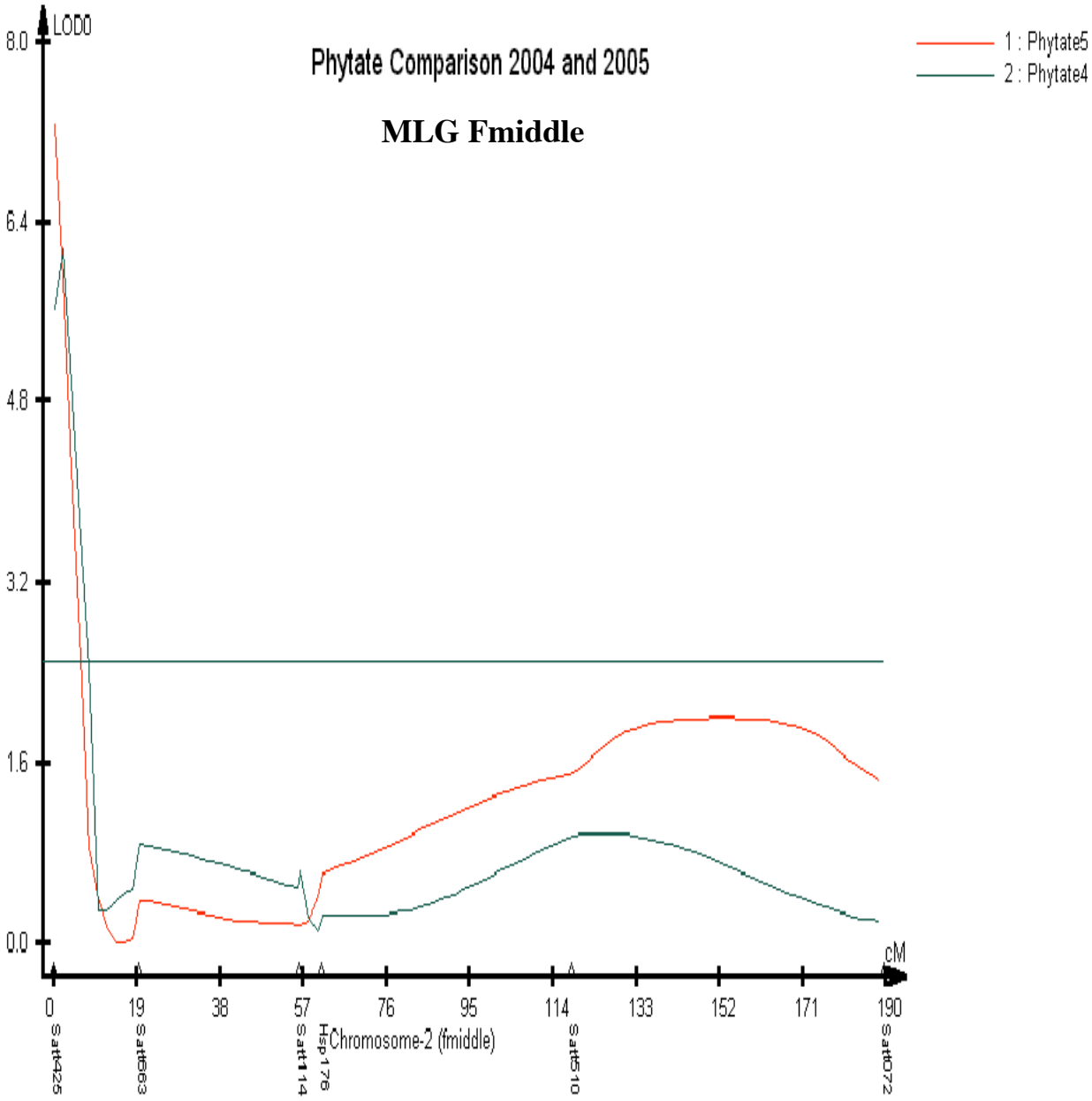


**Figure 17: QTL analysis of Phytate for 2004 and 2005 on MLG F-middle. The line horizontal to the linkage group indicates the set LOD of 2.5.**

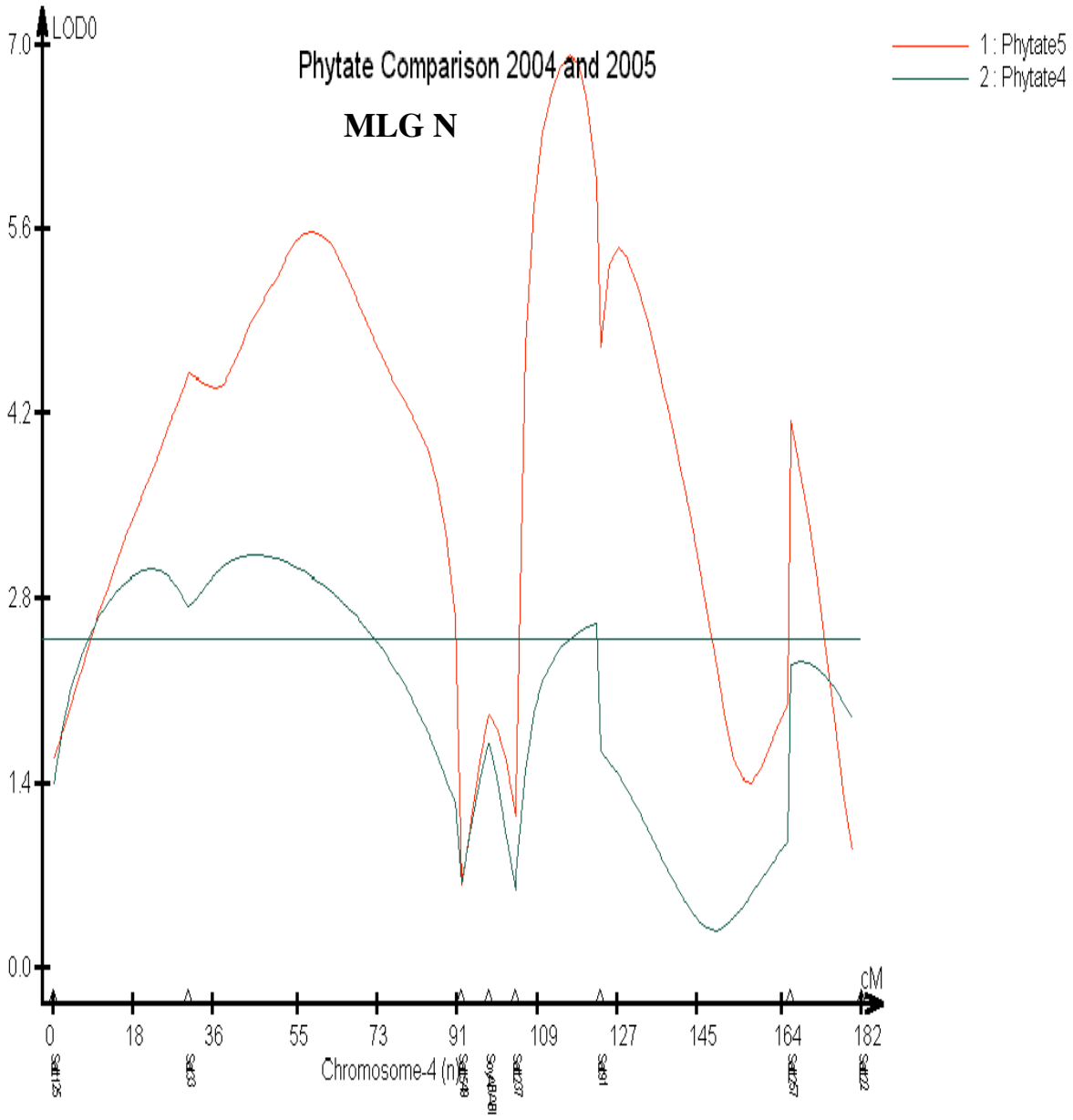


**Figure 18: QTL analysis of Phytate for 2004 and 2005 on MLG N. The line horizontal to the linkage group indicates the set LOD of 2.5.**

**Table 3 One Way ANOVA Analysis of Sugar Traits for the Years 2004 and 2005.**

| Marker | MLG | Map Position (Cregan et al., 1999) | Trait | Year | P-Value | 2 Year Average P-value |
|---|---|---|---|---|---|---|
| Sat272 | B1 | 14.3 | Sucrose | 2005 | .008 | - |
| Satt251 | B1 | 36.5 | Sucrose | - | - | .039 |
| Sat123 | B1 | 100.9 | Sucrose | 2004 | .006 | .041 |
| Sat331 | B1 | 125.7 | Sucrose | 2004 | <.001 | <.001 |
| Sat424 | B2 | 100.1 | Sucrose | 2004 | .004 | .024 |
| Satt281 | C2 | 40.3 | Sucrose | 2004 | .037 | .019 |
| **Satt422❖** | **C2** | **44.7** | **Sucrose** | **2004** | **.007** | **.012** |
| **Satt422** | **C2** | **44.7** | **Sucrose** | **2005** | **.047** | **.012** |
| Satt643 | C2 | 94.6 | Sucrose | - | - | .005 |
| Satt277 | C2 | 107.6 | Sucrose | - | - | .003 |
| Satt548 | D1a | 69.8 | Sucrose | - | - | .033 |
| Satt296 | D1b | 52.6 | Sucrose | 2004 | .043 | - |
| Satt569 | F | 3.4 | Sucrose | - | - | .048 |
| **Satt348** | **F** | **15.3** | **Sucrose** | **2004** | **.042** | **.012** |
| **Satt348** | **F** | **15.3** | **Sucrose** | **2005** | **.007** | **.012** |
| **BE806387** | **F** | **22.9** | **Sucrose** | **2004** | **.006** | **<.001** |
| **BE806387** | **F** | **22.9** | **Sucrose** | **2005** | **.006** | **<.001** |
| **Satt163** | **G** | **0.0** | **Sucrose** | **2004** | **.025** | **.005** |
| **Satt163** | **G** | **0.0** | **Sucrose** | **2005** | **.001** | **.005** |
| Satt235 | G | 21.9 | Sucrose | - | - | .042 |
| Satt594 | G | 52.9 | Sucrose | - | - | .031 |
| Sct65 | J | 32.1 | Sucrose | - | - | .018 |
| Satt414 | J | 37.8 | Sucrose | 2004 | .008 | .022 |
| Satt715 | K | 0.91 | Sucrose | 2004 | .004 | .013 |
| Satt417 | K | 46.2 | Sucrose | 2005 | .049 | - |
| Satt628 | K | 49.6 | Sucrose | 2005 | .045 | - |
| Satt435 | M | 38.9 | Sucrose | - | - | .026 |
| Sat33 | N | 58.4 | Sucrose | 2004 | .011 | .009 |
| Sat91 | N | 79.5 | Sucrose | 2005 | .041 | - |
| BE820148 | A2 | 35.9 | Stachyose | 2004 | .024 | .016 |
| Satt597 | B1 | 73.8 | Stachyose | 2004 | .034 | - |
| Satt453 | B1 | 123.9 | Stachyose | 2004 | .044 | - |
| **Sat331** | **B1** | **125.7** | **Stachyose** | **2004** | **<.001** | **<.001** |
| **Sat331** | **B1** | **125.7** | **Stachyose** | **2005** | **.001** | **<.001** |
| Sat177 | B2 | 7.8 | Stachyose | 2005 | .033 | - |
| **Satt180** | **C1** | **127.8** | **Stachyose** | **2005** | **.023** | **.025** |
| **Satt180** | **C1** | **127.8** | **Stachyose** | **2004** | **.037** | **.025** |
| Satt281 | C2 | 40.3 | Stachyose | 2004 | .021 | .020 |
| **Satt422** | **C2** | **44.7** | **Stachyose** | **2004** | **.002** | **.003** |

| **Satt422** | **C2** | **44.7** | **Stachyose** | **2005** | **.013** | **.003** |
|---|---|---|---|---|---|---|
| Satt643 | C2 | 94.6 | Stachyose | - | - | <.001 |
| Satt277 | C2 | 107.6 | Stachyose | - | - | <.001 |
| Satt316 | C2 | 127.7 | Stachyose | 2004 | .046 | .040 |
| Satt348 | F | 15.3 | Stachyose | 2004 | .028 | .043 |
| **BE806387** | **F** | **22.9** | **Stachyose** | **2004** | **.006** | **.005** |
| **BE806387** | **F** | **22.9** | **Stachyose** | **2005** | **.013** | **.005** |
| Satt663 | F | 56.2 | Stachyose | - | - | .044 |
| Satt163 | G | 0.0 | Stachyose | 2005 | .050 | - |
| Satt715 | K | 0.91 | Stachyose | 2004 | .023 | - |
| Satt540 | M | 35.8 | Stachyose | 2005 | .034 | - |
| Satt540 | M | 35.8 | Stachyose | 2004 | .050 | - |
| **Satt435** | **M** | **38.9** | **Stachyose** | **2004** | **.021** | **.016** |
| **Satt435** | **M** | **38.9** | **Stachyose** | **2005** | **.038** | **.016** |
| Sat33 | N | 58.4 | Stachyose | 2004 | .023 | - |
| Sat307 | O | 123.4 | Stachyose | 2004 | .015 | .024 |

❖ markers in bold are those that are significant in both 2004 and 2005

- signifies a marker was not significant in that year

**Table 4 One Way ANOVA Analysis of Phytate in the Years 2004 and 2005**

| Marker | MLG | Map Position (Cregan et al., 1999) | 2004 P-value | 2005 P-value | 2 Year Average P-value |
|---|---|---|---|---|---|
| Satt50 | A1 | 46.5 | .033 | - | - |
| Satt207 | A2 | 26.5 | .039 | .025 | .035 |
| Sat247 | B1 | 49.7 | - | - | .009 |
| Sat123 | B1 | 100.9 | .004 | - | .023 |
| Satt453 | B1 | 123.9 | - | .038 | .010 |
| Sat331 | B1 | 125.7 | .000 | <.001 | <.001 |
| Satt422 | C2 | 44.7 | .030 | .014 | .011 |
| Satt316 | C2 | 127.7 | .012 | - | - |
| BE021153 | D1b | 30.2 | - | - | .038 |
| Satt296 | D1b | 52.6 | .006 | - | .020 |
| Satt542 | D1b | 53.0 | .009 | - | .036 |
| Satt266 | D1b | 59.6 | .004 | - | - |
| Satt634 | D1b | 46.6 | .005 | .022 | .012 |
| Satt573 | E | 35.8 | .030 | - | - |
| Satt235 | G | 21.9 | - | .048 | - |
| Satt442 | H | 46.9 | .041 | .048 | - |
| Satt496 | I | 36.4 | .030 | - | - |
| Satt285 | J | 25.5 | .012 | - | - |
| Satt715 | K | .91 | .042 | .014 | - |
| Satt617 | K | 50.9 | - | .042 | .040 |
| Satt544 | K | 43.3 | - | .040 | - |
| Satt540 | M | 35.9 | .020 | .050 | .041 |
| Satt125 | N | 40.6 | - | - | .022 |
| Sat33 | N | 58.4 | .005 | .003 | .001 |
| Sat91 | N | 79.5 | .006 | <.001 | .002 |

- signifies a marker was not significant in that year

**Table 5 R-square values for QTL located in the V99-5089 x PI200205 population**

| Year | Trait | QTL Location | $R^2$ Value |
|---|---|---|---|
| 2004 | Stachyose | MLG B1 Sat156 – Satt197 | 37% |
| 2004 | Sucrose | MLG C2 Satt643 – Satt277 | 13.5% |
| 2005 | Sucrose | MLG C2 Satt643 – Satt316 | 21% |
| 2004 | Stachyose | MLG C2 Satt643 – Satt316 | 64% |
| 2005 | Stachyose | MLG C2 Satt643 – Satt316 | 42% |
| 2004 | Phytate | MLG B1 | 9% |
| 2004 | Phytate | MLG D1b | 28% |
| 2004 | Phytate | MLG N Satt125 – Satt549 | 10% |
| 2005 | Phytate | MLG N Satt237 – Satt257 | 28% |
| 2005 | Phytate | MLG N Satt125 – Satt549 | 18% |
| 2005 | Phytate | MLG N Satt257 – Satt22 | 17% |

## DISCUSSION

### *Low Stachyose*

One gene was hypothesized to control stachyose in V99-5089 populations in preliminary work done by Huhn (2003), in order to see if it did, it would need to fit a 15:1 model. This population of 55 families needed 3.4 or 1/16$^{th}$ of the families to have lower stachyose values than the parents. In examining these values for high and low stachyose for this population it was found that there are three families (473-4, 473-10, 473-17) that exhibited stachyose values lower than either parent. This population then fit the proposed model that two recessive genes are controlling stachyose for this population. Our population is an almost exact fit with 3 families being low stachyose. This would suggest that there are 2 recessive genes controlling low stachyose in this population using a $\chi^2$ value of 0.056. The appearance of families with normal stachyose in a cross of two low stachyose parents also supports the genetic hypothesis that the parents had recessive non-allelic genes for low stachyose. If V99-5089 and PI200508 had the same gene for low stachyose all the families would have had low stachyose.

It is important to note that the same four families held the four highest stachyose values and the same four families held the four lowest stachyose values for both years (Tables 6A and 6B). The families that contained the highest amounts of stachyose for both years also contained high amounts of phytate which can be seen in Tables 6A and 6B, which shows an apparent correlation between high stachyose and high phytate. The families that were low stachyose had the lowest phytate values as well, taken together, these observations indicate a positive relationship between stachyose and phytate content. Stachyose and phytate had a positive correlation of 0.35, while phytate and sucrose had a

negative correlation of –0.44. Since we have families exhibiting both low stachyose and low phytate together then this means that it should be possible to develop lines that have these particular traits.

Sat331 on MLG B1 appears to be significant for sucrose, stachyose and phytate for both years with p-values of <0.001. The following markers were associated with both sucrose and stachyose: Satt422, Satt281, Sat331, Satt643, Satt277, Satt348, BE806387, Satt163, Satt715 and Sat33. As stated before, QTL for stachyose were found on MLGs: B1, C2 and M using QTL Cartographer (Wang et al., 2005). R-square values for 2004 stachyose on B1 explained 37% of the phenotypic variation. This large R-square value for stachyose on MLG B1 coincides with the theory that a gene for stachyose is potentially located on B1 (Maroof unpublished). Phenotypic variation between 42 and 64% was also explained for stachyose on MLG C2 and between 13.5 and 21% for sucrose on that same linkage group. These R-square values were taken from QTL Cartographer (Wang et al., 2005) and explain the amount of phenotypic variation between markers.

There were approximately 19 markers that were found to be significant for stachyose and 25 markers were significant for sucrose with a p-value of 0.05. These results were similar to those of Cicek (2001) and Maughan (2000). Cicek (2001) found significant markers for stachyose on A1 and I. Significant markers for sucrose were found by Cicek (2001) on A1, A2, F, G and M and by Maughan (2000) on B1, B2, C2, D1a, D1b, F, G, J, K, M and N. These results for sucrose agree with this population for MLGs: B1, B2, C2, D1a, D1b, F, G, J, K, M and N. Three linkage groups (F, G and M) are common to all three studies.

It is interesting to note that in another population, V99-5089 by Essex, a major gene for stachyose was also found for V99-5089 (Huhn, 2003). Another interesting thing about this population is that even though there were many significant markers, only three of the linkage groups, B1, C2 and M, actually had QTL found in them when analyzed with QTL Cartographer. More markers from these MLGs (B1, C2 and M) need to be mapped to help clearly define where a QTL is located for this population that could potentially be used for marker assisted selection, markers should be concentrated in those areas that had high r-square values such as the area on C2 that had an R-square value of 64%.

***Importance of Low Phytate***

It has been hypothesized in papers such as Oltmans et al. (2005) that soybean lines that have low phytate have lower seedling emergence than those with normal phytate. This fact means that it may be difficult to develop those lines into acceptable cultivars for production of low phytate soybean meal. It is interesting to note that this was not the case with the lines from V99-5089 by PI200508. Phytate content was found to be weakly correlated with emergence (0.36) but this correlation was not strong enough to conclude that phytate would cause emergence problems in all low phytate lines.

This cross of V99-5089 by PI200508 is exciting because it does not appear to have the low seedling emergence problems associated with other low phytate lines that have been published (Oltmans et al., 2005). Despite the positive association between phytate content and emergence, numerous low stachyose lines can be identified that have above average emergence (Tables 1 and 3). It should also be noted that there appears to

be a locus on MLG N that explains a large effect for phytate for this population. Walker et al. (2006) also found one significant QTL on MLG N and another on MLG L; unfortunately, few polymorphic markers for MLG L were available for this population. Although the two sources of low phytate are unrelated, it seems very likely they both have the same low phytate gene on MLG N. It would be good in the future to find more polymorphic markers to more fully cover that linkage group to see if this population also contains a QTL on MLG L. It is important to note that QTL for both phytate and stachyose are present on MLG B1 for this population, which shows a correlation between phytate and stachyose would make sense since V99-5089 is both low stachyose and low phytate.

***Quantitative Trait Loci***

The population studied here, V99-5089 by PI200508, had a total of 141 markers mapped to all 20 molecular linkage groups, which allowed for the objective of creating a skeletal linkage group map. While the data were not adequate to definitely declare where an important QTL for stachyose may be located, there appears to be a good probability that one is located on MLG B1, which is also associated with low phytate. It would be wise in the future to saturate those linkage groups that had significant markers for stachyose on them such as: B1, C2 and M. Linkage groups such as F and O, are also important to saturate since significant markers located there had r-square values between 10 and 15.8%. To get a better estimate of where true QTL are located, this population needs to be expanded to a hundred individuals or more.

According to the ANOVA results, significant markers for sucrose were found on MLGs: B1, B2, C2, D1b, F, J, K and N. QTL Cartographer analysis shows significant LODs for sucrose on MLGs B1, C2, D1b and N, which agrees with the analysis of variance except for MLGs C2, F, J and K. In order to possibly get more comprehensive LOD for any of the above traits then the linkage groups that are in pieces need to be brought together into one group and more markers need to be mapped to fill in the map. Also, more markers need to be mapped on those linkage groups that had significant QTL for sucrose, stachyose and/or phytate.

Significant markers for phytate were found on MLGs: A1, A2, B1, C2, D1b, E, G, H, I, J, K, M and N. QTL Cartographer analysis showed significant LODs for phytate on the following MLGs: B1, D1b, F, and N. It is important to note that the LODs shown in QTL cartographer for phytate for MLG Fmiddle are not complete and this set of markers needs to be brought together with the other markers from this linkage group.

Many markers seem to be associated with the data, both sugar and phytate, when it is analyzed using one way analysis of variance with an alpha value of 0.05, which would suggest that there are many QTL for these traits. This turns out not to be true once the marker data and phenotypic data are analyzed using QTL Cartographer with a set LOD of 2.5. The QTL results from this study should be confirmed using phenotypic data from a larger population and involve more markers from the key MLG regions as identified in this study.

**Table 6A Five highest stachyose values of 2005 and 2004 and their respective phytate values**

| Line | Stachyose % 2004 | Stachyose % 2005 | Phytate (mg/g) 2004 | Phytate (mg/g) 2005 |
|---|---|---|---|---|
| 473-7 | 4.16 | 5.90 | 14.07 | 14.24 |
| 471-4 | 4.25 | 5.72 | 15.09 | 18.67 |
| 471-10 | 4.23 | 5.42 | 14.69 | 15.20 |
| 470-5 | 4.21 | 4.57 | 12.92 | 15.23 |

**Table 6B Five lowest stachyose values of 2005 and 2004 and their respective phytate values**

| Line | Stachyose % 2004 | Stachyose % 2005 | Phytate (mg/g) 2004 | Phytate (mg/g) 2005 |
|---|---|---|---|---|
| 470-8 | 0.45 | 0.26 | 12.53 | 13.25 |
| 473-4 | 0.16 | 0.17 | 9.34 | 9.95 |
| 473-17 | 0.15 | 0.15 | 9.05 | 10.54 |
| 473-10 | 0.22 | 0.15 | 10.45 | 11.00 |

**Literature Cited:**


Adeola, A. 1995. Digestive utilization of minerals by weanling pigs fed copper- and phytase-supplemented diets. Can. J. Anim. Sci. 75: 603-610.

Bentsink, L. C. Alonso-Blanco, D. Vreugdenhil, K. Tesnier, S.P.C. Groot and M. Koornneef. 2000. Genetic analysis of seed-soluble oligosaccharides in relation to seed storability of Arabidopsis. Plant Physiol. 124:1595-1604.

Brookes, A.J. 1999. The essence of SNPs. Gene. 234:177-186.

Cicek, M. 2001. Genetic Marker Analysis of Three Major Carbohydrates in Soybean Seeds. Doctor of Philosophy Dissertation. Virginia Polytechnic Inst. & State Univ. Blacksburg, VA.

Cicek, M.S., P. Chen, M.A. Saghai Maroof and G.R. Buss. Interrelationships among agronomic and seed quality traits in an interspecific soybean recombinant inbred population. Crop Sci. 46:1253-1259.

Cober, E.R., J.A. Fregeau-Reid, L.N. Pietrzak, A.R. McElroy and H.D. Voldeng. 1997. Genotype and environmental effects on natto soybean quality traits. Crop Sci. 37:1151-1154.

Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. Van Toai, D.G. Lohnes, J. Chung and J.E. Specht. 1999. An integrated genetic linkage map of the soybean genome. Crop Sci. 39:1464-1490.

Cromwell, G.L. 2002. Overview of nutritional and environmental benefits of phytases. J. Anim. Sci. 80 (Suppl. 1):54.

Cromwell, G.L., S.L. Traylor, L.A. White, E.G. Xavier, M.D. Lindemann, T.E. Sauber and D.W. Rice. 2000. Effects of low-phytate corn and low-oligosaccharide, low-phytate soybean meal in diets on performance, bone traits and phosphorous excretion by growing pigs. J. Anim. Sci. 78 (Suppl. 2):72.

de Mejia, E.G., T. Bradford and C. Hasler. 2003. The anticarcinogenic potential of soybean lectin and lunasin. Nutr. Rev. 61:239-246.

Dey, P.M. and R.A. Dixon, editors. 1985. Biochemistry of storage carbohydrates in green plants. Academic Press: London.

Erdman, J.W., Jr. 1979. Oilseed phytates: Nutritional implications. J. Am. Oil Chem. Soc. 56:736-741.

Ertl, D.S., K.A. Young and V. Raboy. 1998. Plant genetic approaches to phosphorous management in agricultural production. J. Environ. Qual. 27:299-304.

Goldstein, DB, Schlotterer C, editors 1999. Microsatellites: Evolution and Applications. University Press: Oxford.

Guimaraes, V.M., S.T. de Rezende, M.A. Moreira, E.G. de Barros and C.R. Felix. 2001. Characterization of α-galactosidases from germinating soybean seed and their use for hydrolysis of oligosaccharides. Phytochem. 58:67-73.

Hammond, E.G., L. Johnson, T. Wang, C. Su and P. White. 2004. Soybean oil. *In* F. Shahidi (ed.) Bailey's industrial oil and fat products. Sixth ed. John Wiley & Sons, New York.

Hartwig, E.E., T.M. Kuo and M.M. Kenty. 1997. Seed protein and its relationship to soluble sugars in soybean. Crop Sci. 37:770-773.

Hedley, C.L. ed. 2001. Carbohydrates in Grain Legume Seeds: Improving Nutritional Quality and Agronomic Characteristics. CABI Publishing: New York.

Hegeman, C.E., Good L.L, E.A. Grabau. 2001. Expression of D-*myo*-inositol-3-phosphate synthase in soybean. Implications for phytic acid biosynthesis. Plant Physiol. 125:1941-1948.

Hitz, W.D., T.J. Carlson, P.S. Kerr and S.A. Sebastian. 2002. Biochemical and molecular characterization of a mutation that confers a decreased raffinosaccharide and phytic Acid phenotype on soybean seeds. Plant Physiol. 128:650-660.

Hsia, A., T. Wen, H. Chen, Z. Liu, M. Yandeau-NelsonY. Wei, L. Guo and P. Schnable. 2005. Temperature gradient capillary electrophoresis (TGCE)-a tool for the high-throughput discovery and mapping of SNPs and IDPs. Theor. Appl. Genet. 111:218-225.

Huhn, M.R. 2003. Inheritance of Soluble Oligosaccharide Content of Soybean Seeds. Master's Thesis. Virginia Polytechnic Inst. & State Univ. Blacksburg, VA.

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

Hymowitz, T. and F.I. Collins. 1974. Variability of sugar content in seed of *Glycine max* (L.) Merrill and *G. soja* Sieb. and Zucc. Agron. J. 66:239-240.

Hymowitz, T., F.I. Collins, J. Panczner and W.M. Walker. 1972. Relationship between the content of oil, protein and sugar in soybean seed. Agron. J. 64:613-616.

Karner, U., T. Peterbauer, V. Raboy, D.A. Jones, C.L. Hedley and A. Richter. 2004. *myo*-inositol and sucrose concentrations affect the accumulation of raffinose family oligosaccharides in seeds. J. of Exp. Botany. 55: 1981-1987.

Kennedy, I.R., O.D. Mwandemele and K.S. McWhirter. 1985. Estimation of sucrose, raffinose and stachyose in soybean seeds. Food Chem. 17:85-93.

Kerley, M.S. and G.L. Allee. 2003. Modifications in soybean seed composition to enhance animal feed use and value: moving from a dietary ingredient to a functional dietary component. AgBioForum. 6:14-17.

Klug, W.S. and M.R. Cummings. 2002. Essentials of Genetics 4th Edition. Prentice Hall: New Jersey.

Kuo, T.M., J.F. VanMiddlesworth and W.J. Wolf. 1988. Content of raffinose oligosaccharides and sucrose in various plant seeds. J. Agric. Food Chem. 36:32-36.

Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M.J. Daly, S.E. Lincoln and T. Etoh. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174-181.

Lander, E.S. and D. Botstein. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics. 121:185-199.

Larson, S., K. Young, A. Cook, T. Blake and V. Raboy. 1998. Linkage mapping of two mutations that reduce phytic acid content of barley grain. Theor. Appl. Genet. 97:141-146.

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

Larson, S.R., J.N. Rutger, K.A. Young and V. Raboy. 2000. Isolation and genetic mapping of a non-lethal rice (*Oryza sativa L.*) *low phytic acid* mutation. Crop Sci. 40:1397-1405.

Lei, X.G. and J.M. Porres. 2003. Phytase enzymology, applications, and biotechnology. Biotechnol. Lett. 25:1787-1794.

Lowell, C.A. and T.M. Kuo. 1989. Oligosaccharide metabolism and accumulation in developing soybean seeds. Crop Sci. 29:459-465.

Maughan, P.J., M.A. Saghai Maroof and G.R. Buss. 2000. Identification of quantitative trait loci controlling sucrose content in soybean (*Glycine max*). Mol. Breeding. 6:105-111.

Maughan, P.J., M.A. Saghai Maroof and G.R. Buss. 1995. Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean. Genome. 38:715-723.

Meis, S.J., W.R. Fehr and S.R. Schnebly. 2003. Seed source effect on field emergence of soybean lines with reduced phytate and raffinose saccharides. Crop Sci. 43:1336-1339.

Minitab Inc. 2003. Minitab Release 14 for Windows.

Minorsky, P.V. 2003. Raffinose oligosaccharides. Plant Phys. 131:1159-1160.

Neus, J.D., W.R. Fehr and S.R. Schnebly. 2005. Agronomic and seed characteristics of soybean with reduced raffinose and stachyose. Crop Sci. 45:589-592.

Oltmans, S.E., W.R. Fehr, G.H. Welke, V. Raboy and K.L. Peterson. 2005. Agronomic and seed traits of soybean lines with low-phytate phosphorous. Crop Sci. 45:593-598.

Openshaw, S.J., T. Hymowitz and H.H. Hadley. 1977. Distribution of sugars in the soybean embryo. Crop Sci. 17:975-977.

Parsons, C.M., Y. Zhang and M. Araba. 2000. Nutritional evaluation of soybean meals varying in oligosaccharide content. Poultry Sci. 79:1127-1131.

Primomo, V.S., V. Poysa, G.R. Ablett, C.J. Jackson, M. Gijzen and I. Rajcan. 2005. Mapping QTL for individual and total isoflavone content in soybean seeds. Crop Sci. 45:2454-2464.

Raboy, V., D.B. Dickson and F.E. Below. 1984. Variation of seed total phosphorous, phytic acid, zinc, calcium, magnesium and protein among lines of *Glycine max* and *G. soja*. Crop Sci. 24:431-434.

Raboy, V., S.J. Hudson and D. B. Dickson. 1985. Reduced phytic acid content does not have an adverse effect on germination of soybean seeds. Plant Physiol. 79:323-325.

Raboy, V. and D.B. Dickinson. 1987. The timing and rate of phytic acid accumulation in developing soybean seeds. Plant Physiol. 85:841-844.

Raboy, V., P. Gerbasi, K. Young, S. Stoneberg, S. Pickett, A. Bauman, P. Murthy, W. Sheridan and D. Ertl. 2000. Origin and seed phenotype of maize *low phytic acid 1-1* and *low phytic acid 2-1*. Plant Physiol. 124:355-368.

Raboy, V. 2001. Seeds for a better future: 'low phytate' grains help to overcome malnutrition and reduce pollution. Trends in Plant Sci. 6:458-462

Raboy, V. 2002. Progress in breeding low phytate crops. American Society for Nutritional Sciences. 503S-505S.

Rafalski, J.A., S.V. Tingey and J.G.K. Williams. 1991. RAPD markers-a new technology for genetic mapping and plant breeding. AgBiotech News and Information.

Reddy, N.R., M.D. Pierson, S.K. Sathe and D.K. Salunkhe. 1989. Phytates in Cereals and Legumes. CRC Press Inc.: Boca Raton.

Saghai Maroof, M.A., K.M. Soliman, R.A. Jorgensen and R.W. Allard. 1984. Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance chromosomal location and population dynamics. Proc. Natl. Acad. Sci. USA 81:8014-8018.

Schlotterer, C. and D. Tautz. 1992. Slippage synthesis of simple sequence DNA. Nuc. Acids Res. 20: 211-215.

Sebastian, S.A., P.S. Kerr, R.W. Pearlstein and W.D. Hitz. 2000. Soybean germplasm with novel genes for improved digestibility. *In* JK Drackley, ed, Soy in Animal Nutrition. Federation of Animal Science Societies, Savoy, IL, pp. 56-74.

Song, Q.J., L.F. Marek, R.C. Shoemaker, K.G. Lark, V.C. Concibido, X. Delannay, J.E. Specht and P.B. Cregan. 2004. A new integrated genetic linkage map of the soybean. Theor. Appl. Genet. 109: 122-128.

Tewary, H.K. and H.G. Muller. 1992. The fat of some oligosaccharides during the preparation of wari, an Indian fermented food. Food Chem. 43:107-111.

U.S. Food and Drug Administration. 1999. A food labeling guide. U.S. Dep. Of Health and Human Services, Public Health Service, Washington, D.C.

Valdes, A.M., M. Slatkin and N.B. Freimer. 1993. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. Genetics. 133: 737-749.

Verma, D.P.S. and R.C. Shoemaker eds. 1996. Soybean: Genetics, Molecular Biology and Biotechnology. CAB International: Wallingford.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. Nuc. Acids Res. 23:4407-4414.

Walker, D.R., A.M. Scaboo, V.R. Pantalone, J.R. Wilcox and H.R. Boerma. 2006. Genetic mapping of loci associated with seed phytic acid content in CX1834-1-2 soybean. Crop Sci. 46:390-397.

Wang, S., C.J. Basten and Z.B. Zeng. 2005. Windows QTL Cartographer V2.5. Program in Statistical Genetics, North Carolina State University.

White, P.J. 2000. Fatty acids in oilseeds (vegetable oils). P. 209-237. *In* C.K. Chow (ed.) Fatty acids in foods and their health implications, 2nd ed. Marcel Dekker, New York.

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

Yazdi-Samadi, B., R.W. Rinne and R.D. Seif. 1977. Components of developing soybean seeds: oil, protein, sugars, starch, organic acids and amino acids. Agron. J. 69:481-486.

Yu, Y.G., M.A. Saghai Maroof, G.R. Buss, P.J. Maughan and S.A. Tolin. 1994. RFLP and microsatellite mapping of a gene for soybean mosaic virus resistance. Phytopathology 84:60-64.

Zhang, W.K., Y.J. Wang, G.Z. Luo, J.S. Zhang, C.Y. He, X.L. Wu, J.Y. Gai and S.Y. Chen. 2004. QTL mapping of ten agronomic traits on the soybean (*Glycine Max* L. Merr.) genetic map and their association with EST markers. Theor. Appl. Genet. 108:1131-1139.