Mapping and Identification of QTL Associated with Soybean Seed Protein, Oil, and Yield in 5601T × U99- 310255 RIL Population Using SNP Genotyping

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Mia Justina Cunicelli
May 2020

ACKNOWLEDGEMENTS

First, I would like to thank my major professor Dr. Vince Pantalone for giving me the chance to further my education, while working full time for the Soybean Breeding and Genetics program. I appreciate your great patience and the opportunities you offered me to become a better scientist. Thank you for always making time to meet with me and for answering my many questions. Thank you for consistently challenging me so I could improve my skills. I would also like to thank my committee members, Dr. Carl Sams, Dr. Liesel Schneider, and Dr. Dennis West. I would like to thank everyone in the Plant Science Office for helping make sure I got my degree on time. I would also like to thank The Tennessee Soybean Promotion Board and the United Soybean Board for funding.

Next, I would like to thank all the members of the Bean Team for their assistance and continuing support. I learned something from each of you, and working beside you day by day made the hard work much easier. I would also like to thank all the UT REC members at East Tennessee, Milan, and Highland Rim, especially BJ DeLozier, Chris Bridges, and Brad Fisher, for helping me manage my soybeans across the state of Tennessee and for being continuously informative and supportive.

Lastly, I would like to thank my wonderful parents and brother for putting up with me and for encouraging me to further my education. I would like to thank

them for giving me the love and support I needed to do so. I would not have been able to do this without your faith in me.

ABSTRACT

Soybean [*Glycine max* (L.) Merr.] is the world's leading source of vegetable oil and high quality protein meal. However, soybean protein and oil and protein and yield are negatively genetically correlated. Agronomic and seed quality traits of soybean lines are important because they lead to increased soybean value. Molecular markers have played and will continue to play a major role in the genetic characterization and improvement of soybeans. Quantitative trait loci (QTL) are regions within a genome that contain genes that influence particular quantitative traits. Identification of QTLs through phenotypic evaluation only is not possible; therefore, to locate, identify, and characterize these regions within a genome, DNA or molecular markers are used. A major objective in soybean breeding is to develop high yielding cultivars. Unfortunately, soybean seed yield, as well as protein and oil content, are complex quantitative traits to characterize from the phenotypic and genotypic perspectives. The objectives of this study are to detect soybean genomic regions that increase protein content, while maintaining oil content and seed yield, and to successfully identify soybean QTL associated with these seed quality traits. To achieve these objectives, a population of 138 $F_{8:10}$ [eighth filial generation advanced to the tenth filial generation] and 138 $F_{8:11}$ [eighth filial generation advanced to the eleventh filial generation] recombinant inbred lines (RIL) of soybean were created from a cross between 5601T and U99-310255. These RILs were grown in a replicated six environment field trial across the state of Tennessee in the 2017 and 2018 field

seasons. Data from the 138 RIL were used to perform QTL detection analyses in search of significant genomic regions affecting soybean seed protein, oil, and yield. A total of 21 QTL were successfully identified through QTL analysis of the population. Of these, there were three for yield, two for protein, five for oil, one for methionine, two for threonine, five for maturity, one for lodging, and two for meal. Knowledge of their locations and flanking markers will aid in marker assisted selection for plant breeders. This will lead to a more valuable soybean for farmers, processors, and animal nutritionists.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION AND LITERATURE REVIEW

Soybean [*Glycine max* (L.) Merr.] is a legume that belongs to the *Glycine* genus and *Fabaceae* family, and is a domesticated annual plant species native to Eastern Asia (Hymowitz, 2004). It was first cultivated in China and was introduced to other countries such as Thailand, Nepal, and Japan, where landraces were developed (Aoyagi and Shurtleff, 2004; Hymowitz, 2004). Trade routes were eventually established on land and sea which allowed for the expansion of soybean to other countries (Hymowitz, 2004). Soybeans eventually arrived in North America in the United States (U.S.) in 1765. Until the mid-1930's, soybean was grown primarily as a forage and cover crop in the U.S. (Aoyagi and Shurtleff, 2004); however, it started its rise to prominence as a grain crop in the 1920's. This transition was fueled by many factors, including new technology, interests of private processing companies, and ongoing research conducted by the United States Department of Agriculture (USDA) (Hymowitz, 1990).

Soybean growth has continued expanding across North America since its arrival, and has expanded into South American countries such as Brazil. There are high demands for soybean across the world, making expansion of the crop necessary. Competitive expansion of soybean crops will encourage improvement in soybean agronomic and seed quality traits through breeding and molecular techniques (Aoyagi and Shurtleff, 2004), benefitting consumers of soy products. Additionally, expansion of soybean crops is also necessary to help meet the growing world demand for protein.

Once planted, soybean is able to fix nitrogen due to a symbiotic relationship with *Bradyrhizobia japonicum,* which establishes nodules on the soybean root (Casteel, 2011). These rhizobia live in nodules located on the soybean roots and fix nitrogen for the plant in exchange for sugars. This is one reason why soybeans are utilized in crop rotation (North Carolina Soybean Producers Association, Inc., 2014). Another reason is that planting soybeans provides benefits to farmers by reducing the need for chemical fertilizer (Araújo et al., 2015). The nitrogen fixation performed by the root nodule bacteria helps create the soy proteins present in soybean (Friedman and Brandon, 2001).

Due largely to breeding efforts and improvements in production, equipment, pest management, irrigation, and fertilizer, soybean has become a valuable agricultural crop worldwide. In the U.S. in 2016, 31% of the crop area planted was occupied by soybeans (Soy Stats, 2018). Additionally, in the same year, the U.S. produced 34% of the world's soybeans, making it the leading soybean producer, followed closely by Brazil (Soystats, 2018). The National Agricultural Statistics Service reported that U.S. soybean farmers planted 36.1 million hectares of soybean in 2018, which resulted in a total soybean value of $39.13 billion (USDA, NASS, 2019). Its status as a cash crop is due to the versatility of the seed. Humans consume a modest amount of the harvested seeds, as soybean is a great source of vegetable protein; however, they are primarily used as a source of oil and high protein meal for livestock feed (Aoyagi and Shurtleff, 2004; Araújo et al., 2015; Friedman and Brandon, 2001).

Soybean Protein

Soybean is the leading source of oil and low cost, high quality protein meal in the world (Aoyagi and Shurtleff, 2004; Diers et al., 1992; Joshi et al., 2013; Lin et al., 2011; Miller-Garvin and Naeve, 2015; Pantalone, 2012; Warrington et al., 2015; Wilcox, 1998; Wilcox and Shibles, 2001). Soymeal is the leading source of protein for poultry (*Gallus gallus domesticus*) and swine (*Sus scrofa*) diets throughout the United States. Amino acids comprise protein, and their balance is important in determining soybean nutritional value (Friedman and Brandon, 2001; Khandaker et al., 2015; Thakur and Hurburgh, 2007; Panthee et al., 2006a; Warrington et al., 2015; Wilcox and Shibles, 2001). For this reason, animal diets must include protein in order to supply essential amino acids to the body (Thakur and Hurburgh, 2007; Pantalone, 2012; Panthee et al., 2006a; Warrington et al., 2015). Soybean offers a well-balanced vegetable protein, and for this reason, it has been considered the healthiest protein source for animal feed. However, cysteine and methionine, the sulfur-containing amino acids, are limited in soymeal (Friedman and Brandon, 2001; Khandaker et al., 2015; Panthee et al., 2006b; Warrington et al., 2015; Wilcox and Shibles, 2001) and require animal feed to be supplemented with synthetic amino acids (Warrington et al, 2015; Wilcox and Shibles, 2001). Baker et al. (2011) suggests that there are more digestible amino acids available for growing livestock fed high protein soybean meal rather than conventional, because there is a greater concentration of amino acids in high protein soybean meal than in conventional soybean meal. Producing soybeans with higher quality protein could help solve amino acid

deficiency problems and lower excess costs for farmers caused by supplementing animal feed.

In this experiment, one goal is to detect genomic regions governing soybean amino acids to guide breeders forward with genetic gains. Common goals in soybean breeding have been to develop cultivars with stability across environments, to increase soybean seed protein for livestock feed in order to keep soybean competitive (Hyten et al., 2004), as well as to increase protein concentration and quality within the soybean, while maintaining a high yield and oil concentration (Brim and Burton, 1979; Diers et al., 1992; Lee et al., 2010; Pantalone et al., 2020; Pantalone and Smallwood, 2018; Warrington et al., 2015; Warrington et al., 2014, Wilcox, 1998). However, accomplishing these goals presents a challenge.

With the global population on the rise, world demand for soybean protein is estimated to steadily rise over the years ahead. However, because of the negative genetic correlation between protein and both oil and yield, breeding for increased protein will lead to reduced oil and yield (Chung et al., 2003; Cober and Voldeng, 2000; Hernández-Sebastiá et al., 2005; Krishnan et al, 2007; Lee et al., 2010; Warrington et al. 2015; Wilcox, 1998; Yaklich, 2001). Nevertheless it is possible, although difficult, to develop high yielding, high protein lines (Pantalone, 2018; Pantalone et al., 2020). Much research has taken place in an effort to increase soybean protein and improve its amino acid profile (Fallen et al., 2013; Panthee et al., 2006a; Panthee et al., 2006b). Quantitative trait loci

(QTL) that increase both oil and protein simultaneously, along with QTL that simultaneously increase oil and yield have been identified by Eskandari et al. (2013a). More recently, single nucleotide polymorphism (SNP) genotyping has led to the detection of two QTL for protein concentration using a recombinant inbred line (RIL) population created by crossing PI 43848913 by Hamilton (Kim et al., 2012). Additionally, a protein QTL in the MD96-5722 × Spencer RIL population was detected using a high density map developed from using 5376 SNP markers from the Illumina Infinium BeadChip array (Akond et al., 2014). Liu et al. (2017) identified three protein concentration QTL in an RIL population developed from a cross of Zhonghuang 25 × Huaxia 3. This was done through SNP genotyping and construction of a high-density genetic map (Liu et al., 2017). Panthee et al. (2006b) identified four QTL associated with cysteine and three QTL associated with methionine concentration in soybean seed in an RIL population developed from a cross of N87-984-16 × TN93-99. Their soybean seed samples were ground and analyzed on a near infrared spectroscopy instrument and were further analyzed using SAS software and QTL Cartographer (Panthee et al., 2006b). Fallen et al. (2013) identified ten QTL on chromosomes 5, 7, 9, 10, 13, and 20 which were able to explain anywhere from 5-14% of a particular amino acids total phenotypic variation. These QTL were associated with amino acid's in an RIL population developed from a cross of Essex × Williams 82 (Fallen et al., 2013). Further efforts in characterizing genomic regions for the protein quality trait of interest should continue.

Soybean Oil

The value of soybeans is traditionally determined by their oil and protein content (Charron et al., 2005). The economic importance of soybean was primarily in its oil prior to 1950; however, commercial interest has broadened to include protein meal, a by-product of soybean oil extraction (Chung et al., 2003). When soybeans are processed, they are cleaned, cracked, dehulled, and rolled into flakes causing the rupturing of oil cells. The two products, oil and protein meal, are separated in this process allowing for extraction of the oil (Soy Stats, 2018).

Approximately 20% of the total composition of a soybean seed is oil, which is used in many products (Eskandari et al., 2013a). Products containing soybean oil have both industrial and food applications such as: hair products, biodiesel, ink, plastic, salad dressing, vegetable oil, and mayonnaise (Yesudas et al., 2013). An increase in soybean oil would greatly benefit the aforementioned industries; however, other targeted goals must be kept in mind in soybean production. Oil content of soybean is a quantitative trait that is controlled by many genes (Qi et al., 2011). Extensive research has shown there is a negative genetic correlation between soybean seed oil and protein, and protein and yield (Brummer et al., 1997; Chung et al., 2003; Cober and Voldeng, 2000; Eskandari et al., 2013b; Hyten et al., 2004; Krishnan et al, 2007; Lee et al., 2010; Phansak et al., 2016; Wilcox, 1998; Yaklich, 2001). There is also a well-documented positive relationship between soybean oil and yield (Morrison et al., 2008). For

this reason, increases in soybean oil and yield must be pursued while maintaining adequate protein content and quality.

It is important to note that using marker assisted selection and the identification of QTL, such as in the study by Brummer et al. (1997), could aid in increasing soybean oil and protein levels, while simultaneously increasing soybean seed yield. More recently, SNP genotyping has led to the detection of six soybean oil QTL in an RIL population derived from a cross of PI 43838913 × Hamilton, which was reported by Akond et al. (2012). Additionally, in a study by Akond et al. (2014), 11 oil content QTLs were detected in an RIL population derived from a cross of MD96-5722 × Spencer by developing a high density map using 5376 SNP markers from the Illumina Infinium BeadChip array. Liu et al. (2017) were able to identify five QTL for oil content using SNP genotyping methods on a recombinant inbred line (RIL) population created by a cross of Zhonghuang 25 × Huaxia 3. Identifying QTL in soybean is an important task, and further efforts in characterizing genomic regions for the oil quality trait of interest should continue.

Soybean Yield

An important goal in soybean breeding has been to develop lines with increased yield, while maintaining seed quality traits such as protein and oil content. Soybean yield is a quantitative trait that is controlled by a combination of many genes and environmental factors (Liu et al., 2017; Song et al., 2017). For this reason, it is difficult to pinpoint exactly which genes lend to an increase

in soybean seed yield. There are a limited number of soybean ancestors that lent genes to the genetic base of current North American soybean cultivars (Carter et al., 2004). In North American soybean cultivars, more than 80% of genes can be traced to 17 ancestral plant introductions and their early progeny (Kim et al., 2012). Most breeding for yield improvements in North America has been done using crosses between elite germplasm instead of exotic or wild relatives, which are important sources of new genes (Carter et al., 2004). However, breeders have had difficulty improving yield through crossing with plant introductions as they typically have lower average yield when compared with elite breeding lines. Additionally, there are unreliable methods of predicting if plant introductions carry alleles linked to increasing yield (Kim et al., 2012). Despite this, there has been success creating experimental lines derived from exotic germplasm that outperform established cultivars in yield, which suggests that exotic lines do possess yield genes (Kim et al., 2012).

Newer technologies utilizing SNP genotyping, as will be used in this study, have been successful in the identification of many QTL associated with yield (Liu et al., 2017). For example, Kim et al. (2012) reported three QTL for seed yield that were identified using 1536 SNP markers to analyze two backcrossed soybean populations for seed yield. The first population was created using Elgin as the recurrent parent and PI 436684 as the donor parent, and the second was created using Williams 82 as the recurrent parent and PI 90566-1 as the donor parent (Kim et al., 2012). Additionally, 47 QTL associated with yield related traits

such as plant height, 100-seed weight, number of nodes, number of branches, and number of effective and invalid pods were identified using SNP genotyping in an RIL population created from a cross of Zhonghuang 25 × Huaxia 3 (Liu et al., 2017). Fallen et al. (2015) used an RIL population developed from a cross between Essex × Williams 82 to identify 23 yield QTL based on composite interval mapping and 21 yield QTL using single factor ANOVA. In this study, each RIL was genotyped using 50,000 SNP markers of which 17,232 were polymorphic across the population, and each QTL found was able to explain from 4.5-11.9% of phenotypic variation for yield (Fallen et al., 2015). Further efforts to implement selection of desirable allele forms into high yielding cultivars should continue.

Quantitative Trait Loci

Traits controlled by many genes are known as quantitative traits (Bernardo, 2014; Collard et al., 2005), and most agriculturally important traits are controlled by many genes. Quantitative traits in soybean include but are not limited to yield, disease resistance, and seed protein and oil content. QTL are regions within a genome that contain genes that influence particular quantitative traits (Collard et al., 2005). Identification of QTLs through phenotypic evaluation only is not possible (Collard et al., 2005). In order to locate, identify, and characterize these regions within a genome, DNA or molecular markers are used. DNA markers reveal sites of variation in DNA and have been used to construct linkage maps in soybean, which are used to identify specific regions of

chromosomes containing genes controlling both qualitative and quantitative traits through QTL analysis (Collard et al., 2005). QTL mapping is the process of constructing linkage maps and conducting QTL analysis, and leads to identifying novel QTL in the genome (Collard et al., 2005).

Marker-assisted selection (MAS) in plant breeding is a tool that uses DNA markers that are tightly linked or gene tagged to important agronomic traits (Collard et al., 2005). MAS acts as a substitute for or aids in phenotypic selection by using the presence or absence of a known marker, making it more efficient, reliable, and cost-effective compared to conventional plant breeding methods (Collard et al., 2005). For this reason, DNA markers are accepted as valuable tools for the improvement of agronomic and seed quality traits of soybeans.

Recombinant inbred lines are used to map QTL (Pollard, 2012). When two parental lines, selected based on phenotype, marker availability, and compatibility, are crossed, they create recombinants in the F_2 population which lend to RILs with further inbreeding. These RILs possess a genome that is a mosaic of the parental genomes and are inbred to isogenicity, creating genetically stable recombinant inbreds (Broman, 2005; Pollard, 2012). This results in a population that can be used for trait mapping and QTL analysis (Pollard, 2012).

A common challenge in the improvement of soybean protein, oil, and yield is the quantitative inheritance of these traits. A large number of QTL are reported

in the literature, suggesting that a large number of genes affect these three traits in soybean. The large number of QTL creates a challenge for plant breeders, as they are forced to select for a large number, and many of the QTL are not genetically linked and are not on the same chromosomes. For this reason, a priority should be to narrow the QTL pool and identify QTL with multiple effects. For example, identifying QTL that increase protein and either have no effect on or simultaneously increase soybean oil. Alternatively, identifying high protein QTL with no detriment to seed yield or oil could help breeders achieve 48% or more soybean meal protein. Upon identifying any of these seed quality trait QTL, breeders could incorporate them into high yielding cultivars in hopes of creating lines with stacked QTL.

The objectives of this study are to detect soybean genomic regions that increase protein content, while maintaining oil content and seed yield and to successfully identify soybean QTL. This will lead to a more valuable soybean for farmers, processors, and animal nutritionists.

Objectives

1. Evaluate agronomic traits of 138 field-tested RILs of soybean grown in six environments across Tennessee. Traits will include yield, days to maturity, plant height, and lodging.
2. Evaluate the seed quality traits of 138 field-tested recombinant inbred lines (RIL) of soybean grown in six environments across Tennessee using near-infrared reflectance (NIR) spectroscopy. Traits will include total protein, oil, and amino acid content.
3. Determine genotypes of 5601T × U99-310255 RIL using the Infinium BARCSoySNP6K BeadChip array.
4. Identify and locate important soybean protein, amino acid, oil, and yield QTL.

References

Akond, M., S. Liu, M. Boney, S.K. Kantartzi, K. Meksem, N. Bellaloui, D.A. Lightfoot, M.A. Kassem. 2014. Identification of quantitative trait loci (QTL) underlying protein, oil, and five major fatty acids contents in soybean. Am. J. Plant Sci. 5: 158-167.

Akond, M., B. Ragin, R. Bazzelle, and S.K. Kantartzi. 2012. Quantitative trait loci associated with moisture, protein, and oil content in soybean [*Glycine max* (L.) Merr.]. J. Agric. Sci. 4: 16-25.

Aoyagi, A. and W. Shurtleff. 2004. History of world soybean production and trade- Part 1. Soyinfo Center. http://www.soyinfocenter.com/HSS/production_and_trade1.php

Araújo, S.S., S. Beebe, M. Crespi, B. Delbreil, E.M. González, V. Gruber, I. Lejeune-Henaut, W. Link, M.J. Monteros, E. Prats, I. Rao, V. Vadez, and M.C. Vaz Patto. 2015. Abiotic stress responses in legumes: strategies used to cope with environmental challenges. Crit. Rev. Plant Sci. 34:237-280.

Baker, K.M., P.L. Utterback, C.M. Parsons, and H.H. Stein. 2011. Nutritional value of soybean meal produced from conventional, high-protein, or low-oligosaccharide varieties of soybeans and fed to broiler chicks. Poult. Sci. 90: 390-395.

Bernardo, R. Genetic basis of plant improvement. 2014. p. 23-69. Essentials of plant breeding. Stemma Press, Woodbury, MN.

Brim, C.A., and J.W. Burton. 1979. Recurrent selection in soybeans. II. Selection for increased percent in protein in seeds. Crop Sci. 19:494-498.

Broman, K.W. 2005. The genomes of recombinant inbred lines. Genetics. 169: 1133-1146.

Brummer, E.C., G.L. Graef, J. Orf, J.R. Wilcox, and R.C. Shoemaker. 1997. Mapping QTL for seed protein and oil content in eight soybean populations. Crop Sci. 37: 370-378.

Carter, T.E. Jr., R.L. Nelson, C.H. Sneller, and C. Zhanglin. 2004. Genetic diversity in soybean, p. 303-416. In: H.R. Boerma and J.E. Specht (eds.). Soybeans: improvement, production, and uses. 3rd ed. ASA Inc., CSSA Inc., SSSA Inc., Madison, WI.

Casteel, S. 2011. Soybean physiology: How well do you know soybeans? Purdue University Agronomy Extension. https://www.agry.purdue.edu/ext/soybean/Arrivals/10SoyDevt.pdf

Charron, C.S., F.L. Allen, R.D. Johnson, V.R. Pantalone and C.E. Sams. 2005. Correlations of oil and protein with isoflavone concentration in soybean [Glycine max (L.) Merr.]. J. Agric. Food Chem. 53: 7128-7135.

Chung, J., H.L. Babka, G.L. Graef, P.E. Staswick, D.J. Lee, P.B. Cregan, R.C. Shoemaker, and J.E. Specht. 2003. The seed protein, oil and yield QTL on soybean linkage group I. Crop Sci. 43: 1053-1067.

Cober, E.R. and H.D. Voldeng. 2000. Developing high-protein, high-yield soybean populations and lines. Crop Sci. 40:39-42.

Collard, B.C.Y., Jahufer, M.Z.Z., Brouwer, J.B., and Pang, E.C.K. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica. 142: 169-196.

Diers, B.W., P. Keim, W.R. Fehr, and R.C. Shoemaker. 1992. RFLP analysis of soybean seed protein and oil content. Theor. Appl. Genet. 83: 608-612.

Eskandari, M., E.R. Cober, and I. Rajcan. 2013a. Genetic control of soybean seed oil: I. QTL and genes associated with seed oil concentration in RIL populations derived from crossing moderately high-oil parents. Theor. Appl. Genet. 126:483-495.

Eskandari, M., E.R. Cober, and I. Rajcan. 2013b. Genetic control of soybean seed oil: II. QTL and genes that increase oil concentration without decreasing protein or with increased yield. Theor. Appl. Genet. 126:1677–1687.

Fallen, B.D., C.N. Hatcher, F.L. Allen, D.A. Kopsell, A.M. Saxton, P. Chen, S.K. Kantartzi, P.B. Cregan, D.L. Hyten, and V.R. Pantalone. 2013. Soybean seed amino acid content QTL detected using the universal soy linkage panel 1.0 with 1,536 SNPs. Journal of Plant Genome Sciences 1: 68-79.

Fallen, B.D., F.L. Allen, D.A. Kopsell, A.M. Saxton, L. McHale, J.G. Shannon, S.K. Kantartzi, A.J. Cardinal, P.B. Cregan, D.L. Hyten, and V.R. Pantalone. 2015. Selective genotyping for marker assisted selection strategies for soybean yield improvement. Plant Genetics, Genomics, and Biotechnology 2: 95-119.

Friedman, M. and D.L. Brandon. 2001. Nutritional and health benefits of soy proteins. J. Agric. Food Chem. 49: 1069-1086.

Hernández-Sebastiá, C., F. Marsolais, C. Saravitz, D. Israel, R.E. Dewey, and S.C. Huber. 2005. Free amino acid profiles suggest a possible role for asparagine in the control of storage-product accumulation in developing seeds of low- and high-protein soybean lines. J. Exp. Bot. 56:1951-1963.

Hymowitz, T. 2004. Speciation and cytogenetics, p. 97-136. In: H.R. Boerma and J.E. Specht (eds.). Soybeans: improvement, production, and uses. 3rd ed. ASA Inc., CSSA Inc., SSSA Inc., Madison, WI.

Hymowitz, T. 1990. Soybeans: The success story. P. 159-163. In: J. Janick and J.E. Simon (eds.). Advances in new crops. Timber Press, Portland, OR.

Hyten, D.L., V.R. Pantalone, C.E. Sams, A.M. Saxton, D. Landau-Ellis, T.R. Stefaniak, and M.E. Schmidt. 2004. Seed quality QTL in a prominent soybean population. Theor. Appl. Genet. 109: 552-561.

Joshi, T., B. Valliyodan, J.H. Wu, S.H. Lee, D. Xu, and H.T. Nguyen. 2013. Genomic differences between cultivated soybean, *G. max* and its wild relative *G. soja*. BMC Genomics. 14(Suppl 1):S5.

Kim, K.S., B.W. Diers, D.L. Hyten, M.A.R. Mian, J.G. Shannon, and R.L. Nelson. 2012. Identification of positive yield QTL alleles from exotic soybean germplasm in two backcross populations. Theor. Appl. Genet. 125: 1353-1369.

Khandaker, L., M. Akond, S. Liu, S.K. Kantartzi, K. Meksem, N. Bellaloui, D.A. Lightfoot, and M.A. Kassem. 2015. Mapping of QTL associated with seed amino acids content in “MD96-5722” by “Spencer” RIL population of soybean using SNP markers. Food and Nutrition Sciences. 6: 974-984.

Krishnan, H.B., S.S. Natarajan, A.A. Mahmoud, and R.L. Nelson. 2007. Identification of glycinin and β-conglycinin subunits that contribute to the increased protein content of high-protein soybean lines. J. Agric. Food Chem. 55:1839-1845.

Lee, J-D., G. Shannon, and M-G. Choung. 2010. Selection for protein content in soybean from single F_2 seed by near infrared reflectance spectroscopy. Euphytica. 172: 117-123.

Lin, J., V.R. Pantalone, G. Li, and F. Chen. 2011. Molecular cloning and biochemical characterization of an endo-β-mannanase gene from soybean for soybean meal improvement. J. Agric. Food Chem. 59: 622-628.

Liu, N., M. Li, X. Hu, Q. Ma, Y. Mu, Z. Tan, Q. Xia, G. Zhang, and H. Nian. 2017. Construction of high-density genetic map and QTL mapping of yield-related and two quality traits in soybean RILs population by RAD-sequencing. BMC Genomics. 18: 466-479.

Miller-Garvin, J., and S.L. Naeve. 2015. United States Soybean Quality Annual Report 2015. http://ussec.org/wp-content/uploads/2016/01/2015-Final-US-Commodity-Soybean-Quality-Report.pdf

Morrison, M.J., E.K. Cober, M.F. Saleem, N.B. McLaughlin, J. Fregeau-Reid, B.L. Ma, W. Yan, and L. Woodrow. 2008. Changes in isoflavone concentration with 58 years of genetic improvement of short-season soybean cultivars in Canada. Crop Sci. 48: 2201–2208.

North Carolina Soybean Producers Association, Inc. 2014. Uses of soybeans. http://www.ncsoy.org/ABOUT-SOYBEANS/Uses-of-Soybeans.aspx

United States Department of Agriculture, National Agricultural Statistics Service (USDA, NASS). 2019. National Statistics for Soybeans. https://www.nass.usda.gov/Statistics _by_Subject/result.php?87D31FA7-F2D0-3139-A36A-0ED237720D5F&sector=CROPS &group=FIELD%20CROPS&comm=SOYBEANS

Pantalone, V.R. 2012. Modern breeding approaches for enhancing soybean protein quality. In: R.F. Wilson, editor, Designing soybean for 21st century markets. AOCS Press, Urbana, IL. p. 197-226.

Pantalone, V.R. and C.J. Smallwood. 2018. Registration of 'TN11-5102' soybean cultivar with high yield and high protein meal. Journal of Plant Registrations. 12: 304-308.

Pantalone, V., M. Cunicelli, and C. Wyman. 2020. Registration of soybean cultivar TN15-5007 with high meal protein. Journal of Plant Registrations. 14: in press.

Panthee, D.R., V.R. Pantalone, A.M. Saxton, D.R. West, and C.E. Sams. 2006a. Genomic regions associated with amino acid composition in soybean. Molecular Breeding. 17: 79-89.

Panthee, D.R., V.R. Pantalone, C.E. Sams, A.M. Saxton, D.R. West, J.H. Orf, and A.S. Killam. 2006b. Quantitative trait loci controlling sulfur containing amino acids, methionine and cysteine, in soybean seeds. Theor. Appl. Genet. 112: 546-553.

Phansak, P., W. Soonsuwon, D.L. Hyten, Q. Song, P.B. Cregan, G.L. Graef, and J.E. Specht. 2016. Multi-population selective genotyping to identify soyean [*Glycine max* (L.) Merr.] seed protein and oil QTLs. G3-Genes. Genom. Genet. 6: 1635-1646.

Pollard, D.A. Design and construction of recombinant inbred lines. 2012. Methods Mol Biol. 871:31-39.

Qi, Z., Q. Wu, X. Han, Y. Sun, X. Du, C. Liu, H. Jiang, G. Hu, and Q. Chen. 2011. Soybean oil content QTL mapping and integrating with meta-analysis method for mining genes. Euphytica. 179: 499-514.

Smallwood, C.J., A.M. Saxton, J.D. Gillman, H.S. Bhandari, P.A. Wadl, B.D. Fallen, D.L. Hyten, Q. Song, and V.R. Pantalone. 2019. Context-specific genomic selection strategies outperform phenotypic selection for soybean quantitative traits in the progeny row stage. Crop Sci. 59: 54-67.

Song, Q., L. Yan, C. Quigley, B.D. Jordan, E. Fickus, S. Schroeder, B. Song, Y. C. An, D. Hyten, R. Nelson, K. Rainey, W.D. Beavis, J. Specht, B. Diers, and P. Cregan. 2017. Genetic characterization of the soybean nested association mapping population. Plant Genome. 10: 1-14.

Soy Stats. 2018. 2017 Soystats: A Reference Guide to Important Soybean Facts and Figures. The American Soybean Association. http://soystats.com/2017-soystats/

Thakur, M. and C.R. Hurburgh. 2007. Quality of US soybean meal compared to the quality of soybean meal from other origins. J. Am. Oil Chem. Soc. 84:835-843.

Warrington, C.V., H. Abdel-Haleem, D.L. Hyten, P.B. Cregan, J.H. Orf, A.S. Killam, N. Bajjalieh, Z. Li, and H.R. Boerma. 2015. QTL for seed protein and amino acids in the Benning x Danbaekkong soybean population. Theor. Appl. Genet. 128:2474.

Warrington, C.V., H. Abdel-Haleem, J.H. Orf, A.S. Killam, N. Bajjaleih, Z. Li, and H.R. Boerma. 2014. Resource allocation for selection of seed protein and amino acids in soybean. Crop Sci. 54:963-970.

Wilcox, J.R. 1998. Increasing seed protein in soybean with eight cycles of recurrent selection. Crop Sci. 38: 1536-1540.

Wilcox, J.R. and R.M. Shibles. 2001. Interrelationships among seed quality attributes in soybean. Crop Sci. 41: 11-14.

Yaklich, R.W. 2001. β-Conglycinin and glycinin in high-protein soybean seeds. J. Agric. Food Chem. 49:729-735.

Yesudas, C.R., R. Bashir, M.B Geisler, and D.A. Lightfoot. 2013. Identification of germplasm with stacked QTL underlying seed traits in an inbred soybean population from cultivars Essex and Forrest. Mol. Breeding. 31: 693-703.

CHAPTER I: EVALUATION OF AGRONOMIC TRAITS OF 138 $F_{8:10}$ AND $F_{8:11}$ RECOMBINANT INBRED LINES

Abstract

Soybean [*Glycine max* (L.) Merr.] is the world's leading source of vegetable oil and high quality protein meal. However, soybean protein and oil and protein and yield are negatively genetically correlated. Agronomic traits of soybean lines are important because they lead to increased soybean value. The objective of this study was to evaluate the agronomic traits of 138 field-tested RILs of soybean grown in six environments across Tennessee. Agronomic traits evaluated include yield, days to maturity, plant height, and lodging. A population of 138 $F_{8:10}$ (2017) and $F_{8:11}$ (2018) recombinant inbred lines (RILs) of soybean was created from the crossing of 5601T × U99-310255. This cross is unique because both parents display higher protein with high yields and favorable oil. The experimental population was grown over the 2017 and 2018 field seasons in six environments with three replications each. Notes on agronomic traits were taken on all plots at all locations. Yield and moisture were recorded for all plots after maturation at the time of combine harvest. The yield ranged from 2461.0-4768.5 kg ha^{-1} for the experimental population. Additionally, the heritability estimate on an entry means basis for yield was 0.82, which is high with regard to genetic influence. The range of maturity was 129.2-145.4 DAP. Height and lodging were recorded for all plots prior to combine harvest. Height in the experimental population ranged from 62.5-131.0 cm. Lodging scores ranged from 1.9-4.8.

Introduction

Agronomic traits such as yield, maturity, height, and lodging associated with a particular line of soybean are of high importance to plant breeders, as they focus their attention on improving the agronomic traits of soybeans to increase soybean value and to satisfy farmer preferences. Soybean lines lacking high values for agronomic traits, will not be selected for further development into cultivars. Seed yield plays a major role in influencing farmers to grow a specific cultivar. Along with yield, days to maturity, plant height, and lodging provide a profile for cultivar selection in soybean. Farmers are likely to grow cultivars that have good values in all these categories.

Soybean characteristics are easily broken up into two genetic categories, qualitative and quantitative traits. Qualitative traits are controlled by one to a few genes and often follow Mendelian inheritance (Bernardo, 2014a; Bernardo, 2014b). Some examples of qualitative traits are flower color or pubescence color, which are not typically affected by environmental conditions. Quantitative traits are controlled by many genes, sometimes at different loci or on different chromosomes with differing levels of additive, dominance, and epistatic effects (Bernardo, 2014a; Bernardo, 2014b). Quantitative traits are more likely to be influenced by environmental conditions than qualitative traits (Bernardo, 2014a; Bernardo, 2014b). Quantitative traits are also much more difficult to identify and do not exhibit the same clear-cut differences in how traits are expressed as qualitative traits do (Bernardo, 2014a). Some examples of quantitative traits are the economically important seed protein and oil concentrations and seed yield

(Diers et al., 1992). Hernández-Sebastiá et al. (2005) suggest that there is a direct relationship between seed composition and economic value of the soybeans, suggesting specifically that much of the economic value in soybean right now is due to the protein concentration. This experiment analyzes an RIL population for agronomic traits.

Materials and Methods

Plant Materials

This study features an RIL population from the cross 5601T (Pantalone et al., 2003) × U99-310255. This is a unique cross as both parents display higher protein with high yields and favorable oil. The first parent, 5601T, is a registered cultivar that belongs to maturity group V and has a determinate growth habit, white flowers, and gray pubescence (Pantalone et al., 2003). It was chosen as a parent of this cross due to its high yield potential and above average protein concentration. The second parent, U99-310255, is an unreleased line from the University of Nebraska. It belongs to maturity group III and has an indeterminate growth habit, purple flowers, and gray pubescence. It was chosen as a parent of this cross due to its above average protein concentration as well as its good oil concentration.

The experimental lines were grown in a replicated field trial across Tennessee (TN) in the summers of 2017 and 2018. These lines were grown in Knoxville, Springfield, and Milan, TN. A population of 138 RIL were created in 2006, when the initial cross of 5601T × U99-310255 was made in Knoxville, TN,

at the East Tennessee Research and Education Center (ETREC) (Table 1.1 in Appendix A). Successful crossing led to F_1 seed, which were sent to a winter nursery in Isabela, Puerto Rico. The F_2 seed were harvested in Puerto Rico and planted at ETREC in the spring of 2007. The F_3 seed were harvested in the fall of 2007 at ETREC and were grown in Homestead, FL during the winter of 2007-2008. The F_4 seed were then harvested and planted at ETREC in spring of 2008. In fall of that same year, individual F_4 single plants were selected and harvested. In spring of 2009, the $F_{4:5}$ seed from each single plant were grown at ETREC as individual plant rows. In spring of 2010, the harvested $F_{4:6}$ seed were planted at two locations: ETREC and the Highland Rim Research and Education Center (HRREC) in Springfield, TN. In the spring of 2011, the $F_{4:7}$ seed were planted at ETREC, HRREC, and the Research and Education Center at Milan in Milan, TN (RECM) (Wiggins et al., 2018). In the spring of 2014, the $F_{4:8}$ seed was planted in a yield trial at ETREC. Single plants from this population were harvested in the fall of 2014. The $F_{8:9}$ seed were used to source the 2015 plant rows at ETREC. These plant rows were bulk harvested in the fall of 2015. In the spring of 2017, the $F_{8:10}$ seed were planted in a two replication field trial at ETREC, HRREC, and RECM, and the seed were combine harvested in fall 2017. This $F_{8:11}$ seed was used to source a three replication field trial planted in the spring of 2018 at ETREC, HRREC, and RECM (Table 1.1). The seed were combine harvested in the fall of 2018.

Agronomic Traits

The phenotypic data of the agronomic traits were collected from every plot in all six locations across the 2017 and 2018 field seasons. Each entry in the 2017 and 2018 field trials was planted in a single plot consisting of two adjacent rows 6.1 m in length, and the rows were spaced 0.8 m apart. Along with the RILs and one parent (5601T), four checks were included in the field trials. The checks were 'Ellis' (Pantalone et al., 2017) (MG IV-late), '5002T' (Pantalone et al., 2004) (MG V-early), 'TN11-5104' (MG V-early), and 'Osage' (Chen et al., 2007) (MG V-mid). In both 2017 and 2018 field trials, flower color was checked and recorded during the R_2 growth stage (Fehr and Caviness, 1977). Additionally, the maturity, plant height, lodging, and pubescence color were recorded at the R_8 growth stage (Fehr and Caviness, 1977). Maturity was determined as the number of days after planting until 95% of the soybean pods were mature. Plant height was recorded in centimeters (cm) and was measured from the soil surface to the top of the main stem. Lodging was scored on a scale of 1-5, where 1 is completely upright and 5 is prostrate. Additionally, seed yield was recorded by the combine at harvest and was converted to kg ha^{-1}.

Data Analysis

The experimental design was a randomized complete block model. Data on agronomic and seed quality traits were analyzed using a mixed model analysis of variance using SAS PROC GLIMMIX for SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) to estimate least squares means for the 138 RILs averaged over 2017-2018. For the purpose of analysis, fixed effects were

individual lines. Random effects were location and GxE or location by line interaction. Tukey's LSD was used to compare lines within each test. It was chosen over Fisher's protected LSD method because it offers full control of the experiment wise type I error rate.

A broad sense estimate of heritability on an entry means basis for each trait of interest in the population was calculated as follows (Nyquist, 1991):

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \left(\frac{\sigma_{ge}^2}{e}\right) + \left(\frac{\sigma^2}{re}\right)}$$

where h^2 is the heritability, σ_g^2 is the genotypic variance, σ_{ge}^2 is genotype x environmental interaction variance, σ^2 is error variance, *e* is number of environments, and *r* is number of replications.

Results and Discussion

Soybean yield, height, and lodging are all traits taken into account when selecting cultivars for production. Frequency distributions of these traits' means are presented in Figures 1.1-1.4 in Appendix A. Seed yield was averaged over the 2017 and 2018 growing seasons with two replications in 2017 and three replications in 2018 at six locations for all 138 RIL, checks, and parent 5601T. The checks included in this study and their respective average yields were: 'Ellis' with 4781.0 kg ha^{-1}, '5002T' with 4306.6 kg ha^{-1}, 'TN11-5104' with 4576.0 kg ha^{-1}, and 'Osage' with 4261.3 kg ha^{-1}. The minimum yield across all plots in the population was 2461.0 kg ha^{-1}, and the maximum value was 4768.5 kg ha^{-1}. The

mean yield was 3748.1 kg ha^{-1} (Table 1.2). While 'Ellis' was numerically the highest yielding line, 56U99-092, 56U99-086, and 56U99-089 had similarly high yields and outperformed the remaining check lines and parental line '5601T' for yield (Table 1.3). There were 14 individual RILs whose average yield outperformed the lowest yielding check, 'Osage' (Table 1.4). The yield of these lines is able to compete with that of high yielding check cultivars. The heritability estimate for yield was high (0.82). The high heritability estimate found in this study suggests that genetic improvements can be made in soybean yield.

Maturity of the lines varied across all three locations. Maturity groups grown in Tennessee environments typically range from early group IV to group V. The checks included in this study had relative maturities (RM) of: 'Ellis' RM 4.9, '5002T' RM 5.0, 'TN11-5104' RM 5.2, and 'Osage' RM 5.5. Maturity was averaged over 2017 and 2018 at six locations with two replications in 2017 and three replications in 2018. The minimum maturity from the 2017 and 2018 field trials was 129.2 DAP, and the maximum was 145.4 DAP. The mean maturity was 137.8 DAP (Table 1.2). The heritability estimate for maturity was high (0.92). The high heritability estimate found in this study suggests that genetic improvements can be made in soybean maturity.

Plant height (cm) was averaged over the 2017 and 2018 growing seasons with two replications in 2017 and three replications in 2018 at six locations for all 138 RIL. The minimum height was 62.5 cm, and the maximum was 131.0 cm. The mean height was 101.8 cm (Table 1.2). Lodging and plant height typically

have a positive relationship, as tall plants lacking lodging resistance are often more susceptible to lodging than shorter plants. Lodging scores were averaged over the 2017 and 2018 growing seasons with two replications in 2017 and three replications in 2018 at six locations for all 138 RIL. The minimum lodging score was 1.9, and the maximum score was 4.8. The mean lodging score was a 3.1 (Table 1.2).

Yield, maturity, height, and lodging of soybeans are all agronomic traits considered when selecting cultivars for production. Difference at the stem termination locus in soybeans (Dt1) often produces noticeable variations in soybean plant height. An experiment by Wilcox and Sediyama (1981) showed that lodging of taller indeterminate plants increased as plant density increased; however, shorter determinate plants did not exhibit lodging under similar conditions. Additionally, they reported that all 73 determinate lines in their study were resistant to lodging, while 93 indeterminate lines ranged in their lodging scores (Wilcox and Sediyama, 1981). Lodging can also have negative effects on soybean seed yield, as severely lodged soybeans often have lower yields than upright soybeans (Wilcox and Sediyama, 1981).

References

Bernardo, R. Process of plant breeding. 2014a. p. 1-22. Essentials of plant breeding. Stemma Press, Woodbury, MN.

Bernardo, R. Genetic basis of plant improvement. 2014b. p. 23-69. Essentials of plant breeding. Stemma Press, Woodbury, MN.

Chen, P., C.H. Sneller, L.A. Mozzoni, and J.C. Rupe. 2007. Registration of 'Osage' soybean. Crop Sci. J. Plant Regist. 1: 89-92.

Diers, B.W., P. Keim, W.R. Fehr, and R.C. Shoemaker. 1992. RFLP analysis of soybean seed protein and oil content. Theor. Appl. Genet. 83: 608-612.

Fehr W.R., and C.E. Caviness. Stages of soybean development. Special Report, Agriculture and Home Economics Experiment Station, Iowa State University, 1977, issue 80, p 11.

Hernández-Sebastiá, C., F. Marsolais, C. Saravitz, D. Israel, R.E. Dewey, and S.C. Huber. 2005. Free amino acid profiles suggest a possible role for asparagine in the control of storage-product accumulation in developing seeds of low- and high-protein soybean lines. J. Exp. Bot. 56:1951-1963.

Nyquist, W.E. 1991. Estimation of heritability and prediction of selection response in plant populations. Critical Reviews in Plant Sciences. 10: 235-322.

Pantalone, V.R., C.R. Smallwood and B. Fallen. 2017. Development of 'Ellis' soybean with high soymeal protein, resistance to stem canker, southern rootknot nematode, and frogeye leaf spot. J. Plant Regist. 11: 250-255.

Pantalone, V.R., F.L. Allen, and D. Landau-Ellis. Registration of '5002T' soybean. 2004. Crop Sci. 44: 1483-1484.

Pantalone, V.R., F.L. Allen, and D. Landau-Ellis. Registration of '5601T' soybean. 2003. Crop Sci. 43: 1123-1124.

Wiggins, B., S. Wiggins, M. Cunicelli, C. Smallwood, F. Allen, D. West, and V. Pantalone. Genetic gain for soybean seed protein, oil, and yield in a recombinant inbred line population. 2018. Journal of the American Oil Chemists' Society. 96: 43-50.

Wilcox, J.R. and T. Sediyama. Interrelationships among height, lodging, and yield in determinate and indeterminate soybeans. 1981. Euphytica. 30: 323-326.

Appendix A

Table 1.1 Depiction of population development from inception of cross between 5601T and U99-310255 to summer 2018 field trial.

Year	Season	Location	Generation grown	Procedure
2006	Summer	ETREC[a]	Parentals	Cross P1 x P2
2006-2007	Winter	TARS[b]	F_1	Harvest individual F_1 plants
2007	Summer	ETREC	F_2	Pick one pod per plant
2007-2008	Winter	27 Farms[c]	F_3	Pick one pod per plant
2008	Summer	ETREC	F_4	Harvest single plants
2009	Summer	ETREC	$F_{4:5}$	Bulk harvest individual rows
2010	Summer	ETREC, HRREC[d]	$F_{4:6}$	Combine harvest individual plots
2011	Summer	ETREC, HRREC, RECM[e]	$F_{4:7}$	Combine harvest individual plots
2014	Summer	ETREC	$F_{4:8}$	Harvest single plants
2015	Summer	ETREC	$F_{8:9}$	Bulk harvest individual rows
2017	Summer	ETREC, HRREC, RECM	$F_{8:10}$	Combine harvest individual plots
2018	Summer	ETREC, HRREC, RECM	$F_{8:11}$	Combine harvest individual plots

[a]ETREC, East Tennessee Research and Education Center
[b]TARS, USDA Tropical Agriculture Research Station, Isabela, PR
[c]27 Farms, 27 Farms of Homestead, Inc., Homestead, FL
[d]HRREC, Highland Rim Research and Education Center
[e]RECM, Research and Education Center at Milan

Table 1.2 Minimum, mean, and maximum agronomic trait values for yield, maturity, height, and lodging for 138 $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 compared with four checks and a parent grown over six environments with three replications each in 2017 and 2018.

	Agronomic trait means			
	Yield	Maturity	Height	Lodging
	kg ha^{-1}	Days	cm	1-5 scale
Minimum	2461.0	129.2	62.5	1.9
Mean	3748.1	137.8	101.8	3.1
Maximum	4768.5	145.4	131.0	4.8
Checks[a]	4481.2	138.5	77.2	2.0
Parent[b]	4360.4	141.3	92.7	2.3
LSD (0.05)	647.4	3.6	10.8	0.8

[a] Checks included in field trials were Ellis, 5002T, TN11-5104, and Osage.
[b] Parent included in field trials was 5601T.

Table 1.3 Mean agronomic trait values for 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255, four checks, and one parent grown over six environments with three replications each in 2017 and 2018.

Line	Yield	Maturity	Height	Lodging
	kg ha^{-1}	Days	cm	1-5 scale
Ellis[a]	4781.0	138.3	74.0	1.6
5002T[a]	4306.6	139.1	75.5	2.3
TN11-5104[a]	4576.0	138.4	83.5	2.1
Osage[a]	4261.3	138.0	75.7	1.9
5601T[b]	4360.4	141.3	92.7	2.3
56U99-006	3852.5	138.8	74.8	2.7
56U99-007	3915.1	136.6	74.5	2.3
56U99-008	3724.8	133.9	70.6	2.1
56U99-011	4188.2	139.4	79.1	2.4
56U99-012	4204.7	134.1	74.5	2.9
56U99-013	3952.0	138.0	110.9	2.9
56U99-014	3722.0	140.6	112.1	4.5
56U99-015	3519.5	137.8	113.1	3.9
56U99-016	3659.9	141.1	103.2	2.5
56U99-017	3446.3	136.9	109.5	2.9
56U99-018	3732.6	136.7	119.2	4.0
56U99-019	3592.9	137.5	102.4	2.6
56U99-020	3083.7	129.2	104.9	2.8
56U99-021	3434.5	141.4	110.1	4.6
56U99-022	3105.7	135.5	100.6	4.1
56U99-023	3754.7	134.8	112.2	2.6
56U99-025	3557.4	137.3	109.3	2.8
56U99-026	3559.9	136.9	110.0	2.9
56U99-027	3750.4	134.2	109.2	2.7
56U99-028	4092.4	142.6	116.4	2.9
56U99-029	3625.5	138.2	74.9	2.7
56U99-030	3438.1	137.6	121.7	4.2
56U99-031	3270.9	137.4	69.6	2.4
56U99-032	3463.6	138.2	118.0	3.3
56U99-034	3910.8	134.3	110.8	2.8
56U99-035	4034.4	140.9	122.4	2.9
56U99-036	3024.2	137.3	114.4	4.0
56U99-037	3285.7	138.6	114.9	4.4

Table 1.3 Continued.

Line	Yield	Maturity	Height	Lodging
	kg ha^{-1}	Days	cm	1-5 scale
56U99-038	3388.3	136.1	101.8	2.5
56U99-039	3497.6	134.8	94.0	2.3
56U99-040	3885.8	134.3	116.1	2.5
56U99-041	3383.5	136.2	105.7	2.6
56U99-043	3711.6	140.1	76.4	2.2
56U99-044	3437.5	139.2	105.6	3.2
56U99-045	3342.8	139.0	100.1	3.1
56U99-046	2855.4	137.9	112.6	3.1
56U99-047	3162.2	136.0	117.0	3.7
56U99-048	3935.6	135.8	79.4	2.8
56U99-049	3494.4	139.5	113.7	4.2
56U99-052	3955.2	132.8	115.4	3.2
56U99-053	4141.2	133.4	114.2	1.9
56U99-054	3175.8	137.3	115.9	4.0
56U99-055	2856.9	135.9	119.5	4.6
56U99-056	4066.7	140.1	105.6	2.3
56U99-057	3153.4	138.9	106.7	4.5
56U99-061	3806.0	140.9	75.6	2.2
56U99-063	3153.3	139.5	114.1	4.1
56U99-067	2903.2	133.9	120.6	3.7
56U99-068	3656.7	136.9	78.0	2.7
56U99-069	4106.8	139.5	111.1	2.9
56U99-071	3411.2	137.2	108.4	3.2
56U99-075	4043.6	137.6	75.4	2.3
56U99-076	3580.2	138.0	106.9	2.3
56U99-079	3594.5	138.1	121.0	3.7
56U99-080	3501.1	142.4	111.1	3.4
56U99-081	3745.7	137.6	116.9	3.7
56U99-082	4040.1	142.5	116.1	2.9
56U99-086	4679.0	136.2	114.1	2.4
56U99-087	3565.5	143.1	111.4	3.8
56U99-089	4645.1	135.5	112.1	2.4
56U99-092	4768.5	131.8	115.0	2.5
56U99-093	3696.3	132.3	64.5	1.9
56U99-095	3051.6	132.9	109.4	3.4
56U99-096	3802.3	136.0	117.2	2.7

Table 1.3 Continued.

Line	Yield	Maturity	Height	Lodging
	kg ha^{-1}	Days	cm	1-5 scale
56U99-097	3499.4	133.9	75.6	2.3
56U99-099	2461.0	133.1	99.5	4.2
56U99-100	4118.0	140.3	122.8	3.3
56U99-102	3588.7	133.7	64.9	2.4
56U99-104	3747.4	138.8	116.3	3.0
56U99-105	3815.6	134.0	62.5	1.9
56U99-107	3911.8	138.2	73.3	1.9
56U99-108	3064.3	134.7	101.2	4.4
56U99-109	3987.1	140.1	121.3	3.8
56U99-111	4079.8	137.5	75.7	2.1
56U99-115	3474.5	134.1	109.1	3.0
56U99-118	3401.8	139.8	114.5	3.5
56U99-121	3516.4	135.6	108.1	3.7
56U99-123	3577.0	137.7	84.1	2.8
56U99-125	4158.0	141.5	80.5	2.1
56U99-126	4074.3	138.2	84.4	2.9
56U99-130	4333.7	142.3	75.3	2.2
56U99-131	3366.3	131.6	68.7	2.2
56U99-133	3958.2	137.4	121.4	2.7
56U99-135	3335.4	134.5	67.6	2.5
56U99-136	4269.8	140.8	79.7	2.3
56U99-137	4116.7	143.9	85.9	3.2
56U99-139	3847.1	140.4	69.2	2.4
56U99-138	3011.7	137.1	115.2	4.6
56U99-146	3979.2	139.9	114.5	3.0
56U99-151	3450.1	138.1	112.8	4.2
56U99-152	3706.0	139.4	112.6	3.3
56U99-154	3843.5	145.0	119.0	3.8
56U99-155	4411.7	142.3	129.3	3.0
56U99-157	4225.2	143.5	131.0	3.7
56U99-159	3828.8	143.1	130.9	3.9
56U99-162	3433.2	136.8	82.9	3.1
56U99-164	4327.3	134.1	68.4	2.5
56U99-165	3291.0	135.2	112.1	4.8
56U99-166	3703.3	131.8	107.1	2.5
56U99-168	3656.5	138.4	122.2	4.2
56U99-170	3804.8	145.4	113.5	4.0

Table 1.3 Continued.

Line	Yield	Maturity	Height	Lodging
	kg ha^{-1}	Days	cm	1-5 scale
56U99-175	3922.3	138.5	116.3	3.5
56U99-176	4280.0	135.6	107.2	2.1
56U99-177	3874.3	140.6	107.4	3.4
56U99-179	3790.8	139.7	78.5	2.4
56U99-181	4110.0	135.3	112.9	2.7
56U99-183	3634.5	140.8	83.7	3.2
56U99-184	4065.3	132.3	81.1	2.6
56U99-185	4260.7	135.8	66.9	2.0
56U99-187	4059.8	139.5	128.1	3.3
56U99-189	3695.7	137.4	122.5	4.1
56U99-190	3441.6	133.7	111.6	3.6
56U99-191	3973.3	136.7	118.4	3.8
56U99-199	4410.0	138.9	118.7	2.8
56U99-200	3594.9	139.0	116.4	3.8
56U99-201	3323.2	136.0	103.0	3.3
56U99-202	4154.9	133.9	74.3	2.9
56U99-203	4251.8	139.0	77.3	2.3
56U99-206	4209.1	139.6	112.2	2.2
56U99-207	3945.2	135.2	101.3	2.7
56U99-208	4026.9	135.2	65.2	2.2
56U99-211	3607.5	139.8	108.8	3.2
56U99-215	4249.9	139.4	118.1	3.2
56U99-216	3797.2	143.2	111.6	3.0
56U99-219	4335.9	142.0	81.3	2.2
56U99-220	3272.5	139.2	96.9	3.5
56U99-221	3992.7	138.2	115.0	3.4
56U99-222	4301.5	138.7	77.2	3.1
56U99-223	4311.7	135.8	109.3	3.0
56U99-227	4177.6	137.7	75.8	2.0
56U99-228	4366.6	141.0	104.4	2.8
56U99-232	3767.5	139.2	115.7	3.9
56U99-233	2908.3	137.1	113.3	4.2
56U99-235	3409.6	138.0	106.5	3.5
56U99-237	3908.3	142.5	118.9	3.4
56U99-238	4109.5	137.6	115.0	3.4
56U99-242	4279.9	138.7	73.0	2.2
56U99-243	3839.2	138.4	110.9	2.4

Table 1.3 Continued.

Line	Yield	Maturity	Height	Lodging
	kg ha^{-1}	Days	cm	1-5 scale
56U99-244	3860.5	138.7	104.8	2.3
LSD (0.05)	647.4	3.6	10.8	0.8

[a] Check line

[b] Parental line 5601T. Seeds of parental line U99-310255 were unavailable.

Table 1.4 Mean agronomic trait values for 14 high yielding individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255, four checks, and one parent grown over six environments with three replications each in 2017 and 2018.

Line	Yield	Maturity	Height	Lodging
	kg ha^{-1}	Days	cm	1-5 scale
Ellis[a]	4781.0	138.3	74.0	1.6
56U99-092	4768.5	131.8	115.0	2.5
56U99-086	4679.0	136.2	114.1	2.4
56U99-089	4645.1	135.5	112.1	2.4
TN11-5104[a]	4576.0	138.4	83.5	2.1
56U99-155	4411.7	142.3	129.3	3.0
56U99-199	4410.0	138.9	118.7	2.8
56U99-228	4366.6	141.0	104.4	2.8
5601T[b]	4360.4	141.3	92.7	2.3
56U99-219	4335.9	142.0	81.3	2.2
56U99-130	4333.7	142.3	75.3	2.2
56U99-164	4327.3	134.1	68.4	2.5
56U99-223	4311.7	135.8	109.3	3.0
5002T[a]	4306.6	139.1	75.5	2.3
56U99-222	4301.5	138.7	77.2	3.1
56U99-176	4280.0	135.6	107.2	2.1
56U99-242	4279.9	138.7	73.0	2.2
56U99-136	4269.8	140.8	79.7	2.3
Osage[a]	4261.3	138.0	75.7	1.9
LSD (0.05)	647.4	3.6	10.8	0.8

[a] Check line

[b] Parental line 5601T. Seeds of parental line U99-310255 were unavailable.

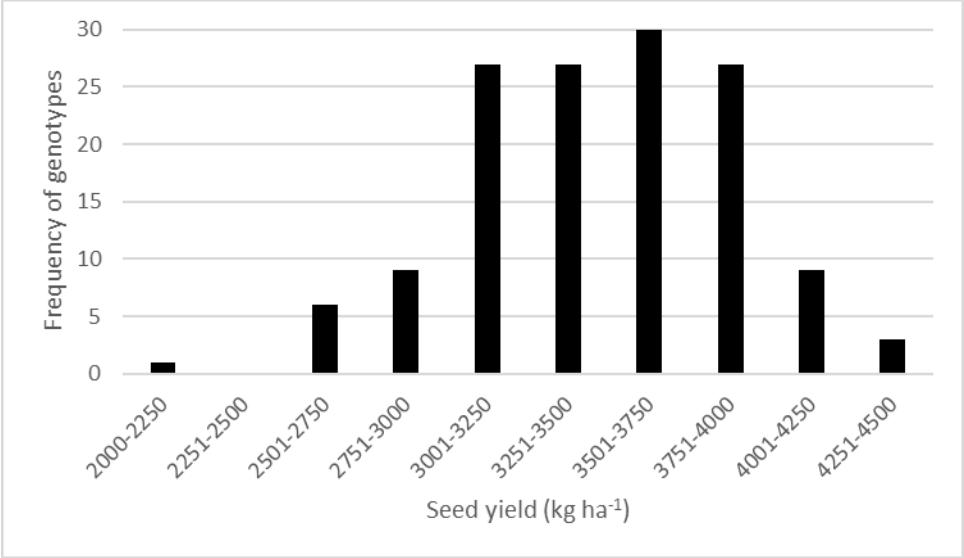

Figure 1.1 Frequency distributions for yield means for a population of 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 grown over six environments with three replications each in 2017 and 2018.

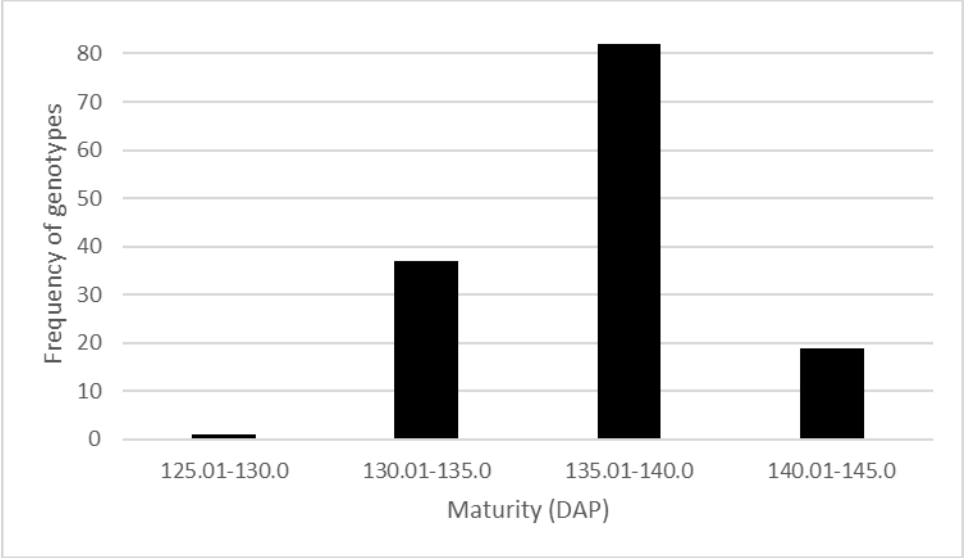

Figure 1.2 Frequency distributions for maturity means for a population of 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 grown over six environments with three replications each in 2017 and 2018.

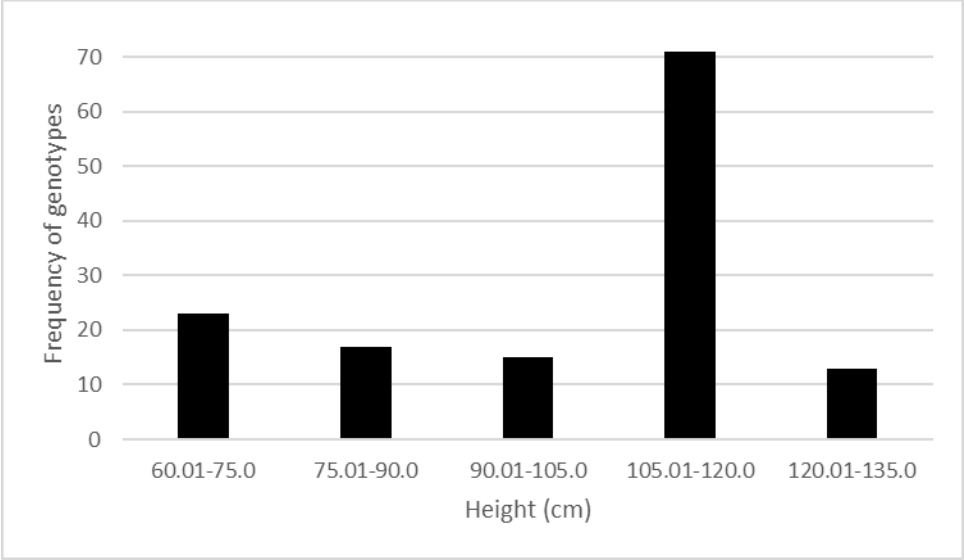

Figure 1.3 Frequency distributions for height means for a population of 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 grown over six environments with three replications each in 2017 and 2018.

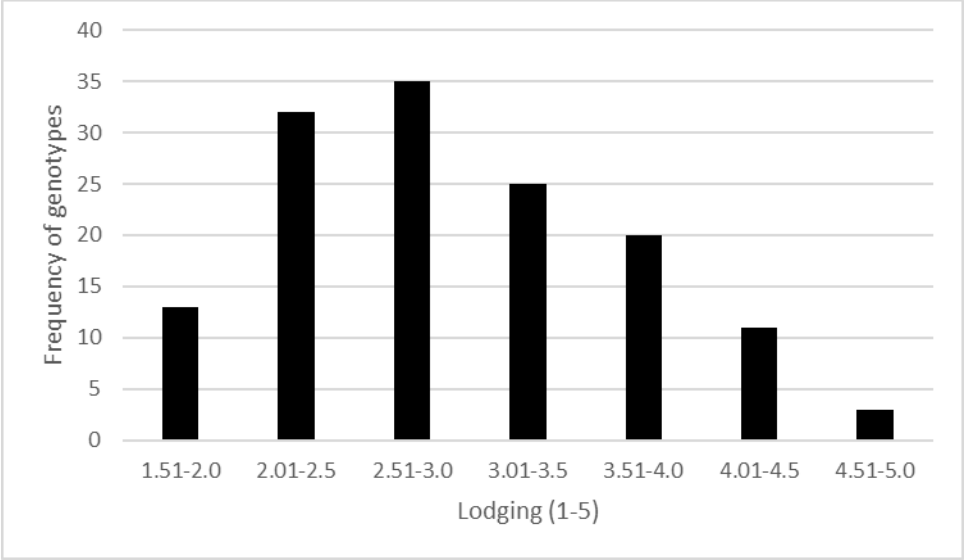

Figure 1.4 Frequency distributions for lodging means for a population of 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 grown over six environments with three replications each in 2017 and 2018.

CHAPTER II: EVALUATION OF SEED QUALITY TRAITS OF 138 $F_{8:10}$ AND $F_{8:11}$ RECOMBINANT INBRED LINES

Abstract

A population of 138 $F_{8:10}$ (2017) and $F_{8:11}$ (2018) recombinant inbred lines (RILs) of soybean was created from a cross between 5601T × U99-310255. This cross is unique because both parents display higher protein with high yields and favorable oil. The RILs were grown in a replicated six environment field trial across Tennessee in the 2017 and 2018 growing seasons. The objectives of this study were to evaluate the seed protein, oil, meal protein, and amino acid content of the RILs using near-infrared reflectance (NIR) spectroscopy. In total, five amino acids: cysteine, methionine, tryptophan, lysine, and threonine were analyzed. Protein, oil, and meal protein are reported in g kg^{-1}. Amino acids are reported as g amino acid (aa) kg^{-1} crude protein (cp). Measurements for seed quality traits including protein, oil, and the five amino acids of interest were taken on all plots at all locations. The protein content ranged from 408.0-446.4 g kg^{-1} for the experimental population. The oil content ranged from 205.3-229.2 g kg^{-1}. The meal protein content ranged from 473.3-518.0 g kg^{-1}. Additionally, heritability estimates on an entry means basis for protein (0.97), oil (0.96), and meal protein (0.90) varied by trait with regard to genetic influence. All experimental lines outperformed ‘Ellis’, ‘5002T’, and ‘TN11-5104’check lines, and all experimental lines with the exception of 56U99-097 outperformed parental line ‘5601T’ for meal protein content.

Introduction

Soybean is the leading source of oil and high quality protein meal in the world (Aoyagi and Shurtleff, 2004; Diers et al., 1992; Joshi et al., 2013; Lin et al., 2011; Miller-Garvin and Naeve, 2015; Pantalone, 2012; Warrington et al., 2015; Wilcox, 1998; Wilcox and Shibles, 2001). For this reason, the value of soybeans was traditionally determined by their oil and protein content (Charron et al., 2005; Wilson, 2004). Prior to 1950, the economic importance of soybean was primarily in its oil; however, commercial interest has broadened to include protein meal, a by-product of soybean oil extraction (Chung et al., 2003). Conventional (43-44% protein) and dehulled high protein (47-49% protein) soybean meal are the two primary kinds of soybean meal available in the market (Cromwell, 2012). The hulls are blended back into the meal in conventional soybean meal to standardize crude protein (cp) to approximately 44% (Cromwell, 2012). In dehulled soybean meal, the hulls are left out, leading to an overall higher protein content (Cromwell, 2012). During processing, soybeans are rolled into flakes, which ruptures the oil cells. This separates the oil from the protein meal allowing for extraction of the soybean oil (Soy Stats, 2018). Most of the flakes are then used in animal feed, while the remaining meal is processed into products containing soy protein (Soy Stats, 2018), including products consumed directly by humans, such as soy milk, tofu, and soy flour (Friedman and Brandon, 2001; North Carolina Soybean Producers Association, Inc., 2014).

Amino acids have been described as the building blocks that form proteins (Cromwell, 2012). Having a balance of essential and non-essential amino acids

in protein is essential in determining the nutritional value of soybeans (Cromwell, 2012; Friedman and Brandon, 2001; Warrington et al., 2015). Essential amino acids are supplied to the consumer through ingesting animal and plant based protein (Miller-Garvin and Naeve, 2015; Thakur and Hurburgh, 2007; Warrington et al., 2015; Wilcox and Shibles, 2001). The balance of soybean amino acids provides almost all the essential amino acids required by both swine (*Sus scrofa*) and poultry (*Gallus gallus domesticus*) (Wilson, 2004). Swine and poultry consume the amino acids needed through feed proteins such as highly digestible soybean meal (Cromwell, 2012; Miller-Garvin and Naeve, 2015; Wilcox and Shibles, 2001). Therefore, any soybean deficiency in essential amino acids for swine and poultry can result in increased costs for producers, as they have to supplement animal feed (Friedman and Brandon, 2001; Warrington et al., 2015). Since swine and poultry cannot produce methionine, cysteine, threonine, and lysine naturally and cannot obtain them through consumption of deficient soybean meal, synthetic amino acids must be supplemented in the animal feed to supply these essential amino acids (Warrington et al, 2015; Wilcox and Shibles, 2001). Therefore, animal producers incur additional expenses resulting from supplementation of amino acids in their animal feed. Research conducted by Baker et al. (2011) suggests a greater concentration of amino acids in high protein soybean meal than conventional soybean meal. Therefore, there are more amino acids available for swine and chickens fed high protein soybean meal. Producing soybeans with higher quality protein could help solve some of

these amino acid deficiency problems and, therefore, lower feed supplementation costs for swine and poultry producers.

Oil content of soybean is a quantitative trait, which means that it is controlled by a number of genes (Qi et al., 2011). Soybean oil comprises approximately 20% of the soybean seed and is used in a number of products (Eskandari et al., 2013a). Products containing soybean oil have industrial and food applications such as: biodiesel, ink, plastic, salad dressing, and vegetable oil (Yesudas et al., 2013). An increase in soybean oil would greatly benefit these industries; however, targeted goals must be kept in mind during soybean production. Research has shown there is a negative genetic correlation between soybean seed oil and protein, and between protein and yield (Brummer et al., 1997; Chung et al., 2003; Cober and Voldeng, 2000; Eskandari et al., 2013b; Hyten et al., 2004; Krishnan et al, 2007; Lee et al., 2010; Phansak et al., 2016; Wilcox, 1998; Yaklich, 2001). However, research has also shown there is a positive relationship between soybean oil and yield (Morrison et al., 2008). For this reason, increases in both soybean oil and yield must be simultaneously pursued while maintaining adequate protein content and quality.

In this experiment, one objective is to evaluate the seed protein, oil, meal protein, and amino acid content of the RILs using near-infrared reflectance (NIR) spectroscopy. Seed quality traits such as protein, oil, and amino acid content associated with a particular line of soybean are important to plant breeders. The improvement of these traits through selection allows for an increase in soybean

value. However, these are all quantitative traits and are therefore controlled by many genes, sometimes at different loci or on different chromosomes with differing levels of additive, dominance, and epistatic effects (Bernardo, 2014a; Bernardo, 2014b). Additionally, quantitative traits are more likely to be influenced by environmental conditions than qualitative traits which are controlled by a single gene (Bernardo, 2014a; Bernardo, 2014b). Quantitative traits are also much more difficult to identify and do not exhibit the same clear-cut differences in how traits are expressed as qualitative traits (Bernardo, 2014a). Some examples of quantitative traits are the economically important seed protein and oil concentrations and seed yield (Diers et al., 1992). Along with agronomic traits such as yield, these seed quality traits provide a profile for cultivar selection in soybean. However, farmers are not likely to grow cultivars that have good protein, oil, and protein meal content as they get paid only for yield. These traits are of high importance to processors, however. This experiment analyzes an RIL population for the seed quality traits of soybean seed protein, oil, and amino acid concentrations.

Materials and Methods

Plant Materials

An experimental recombinant inbred line (RIL) population was created from the cross 5601T (Pantalone et al., 2003) × U99-310255. Both parents display higher protein with high yields and favorable oil, making this a unique cross, as there is a known negative genetic correlation between protein and yield

as well as between protein and oil. The female parent, 5601T, is a registered cultivar that has a determinate growth habit, white flowers, and gray pubescence and belongs to maturity group V (Pantalone et al., 2003). 5601T has a high yield potential and above average protein concentration, making it an ideal parent in this cross. The male parent, U99-310255, is an unreleased line from the University of Nebraska. It has an indeterminate growth habit, purple flowers, gray pubescence and belongs to maturity group III. U99-310255 was chosen as a parent of this cross due to its above average protein concentration as well as its good oil concentration.

The experimental lines in this RIL population were grown in a replicated field trial across Tennessee (TN) in the summers of 2017 and 2018. The lines were grown in Knoxville, Springfield, and Milan, TN. A population of 138 individual RIL were created in 2006, when the initial cross of 5601T × U99-310255 was made in Knoxville, TN, at the East Tennessee Research and Education Center (ETREC) (Table 1.1). An RIL is a homozygous or nearly homozygous line that is created when generations are continually allowed to self-pollinate after a cross between two parental inbreds, creating a new inbred line with a mosaic of the parental genomes (Bernardo, 2014a; Bernardo, 2014b; Bernardo, 2014c; Broman, 2005). Resultant individual RILs are all different, and therefore a population of RILs is advantageous in genetic mapping for quantitative trait loci (QTL) analysis (Broman, 2005).

Successful crossing in 2006 led to F_1 seed, which were continually allowed to self-pollinate over nine years until the experimental population reached the $F_{8:10}$ and $F_{8:11}$ generations in 2017 and 2018, respectively (Table 1.1). In the spring of 2017, the $F_{8:10}$ seed were planted in a two replication field trial at ETREC, HRREC, and RECM, and the seed were combine harvested in fall 2017. This $F_{8:11}$ seed was used to source a second year three replication field trial planted in the spring of 2018 at ETREC, HRREC, and RECM. The seed were combine harvested in the fall of 2018.

Near Infrared Reflectance Spectroscopy

Total protein, oil, and amino acid concentrations were determined using a Perten DA 7250 near-infrared reflectance (NIR) spectroscopy analyzer. Perten and the University of Minnesota developed calibration equations to enable soybean seed quality traits to be analyzed. The equations were developed using HPLC in the analytical laboratory at the University of Missouri (Warrington et al., 2015). Calibrations of the machine were performed using wet chemistry from approximately 900 breeding lines and established North American soybean cultivars. Samples of approximately 20 g of ground soybean seed were sampled from each replication in every location for each of the 138 lines grown in a 2017 replicated field trial in three Tennessee locations. Thus, there were 138 lines × 2 replications × 3 locations for a total of 864 samples run on the NIR in 2017. The second replicated field trial was grown in 2018. For the 2018 field trial, there were supposed to be 138 lines × 3 replications × 3 locations for a total of 1,296

samples run on the NIR. However, there was an error at HRREC and the entire first replication had to be dropped, reducing the number of samples to 1,152. Soybean seed samples were ground for 20 seconds to create uniform particle size using a Foss water-cooled Knifetec 1095 Sample Mill (Tecator, Hoganas, Sweden). Protein, oil, and amino acid concentrations were determined based on 13% seed moisture and were converted to g kg^{-1} dry weight. Amino acid concentrations were converted again by dividing by the protein concentration to report as g amino acid kg^{-1} of crude protein (cp).

Data Analysis

The experimental design was a randomized complete block model. Data on seed quality traits were analyzed using a mixed model analysis of variance using SAS PROC GLIMMIX for SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) to estimate least squares means for the 138 RILs averaged over 2017-2018. For the purpose of analysis, fixed effects were individual lines. Random effects were location and GxE or location by line interaction. Tukey's LSD was used to compare lines within each test. It was chosen over Fisher's protected LSD method because it offers full control of the experiment wise type I error rate.

A broad sense estimate of heritability on an entry means basis for each trait of interest in the population was calculated as follows (Nyquist, 1991):

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \left(\frac{\sigma_{ge}^2}{e}\right) + \left(\frac{\sigma^2}{re}\right)}$$

where h^2 is the heritability, σ_g^2 is the genotypic variance, σ_{ge}^2 is genotype x environmental interaction variance, σ^2 is error variance, *e* is number of environments, and *r* is number of replications.

Results and Discussion

Mean concentrations of protein, oil, and amino acids for 138 RILs are found in Table 2.1 in Appendix B. Frequency distributions of protein, oil, and meal are presented in Figures 2.1-2.3 in Appendix B. Frequency distributions for amino acids are presented in Figures 2.4-2.8. Protein concentration was averaged over the 2017 and 2018 growing seasons with two replications in 2017 and three replications in 2018 at six locations for all 138 RIL, checks, and parent 5601T. The checks included in this study and their respective average protein concentrations were: 'Ellis' with 393.8 g kg^{-1}, '5002T' with 398.1 g kg^{-1}, 'TN11-5104' with 407.6 g kg^{-1}, and 'Osage' with 420.1 g kg^{-1}. The minimum protein content from the combined 2017 and 2018 field trials was 407.6 g kg^{-1}, and the maximum value was 446.4 g kg^{-1}. The mean protein content was 423.5 g kg^{-1} (Table 2.1). The heritability estimate on an entry-means basis for protein was high at 0.97. The high heritability estimate found in this study suggests that genetic improvements can be made in soybean protein. The mean protein values for the 14 highest RILs in protein content can be found in Table 2.2.

Additionally, protein meal is of high importance for the livestock feed industry. Meal protein concentration was averaged over the 2017 and 2018 growing seasons with two replications in 2017 and three replications in 2018 at

six locations for all 138 RIL, checks, and parent 5601T. The checks included in this study and their respective average meal protein concentrations were: 'Ellis' with 455.6 g kg^{-1}, '5002T' with 465.0 g kg^{-1}, 'TN11-5104' with 471.9 g kg^{-1}, and 'Osage' with 484.4 g kg^{-1}. The minimum soybean meal content from the combined 2017 and 2018 field trials was 473.3 g kg^{-1}, and the maximum value was 517.5 g kg^{-1}. The mean soybean meal protein content was 492.3 g kg^{-1} (Table 2.1). The mean soybean meal values for the 14 highest RILs in meal content can be found in Table 2.2.

Soybean oil concentration was averaged over the 2017 and 2018 growing seasons with two replications in 2017 and three replications in 2018 at six locations for all 138 RIL, checks, and parent 5601T. The checks included in this study and their respective average oil concentrations were: 'Ellis' with 210.0 g kg^{-1}, '5002T' with 218.9 g kg^{-1}, 'TN11-5104' with 210.7 g kg^{-1}, and 'Osage' with 206.8 g kg^{-1}. The minimum oil content from the combined 2017 and 2018 field trials was 205.3 g kg^{-1}, and the maximum value was 229.2 g kg^{-1}. The mean soybean oil content was 214.3 g kg^{-1} (Table 2.1). The heritability estimate on an entry means basis for oil were high at 0.96. The high heritability estimate found in this study suggests that genetic improvements can be made in soybean oil. The mean soybean oil values for the 14 highest RILs in oil content can be found in Table 2.2.

NIR was used to evaluate soybean seed amino acid concentrations as well. Amino acids play a big part in the protein quality of soybeans. Amino acid

values were divided by the protein concentrations to report as g amino acid kg^{-1} crude protein (cp). Amino acids of interest in this experiment were: cysteine (Cys), methionine (Met), tryptophan (Try), lysine (Lys), and threonine (Thr), due to their importance for poultry and swine nutrition. Amino acid values were averaged over the 2017 and 2018 growing seasons with two replications in 2017 and three replications in 2018 at six locations for all 138 RIL, checks, and parent 5601T. The checks included in this study were: 'Ellis', '5002T', 'TN11-5104', and 'Osage'. The average amino acid concentrations for 'Ellis' were 15.1 g cys kg^{-1} cp, 14.6 g met kg^{-1} cp, 10.4 g trp kg^{-1} cp, 67.4 g lys kg^{-1} cp, and 38.3 g thr kg^{-1} cp (Table 2.2). The average amino acid concentrations for '5002T' were 21.9 g cys kg^{-1} cp, 14.3 g met kg^{-1} cp, 10.4 g trp kg^{-1} cp, 60.1 g lys kg^{-1} cp, and 37.8 g thr kg^{-1} cp (Table 2.2). The average amino acid concentrations for 'TN11-5104' were 18.5 g cys kg^{-1} cp, 14.4 g met kg^{-1} cp, 10.4 g trp kg^{-1} cp, 63.1 g lys kg^{-1} cp, and 37.9 g thr kg^{-1} cp (Table 2.2). The average amino acid concentrations for 'Osage' were 18.5 g cys kg^{-1} cp, 14.2 g met kg^{-1} cp, 10.4 g trp kg^{-1} cp, 62.3 g lys kg^{-1} cp, and 37.5 g thr kg^{-1} cp (Table 2.3). One parent, '5601T', was planted with the trials as well, and its average amino acid concentrations were 18.6 g cys kg^{-1} cp, 14.3 g met kg^{-1} cp, 10.6 g trp kg^{-1} cp, 62.6 g lys kg^{-1} cp, and 37.7 g thr kg^{-1} cp (Table 2.3).

The minimum amino acid values from the combined 2017 and 2018 field trials were 13.7 g cys kg^{-1} cp, 13.7 g met kg^{-1} cp, 10.0 g trp kg^{-1} cp, 57.5 g lys kg^{-1} cp, and 36.5 g thr kg^{-1} cp (Table 2.3). The maximum values were 22.6 g cys kg^{-1}

cp, 14.3 g met kg^{-1} cp, 10.6 g trp kg^{-1} cp, 66.6 g lys kg^{-1} cp, and 37.6 g thr kg^{-1} cp (Table 2.4). The mean values were 16.7 g cys kg^{-1} cp, 14.0 g met kg^{-1} cp, 10.3 g trp kg^{-1} cp, 63.8 g lys kg^{-1} cp, and 37.1 g thr kg^{-1} cp (Table 2.4). The heritability estimates for the amino acids of interest were 0.00 for cys, 0.32 for met, 0.01 for trp, 0.00 for lys, and 0.49 for thr. Met and thr had low to moderate heritability, while cys, trp, and lys had low heritability. The heritability for both cys and lys was 0.00, which was due to the lack of variation in their observed concentrations. Only the 5601T parent was grown in 2017 and 2018 field trials. If both parents were grown in the experiment, we may have seen that both parents had similar concentrations of cys and lys, which would have led to the low heritability. The low to moderate heritability estimates found in this study suggest that genetic improvements can be made with some difficulty in met and thr soybean amino acids.

References

Aoyagi, A. and W. Shurtleff. 2004. History of world soybean production and trade- Part 1. Soyinfo Center. http://www.soyinfocenter.com/HSS/production_and_trade1.php

Baker, K.M., P.L. Utterback, C.M. Parsons, and H.H. Stein. 2011. Nutritional value of soybean meal produced from conventional, high-protein, or low-oligosaccharide varieties of soybeans and fed to broiler chicks. Poult. Sci. 90: 390-395.

Bernardo, R. Process of plant breeding. 2014a. p. 23-69. Essentials of plant breeding. Stemma Press, Woodbury, MN.

Bernardo, R. Genetic basis of plant improvement. 2014b. p. 71-93. Essentials of plant breeding. Stemma Press, Woodbury, MN.

Bernardo, R. Genetic basis of plant improvement. 2014c. p. 119-146. Essentials of plant breeding. Stemma Press, Woodbury, MN.

Broman, K.W. 2005. The genomes of recombinant inbred lines. Genetics. 169: 1133-1146.

Brummer, E.C., G.L. Graef, J. Orf, J.R. Wilcox, and R.C. Shoemaker. 1997. Mapping QTL for seed protein and oil content in eight soybean populations. Crop Sci. 37: 370-378.

Charron, C.S., F.L. Allen, R.D. Johnson, V.R. Pantalone and C.E. Sams. 2005. Correlations of oil and protein with isoflavone concentration in soybean [Glycine max (L.) Merr.]. J. Agric. Food Chem. 53: 7128-7135).

Chung, J., H.L. Babka, G.L. Graef, P.E. Staswick, D.J. Lee, P.B. Cregan, R.C. Shoemaker, and J.E. Specht. 2003. The seed protein, oil and yield QTL on soybean linkage group I. Crop Sci. 43: 1053-1067.

Cober, E.R. and H.D. Voldeng. 2000. Developing high-protein, high-yield soybean populations and lines. Crop Sci. 40:39-42.

Cromwell, G.L. 2012. Soybean meal-An exceptional protein source. Soybean Meal INFOcenter. http://www.soymeal.org/ReviewPapers/SBMExceptionalProteinSource.pdf

Diers, B.W., P. Keim, W.R. Fehr, and R.C. Shoemaker. 1992. RFLP analysis of soybean seed protein and oil content. Theor. Appl. Genet. 83: 608-612.

Eskandari, M., E.R. Cober, and I. Rajcan. 2013a. Genetic control of soybean seed oil: I. QTL and genes associated with seed oil concentration in RIL populations derived from crossing moderately high-oil parents. Theor. Appl. Genet. 126:483-495.

Eskandari, M., E.R. Cober, and I. Rajcan. 2013b. Genetic control of soybean seed oil: II. QTL and genes that increase oil concentration without decreasing protein or with increased yield. Theor. Appl. Genet. 126:1677–1687.

Friedman, M. and D.L. Brandon. 2001. Nutritional and health benefits of soy proteins. J. Agric. Food Chem. 49: 1069-1086.

Hyten, D.L., V.R. Pantalone, C.E. Sams, A.M. Saxton, D. Landau-Ellis, T.R. Stefaniak, and M.E. Schmidt. 2004. Seed quality QTL in a prominent soybean population. Theor. Appl. Genet. 109: 552-561.

Joshi, T., B. Valliyodan, J.H. Wu, S.H. Lee, D. Xu, and H.T. Nguyen. 2013. Genomic differences between cultivated soybean, *G. max* and its wild relative *G. soja*. BMC Genomics. 14(Suppl 1):S5.

Krishnan, H.B., S.S. Natarajan, A.A. Mahmoud, and R.L. Nelson. 2007. Identification of glycinin and β-conglycinin subunits that contribute to the increased protein content of high-protein soybean lines. J. Agric. Food Chem. 55:1839-1845.

Lee, J-D., G. Shannon, and M-G. Choung. 2010. Selection for protein content in soybean from single F_2 seed by near infrared reflectance spectroscopy. Euphytica. 172: 117-123.

Lin, J., V.R. Pantalone, G. Li, and F. Chen. 2011. Molecular cloning and biochemical characterization of an endo-β-mannanase gene from soybean for soybean meal improvement. J. Agric. Food Chem. 59: 622-628.

Miller-Garvin, J., and S.L. Naeve. 2015. United States Soybean Quality Annual Report 2015. http://ussec.org/wp-content/uploads/2016/01/2015-Final-US-Commodity-Soybean-Quality-Report.pdf

Morrison, M.J., E.K. Cober, M.F. Saleem, N.B. McLaughlin, J. Fregeau-Reid, B.L. Ma, W. Yan, and L. Woodrow. 2008. Changes in isoflavone concentration with 58 years of genetic improvement of short-season soybean cultivars in Canada. Crop Sci. 48: 2201–2208.

North Carolina Soybean Producers Association, Inc. 2014. Uses of soybeans. http://www.ncsoy.org/ABOUT-SOYBEANS/Uses-of-Soybeans.aspx

Nyquist, W.E. 1991. Estimation of heritability and prediction of selection response in plant populations. Critical Reviews in Plant Sciences. 10: 235-322.

Pantalone, V.R. 2012. Modern breeding approaches for enhancing soybean protein quality. In: R.F. Wilson, editor, Designing soybean for 21st century markets. AOCS Press, Urbana, IL. p. 197-226.

Pantalone, V.R., F.L. Allen, and D. Landau-Ellis. Registration of '5601T' soybean. 2003. Crop Sci. 43: 1123-1124.

Phansak, P., W. Soonsuwon, D.L. Hyten, Q. Song, P.B. Cregan, G.L. Graef, and J.E. Specht. 2016. Multi-population selective genotyping to identify soybean [*Glycine max* (L.) Merr.] seed protein and oil QTLs. G3-Genes. Genom. Genet. 6: 1635-1646.

Qi, Z., Q. Wu, X. Han, Y. Sun, X. Du, C. Liu, H. Jiang, G. Hu, and Q. Chen. 2011. Soybean oil content QTL mapping and integrating with meta-analysis method for mining genes. Euphytica. 179: 499-514.

Soy Stats. 2018. 2017 Soystats: A Reference Guide to Important Soybean Facts and Figures. The American Soybean Association. http://soystats.com/2017-soystats/

Thakur, M. and C.R. Hurburgh. 2007. Quality of US soybean meal compared to the quality of soybean meal from other origins. J. Am. Oil Chem. Soc. 84:835-843.

Warrington, C.V., H. Abdel-Haleem, D.L. Hyten, P.B. Cregan, J.H. Orf, A.S. Killam, N. Bajjalieh, Z. Li, and H.R. Boerma. 2015. QTL for seed protein and amino acids in the Benning x Danbaekkong soybean population. Theor. Appl. Genet. 128:2474.

Wilcox, J.R. 1998. Increasing seed protein in soybean with eight cycles of recurrent selection. Crop Sci. 38: 1536-1540.

Wilcox, J.R. and R.M. Shibles. 2001. Interrelationships among seed quality attributes in soybean. Crop Sci. 41: 11-14.

Wilson, R.F. 2004. Seed Composition, p. 621-669. In: H.R. Boerma and J.E. Specht (eds.). Soybeans: improvement, production, and uses. 3rd ed. ASA Inc., CSSA Inc., SSSA Inc., Madison, WI.

Yaklich, R.W. 2001. β-Conglycinin and glycinin in high-protein soybean seeds. J. Agric. Food Chem. 49:729-735.

Yesudas, C.R., R. Bashir, M.B Geisler, and D.A. Lightfoot. 2013. Identification of germplasm with stacked QTL underlying seed traits in an inbred soybean population from cultivars Essex and Forrest. Mol. Breeding. 31: 693-703.

Appendix B

Table 2.1 Minimum, mean, and maximum seed quality trait values for protein, oil, and meal for 138 $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 compared with four checks and a parent grown over six environments with three replications each in 2017 and 2018.

Protein, oil, and meal concentration means			
	Protein	Oil	Meal
		g kg^{-1}	
Minimum	407.6	205.3	473.3
Mean	423.5	214.3	492.2
Maximum	446.4	229.2	517.5
Checks[a]	404.9	211.6	469.2
Parent[b]	409.6	211.6	474.7
LSD (0.05)	8.1	4.6	8.2

[a] Checks included in field trials were Ellis, 5002T, TN11-5104, and Osage.
[b] Parent included in field trials was 5601T.

Table 2.2 Mean protein, oil, and meal values for the 14 highest individual $F_{8:10}$ recombinant inbred lines (RIL) in each trait. RILs were created from a cross of 5601T × U99-310255 and were grown over six environments with three replications each in 2017 and 2018.

Line	Protein	Line	Oil	Line	Meal
	g kg^{-1}		g kg^{-1}		g kg^{-1}
56U99-095	446.4	56U99-144	229.2	56U99-095	517.5
56U99-047	440.8	56U99-115	223.9	56U99-047	509.6
56U99-016	439.2	56U99-018	223.6	56U99-216	506.8
56U99-019	437.4	56U99-108	222.4	56U99-175	506.5
56U99-216	436.2	56U99-185	221.9	56U99-099	506.1
56U99-175	435.8	56U99-199	221.2	56U99-016	505.6
56U99-043	434.7	56U99-189	221.1	56U99-043	504.6
56U99-139	434.7	56U99-165	221.0	56U99-019	504.4
56U99-099	434.0	56U99-052	221.0	56U99-039	504.0
56U99-087	433.8	56U99-055	221.0	56U99-181	503.7
56U99-154	433.7	56U99-079	220.9	56U99-154	503.2
56U99-181	433.5	56U99-191	220.5	56U99-139	503.2
56U99-006	433.0	56U99-215	220.0	56U99-096	502.2
56U99-123	433.0	56U99-242	220.0	56U99-201	502.1
LSD (0.05)	8.1	LSD (0.05)	4.6	LSD (0.05)	8.2

Table 2.3 Mean seed quality trait values for 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255, four checks, and one parent grown over six environments with three replications each in 2017 and 2018.

Line	Protein	Oil	Meal	Cysteine	Methionine	Tryptophan	Lysine	Threonine
	g kg^{-1}			g aa kg^{-1} cp				
Ellis[a]	393.8	210.0	455.6	15.1	14.6	10.4	67.4	38.3
5002T[a]	398.1	218.9	465.0	21.9	14.3	10.4	60.1	37.8
TN11-5104[a]	407.6	210.7	471.9	18.5	14.4	10.4	63.1	37.9
Osage[a]	420.1	206.8	484.4	18.5	14.2	10.4	62.3	37.5
5601T[b]	409.6	211.6	474.7	18.6	14.3	10.6	62.6	37.7
56U99-006	433.0	210.3	501.1	15.1	13.9	10.2	64.6	36.8
56U99-007	427.1	210.6	494.4	18.4	14.0	10.4	62.1	37.2
56U99-008	417.1	217.8	486.6	18.4	14.2	10.3	62.3	37.3
56U99-011	421.4	219.7	492.5	14.8	14.2	10.4	66.0	37.4
56U99-012	421.8	211.4	488.7	18.2	13.9	10.3	62.1	37.2
56U99-013	424.4	215.9	494.2	14.6	14.1	10.3	65.8	37.1
56U99-014	421.4	219.4	492.4	18.4	14.1	10.5	62.2	37.1
56U99-015	423.9	211.0	491.0	14.5	13.9	10.3	65.7	37.1
56U99-016	439.2	205.5	505.6	18.0	13.8	10.2	61.8	36.8
56U99-017	426.9	215.8	497.0	18.2	14.1	10.2	62.4	37.2
56U99-018	409.5	223.6	480.8	18.1	14.1	10.4	62.8	37.4
56U99-019	437.4	207.0	504.4	18.0	13.8	10.3	61.9	36.8
56U99-020	432.3	210.9	500.6	14.6	13.7	10.4	65.0	36.8
56U99-021	429.9	212.3	498.5	18.2	13.9	10.3	61.9	37.0
56U99-022	423.7	219.7	495.3	17.8	13.8	10.5	62.5	37.0
56U99-023	421.2	216.1	490.6	14.8	14.0	10.3	65.7	37.2
56U99-025	417.9	216.7	486.9	18.3	13.8	10.5	62.3	37.1
56U99-026	421.3	214.1	489.5	18.2	14.0	10.4	62.2	37.2
56U99-027	426.4	216.1	496.6	18.1	13.8	10.5	61.9	36.9
56U99-028	423.4	211.7	490.8	18.1	14.0	10.3	62.1	37.1
56U99-029	423.2	210.0	489.5	18.1	14.0	10.4	62.4	37.1
56U99-030	431.4	212.6	500.6	18.1	13.9	10.2	61.8	36.7

Table 2.3 Continued.

Line	Protein	Oil	Meal	Cysteine	Methionine	Tryptophan	Lysine	Threonine
	g kg^{-1}			g aa kg^{-1} cp				
56U99-031	419.8	213.6	487.4	18.3	13.9	10.2	62.5	37.2
56U99-032	417.0	212.9	484.0	18.0	13.9	10.4	62.1	37.1
56U99-034	431.0	213.0	500.2	22.6	13.9	10.1	57.5	36.7
56U99-035	423.7	213.6	492.0	18.2	14.0	10.3	62.3	37.2
56U99-036	418.6	209.7	484.0	18.3	13.8	10.3	61.9	37.1
56U99-037	422.8	210.8	489.4	14.4	13.7	10.4	65.8	37.1
56U99-038	426.3	216.4	496.6	14.6	13.9	10.4	65.7	37.1
56U99-039	432.1	217.6	504.0	14.6	13.9	10.5	65.5	37.0
56U99-040	422.4	217.1	492.5	18.3	14.0	10.0	62.5	37.1
56U99-041	411.1	214.8	478.1	21.7	14.0	10.3	59.3	37.3
56U99-043	434.7	213.1	504.6	14.6	14.0	10.2	65.8	37.1
56U99-044	417.1	215.6	485.4	14.6	14.0	10.3	66.3	37.2
56U99-045	431.5	214.4	501.6	14.4	13.7	10.2	65.6	37.0
56U99-046	414.1	209.3	478.7	18.3	14.1	10.2	62.4	37.2
56U99-047	440.8	209.2	509.6	18.2	13.8	10.2	61.5	36.5
56U99-048	421.1	213.1	488.7	14.7	14.1	10.4	66.1	37.3
56U99-049	422.8	215.4	492.1	18.0	13.9	10.3	62.3	37.0
56U99-052	425.1	221.0	497.6	18.5	14.1	10.4	62.5	37.1
56U99-053	428.8	218.4	500.6	18.3	14.1	10.2	62.3	36.9
56U99-054	430.4	215.0	500.7	14.6	14.0	10.0	65.6	36.9
56U99-055	423.2	221.0	495.5	18.0	13.9	10.4	62.5	37.1
56U99-056	412.1	218.2	480.9	14.8	14.2	10.6	66.2	37.4
56U99-057	423.8	213.6	492.1	17.8	13.7	10.4	62.0	37.0
56U99-061	419.2	213.3	486.6	14.8	14.0	10.6	65.8	37.1
56U99-063	417.5	216.8	486.5	18.0	13.8	10.4	62.3	37.1
56U99-067	423.1	213.5	491.4	21.6	14.0	10.3	59.0	37.0
56U99-068	430.7	211.5	499.1	17.9	13.7	10.2	62.1	37.0
56U99-069	413.5	215.2	481.0	14.7	14.2	10.5	66.4	37.4
56U99-071	425.5	213.1	494.0	17.9	14.0	10.3	62.2	37.0
56U99-075	424.9	208.5	490.8	14.5	14.0	10.3	66.0	37.3

Table 2.3 Continued.

Line	Protein	Oil	Meal	Cysteine	Methionine	Tryptophan	Lysine	Threonine
	g kg^{-1}			g aa kg^{-1} cp				
56U99-076	430.7	212.0	499.4	18.4	14.0	10.0	62.2	37.1
56U99-079	421.8	220.9	493.8	18.2	14.0	10.3	62.4	37.0
56U99-080	415.2	217.1	484.0	18.2	14.0	10.4	62.4	37.3
56U99-081	424.6	214.0	493.3	18.2	14.1	10.2	62.3	37.2
56U99-082	416.0	219.4	486.1	18.0	14.0	10.4	62.4	37.2
56U99-086	422.8	211.6	490.0	14.4	13.9	10.2	66.0	37.2
56U99-087	433.8	209.4	501.5	14.1	13.7	10.2	65.6	36.9
56U99-089	429.0	212.3	497.6	14.3	13.8	10.2	65.6	36.9
56U99-092	425.9	215.7	495.8	14.5	13.9	10.1	65.9	36.9
56U99-093	422.0	214.6	490.6	14.8	14.1	10.1	65.9	37.2
56U99-095	446.4	212.1	517.5	17.7	13.8	10.2	61.8	36.7
56U99-096	431.6	215.4	502.2	14.5	13.9	10.4	65.6	37.0
56U99-097	408.5	211.4	473.3	18.4	14.2	10.3	62.9	37.5
56U99-099	434.0	217.4	506.1	14.8	13.9	10.2	65.5	36.8
56U99-100	422.7	212.7	490.5	18.2	14.1	10.5	62.1	37.1
56U99-102	415.8	219.4	485.8	14.5	14.1	10.4	66.4	37.2
56U99-104	413.3	214.0	480.2	14.8	14.1	10.3	66.6	37.5
56U99-105	415.9	218.8	485.7	15.0	14.2	10.6	66.2	37.5
56U99-107	421.0	205.3	484.6	18.1	13.9	10.3	62.2	37.3
56U99-108	418.3	222.4	490.5	18.5	14.0	10.3	62.5	37.4
56U99-109	433.0	207.0	499.5	18.0	13.8	10.3	61.8	36.8
56U99-111	425.7	207.4	491.1	14.4	13.8	10.4	65.8	37.2
56U99-115	416.8	223.9	489.5	14.9	14.1	10.4	66.1	37.2
56U99-118	427.5	213.2	496.2	18.1	13.7	10.3	61.9	36.9
56U99-121	431.1	211.6	499.6	14.6	13.9	10.4	65.4	37.0
56U99-123	433.0	209.3	500.6	14.4	13.8	10.4	65.4	36.9
56U99-125	423.0	205.3	486.8	14.6	13.9	10.3	65.8	37.3
56U99-126	432.8	208.8	500.0	14.3	13.7	10.3	65.4	36.9
56U99-130	420.5	210.1	486.5	14.8	14.1	10.4	66.0	37.3
56U99-131	418.9	215.4	487.3	14.9	14.0	10.6	65.7	37.3

Table 2.3 Continued.

Line	Protein	Oil	Meal	Cysteine	Methionine	Tryptophan	Lysine	Threonine
	g kg^{-1}			g aa kg^{-1} cp				
56U99-133	413.4	217.4	482.1	14.6	14.0	10.5	66.0	37.2
56U99-135	412.6	217.3	481.0	18.3	14.0	10.5	62.6	37.4
56U99-136	423.9	209.2	490.0	14.6	14.0	10.4	65.8	37.2
56U99-137	431.0	207.4	497.1	21.6	14.2	10.3	59.0	36.9
56U99-139	434.7	210.6	503.2	13.7	14.0	10.1	66.3	37.1
56U99-138	414.7	229.2	489.9	15.1	14.2	10.6	66.3	37.2
56U99-146	422.7	217.4	492.9	18.1	13.8	10.4	62.2	37.1
56U99-151	428.7	217.4	500.0	14.9	14.1	10.4	65.6	37.1
56U99-152	426.8	209.0	493.3	18.2	13.9	10.2	61.9	36.7
56U99-154	433.7	212.7	503.2	14.6	14.1	10.1	65.7	36.9
56U99-155	417.8	218.5	487.7	18.3	14.3	10.4	62.8	37.2
56U99-157	412.7	214.5	479.7	18.4	14.2	10.4	62.5	37.4
56U99-159	416.3	217.8	485.7	14.8	14.2	10.5	66.0	37.2
56U99-162	428.1	208.5	494.4	14.7	13.9	10.4	65.6	37.1
56U99-164	424.9	214.3	493.9	18.4	14.1	10.5	62.2	37.1
56U99-165	407.6	221.0	477.2	19.1	14.1	10.3	62.2	37.3
56U99-166	415.2	216.0	483.4	15.2	14.0	10.3	65.7	37.4
56U99-168	429.2	210.6	496.9	21.3	13.8	10.4	58.6	36.8
56U99-170	430.8	214.2	500.7	14.7	14.1	10.2	65.7	36.9
56U99-175	435.8	214.4	506.5	18.0	14.0	10.0	62.2	36.8
56U99-176	431.2	208.9	498.4	18.0	13.7	10.4	61.8	36.9
56U99-177	425.9	214.7	495.2	14.6	13.9	10.4	65.6	36.9
56U99-179	421.0	214.9	489.5	18.5	14.1	10.3	62.5	37.1
56U99-181	433.5	213.9	503.7	14.6	13.8	10.2	65.5	36.9
56U99-183	430.0	207.8	496.4	14.6	13.7	10.5	65.2	36.9
56U99-184	427.3	215.8	497.4	18.0	13.8	10.2	62.1	36.9
56U99-185	417.7	221.9	489.4	14.8	14.2	10.3	66.4	37.4
56U99-187	423.0	216.8	492.9	15.2	14.2	10.3	65.5	37.1
56U99-189	416.5	221.1	487.7	18.1	14.1	10.4	62.4	37.4
56U99-190	415.1	219.8	485.3	14.8	14.2	10.4	65.9	37.3

Table 2.3 Continued.

Line	Protein	Oil	Meal	Cysteine	Methionine	Tryptophan	Lysine	Threonine
	g kg^{-1}			g aa kg^{-1} cp				
56U99-191	411.0	220.5	480.9	14.8	14.2	10.3	66.5	37.4
56U99-199	411.0	221.2	481.3	18.5	14.1	10.3	62.6	37.4
56U99-200	423.3	212.3	490.9	15.1	14.1	10.6	65.3	37.0
56U99-201	432.7	212.6	502.1	18.3	14.0	10.2	62.2	36.9
56U99-202	422.5	213.8	490.8	18.7	14.0	10.3	61.9	37.2
56U99-203	421.0	216.9	490.7	14.7	14.0	10.3	66.0	37.2
56U99-206	414.6	210.7	480.1	18.3	13.9	10.4	62.2	37.3
56U99-207	426.9	217.3	497.8	18.1	14.0	10.1	62.2	37.0
56U99-208	423.5	211.9	490.9	18.3	14.0	10.3	62.3	37.2
56U99-211	417.5	219.3	487.8	14.5	13.9	10.5	65.8	37.0
56U99-215	411.2	220.0	480.9	15.0	14.3	10.5	66.5	37.6
56U99-216	436.2	214.0	506.8	18.2	14.0	10.1	62.0	36.8
56U99-219	431.0	207.7	497.3	14.6	13.9	10.4	65.5	37.2
56U99-220	420.3	219.5	491.3	15.4	14.1	10.3	65.5	37.3
56U99-221	408.8	216.5	476.3	14.7	14.1	10.6	66.2	37.3
56U99-222	429.7	212.1	498.2	14.8	14.1	10.3	65.7	37.1
56U99-223	417.5	216.5	486.4	14.8	14.1	10.3	66.1	37.3
56U99-227	427.0	206.4	492.0	18.4	13.8	10.4	61.7	37.1
56U99-228	426.1	209.6	492.7	14.5	14.0	10.3	65.8	37.0
56U99-232	416.2	216.1	484.5	18.5	14.1	10.4	62.5	37.3
56U99-233	430.7	212.0	499.4	18.0	13.7	10.4	61.7	36.9
56U99-235	431.3	211.9	500.0	17.9	13.7	10.3	61.9	37.0
56U99-237	413.3	216.8	481.5	18.6	14.0	10.6	62.5	37.1
56U99-238	417.9	216.1	486.6	14.7	14.1	10.3	66.2	37.1
56U99-242	413.5	220.0	483.5	18.5	14.2	10.6	62.3	37.3
56U99-243	415.0	212.0	481.1	14.8	14.1	10.4	66.2	37.2
56U99-244	424.9	212.0	492.6	18.0	13.8	10.4	61.8	37.1
LSD (0.05)	8.1	4.6	8.2	6.9	0.2	0.3	6.8	0.3

[a] Checks included in field trials were Ellis, 5002T, TN11-5104, and Osage.
[b] Parent included in field trials was 5601T.

Table 2.4 Minimum, mean, and maximum amino acid values for 138 $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 compared with four checks and a parent grown over six environments with three replications each in 2017 and 2018.

	Amino acid concentration means				
	Cysteine	Methionine	Tryptophan	Lysine	Threonine
	g aa kg^{-1} cp				
Minimum	13.7	13.7	10.0	57.5	36.5
Mean	16.7	14.0	10.3	63.8	37.1
Maximum	22.6	14.3	10.6	66.6	37.6
Checks[a]	18.5	14.4	10.4	63.2	37.9
Parent[b]	18.6	14.3	10.6	62.6	37.7
LSD (0.05)	6.9	0.2	0.3	6.8	0.3

[a] Checks included in field trials were Ellis, 5002T, TN11-5104, and Osage.
[b] Parent included in field trials was 5601T.

Table 2.5 Mean agronomic and seed quality trait values for 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255, four checks, and one parent grown over six environments with three replications each in 2017 and 2018, sorted by yield from highest to lowest.

Line	Yield	Maturity	Height	Lodging	Protein	Oil	Meal	Cysteine	Methionine	Tryptophan	Lysine	Threonine
	kg ha^{-1}	Days	cm	1-5 scale	g kg^{-1}			g aa kg^{-1} cp				
Ellis[a]	4781.0	138.3	74.0	1.6	393.8	210.0	455.6	15.1	14.6	10.4	67.4	38.3
56U99-092	4768.5	131.8	115.0	2.5	425.9	215.7	495.8	14.5	13.9	10.1	65.9	36.9
56U99-086	4679.0	136.2	114.1	2.4	422.8	211.6	490.0	14.4	13.9	10.2	66.0	37.2
56U99-089	4645.1	135.5	112.1	2.4	429.0	212.3	497.6	14.3	13.8	10.2	65.6	36.9
TN11-5104[a]	4576.0	138.4	83.5	2.1	407.6	210.7	471.9	18.5	14.4	10.4	63.1	37.9
56U99-155	4411.7	142.3	129.3	3.0	417.8	218.5	487.7	18.3	14.3	10.4	62.8	37.2
56U99-199	4410.0	138.9	118.7	2.8	411.0	221.2	481.3	18.5	14.1	10.3	62.6	37.4
56U99-228	4366.6	141.0	104.4	2.8	426.1	209.6	492.7	14.5	14.0	10.3	65.8	37.0
5601T[b]	4360.4	141.3	92.7	2.3	409.6	211.6	474.7	18.6	14.3	10.6	62.6	37.7
56U99-219	4335.9	142.0	81.3	2.2	431.0	207.7	497.3	14.6	13.9	10.4	65.5	37.2
56U99-130	4333.7	142.3	75.3	2.2	420.5	210.1	486.5	14.8	14.1	10.4	66.0	37.3
56U99-164	4327.3	134.1	68.4	2.5	424.9	214.3	493.9	18.4	14.1	10.5	62.2	37.1
56U99-223	4311.7	135.8	109.3	3.0	417.5	216.5	486.4	14.8	14.1	10.3	66.1	37.3
5002T[a]	4306.6	139.1	75.5	2.3	398.1	218.9	465.0	21.9	14.3	10.4	60.1	37.8
56U99-222	4301.5	138.7	77.2	3.1	429.7	212.1	498.2	14.8	14.1	10.3	65.7	37.1
56U99-176	4280.0	135.6	107.2	2.1	431.2	208.9	498.4	18.0	13.7	10.4	61.8	36.9
56U99-242	4279.9	138.7	73.0	2.2	413.5	220.0	483.5	18.5	14.2	10.6	62.3	37.3
56U99-136	4269.8	140.8	79.7	2.3	423.9	209.2	490.0	14.6	14.0	10.4	65.8	37.2
Osage[a]	4261.3	138.0	75.7	1.9	420.1	206.8	484.4	18.5	14.2	10.4	62.3	37.5
56U99-185	4260.7	135.8	66.9	2.0	417.7	221.9	489.4	14.8	14.2	10.3	66.4	37.4
56U99-203	4251.8	139.0	77.3	2.3	421.0	216.9	490.7	14.7	14.0	10.3	66.0	37.2
56U99-215	4249.9	139.4	118.1	3.2	411.2	220.0	480.9	15.0	14.3	10.5	66.5	37.6
56U99-157	4225.2	143.5	131.0	3.7	412.7	214.5	479.7	18.4	14.2	10.4	62.5	37.4
56U99-206	4209.1	139.6	112.2	2.2	414.6	210.7	480.1	18.3	13.9	10.4	62.2	37.3
56U99-012	4204.7	134.1	74.5	2.9	421.8	211.4	488.7	18.2	13.9	10.3	62.1	37.2
56U99-011	4188.2	139.4	79.1	2.4	421.4	219.7	492.5	14.8	14.2	10.4	66.0	37.4
56U99-227	4177.6	137.7	75.8	2.0	427.0	206.4	492.0	18.4	13.8	10.4	61.7	37.1
56U99-125	4158.0	141.5	80.5	2.1	423.0	205.3	486.8	14.6	13.9	10.3	65.8	37.3
56U99-202	4154.9	133.9	74.3	2.9	422.5	213.8	490.8	18.7	14.0	10.3	61.9	37.2
56U99-053	4141.2	133.4	114.2	1.9	428.8	218.4	500.6	18.3	14.1	10.2	62.3	36.9
56U99-100	4118.0	140.3	122.8	3.3	422.7	212.7	490.5	18.2	14.1	10.5	62.1	37.1
56U99-137	4116.7	143.9	85.9	3.2	431.0	207.4	497.1	21.6	14.2	10.3	59.0	36.9
56U99-181	4110.0	135.3	112.9	2.7	433.5	213.9	503.7	14.6	13.8	10.2	65.5	36.9
56U99-238	4109.5	137.6	115.0	3.4	417.9	216.1	486.6	14.7	14.1	10.3	66.2	37.1
56U99-069	4106.8	139.5	111.1	2.9	413.5	215.2	481.0	14.7	14.2	10.5	66.4	37.4
56U99-028	4092.4	142.6	116.4	2.9	423.4	211.7	490.8	18.1	14.0	10.3	62.1	37.1
56U99-111	4079.8	137.5	75.7	2.1	425.7	207.4	491.1	14.4	13.8	10.4	65.8	37.2
56U99-126	4074.3	138.2	84.4	2.9	432.8	208.8	500.0	14.3	13.7	10.3	65.4	36.9
56U99-056	4066.7	140.1	105.6	2.3	412.1	218.2	480.9	14.8	14.2	10.6	66.2	37.4
56U99-184	4065.3	132.3	81.1	2.6	427.3	215.8	497.4	18.0	13.8	10.2	62.1	36.9
56U99-187	4059.8	139.5	128.1	3.3	423.0	216.8	492.9	15.2	14.2	10.3	65.5	37.1
56U99-075	4043.6	137.6	75.4	2.3	424.9	208.5	490.8	14.5	14.0	10.3	66.0	37.3
56U99-082	4040.1	142.5	116.1	2.9	416.0	219.4	486.1	18.0	14.0	10.4	62.4	37.2
56U99-035	4034.4	140.9	122.4	2.9	423.7	213.6	492.0	18.2	14.0	10.3	62.3	37.2
56U99-208	4026.9	135.2	65.2	2.2	423.5	211.9	490.9	18.3	14.0	10.3	62.3	37.2

Table 2.5 Continued.

Line	Yield	Maturity	Height	Lodging	Protein	Oil	Meal	Cysteine	Methionine	Tryptophan	Lysine	Threonine
	kg ha^{-1}	Days	cm	1-5 scale	g kg^{-1}			g aa kg^{-1} cp				
56U99-221	3992.7	138.2	115.0	3.4	408.8	216.5	476.3	14.7	14.1	10.6	66.2	37.3
56U99-109	3987.1	140.1	121.3	3.8	433.0	207.0	499.5	18.0	13.8	10.3	61.8	36.8
56U99-146	3979.2	139.9	114.5	3.0	422.7	217.4	492.9	18.1	13.8	10.4	62.2	37.1
56U99-191	3973.3	136.7	118.4	3.8	411.0	220.5	480.9	14.8	14.2	10.3	66.5	37.4
56U99-133	3958.2	137.4	121.4	2.7	413.4	217.4	482.1	14.6	14.0	10.5	66.0	37.2
56U99-052	3955.2	132.8	115.4	3.2	425.1	221.0	497.6	18.5	14.1	10.4	62.5	37.1
56U99-013	3952.0	138.0	110.9	2.9	424.4	215.9	494.2	14.6	14.1	10.3	65.8	37.1
56U99-207	3945.2	135.2	101.3	2.7	426.9	217.3	497.8	18.1	14.0	10.1	62.2	37.0
56U99-048	3935.6	135.8	79.4	2.8	421.1	213.1	488.7	14.7	14.1	10.4	66.1	37.3
56U99-175	3922.3	144.5	116.3	3.5	435.8	214.4	506.5	18.0	14.0	10.0	62.2	36.8
56U99-007	3915.1	136.6	74.5	2.3	427.1	210.6	494.4	18.4	14.0	10.4	62.1	37.2
56U99-107	3911.8	138.2	73.3	1.9	421.0	205.3	484.6	18.1	13.9	10.3	62.2	37.3
56U99-034	3910.8	134.3	110.8	2.8	431.0	213.0	500.2	22.6	13.9	10.1	57.5	36.7
56U99-237	3908.3	142.5	118.9	3.4	413.3	216.8	481.5	18.6	14.0	10.6	62.5	37.1
56U99-040	3885.8	134.3	116.1	2.5	422.4	217.1	492.5	18.3	14.0	10.0	62.5	37.1
56U99-177	3874.3	140.6	107.4	3.4	425.9	214.7	495.2	14.6	13.9	10.4	65.6	36.9
56U99-244	3860.5	138.7	104.8	2.3	424.9	212.0	492.6	18.0	13.8	10.4	61.8	37.1
56U99-006	3852.5	138.8	74.8	2.7	433.0	210.3	501.1	15.1	13.9	10.2	64.6	36.8
56U99-139	3847.1	140.4	69.2	2.4	434.7	210.6	503.2	13.7	14.0	10.1	66.3	37.1
56U99-154	3843.5	145.0	119.0	3.8	433.7	212.7	503.2	14.6	14.1	10.1	65.7	36.9
56U99-243	3839.2	138.4	110.9	2.4	415.0	212.0	481.1	14.8	14.1	10.4	66.2	37.2
56U99-159	3828.8	143.1	130.9	3.9	416.3	217.8	485.7	14.8	14.2	10.5	66.0	37.2
56U99-105	3815.6	134.0	62.5	1.9	415.9	218.8	485.7	15.0	14.2	10.6	66.2	37.5
56U99-061	3806.0	140.9	75.6	2.2	419.2	213.3	486.6	14.8	14.0	10.6	65.8	37.1
56U99-170	3804.8	145.4	113.5	4.0	430.8	214.2	500.7	14.7	14.1	10.2	65.7	36.9
56U99-096	3802.3	136.0	117.2	2.7	431.6	215.4	502.2	14.5	13.9	10.4	65.6	37.0
56U99-216	3797.2	143.2	111.6	3.0	436.2	214.0	506.8	18.2	14.0	10.1	62.0	36.8
56U99-179	3790.8	139.7	78.5	2.4	421.0	214.9	489.5	18.5	14.1	10.3	62.5	37.1
56U99-232	3767.5	139.2	115.7	3.9	416.2	216.1	484.5	18.5	14.1	10.4	62.5	37.3
56U99-023	3754.7	134.8	112.2	2.6	421.2	216.1	490.6	14.8	14.0	10.3	65.7	37.2
56U99-027	3750.4	134.2	109.2	2.7	426.4	216.1	496.6	18.1	13.8	10.5	61.9	36.9
56U99-104	3747.4	138.8	116.3	3.0	413.3	214.0	480.2	14.8	14.1	10.3	66.6	37.5
56U99-081	3745.7	137.6	116.9	3.7	424.6	214.0	493.3	18.2	14.1	10.2	62.3	37.2
56U99-018	3732.6	136.7	119.2	4.0	409.5	223.6	480.8	18.1	14.1	10.4	62.8	37.4
56U99-008	3724.8	133.9	70.6	2.1	417.1	217.8	486.6	18.4	14.2	10.3	62.3	37.3
56U99-014	3722.0	140.6	112.1	4.5	421.4	219.4	492.4	18.4	14.1	10.5	62.2	37.1
56U99-043	3711.6	140.1	76.4	2.2	434.7	213.1	504.6	14.6	14.0	10.2	65.8	37.1
56U99-152	3706.0	139.4	112.6	3.3	426.8	209.0	493.3	18.2	13.9	10.2	61.9	36.7
56U99-166	3703.3	131.8	107.1	2.5	415.2	216.0	483.4	15.2	14.0	10.3	65.7	37.4
56U99-093	3696.3	132.3	64.5	1.9	422.0	214.6	490.6	14.8	14.1	10.1	65.9	37.2
56U99-189	3695.7	137.4	122.5	4.1	416.5	221.1	487.7	18.1	14.1	10.4	62.4	37.4
56U99-016	3659.9	141.1	103.2	2.5	439.2	205.5	505.6	18.0	13.8	10.2	61.8	36.8
56U99-068	3656.7	136.9	78.0	2.7	430.7	211.5	499.1	17.9	13.7	10.2	62.1	37.0
56U99-168	3656.5	138.4	122.2	4.2	429.2	210.6	496.9	21.3	13.8	10.4	58.6	36.8
56U99-183	3634.5	140.8	83.7	3.2	430.0	207.8	496.4	14.6	13.7	10.5	65.2	36.9
56U99-029	3625.5	138.2	74.9	2.7	423.2	210.0	489.5	18.1	14.0	10.4	62.4	37.1
56U99-211	3607.5	139.8	108.8	3.2	417.5	219.3	487.8	14.5	13.9	10.5	65.8	37.0
56U99-200	3594.9	139.0	116.4	3.8	423.3	212.3	490.9	15.1	14.1	10.6	65.3	37.0
56U99-079	3594.5	138.1	121.0	3.7	421.8	220.9	493.8	18.2	14.0	10.3	62.4	37.0
56U99-019	3592.9	137.5	102.4	2.6	437.4	207.0	504.4	18.0	13.8	10.3	61.9	36.8
56U99-102	3588.7	133.7	64.9	2.4	415.8	219.4	485.8	14.5	14.1	10.4	66.4	37.2
56U99-076	3580.2	138.0	106.9	2.3	430.7	212.0	499.4	18.4	14.0	10.0	62.2	37.1
56U99-123	3577.0	137.7	84.1	2.8	433.0	209.3	500.6	14.4	13.8	10.4	65.4	36.9
56U99-087	3565.5	143.1	111.4	3.8	433.8	209.4	501.5	14.1	13.7	10.2	65.6	36.9
56U99-026	3559.9	136.9	110.0	2.9	421.3	214.1	489.5	18.2	14.0	10.4	62.2	37.2

Table 2.5 Continued.

Line	Yield	Maturity	Height	Lodging	Protein	Oil	Meal	Cysteine	Methionine	Tryptophan	Lysine	Threonine
	kg ha^{-1}	Days	cm	1-5 scale	g kg^{-1}			g aa kg^{-1} cp				
56U99-025	3557.4	137.3	109.3	2.8	417.9	216.7	486.9	18.3	13.8	10.5	62.3	37.1
56U99-015	3519.5	137.8	113.1	3.9	423.9	211.0	491.0	14.5	13.9	10.3	65.7	37.1
56U99-121	3516.4	135.6	108.1	3.7	431.1	211.6	499.6	14.6	13.9	10.4	65.4	37.0
56U99-080	3501.1	142.4	111.1	3.4	415.2	217.1	484.0	18.2	14.0	10.4	62.4	37.3
56U99-097	3499.4	133.9	75.6	2.3	408.5	211.4	473.3	18.4	14.2	10.3	62.9	37.5
56U99-039	3497.6	134.8	94.0	2.3	432.1	217.6	504.0	14.6	13.9	10.5	65.5	37.0
56U99-049	3494.4	139.5	113.7	4.2	422.8	215.4	492.1	18.0	13.9	10.3	62.3	37.0
56U99-115	3474.5	134.1	109.1	3.0	416.8	223.9	489.5	14.9	14.1	10.4	66.1	37.2
56U99-032	3463.6	138.2	118.0	3.3	417.0	212.9	484.0	18.0	13.9	10.4	62.1	37.1
56U99-151	3450.1	138.1	112.8	4.2	428.7	217.4	500.0	14.9	14.1	10.4	65.6	37.1
56U99-017	3446.3	136.9	109.5	2.9	426.9	215.8	497.0	18.2	14.1	10.2	62.4	37.2
56U99-190	3441.6	133.7	111.6	3.6	415.1	219.8	485.3	14.8	14.2	10.4	65.9	37.3
56U99-030	3438.1	137.6	121.7	4.2	431.4	212.6	500.6	18.1	13.9	10.2	61.8	36.7
56U99-044	3437.5	139.2	105.6	3.2	417.1	215.6	485.4	14.6	14.0	10.3	66.3	37.2
56U99-021	3434.5	141.4	110.1	4.6	429.9	212.3	498.5	18.2	13.9	10.3	61.9	37.0
56U99-162	3433.2	136.8	82.9	3.1	428.1	208.5	494.4	14.7	13.9	10.4	65.6	37.1
56U99-071	3411.2	137.2	108.4	3.2	425.5	213.1	494.0	17.9	14.0	10.3	62.2	37.0
56U99-235	3409.6	138.0	106.5	3.5	431.3	211.9	500.0	17.9	13.7	10.3	61.9	37.0
56U99-118	3401.8	139.8	114.5	3.5	427.5	213.2	496.2	18.1	13.7	10.3	61.9	36.9
56U99-038	3388.3	136.1	101.8	2.5	426.3	216.4	496.6	14.6	13.9	10.4	65.7	37.1
56U99-041	3383.5	136.2	105.7	2.6	411.1	214.8	478.1	21.7	14.0	10.3	59.3	37.3
56U99-131	3366.3	131.6	68.7	2.2	418.9	215.4	487.3	14.9	14.0	10.6	65.7	37.3
56U99-045	3342.8	139.0	100.1	3.1	431.5	214.4	501.6	14.4	13.7	10.2	65.6	37.0
56U99-135	3335.4	134.5	67.6	2.5	412.6	217.3	481.0	18.3	14.0	10.5	62.6	37.4
56U99-201	3323.2	136.0	103.0	3.3	432.7	212.6	502.1	18.3	14.0	10.2	62.2	36.9
56U99-165	3291.0	135.2	112.1	4.8	407.6	221.0	477.2	19.1	14.1	10.3	62.2	37.3
56U99-037	3285.7	138.6	114.9	4.4	422.8	210.8	489.4	14.4	13.7	10.4	65.8	37.1
56U99-220	3272.5	139.2	96.9	3.5	420.3	219.5	491.3	15.4	14.1	10.3	65.5	37.3
56U99-031	3270.9	137.4	69.6	2.4	419.8	213.6	487.4	18.3	13.9	10.2	62.5	37.2
56U99-054	3175.8	137.3	115.9	4.0	430.4	215.0	500.7	14.6	14.0	10.0	65.6	36.9
56U99-047	3162.2	136.0	117.0	3.7	440.8	209.2	509.6	18.2	13.8	10.2	61.5	36.5
56U99-057	3153.4	138.9	106.7	4.5	423.8	213.6	492.1	17.8	13.7	10.4	62.0	37.0
56U99-063	3153.3	139.5	114.1	4.1	417.5	216.8	486.5	18.0	13.8	10.4	62.3	37.1
56U99-022	3105.7	135.5	100.6	4.1	423.7	219.7	495.3	17.8	13.8	10.5	62.5	37.0
56U99-020	3083.7	129.2	104.9	2.8	432.3	210.9	500.6	14.6	13.7	10.4	65.0	36.8
56U99-108	3064.3	134.7	101.2	4.4	418.3	222.4	490.5	18.5	14.0	10.3	62.5	37.4
56U99-095	3051.6	132.9	109.4	3.4	446.4	212.1	517.5	17.7	13.8	10.2	61.8	36.7
56U99-036	3024.2	137.3	114.4	4.0	418.6	209.7	484.0	18.3	13.8	10.3	61.9	37.1
56U99-144	3011.7	137.1	115.2	4.6	414.7	229.2	489.9	15.1	14.2	10.6	66.3	37.2
56U99-233	2908.3	137.1	113.3	4.2	430.7	212.0	499.4	18.0	13.7	10.4	61.7	36.9
56U99-067	2903.2	133.9	120.6	3.7	423.1	213.5	491.4	21.6	14.0	10.3	59.0	37.0
56U99-055	2856.9	135.9	119.5	4.6	423.2	221.0	495.5	18.0	13.9	10.4	62.5	37.1
56U99-046	2855.4	137.9	112.6	3.1	414.1	209.3	478.7	18.3	14.1	10.2	62.4	37.2
56U99-099	2461.0	133.1	99.5	4.2	434.0	217.4	506.1	14.8	13.9	10.2	65.5	36.8
LSD (0.05)	647.4	3.6	10.8	0.8	8.1	4.6	8.2	6.9	0.2	0.3	6.8	0.3

[a] Checks included in field trials were Ellis, 5002T, TN11-5104, and Osage.
[b] Parent included in field trials was 5601T.

Table 2.6 Mean meal protein concentration and yield values for individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255, with meal protein over 47.5% and yield over 3500 kg ha^{-1}. Data is sorted by percent meal protein from highest to lowest.

Line	Meal	Yield
	%	kg ha^{-1}
56U99-216	50.7	3797.2
56U99-175	50.7	3922.3
56U99-016	50.6	3659.9
56U99-043	50.5	3711.6
56U99-019	50.4	3592.9
56U99-181	50.4	4110.0
56U99-154	50.3	3843.5
56U99-139	50.3	3847.1
56U99-096	50.2	3802.3
56U99-087	50.2	3565.5
56U99-006	50.1	3852.5
56U99-170	50.1	3804.8
56U99-053	50.1	4141.2
56U99-123	50.1	3577.0
56U99-034	50.0	3910.8
56U99-126	50.0	4074.3
56U99-121	50.0	3516.4
56U99-109	49.9	3987.1
56U99-076	49.9	3580.2
56U99-068	49.9	3656.7
56U99-176	49.8	4280.0
56U99-222	49.8	4301.5
56U99-207	49.8	3945.2
56U99-089	49.8	4645.1
56U99-052	49.8	3955.2
56U99-184	49.7	4065.3
56U99-219	49.7	4335.9
56U99-137	49.7	4116.7
56U99-168	49.7	3656.5
56U99-027	49.7	3750.4
56U99-183	49.6	3634.5
56U99-092	49.6	4768.5

Table 2.6 Continued.

Line	Meal	Yield
	%	kg ha^{-1}
56U99-177	49.5	3874.3
56U99-007	49.4	3915.1
56U99-013	49.4	3952.0
56U99-164	49.4	4327.3
56U99-079	49.4	3594.5
56U99-152	49.3	3706.0
56U99-081	49.3	3745.7
56U99-146	49.3	3979.2
56U99-187	49.3	4059.8
56U99-228	49.3	4366.6
56U99-244	49.3	3860.5
56U99-011	49.3	4188.2
56U99-040	49.3	3885.8
56U99-014	49.2	3722.0
56U99-035	49.2	4034.4
56U99-227	49.2	4177.6
56U99-111	49.1	4079.8
56U99-015	49.1	3519.5
56U99-200	49.1	3594.9
56U99-208	49.1	4026.9
56U99-028	49.1	4092.4
56U99-075	49.1	4043.6
56U99-202	49.1	4154.9
56U99-203	49.1	4251.8
56U99-093	49.1	3696.3
56U99-023	49.1	3754.7
56U99-100	49.0	4118.0
56U99-086	49.0	4679.0
56U99-136	49.0	4269.8
56U99-179	49.0	3790.8
56U99-026	48.9	3559.9
56U99-029	48.9	3625.5
56U99-185	48.9	4260.7
56U99-048	48.9	3935.6
56U99-012	48.9	4204.7
56U99-211	48.8	3607.5

Table 2.6 Continued.

Line	Meal	Yield
	%	kg ha^{-1}
56U99-155	48.8	4411.7
56U99-189	48.8	3695.7
56U99-025	48.7	3557.4
56U99-125	48.7	4158.0
56U99-008	48.7	3724.8
56U99-061	48.7	3806.0
56U99-238	48.7	4109.5
56U99-130	48.6	4333.7
56U99-223	48.6	4311.7
56U99-082	48.6	4040.1
56U99-102	48.6	3588.7
56U99-105	48.6	3815.6
56U99-159	48.6	3828.8
56U99-107	48.5	3911.8
56U99-232	48.5	3767.5
Osage[a]	48.4	4261.3
56U99-080	48.4	3501.1
56U99-242	48.4	4279.9
56U99-166	48.3	3703.3
56U99-133	48.2	3958.2
56U99-237	48.2	3908.3
56U99-199	48.1	4410.0
56U99-243	48.1	3839.2
56U99-069	48.1	4106.8
56U99-056	48.1	4066.7
56U99-191	48.1	3973.3
56U99-215	48.1	4249.9
56U99-018	48.1	3732.6
56U99-104	48.0	3747.4
56U99-206	48.0	4209.1
56U99-157	48.0	4225.2
56U99-221	47.6	3992.7
5601T[b]	47.5	4360.4
LSD (0.05)	8.2	647.4

[a] Checks included in field trials were Ellis, 5002T, TN11-5104, and Osage.
[b] Parent included in field trials was 5601T.

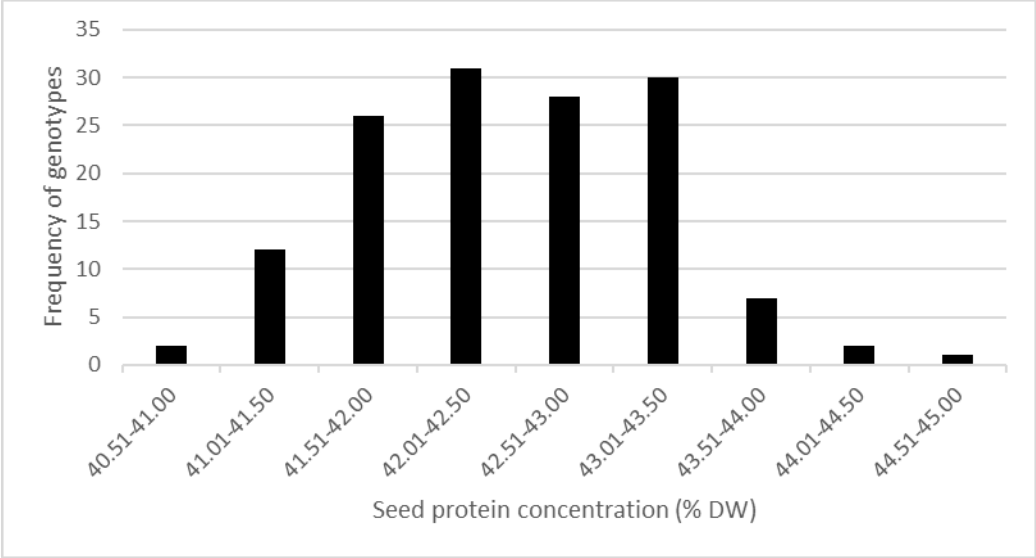

Figure 2.1 Frequency distributions for protein concentration means for a population of 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 grown over six environments with three replications each in 2017 and 2018.

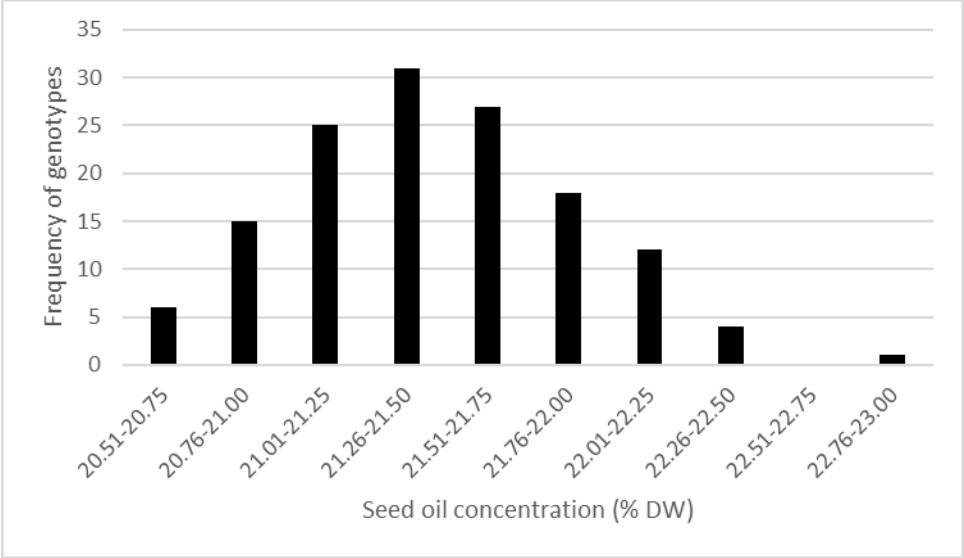

Figure 2.2 Frequency distributions for oil concentration means for a population of 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 grown over six environments with three replications each in 2017 and 2018.

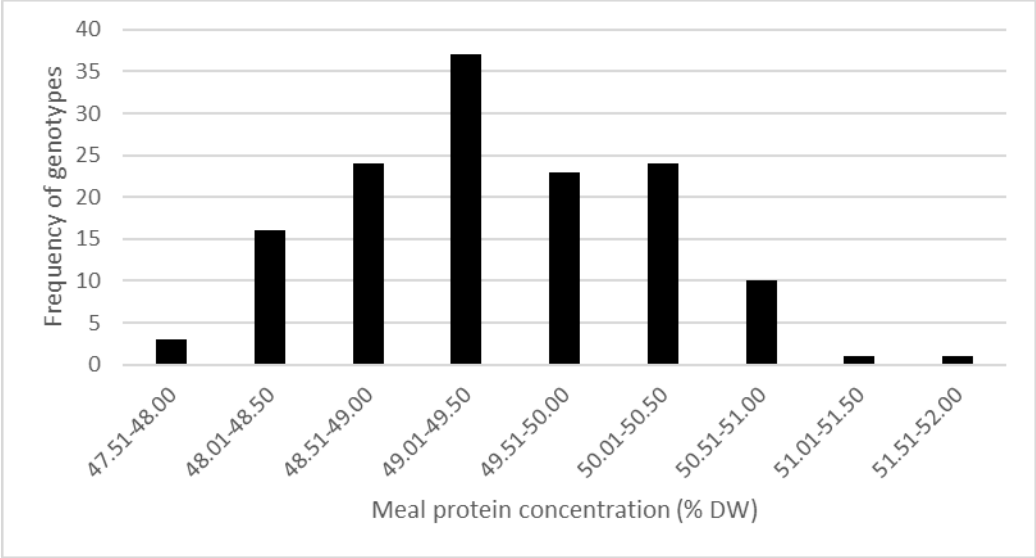

Figure 2.3 Frequency distributions for meal protein concentration means for a population of 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 grown over six environments with three replications each in 2017 and 2018.

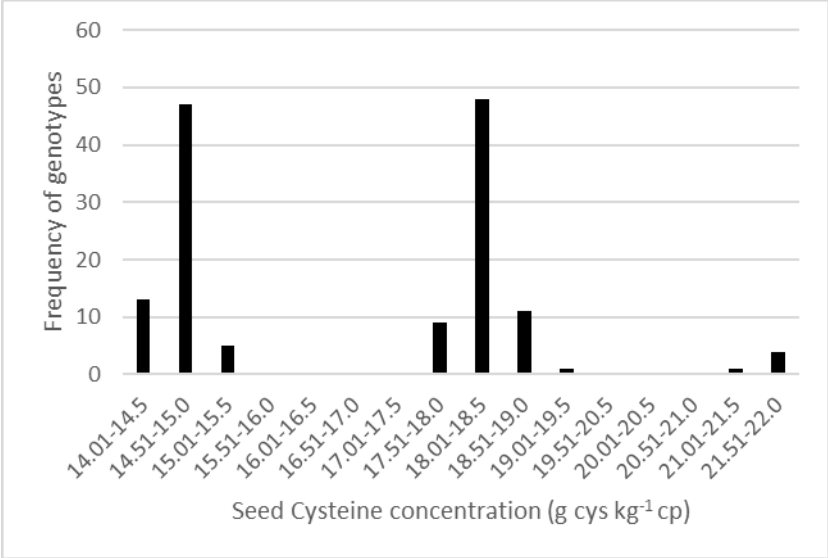

Figure 2.4 Frequency distributions for cysteine concentration means for a population of 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 grown over six environments with three replications each in 2017 and 2018.

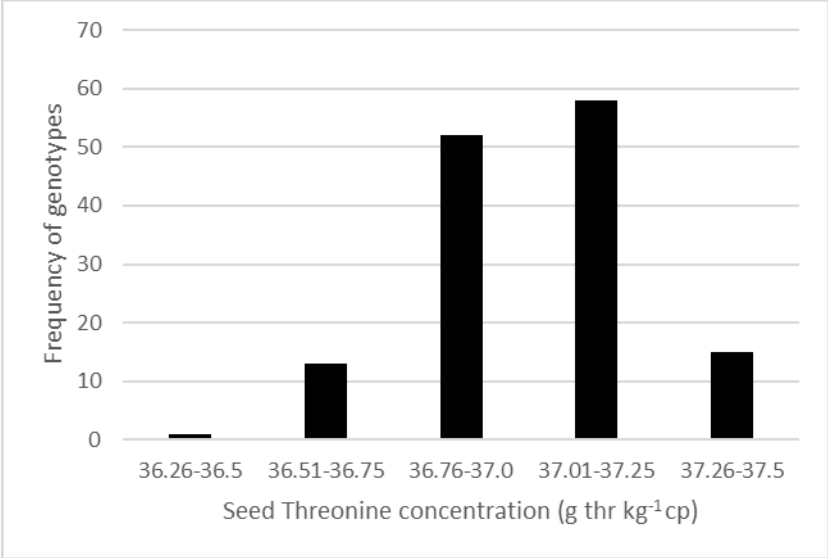

Figure 2.5 Frequency distributions for threonine concentration means for a population of 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 grown over six environments with three replications each in 2017 and 2018.

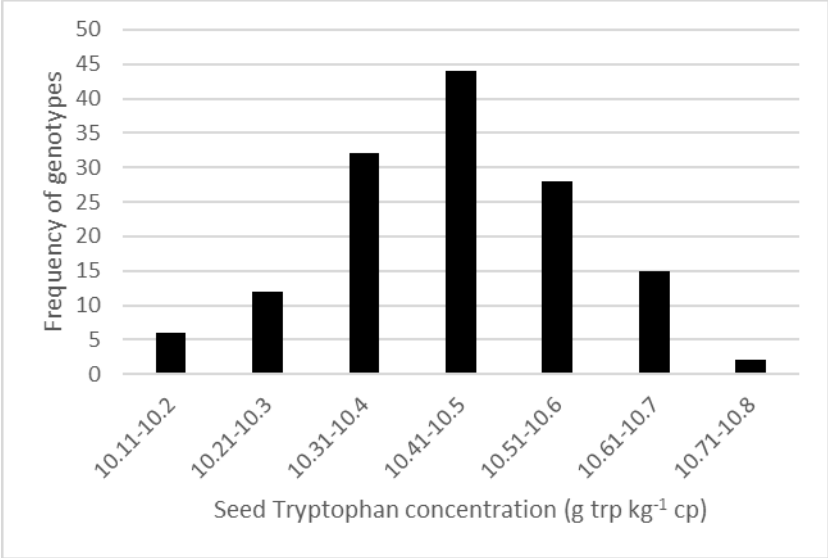

Figure 2.6 Frequency distributions for tryptophan concentration means for a population of 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 grown over six environments with three replications each in 2017 and 2018.

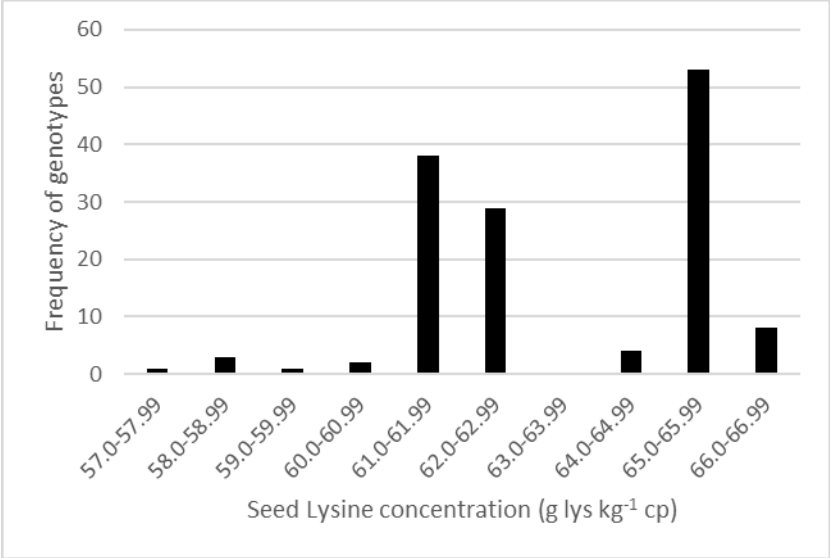

Figure 2.7 Frequency distributions for lysine concentration means for a population of 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 grown over six environments with three replications each in 2017 and 2018.

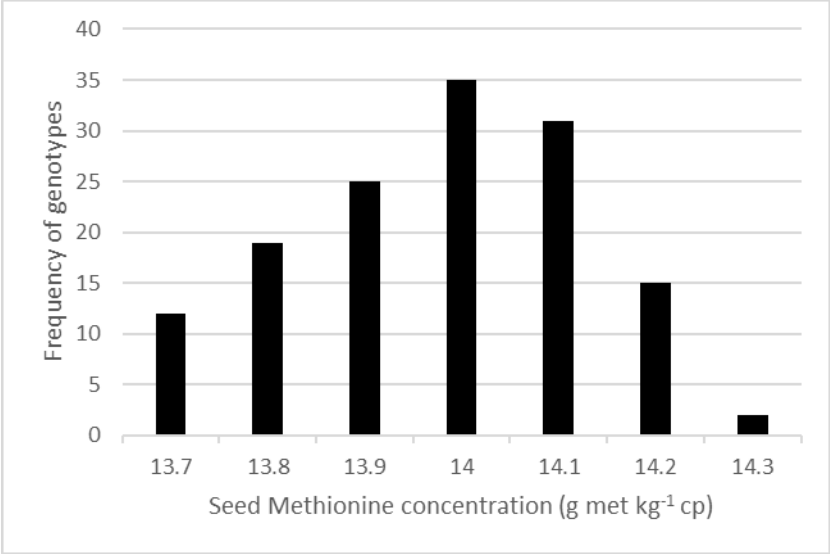

Figure 2.8 Frequency distributions for methionine concentration means for a population of 138 individual $F_{8:10}$ recombinant inbred lines (RIL) of cross 5601T × U99-310255 grown over six environments with three replications each in 2017 and 2018.

CHAPTER III: QUANTITATIVE TRAIT LOCI DETECTION AND ANALYSIS

Abstract

Molecular markers have played and will continue to play a major role in the genetic characterization and improvement of soybeans. They have helped identify major loci for tolerance to abiotic stressors, disease resistance, herbicide resistance, soybean seed quality traits, and yield. However, most yield quantitative trait loci (QTL) are specific to a certain population, and the genetic variation found in the specific bi-parental population is not always shared in other populations. A major objective in soybean breeding is to develop high yielding cultivars. Unfortunately, soybean seed yield, as well as protein and oil content, are complex quantitative traits to characterize from the phenotypic and genotypic perspectives. The objectives of this study are to detect soybean genomic regions that increase protein content, while maintaining oil content and seed yield and to successfully identify soybean QTL associated with these seed quality traits. To achieve these objectives, data from the 138 recombinant inbred lines (RIL) grown in six environments were used to perform QTL detection analyses in search of significant genomic regions affecting soybean seed protein, oil, and yield. A total of twenty-one QTL were successfully identified for yield, protein, oil, methionine, threonine, lodging, maturity, and meal. Knowledge of their locations and flanking markers will aid in marker assisted selection for plant breeders. This will lead to a more valuable soybean for farmers, processors, and animal nutritionists.

Introduction

Soybean yield, oil, protein, and amino acids are commercially important traits. For this reason, it is important to develop an improved understanding of significant genomic regions dealing with these traits. Most agronomically important traits are controlled by multiple genes. Traits controlled by many genes are known as quantitative traits (Bernardo, 2014; Collard et al., 2005). Yield, disease resistance, and seed protein and oil content are a few examples of quantitative traits in soybean. Quantitative trait loci (QTL) are regions within a genome that contain genes that influence particular quantitative traits (Collard et al., 2005). Identification of QTLs through phenotypic evaluation only is not possible (Collard et al., 2005). In order to locate, identify, and characterize these regions within a genome, DNA or molecular markers are used. DNA markers are able to reveal sites of variation within DNA. These markers have been used to construct linkage maps in soybean, which are used to identify specific regions of chromosomes containing genes controlling both qualitative and quantitative traits through QTL analysis (Collard et al., 2005). QTL mapping is the process of constructing linkage maps and conducting QTL analysis, and leads to identifying novel QTL in the genome (Collard et al., 2005).

In plant breeding, marker-assisted selection (MAS) is a tool that uses DNA markers that are either tightly linked or gene tagged to important agronomic traits such as yield, protein, and oil (Collard et al., 2005). MAS acts as a substitute for or aids in phenotypic selection by using the presence or absence of a known marker, making it more efficient, reliable, and cost-effective compared to

conventional plant breeding methods (Collard et al., 2005). For this reason, DNA markers are accepted as valuable tools for the improvement of agronomic and seed quality traits of soybeans. The single nucleotide polymorphism (SNP) marker is the ideal marker for orienting the soybean genome, because SNPs are the most abundant marker available (Hyten et al., 2010). The cultivated soybean has approximately a 0.001 nucleotide diversity (theta), which is an approximate SNP frequency average of 1000 bp of contiguous sequence (Hyten et al., 2010).

Recombinant inbred lines are commonly used to map QTL (Pollard, 2012). When two parental lines, selected based on phenotype, marker availability, and compatibility, are crossed, they create recombinants in the F_2 population which lead to RILs with further inbreeding. The resulting RILs possess a genome that is a mosaic of the parental genomes. They are then inbred to isogenicity, creating genetically stable recombinant inbreds (Broman, 2005; Pollard, 2012). The result is a population that can be used to map traits for QTL analysis (Pollard, 2012).

The quantitative inheritance of soybean protein, oil, and yield makes improving these traits a challenge. A large number of QTL are reported in the literature, suggesting that a large number of genes affect these three traits in soybean. This creates a challenge for plant breeders, as they are forced to select for a large number of QTL, and many of the QTL are not genetically linked or are not on the same chromosomes. Therefore, narrowing the QTL pool and identifying QTL with multiple effects should be made a priority, such as

identifying QTL that increase protein and either have no effect on or simultaneously increase soybean oil or identifying QTL that increase oil and yield simultaneously. Alternatively, identifying high protein QTL with no detriment to either seed yield or oil could help breeders achieve the goal of at least 48% meal protein. Upon identifying any of these seed quality trait QTL, breeders could incorporate them into high yielding cultivars in hopes of stacking QTL when creating novel lines.

An increase in knowledge of genomic regions governing soybean yield, protein, amino acids, and oil would be helpful to breeders. A total of 21 QTL were identified in this study after QTL analysis for these traits was performed on the population. The resulting QTL from this research should be beneficial for breeders interested in improving soybean agronomic and seed quality traits. The knowledge of genomic regions influencing these traits can be applied towards efforts in improving soybean. The objectives of this study are to detect soybean genomic regions that increase protein content, while maintaining oil content and seed yield and to successfully identify soybean seed quality QTL. To accomplish this, a soybean population of 138 recombinant inbred lines (RIL) was developed and genotyped using 551 polymorphic SNPs using the Illumina Infinium beadchip on a 6K BARCSoy SNP Chip. This will lead to a more valuable soybean for farmers, processors, and animal nutritionists. Future research implementing molecular breeding strategies and identifying genes for these and other QTL

impacting targeted soybean agronomic and seed quality traits will be important for soybean breeders.

Materials and Methods

Plant Materials

One hundred and thirty-eight $F_{8:11}$ recombinant inbred lines (RIL) were created at the University of Tennessee, Knoxville, East Tennessee Research and Education Center (ETREC) from a cross between 5601T (Pantalone et al., 2003) and U99-310255 in the summer of 2006 (Table 1.1). Successful crossing in 2006 led to F_1 seed, which were continually allowed to self-pollinate over nine years until the experimental population reached the $F_{8:10}$ and $F_{8:11}$ generations in 2017 and 2018, respectively. The $F_{8:10}$ RILs were planted in a randomized complete block design with two replications in Knoxville, Springfield, and Milan, TN during the 2017 growing season. The seeds were combine harvested in the fall of 2017 and led to the planting of the $F_{8:11}$ RILs in three replications in the same three environments in the 2018 growing season. The seeds were combine harvested in the fall of 2018 (Table 1.1). An RIL is a nearly homozygous line created when generations are continually self-pollinated following a cross between two parental inbreds, creating a new inbred line with a mosaic of the parental genomes (Bernardo, 2014a; Bernardo, 2014b; Bernardo, 2014c; Broman, 2005). The resulting RILs all vary in their genomes, and are therefore advantageous in genetic mapping for quantitative trait loci (QTL) analysis (Broman, 2005).

DNA Extraction and SNP Genotyping

Leaf tissue samples for experimental RILs grown in the greenhouse in Knoxville, TN were collected in the fall of 2018. Approximately 3 newly emerged leaflets per RIL were pulled. Additionally, leaves were pulled from the two parents of the population, 5601T and U99-310255, to sample DNA from those lines as well. When the leaflets were pulled, they were placed into 2µl labeled tubes with the appropriate plot number and were temporarily placed on ice. The samples were flash frozen in liquid nitrogen before being stored at -80° C until DNA extraction. Genomic DNA was extracted using the DNeasy® Plant Mini Kit (QIAGEN®, Venlo, Netherlands) according to manufacturer's directions. Following DNA extraction, the DNA was tested on a NanoDrop to check for concentration. Following this, a 1% Agarose gel was poured, and the extracted DNA was run in the gel to ensure a good band for the genomic DNA with no degradation. Once the DNA samples were checked and run on the agarose gel electrophoresis, a 50 µL sample from each of the RILs and the two parents was placed into templated 96-well plates. The plates were placed on dry ice and shipped to Dr. Qijian Song with the USDA-ARS in Beltsville, MD. In Maryland, the DNA samples were analyzed using the Illumina Infinium beadchip on a 6K BARCSoy SNP Chip. Since soybean has 20 chromosomes and 20 linkage groups assigned to them, a large number of molecular markers is essential in order to span the entire genome (Akond et al., 2013; Contreras-Soto et al., 2017). The 6K BARCSoy SNP Chip is a medium scale array that uses 6,000 SNPs selected from the SoySNP50K array (Song et al., 2013) to accurately

represent haplotype blocks (Song et al., 2014). This allows for faster but still accurate genotyping of the soybean lines (Lee et al., 2015).

The GenomeStudio Genotyping Module v2.0 (Illumina, Inc. San Diego, CA) program was used to genotype the experimental RIL population after being run on the 6K BARCSoy SNP Chip. The 138 RILs were auto-called by GenomeStudio for each of 6,000 SNP markers. These results for each RIL were then manually checked and adjusted at all marker locations as described by Song et al. (2015). Of the 6,000 SNPs analyzed, 5,683 were initially used for linkage mapping. The rest were thrown out due to missing data or inaccurate software calls.

Linkage Mapping

Genetic linkage maps have been used widely as tools for breeding and genetic research (Song et al., 2015). They are useful in discovering the positions of genes controlling agronomic and seed quality traits as well as in facilitating marker-assisted selection of low heritability traits and are commonly used in quantitative trait loci (QTL) analysis (Song et al., 2015). Genotyped marker data was obtained through the visual scoring of the 6K BARCSoy SNP Chip data using GenomeStudio Genotyping Module v2.0 software. One hundred and thirty-eight lines were mapped using 551 SNP markers. Joinmap 4.1 (Van Ooijen, 2006) was used to estimate genetic distances between the SNP markers. Since the linkage phases were not known prior to linkage mapping, two rounds of linkage analysis were completed. The first was to determine if any loci were out

of linkage phase and the second phase included dummy variables in alternative linkage phases of loci that were not in linkage phase (Olukolu et al., 2009). A minimum likelihood of odds (LOD) ≥3.0 and a maximum distance of ≤50 centimorgan (cM) were used to test linkages among markers. Prior to linkage map analysis, any loci with segregation distortion ($p<0.01$) in the population were removed by chi squared analysis. Additionally, any loci with >10% missing data were eliminated as well, according to Song et al. (2015). Marker distances were calculated using the Kosambi mapping function (Kosambi, 1944) in JoinMap 4.1 (Van Ooijen, 2006). Linkage mapping must be completed prior to QTL analysis. Linkage maps were created using the JoinMap 4.1 (Van Ooijen, 2006) program. The population of 138 recombinant inbred lines was initially genotyped in 2019 using the llumina Infinium beadchip on a 6K BARCSoy SNP Chip with 6000 SNP markers. Of the 6000 SNP markers, 551 total markers were used to create linkage maps across the 20 soybean chromosomes (Figure 3.1 in Appendix C). Linkage maps were refined to 551 markers through the use of heat maps for each chromosome's linkage map. A sample heat map for chromosome 3 in 5601T × U99-310255 population can be seen in Figure 3.2. Heat maps were created upon completion of linkage mapping and helped to refine the maps.

QTL Analysis

QTL analysis was performed using the phenotypic trait value means and SNP genotypic means on 138 RIL planted in six environments in search of significant genomic regions affecting 12 soybean traits. Linkage maps and least

squares means (LSMEAN) combined over locations for 12 traits of interest: yield, protein, oil, cysteine, methionine, tryptophan, lysine, threonine, maturity, lodging, height, and meal were used for QTL identification with QTL Cartographer software v. 2.5 (Wang et al., 2010) with the CIM procedure. Composite interval mapping (CIM) was performed to identify significant QTL when two or more markers were linked. Through QTL analysis in an appropriate population, we can locate genes of interest in specific chromosome regions, estimate the size of their effects, and determine whether their gene action is dominant or additive. Analyses were conducted with the CIM procedure of QTL Cartographer. The standard model in the CIM procedure was run with a control marker number left empty, a 0.1 cM window, a 1 cM walk speed, using a forward regression method. All heterozygote marker loci were excluded from analyses as we were primarily interested in additive genetic effects for the $F_{8:10}$ and $F_{8:11}$ populations. For all 12 traits of interest, 1000 permutation tests were performed to establish empirical logarithm of odds (LOD) thresholds at the 0.05 alpha level (Chruchill and Doerge, 1994). Trait dependent LOD scores were determined based on these permutations and were applied in QTL Cartographer. QTLs with a LOD score exceeding the trait dependent LOD threshold, were considered significant.

QTL Detection

QTL studies entail population development, phenotypic evaluation, and molecular marker evaluation. There are different kinds of population development commonly used for making QTL mapping populations: F_2,

recombinant inbred line (RIL), and backcross populations (Byrne, 2020). In this study, an RIL population was used. RIL populations were developed through a modified single seed descent method from an F_2 population, and lead to homogeneous, homozygous lines able to produce a lot of seed, which is necessary for replicated field trials. The two parents leading to the F_2 population must differ in phenotypic traits and must be polymorphic enough for chromosomal recombination events to be tracked. If the parents of the RILs are polymorphic at a specific locus, meaning they have alleles that vary resulting in phenotypic differences, then we are able to detect the QTL at this locus (Byrne, 2020).

The dataset created by Song et al. (2015) for the SoySNP50K Illumina Infinium BeadChip and, subsequently, the BARCSoySNP6K Beadchip used in this study, is available at Soybase, the USDA, ARS Soybean Genetics and Genomics Database, http://www.soybase.org/snps/download.php and is used in this study for QTL detection and confirmation.

Results and Discussion

The QTL position shown is an interval near the marker that is present on the SoyBase map. A total of 551 SNP markers were used to create a linkage map for the 5601T × U99-310255 population. The resulting genetic linkage map covered 1022.9 cM of the reported 2296 cM reported in the integrated soybean genetic linkage map (Hyten et al., 2010). There were a total of 21 QTL (Table 3.1 in Appendix C) identified through CIM of the RIL population in this study. Of

these 21 QTL, there were three for yield, two for protein, five for oil, one for methionine, two for threonine, five for maturity, one for lodging, and two for meal.

There were a total of three soybean yield QTL that were identified through CIM of the RIL in this study (Figure 3.3). The yield QTL were all located on chromosome (Gm) 16. Currently, there are 165 soybean yield QTL reported on Soybase (Grant et al., 2010). Of these 165, two QTL reported on Soybase were located in the same position range as QTL identified in this study. We consider these as positionally confirmed QTL. The QTL Seed yield 29-2 with a position ranging from 28.38-39.18 cM (Eskandari et al., 2013) overlapped with the first yield QTL found in this study located at 33.3 cM on Gm 16 with flanking marker ss715642535. Additionally, the QTL Seed yield 31-9 with a position ranging from 37.40-67.79 cM (Wang et al., 2014) overlapped with the remaining two QTL found in this study. The first was located at 49.4 cM on Gm 16 with flanking markers ss715624870 and ss715624906. The second was located at 50.9 cM on Gm 16 with flanking markers ss715624888 and ss715624913. QTL positions as well as trait dependent LOD scores can be found in Table3.1.

There were a total of two seed protein concentration QTL that were identified through CIM the RIL in this study (Figure 3.4). The protein QTL were located on chromosome (Gm) 15 and Gm 17. Currently there are 241 seed protein concentration QTL reported in SoyBase (Grant et al., 2010). Of these 241, four QTL reported on Soybase were located in the same position range as QTL identified in this study. Three of these QTL cqseed protein-001 with a

position ranging from 29.8-31.8 cM (Fasoula et al., 2004), Seed protein 5-1 with a position ranging from 29.9-31.9 cM (Lee et al., 1996), and Seed protein 39-2 with a position ranging from 25.74-47.50 cM (Warrington et al., 2015) overlapped with the first seed protein QTL found in this study located at 30.2 cM on Gm 15 with flanking markers ss715620752 and ss715621234. Breeders would have increased confidence in MAS for protein as this is a confirmed QTL. Another QTL found on Soybase, Seed protein 36-17 with a position ranging from 67.7-87.38 cM (Mao et al., 2013), overlapped with the second seed protein QTL found in this study located at 74.2 cM on Gm 17 with flanking marker ss715627848. This is a positionally confirmed QTL. QTL positions as well as trait dependent LOD scores can be found in Table 3.1.

Additionally, five seed oil concentration QTL were identified through CIM in this RIL population (Figure 3.5). Of these, two were located on Gm 2, and the remainder were located on Gms 6, 15, and 18. Currently there are 315 seed oil concentration QTL reported in SoyBase (Grant et al., 2010). Of these 315, five QTL reported on Soybase were located in the same position range as QTL identified in this study. Four of these QTL, mqseed oil-013 with a position ranging from 28.27-35.80 cM (Sun et al., 2011), Seed oil 2-8 with a position ranging from 29.75-33.95 cM (Diers et al., 1992), Seed oil 24-24 with a position ranging from 28.27-35.79 cM (Qi et al., 2011) and Seed oil 27-2 with a position ranging from 31.27-33.27 cM (Reinprecht et al., 2006), overlapped with one seed oil QTL found in this study located at 31.2 cM on Gm 15 with flanking markers

ss7156420752 and ss715621234. Another QTL found on Soybase, Seed oil 26-1 with a position ranging from 7.63-9.63 cM (Zhang et al., 2004) overlapped with another seed oil QTL found in this study located at 8.2 cM on Gm 2 with flanking marker ss715583377 and ss715583392. The remaining three QTL found in this study were not found in Soybase. The first was located at position 1.9 cM on Gm2 with flanking marker ss715583482. The second was located at position 3.1 cM on Gm 6 with flanking markers ss715592984 and ss715593005. The final seed oil QTL was located at position 21.3 cM on Gm 18 with flanking markers ss715631515 and ss715631579 (Table 3.1).

There were three seed amino acid QTL identified in this study. One methionine QTL was identified at position 23.5 cM on Gm 12 with flanking markers ss715612244 and ss715612263 through CIM (Figure 3.6). Additionally, two threonine QTL were identified (Figure 3.7). The first was found at position 30.2 cM on Gm 15 with flanking markers ss715620752 and ss71521234. The second was found at position 42.2 cM on Gm 1 with flanking markers ss715578957 and ss715579141 (Table3.1). To date, there are only 19 soybean seed methionine QTL and nine soybean seed threonine QTL reported on Soybase (Grant et al., 2010).

There were a total of five maturity QTL identified through CIM of the RIL in this study (Figure 3.8). One maturity QTL was located at position 24.6 cM on Gm 18 with flanking markers ss715631579 and ss715631634. The remaining four QTL identified in this study were located on Gm 20. The first was at position 45.5

cM with flanking markers ss715638624 and ss715638667. The next was at position 49.1 cM with flanking markers ss715638812 and ss715638752. The next was found at position 50.4 cM with flanking marker ss715638849. The last was found at position 51.7 cM with flanking markers ss715638885 and ss715638918 (Table3.1). Currently, there are 173 maturity QTL reported on Soybase (Grant et al., 2010).

There was one lodging QTL, located on Gm 15, that was identified in this study (Figure 3.9). The QTL was located at position 32.2 cM with flanking markers ss715620752 and ss715621234 (Table 3.1). There are currently 87 known lodging QTL reported on Soybase (Grant et al., 2010).

Finally, there were two soybean meal QTL identified through CIM of the RIL in this study (Figure 3.10). The first meal QTL was located at position 30.2 cM on Gm 15 with flanking markers ss715620752 and ss715621234. The second was located at position 74.2 cM on Gm 17 with flanking marker ss715627848 (Table3.1).

This study reports significant QTL identified through the CIM function of QTL Cartographer. However, many of the soybean agronomic and seed quality QTL reported in SoyBase (Grant et al., 2010) were not detected through interval mapping, but were identified through simple linear regression methods.

In the future, it would be beneficial to include more RIL in a study like this. While less RIL are more manageable, a lower number of RIL in a population can

lead to inversions of marker order and breakage in linkage groups (Song et al., 2016). Additionally, more RIL lead to higher precision genetic linkage maps.

References

Akond, M., S. Liu, L. Schoener, J.A. Anderson, S.K. Kantartzi, K. Meksem, Q. Song, D. Wang, Z. Wen, D.A. Lightfoot, and M.A. Kassem. 2013. A SNP-based genetic linkage map of soybean using the SoySNP6K Illumina Infinium BeadChip genotyping array. J. Plant Genome Sci. 3: 80-89.

Bernardo, R. Process of plant breeding. 2014a. p. 23-69. Essentials of plant breeding. Stemma Press, Woodbury, MN.

Bernardo, R. Genetic basis of plant improvement. 2014b. p. 71-93. Essentials of plant breeding. Stemma Press, Woodbury, MN.

Bernardo, R. Genetic basis of plant improvement. 2014c. p. 119-146. Essentials of plant breeding. Stemma Press, Woodbury, MN.

Broman, K.W. 2005. The genomes of recombinant inbred lines. Genetics. 169: 1133-1146.

Byrne, P. 2020. Quantitative Trait Locus (QTL) Analysis 1. Plant and Soil Sciences eLibrary. https://passel.unl.edu/pages/informationmodule.php?idinformationmodule=1031263034&topicorder=4&maxto=7&minto=1

Churchill, G.A., and R.W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Collard, B.C.Y., Jahufer, M.Z.Z., Brouwer, J.B., and Pang, E.C.K. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica. 142: 169-196.

Contreras-Soto, R.I., M. Berwanger de Oliveira, D. Costenaro-da-Silva, C.A. Scapim, and I. Schuster. 2017. Population structure, genetic relatedness and linkage disequilibrium blocks in cultivars of tropical soybean (*Glycine max*). Euphytica. 213: 173-184.

Diers, B.W., P. Keim, W.R. Fehr, and R.C. Shoemaker. 1992. RFLP analysis of soybean seed protein and oil content. Theor. Appl. Genet. 83: 608-612.

Eskandari, M., E.R. Cober, and I. Rajcan. 2013. Genetic control of soybean seed oil: II. QTL and genes that increase oil concentration without decreasing protein or with increased yield. Theor. Appl. Genet. 126:1677–1687.

Fasoula, V.A., D.K. Harris, and H.R. Boerma. 2004. Validatioon and designation of quantitative trait loci for seed protein, seed oil, and seed weight from two soybean populations. Crop sci. 2004. 44:1218-1225.

Grant, D., Nelson, R.T., Cannon, S.B. and Shoemaker, R.C. (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. Nucl. Acids Res. (2010) 38 (suppl 1): D843-D846. doi: 10.1093/nar/gkp798

Hyten D.L., I.Y. Choy, Q. Song, J.E. Specht, T.E. Carter Jr., R.C. Shoemaker, E.Y. Hwang, L.K. Matukumalli, and P.B. Cregan. 2010. A high density integrated genetic linkage map of soybean and the development of a 1536 universal soy linkage panel for quantitative trait locus mapping. Crop Sci. 50: 960-968.

Kosambi, D.D. 1944. The estimation of map distance from recombination values. Ann. Eugen. 12:172-175.

Lee, S., K.R. Freewalt, L.K. McHale, Q. Song, T. Jun, A.P. Michel, A.E. Dorrance, M.A.R. Mian. 2015. A high-resolution genetic linkage map of soybean based on 357 recombinant inred lines genotyped with BARCSoySNP6K. Mol. Breeding. 35: 58-65.

Lee, S.H., M.A. Bailey, M.A.R. Mian, T.E. Carter Jr., E.R. Shipe, D.A. Ashley, W.A. Parrott, R.S. Hussey, and H.R. Boerma. 1996. RFLP loci associated with soybean seed protein and oil content across populations and locations. Theor. Appl. Genet. 93: 649-657.

Mao, T., Z. Jiang, Y. Han, W. Teng, X. Zhao, and W. Li. 2013. Identification of quantitative trait loci underlying seed protein and oil contents of soybean across multi-genetic backgrounds and environments. Plant Breed. 132: 630-641.

Olukolu, B.A., T. Trainin, S. Fan, C. Kole, D.G. Bielenberg, G.L. Reighard, A.G. Abbott, and D. Holland. 2009. Genetic linkage mapping for molecular dissection of chilling requirement and budbreak in apricot (*Prunus armeniaca* L.). Genome. 52: 819-828.

Pantalone, V.R., F.L. Allen, and D. Landau-Ellis. Registration of ‘5601T’ soybean. 2003. Crop Sci. 43: 1123-1124.

Pollard, D.A. Design and construction of recombinant inbred lines. 2012. Methods Mol Biol. 871:31-39.

Qi, Z., Q. Wu, X. Han, Y. Sun, X. Du, C. Liu, H. Jiang, G. Hu, and Q. Chen. 2011. Soybean oil content QTL mapping and integrating with meta-analysis method for mining genes. Euphytica. 179: 499-514.

Reinprecht, Y., V. Poysa, K. Yu, I. Rajcan, G. Ablett, and K. Pauls. 2006. Seed and agronomic QTL in low linolenic acid, lipoxygenase-free soybean (Glycine max (L.) Merrill) germplasm. Genome. 49: 1510-1527.

Song, Q., J. Jenkins, G. Jia, D.L. Hyten, V. Pantalone, S.A. Jackson, J. Schmutz, and P.B. Cregan. 2016. Construction of high resolution genetic linkage maps to improve the soybean genome sequence assembly Glyma1.01. BMC Genomics. 17:33-43.

Song, Q., Hyten D.L., Jia G., Quigley C.V., Fickus E.W., Nelson R.L. 2015. Fingerprinting soybean germplasm and its utility in genomic research. G3: Genes Genomes Genetics. 50:1999–2006.

Song, Q., G. Jia, C.V. Quigley, E.W. Fickus, D. Hyten, R.L. Nelson, P.B. Cregan. 2014. Soybean BARCSoySNP6K Beadchip—a tool for soybean genetics research. Plant & Animal Genome XXII, Jan 10-15, 2014, San Diego, CA, Abstract No. P306. http://pag.confex.com/pag/xxii/webprogram/Paper10932.html

Song, Q., D.L. Hyten, G. Jia, C.V. Quigley, E.W. Fickus, R.L. Nelson, P.B. Cregan. 2013. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLOS ONE. 8: 1-12.

Sun, Y., H. Luan, Z. Qi, D. Shan, C. Liu, G. Hu, Q. Chen, Z. Qi, Q. Wu, X. Han, Y. Sun, X. Du, C. Liu, H. Jiang, G. Hu, Q. Chen, and D. Shan. 2011. An integrated quantitative trait locus map of oil content in soybean, Glycine max (L.) Merr., generated using a meta-analysis method for mining genes. Euphytica. 179: 499-514.

Van Ooijen, J.W. 2006. JoinMap ® 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen, Netherlands.

Wang, X., G. Jiang, M. Green, R. Scott, Q. Song, D. Hyten, and P. Cregan. 2014. Identification and validation of quantitative trait loci for seed yield, oil, and protein contents in two recombinant inbred line populations of soybean. Mol. Genet. Genomics. 289: 935-949.

Wang S, Basten CJ, Gaffney P, Zeng ZB: Windows QTL Cartographer 2.5 user manual. 2010, Bioinformatics Research Center, North Carolina State Univ., Raleigh, NC

https://brcwebportal.cos.ncsu.edu/qtlcart/WinQTLCart.pdf

Warrington, C.V., H. Abdel-Haleem, D.L. Hyten, P.B. Cregan, J.H. Orf, A.S. Killam, N. Bajjalieh, Z. Li, and H.R. Boerma. 2015. QTL for seed protein and amino acids in the Benning x Danbaekkong soybean population. Theor. Appl. Genet. 128:2474.

Zhang, W., Y. Wang, G. Luo, J. Zhang, C. He, X. Wu, J. Gai, and S. 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. 2004. 108: 1131-1139.

Appendix C

Table 3.1 QTL found significant at the 5% alpha level associated with soybean seed yield, protein, oil, amino acid, lodging, maturity, and meal in the 5601T × U99-310255 RIL population based on QTL Cartographer analysis.

QTL	Gm	Flanking Markers	QTL position (cM)	r^2	LOD	Effect[a]
Threonine	1	ss715578957, ss715579141	42.2	39.0	3.6	-0.1
Oil	2	715583482	1.9	11.7	3.6	-1.5
Oil	2	ss715583377, ss715583392	8.2	13.1	3.7	1.6
Oil	6	ss715592984, ss715593005	3.1	11.9	3.6	-1.5
Methionine	12	ss715612244, ss715612263	23.5	39.1	4.5	-0.1
Protein	15	ss715620752, ss715621234	30.2	28.5	5.2	-4.1
Oil	15	ss715620752, ss715621234	31.2	32.4	3.3	2.5
Threonine	15	ss715620752, ss715621234	30.2	26.2	5.3	0.1
Lodging	15	ss715620752, ss715621234	32.2	58.1	5.9	-0.6
Meal	15	ss715620752, ss715621234	30.2	23.1	4.2	-3.8
Yield	16	ss715624888, ss715624913	50.9	11.7	3.3	-143.0
Yield	16	ss715624870, ss715624906	49.4	13.7	4.4	-154.4
Yield	16	715624535	33.3	20.6	6.9	-188.6
Protein	17	715627848	74.2	12.7	3.9	2.7
Meal	17	715627848	74.2	11.4	3.5	2.7
Oil	18	ss715631515, ss715631579	21.3	13.2	3.3	1.6
Maturity	18	ss715631579, ss715631634	24.6	13.9	3.6	-1.1
Maturity	20	ss715638624, ss715638667	45.5	10.3	3.1	1.0
Maturity	20	ss715638812, ss715638752	49.1	13.2	3.3	-1.1
Maturity	20	715638849	50.4	22.1	4.4	1.4
Maturity	20	ss715638885, ss715638918	51.7	17.1	5.5	1.3

[a] negative values correspond to parent 5601T, positive values correspond to parent U99-310255.

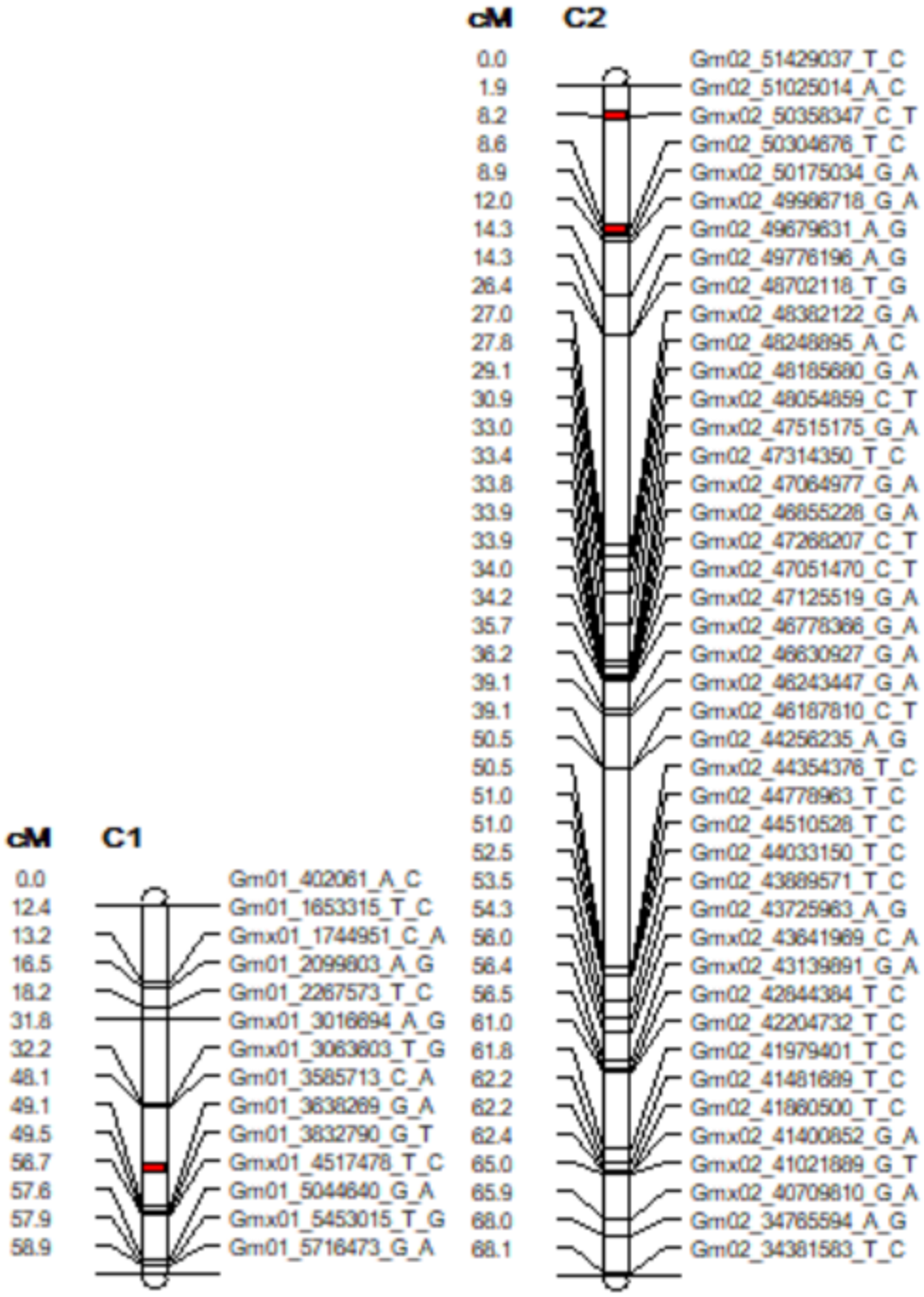

Figure 3.1 Linkage maps for each of the soybean chromosomes with QTLs present from cross 5601T × U99-310255, created using 551 single nucleotide polymorphism (SNP) markers. QTL are highlighted in red on the linkage maps.

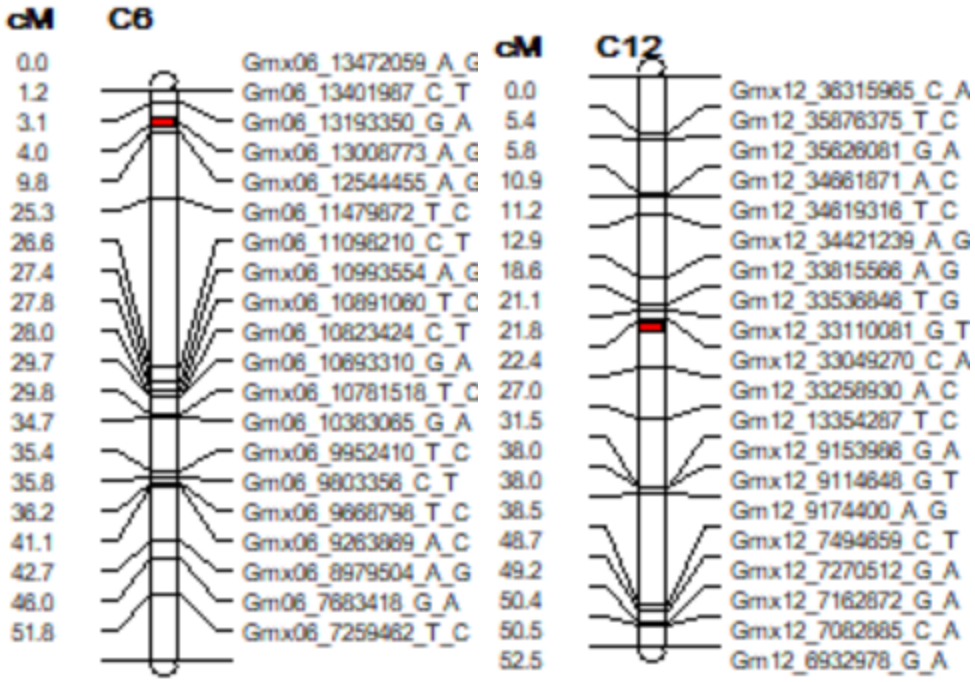

Figure 3.1 Continued.

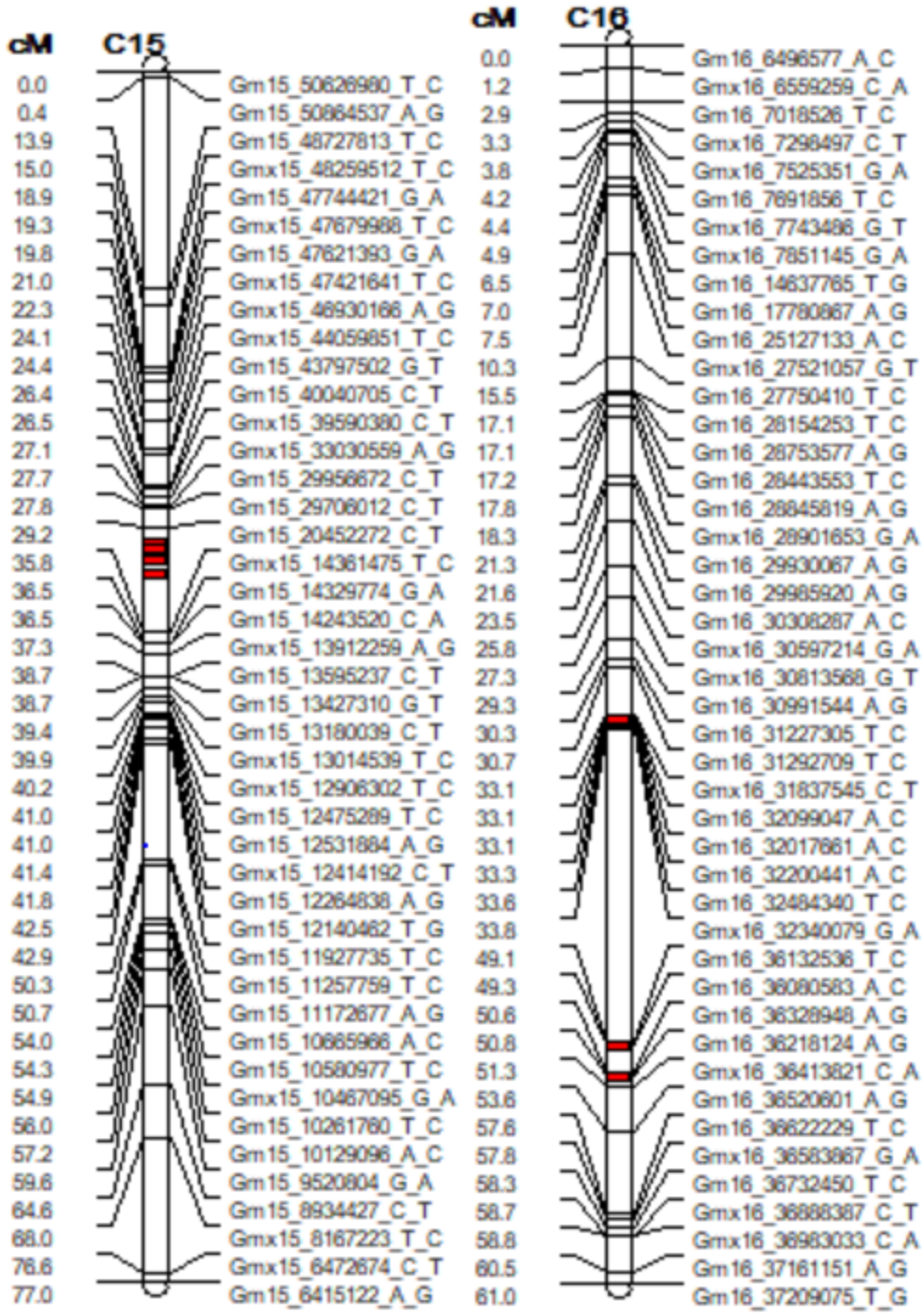

Figure 3.1 Continued.

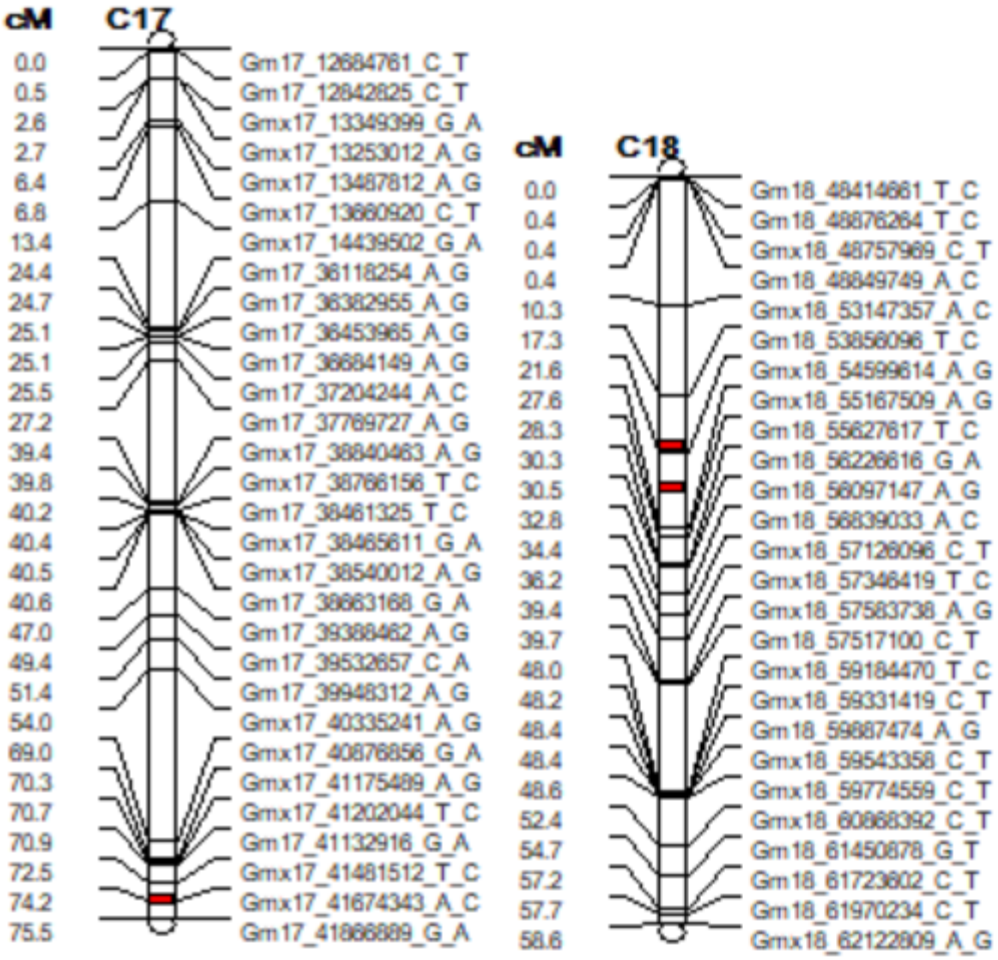

Figure 3.1 Continued.

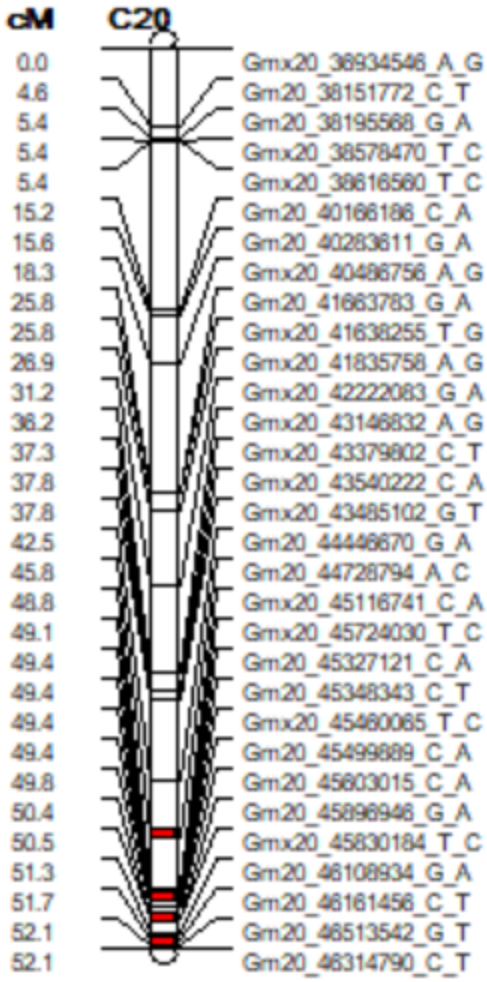

Figure 3.1 Continued.

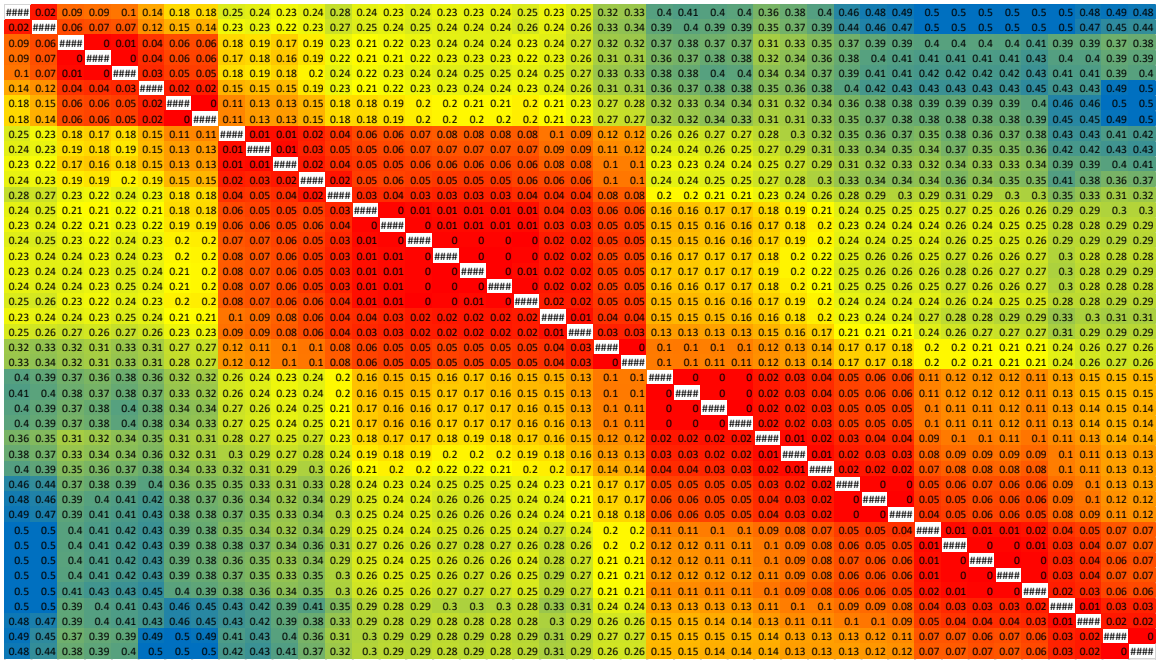

Figure 3.2 Sample heatmap of soybean chromosome 3 from a cross of 5601T × U99-310255 with 48 markers represented used to refine linkage map for quantitative trait loci (QTL) analysis.

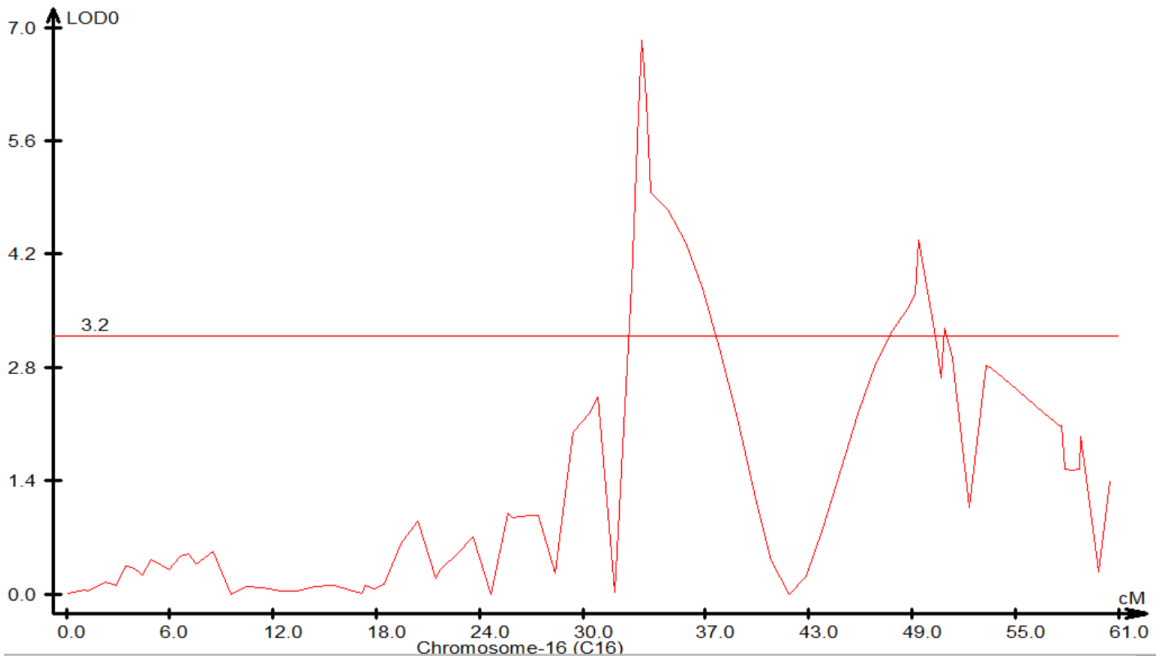

Figure 3.3 Yield QTL discovered in the 2017 $F_{8:10}$ and 2018 $F_{8:11}$ 5601T × U99-310255 population. Trait dependent LOD thresholds were determined at the 0.05 alpha level through 1000 permutation test.

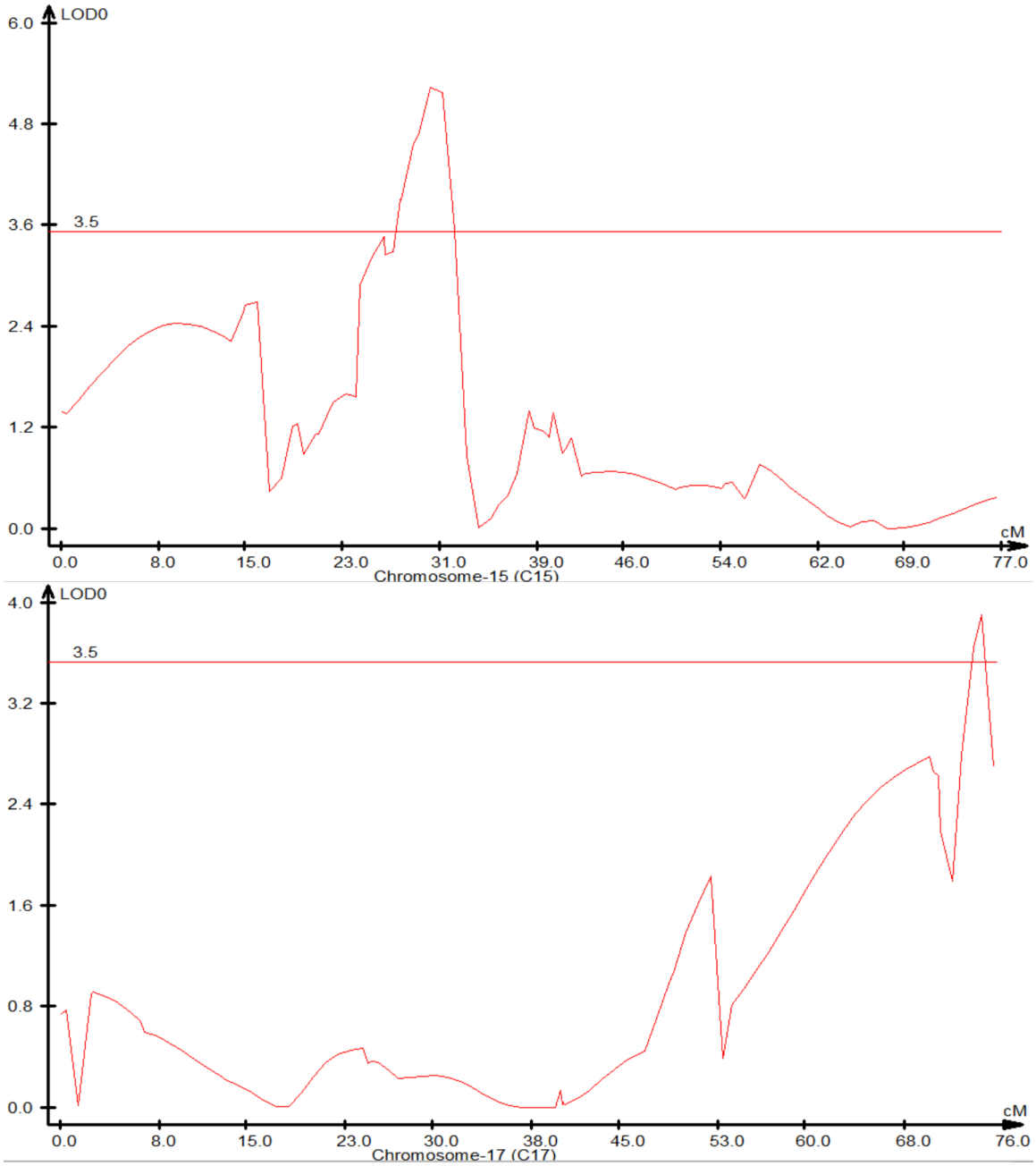

Figure 3.4 Protein QTL discovered in the 2017 $F_{8:10}$ and 2018 $F_{8:11}$ 5601T × U99-310255 population. Trait dependent LOD thresholds were determined at the 0.05 alpha level through 1000 permutation test.

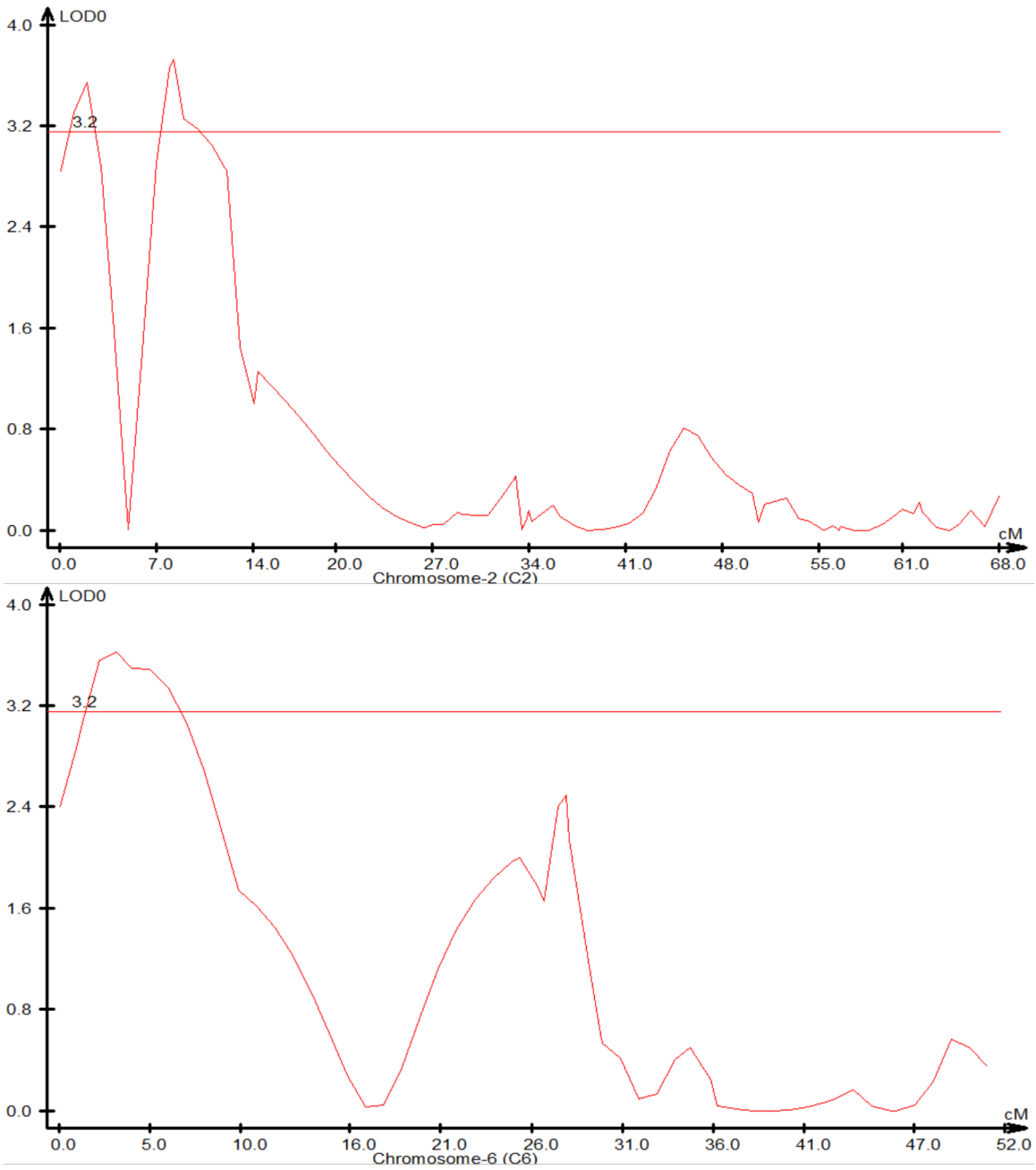

Figure 3.5 Oil QTL discovered in the 2017 $F_{8:10}$ and 2018 $F_{8:11}$ 5601T × U99-310255 population. Trait dependent LOD thresholds were determined at the 0.05 alpha level through 1000 permutation test.

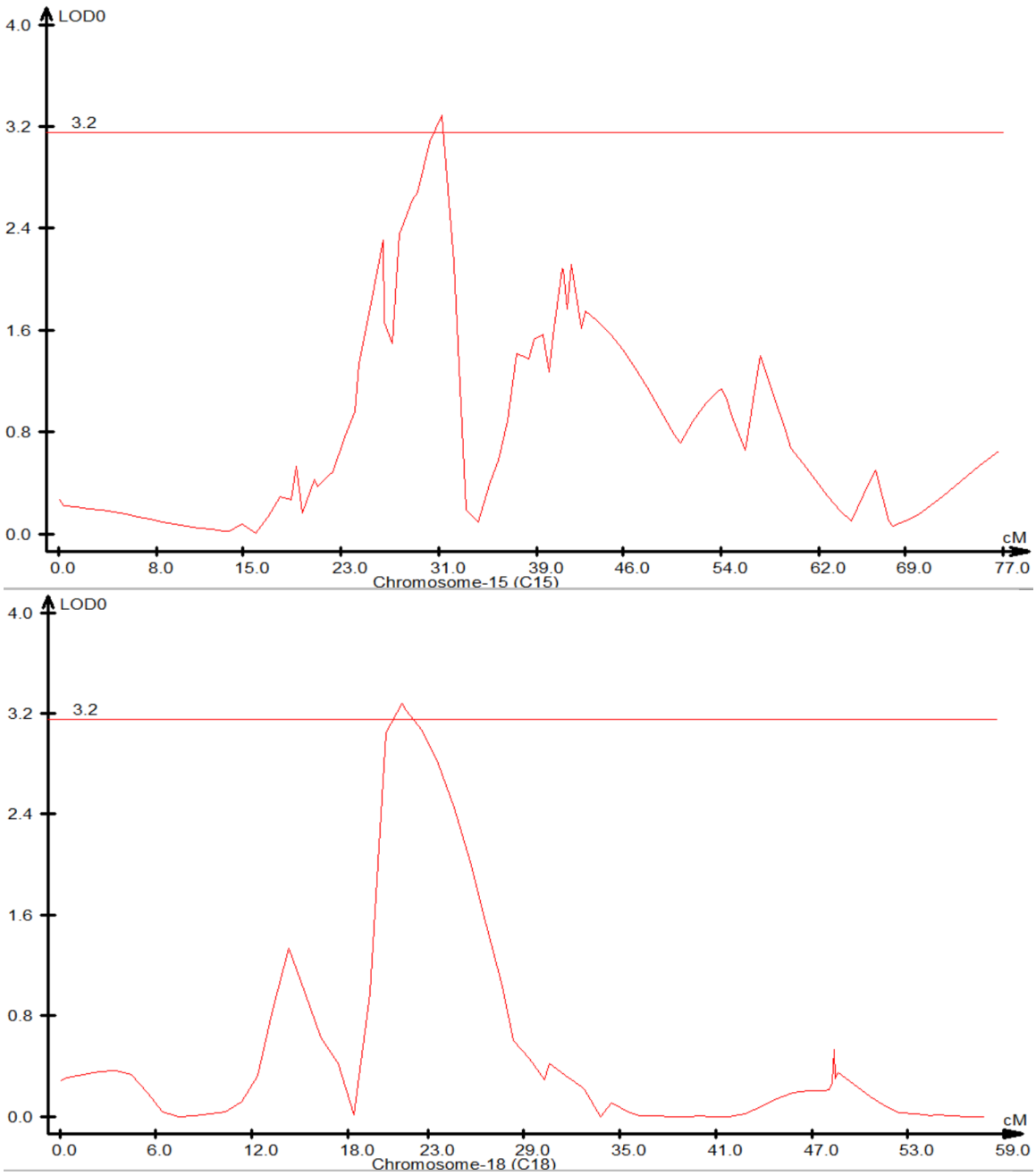

Figure 3.5 Continued.

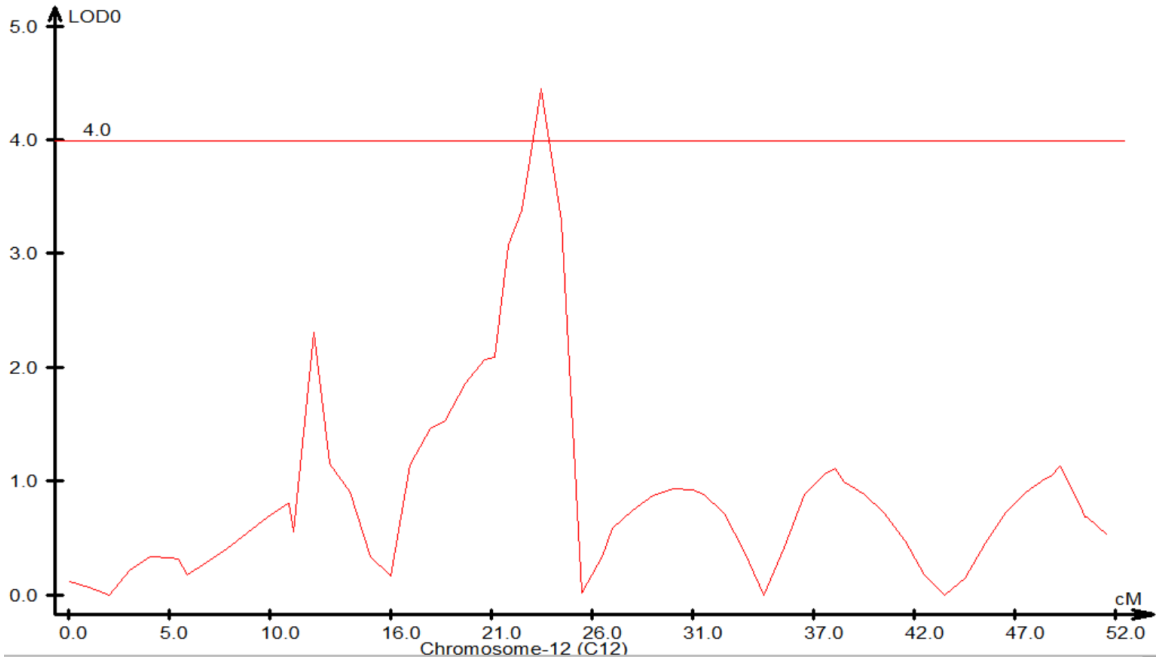

Figure 3.6 Methionine QTL discovered in the 2017 $F_{8:10}$ and 2018 $F_{8:11}$ 5601T × U99-310255 population. Trait dependent LOD thresholds were determined at the 0.05 alpha level through 1000 permutation test.

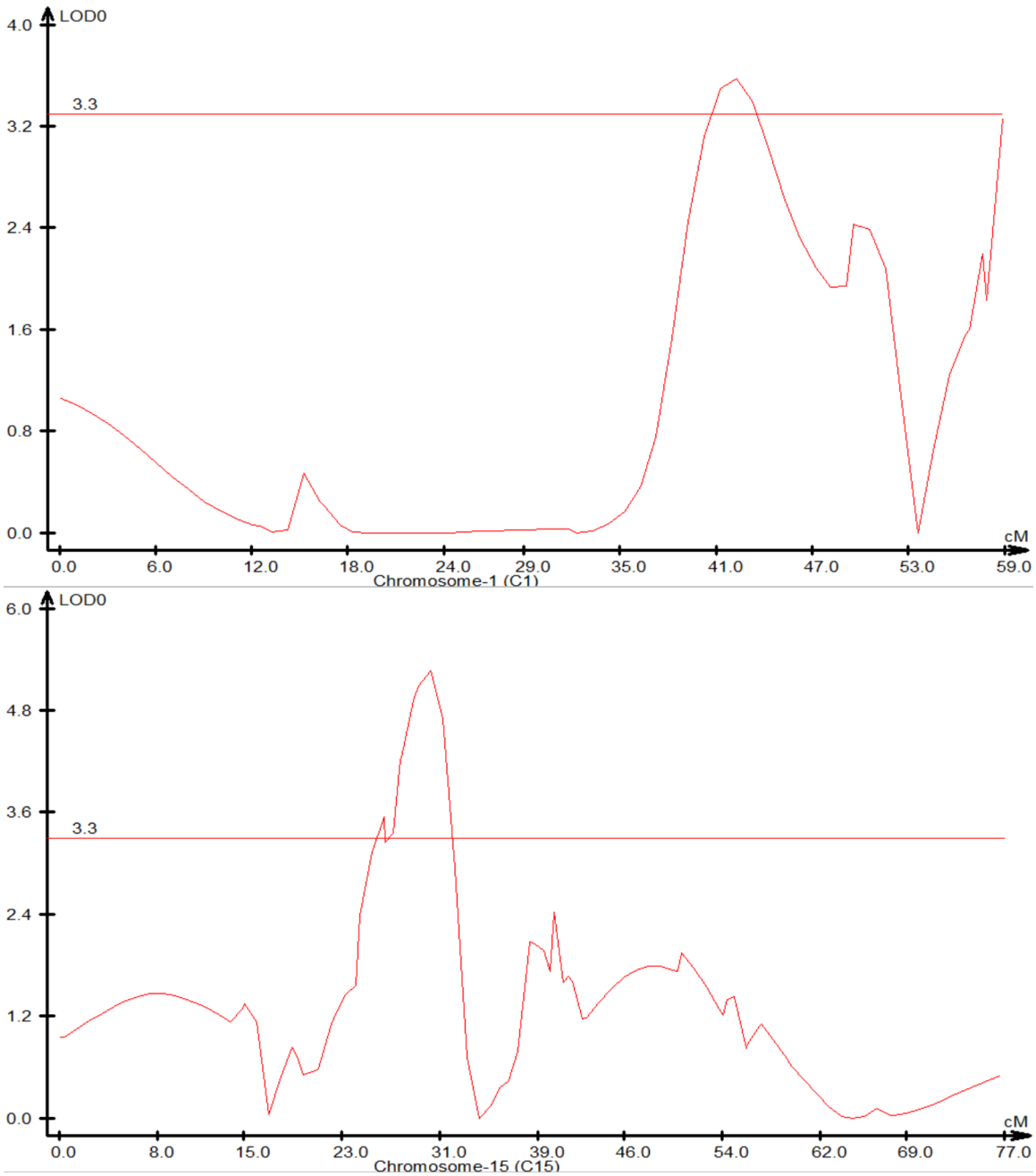

Figure 3.7 Threonine QTL discovered in the 2017 $F_{8:10}$ and 2018 $F_{8:11}$ 5601T × U99-310255 population. Trait dependent LOD thresholds were determined at the 0.05 alpha level through 1000 permutation test.

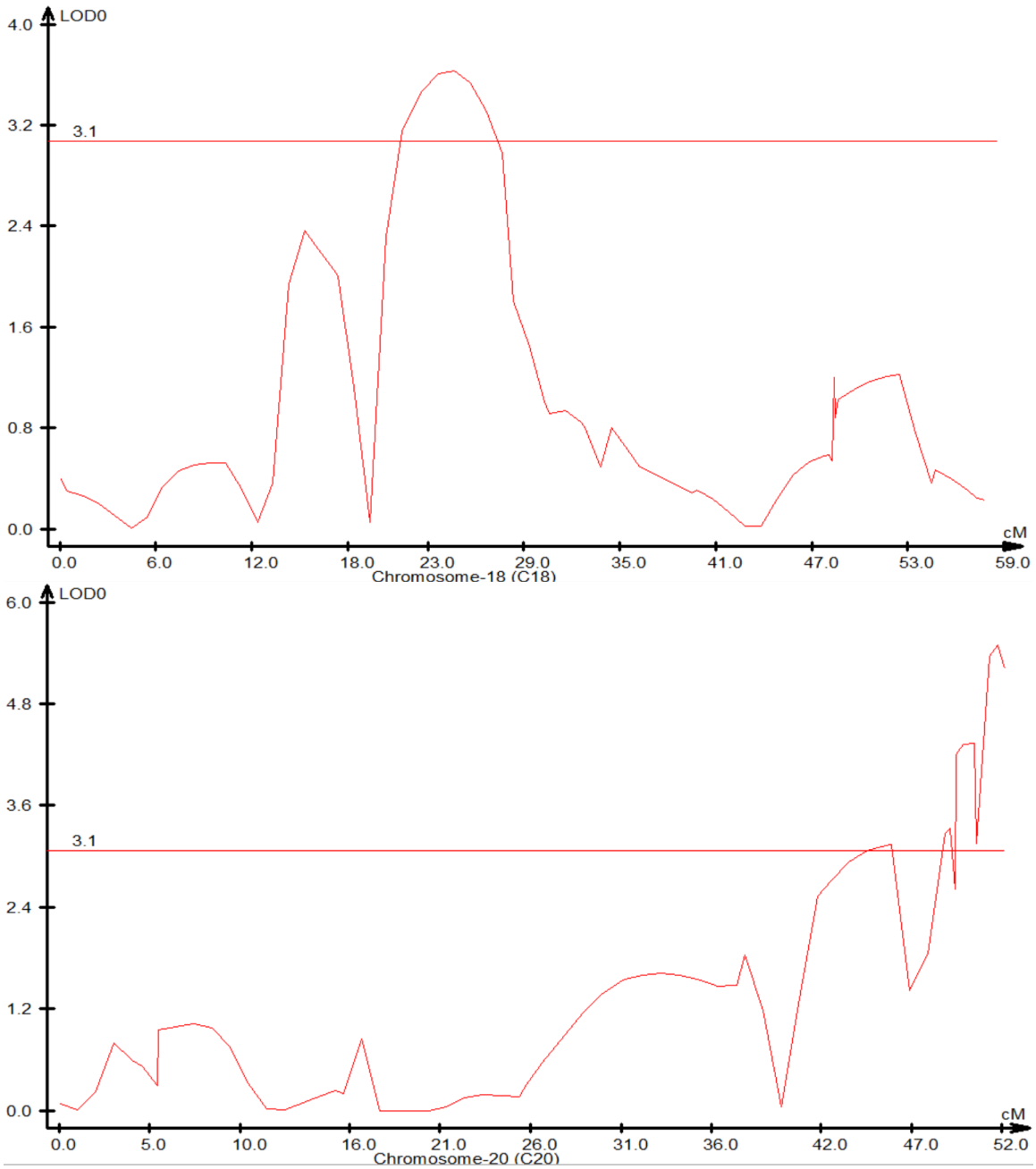

Figure 3.8 Maturity QTL discovered in the 2017 $F_{8:10}$ and 2018 $F_{8:11}$ 5601T × U99-310255 population. Trait dependent LOD thresholds were determined at the 0.05 alpha level through 1000 permutation test.

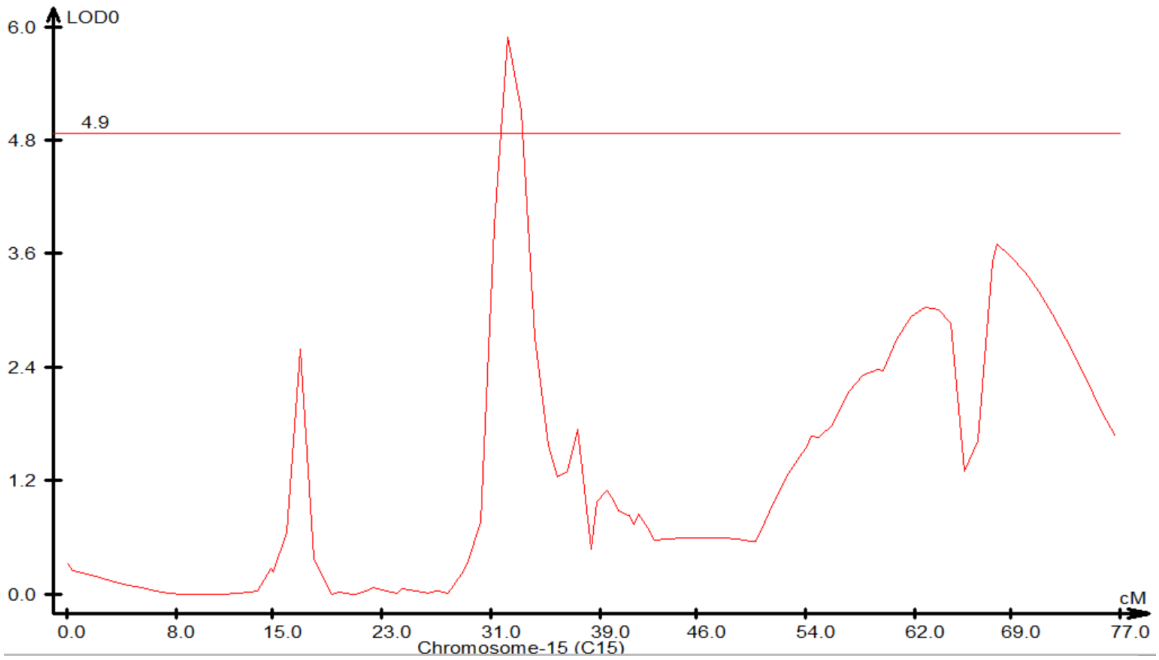

Figure 3.9 Lodging QTL discovered in the 2017 $F_{8:10}$ and 2018 $F_{8:11}$ 5601T × U99-310255 population. Trait dependent LOD thresholds were determined at the 0.05 alpha level through 1000 permutation test.

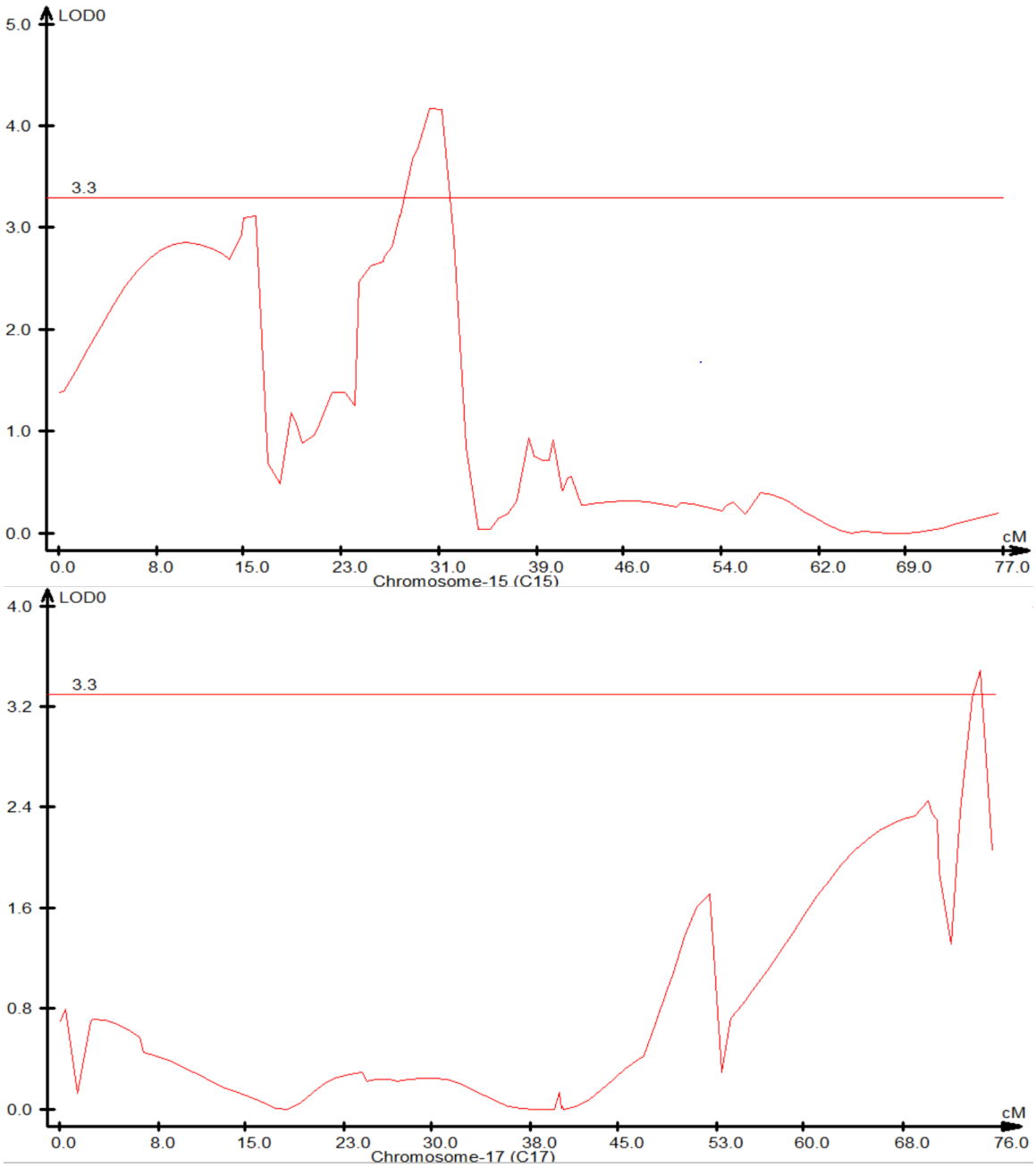

Figure 3.10 Meal QTL discovered in the 2017 $F_{8:10}$ and 2018 $F_{8:11}$ 5601T × U99-310255 population. Trait dependent LOD thresholds were determined at the 0.05 alpha level through 1000 permutation test.

CONCLUSION

Soybean [*Glycine max* (L.) Merr.] is the leading source of oil and low cost, high quality protein meal in the world (Aoyagi and Shurtleff, 2004; Diers et al., 1992; Joshi et al., 2013; Lin et al., 2011; Miller-Garvin and Naeve, 2015; Pantalone, 2012; Warrington et al., 2015; Wilcox, 1998; Wilcox and Shibles, 2001). A common goal in soybean breeding has been to increase protein concentration and quality within the soybean, while maintaining a high yield and oil concentration (Brim and Burton, 1979; Diers et al., 1992; Lee et al., 2010; Pantalone et al., 2020; Pantalone and Smallwood, 2018; Warrington et aI., 2015; Warrington et al., 2014, Wilcox, 1998). In this experiment, one goal is to detect genomic regions governing soybean amino acids to guide breeders forward with genetic gains. Accomplishing these goals presents a challenge, and as there is a negative genetic correlation between protein and both oil and yield, breeding for increased protein will lead to reduced oil and yield (Chung et al., 2003; Cober and Voldeng, 2000; Hernández-Sebastiá et al., 2005; Krishnan et al, 2007; Lee et al., 2010; Warrington et al. 2015; Wilcox, 1998; Yaklich, 2001). Nevertheless it is possible, although difficult, to develop high yielding, high protein lines (Pantalone and Smallwood, 2018; Pantalone et al., 2020). Much research has taken place in an effort to increase soybean protein and improve its amino acid profile (Fallen et al., 2013; Panthee et al., 2006a; Panthee et al., 2006b), and quantitative trait loci (QTL) that increase both oil and protein simultaneously, along with QTL that simultaneously increase oil and yield have been identified by Eskandari et al. (2013). The quantitative inheritance of soybean protein, oil, and

yield challenges the improvement of the crop. There are many QTL reported in the literature and on Soybase (Grant et al., 2010), which suggests that a large number of genes affect these traits in soybean. This creates a challenge for plant breeders, forcing them to select for a large number of QTL, and many of the QTL are not genetically linked or on the same chromosomes. In the future, a priority should be to narrow the QTL pool and to identify QTL with multiple effects, such as identifying QTL that increase protein and either have no effect on or simultaneously increase soybean oil. Alternatively, it would be helpful to identifying high protein QTL with no detriment to seed yield or oil to help breeders achieve 48% or higher soybean meal protein. Upon identifying any of these seed quality trait QTL, breeders could create lines with stacked QTL by incorporating them into high yielding cultivars. The objectives of this study were to detect soybean genomic regions that increase protein content, while maintaining oil content and seed yield and to successfully identify soybean QTL. Twenty-one QTL were identified by this study. Of these twenty-one, there were three for yield, two for protein, five for oil, one for methionine, two for threonine, five for maturity, one for lodging, and two for meal. Many of these QTL were candidates for positional confirmation on Soybase (Grant, 2010). The identification and addition of these QTL to the confirmed soybean QTL will help breeders make selections leading to a more valuable soybean for farmers, processors, and animal nutritionists.

REFERENCES

Aoyagi, A. and W. Shurtleff. 2004. History of world soybean production and trade- Part 1. Soyinfo Center. http://www.soyinfocenter.com/HSS/production_and_trade1.php

Brim, C.A., and J.W. Burton. 1979. Recurrent selection in soybeans. II. Selection for increased percent in protein in seeds. Crop Sci. 19:494-498.

Chung, J., H.L. Babka, G.L. Graef, P.E. Staswick, D.J. Lee, P.B. Cregan, R.C. Shoemaker, and J.E. Specht. 2003. The seed protein, oil and yield QTL on soybean linkage group I. Crop Sci. 43: 1053-1067.

Cober, E.R. and H.D. Voldeng. 2000. Developing high-protein, high-yield soybean populations and lines. Crop Sci. 40:39-42.

Diers, B.W., P. Keim, W.R. Fehr, and R.C. Shoemaker. 1992. RFLP analysis of soybean seed protein and oil content. Theor. Appl. Genet. 83: 608-612.

Eskandari, M., E.R. Cober, and I. Rajcan. 2013. Genetic control of soybean seed oil: I. QTL and genes associated with seed oil concentration in RIL populations derived from crossing moderately high-oil parents. Theor. Appl. Genet. 126:483-495.

Fallen, B.D., C.N. Hatcher, F.L. Allen, D.A. Kopsell, A.M. Saxton, P. Chen, S.K. Kantartzi, P.B. Cregan, D.L. Hyten, and V.R. Pantalone. 2013. Soybean seed amino acid content QTL detected using the universal soy linkage panel 1.0 with 1,536 SNPs. Journal of Plant Genome Sciences 1: 68-79.

Grant, D., Nelson, R.T., Cannon, S.B. and Shoemaker, R.C. (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. Nucl. Acids Res. (2010) 38 (suppl 1): D843-D846. doi: 10.1093/nar/gkp798

Hernández-Sebastiá, C., F. Marsolais, C. Saravitz, D. Israel, R.E. Dewey, and S.C. Huber. 2005. Free amino acid profiles suggest a possible role for asparagine in the control of storage-product accumulation in developing seeds of low- and high-protein soybean lines. J. Exp. Bot. 56:1951-1963.

Joshi, T., B. Valliyodan, J.H. Wu, S.H. Lee, D. Xu, and H.T. Nguyen. 2013. Genomic differences between cultivated soybean, *G. max* and its wild relative *G. soja*. BMC Genomics. 14(Suppl 1):S5.

Krishnan, H.B., S.S. Natarajan, A.A. Mahmoud, and R.L. Nelson. 2007. Identification of glycinin and β-conglycinin subunits that contribute to the increased protein content of high-protein soybean lines. J. Agric. Food Chem. 55:1839-1845.

Lee, J-D., G. Shannon, and M-G. Choung. 2010. Selection for protein content in soybean from single F_2 seed by near infrared reflectance spectroscopy. Euphytica. 172: 117-123.

Lin, J., V.R. Pantalone, G. Li, and F. Chen. 2011. Molecular cloning and biochemical characterization of an endo-β-mannanase gene from soybean for soybean meal improvement. J. Agric. Food Chem. 59: 622-628.

Miller-Garvin, J., and S.L. Naeve. 2015. United States Soybean Quality Annual Report 2015. http://ussec.org/wp-content/uploads/2016/01/2015-Final-US-Commodity-Soybean-Quality-Report.pdf

Pantalone, V.R. 2012. Modern breeding approaches for enhancing soybean protein quality. In: R.F. Wilson, editor, Designing soybean for 21st century markets. AOCS Press, Urbana, IL. p. 197-226.

Pantalone, V.R. and C.J. Smallwood. 2018. Registration of 'TN11-5102' soybean cultivar with high yield and high protein meal. Journal of Plant Registrations. 12: 304-308.

Pantalone, V., M. Cunicelli, and C. Wyman. 2020. Registration of soybean cultivar TN15-5007 with high meal protein. Journal of Plant Registrations. 14: in press.

Panthee, D.R., V.R. Pantalone, A.M. Saxton, D.R. West, and C.E. Sams. 2006a. Genomic regions associated with amino acid composition in soybean. Molecular Breeding. 17: 79-89.

Panthee, D.R., V.R. Pantalone, C.E. Sams, A.M. Saxton, D.R. West, J.H. Orf, and A.S. Killam. 2006b. Quantitative trait loci controlling sulfur containing amino acids, methionine and cysteine, in soybean seeds. Theor. Appl. Genet. 112: 546-553.

Warrington, C.V., H. Abdel-Haleem, D.L. Hyten, P.B. Cregan, J.H. Orf, A.S. Killam, N. Bajjalieh, Z. Li, and H.R. Boerma. 2015. QTL for seed protein and amino acids in the Benning x Danbaekkong soybean population. Theor. Appl. Genet. 128:2474.

Warrington, C.V., H. Abdel-Haleem, J.H. Orf, A.S. Killam, N. Bajjaleih, Z. Li, and H.R. Boerma. 2014. Resource allocation for selection of seed protein and amino acids in soybean. Crop Sci. 54:963-970.

Wilcox, J.R. 1998. Increasing seed protein in soybean with eight cycles of recurrent selection. Crop Sci. 38: 1536-1540.

Wilcox, J.R. and R.M. Shibles. 2001. Interrelationships among seed quality attributes in soybean. Crop Sci. 41: 11-14.

Yaklich, R.W. 2001. β-Conglycinin and glycinin in high-protein soybean seeds. J. Agric. Food Chem. 49:729-735.

VITA

Mia J. Cunicelli was born in Pittsburgh, PA in 1992. She received a Bachelor of Science degree in Biology and a minor in Chemistry from the Virginia Military Institute (VMI) in 2014. After completing her B.S. degree, Mia returned to VMI to work for a year as a researcher before applying to graduate school. In the summer of 2015, she started a Master of Science program at the University of Tennessee, Knoxville. She received her M.S. degree in Plant Sciences, with a concentration in Plant Breeding and Genetics. The title of her thesis was “Effect of a Mutant Danbaekkong Allele on Soybean Seed Yield, Protein, and Oil Concentration.” Mia completed her M.S. degree in the spring of 2017, and transitioned immediately into the Ph.D. program in Plant, Soil, and Environmental Sciences at the University of Tennessee. While completing her dissertation, Mia also worked for the University of Tennessee Soybean Breeding Program as a Research Associate, managing the molecular marker lab. With these combined responsibilities, Mia gained an understanding of both the theoretical and applied components of plant breeding. Upon completion of her dissertation, Mia would like to pursue a career as a breeder in an applied breeding program.