University of Tennessee, Knoxville
TRACE: Tennessee Research and Creative Exchange

Masters Theses | Graduate School

5-2017

Effect of the mutant Danbaekkong or stem termination alleles on soybean seed protein concentration, amino acid composition, and other seed quality and agronomic traits

Mia Justina Cunicelli
University of Tennessee, Knoxville, mcunicel@vols.utk.edu

Follow this and additional works at: https://trace.tennessee.edu/utk_gradthes

Part of the Plant Breeding and Genetics Commons

Recommended Citation

Cunicelli, Mia Justina, "Effect of the mutant Danbaekkong or stem termination alleles on soybean seed protein concentration, amino acid composition, and other seed quality and agronomic traits. " Master's Thesis, University of Tennessee, 2017.
https://trace.tennessee.edu/utk_gradthes/4730

This Thesis is brought to you for free and open access by the Graduate School at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Masters Theses by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

To the Graduate Council:

I am submitting herewith a thesis written by Mia Justina Cunicelli entitled "Effect of the mutant Danbaekkong or stem termination alleles on soybean seed protein concentration, amino acid composition, and other seed quality and agronomic traits." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Plant Sciences.

Vince Pantalone, Major Professor

We have read this thesis and recommend its acceptance:

Hem Bhandari, Carl Sams

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Effect of the mutant Danbaekkong or stem termination alleles on soybean seed protein concentration, amino acid composition, and other seed quality and agronomic traits

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

Mia Justina Cunicelli
May 2017

Copyright © 2017 by Mia J. Cunicelli
All rights reserved.

DEDICATION

I would like to dedicate this thesis to my family and friends who have supported me throughout the entirety of this challenging, yet rewarding journey. I could not have done any of this without you all.

ACKNOWLEDGEMENTS

First I would like to thank my major professor Dr. Vince Pantalone for accepting me to this program and giving me the chance to further my education, while also allowing me to gain hands on experience daily. 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. Hem Bhandari and Dr. Carl Sams. I would like to thank the ladies 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 especially like to thank Chris Smallwood for encouraging me and consistently pushing me outside of my comfort zone to better myself. You are an extraordinary wealth of knowledge, thank you for guiding me through my graduate school career. Your great patience and understanding are two qualities I hope to develop as I lead others as you did. I would also like to thank all the UT REC members at Milan and Highland Rim, especially Chris Bridges and Brad Fisher, for helping me manage my soybeans across the state of Tennessee and for being continuously informative and supportive. I would like to thank everyone in the grad office as well for their support, friendship, and insight into the life of a grad student and for keeping me sane throughout this process.

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. Increasing soybean protein concentration through selection while maintaining oil concentration and yield has been a constant goal for plant breeders, as there is a negative correlation between protein and oil and protein and yield. The objective of this study was to determine if marker assisted selection (MAS) for the Danbaekkong (Dan) protein allele influences agronomic and seed quality traits. A population of 24 $F_{8:10}$ [eighth filial generation advanced to the tenth filial generation] near isogenic lines (NILs) of soybean was created from a cross between G03-3101 and LD00-2817P. Of these 24 NILs, 12 were wild type (WT) and 12 were mutant Dan type. These NILs were grown in a replicated three location field trial across Tennessee. There were significant differences in protein and oil concentrations and yield between the two experimental groups ($p < 0.05$). The Dan experimental group had significantly more protein (414 g kg^{-1}) and less oil (206.9 g kg^{-1}) than the WT and check groups ($p < 0.05$). Additionally, the Dan experimental group was numerically the lowest for yield ranging from 2713-3183 kg ha^{-1}. This result supports previous findings and further solidifies the fact that protein and oil and protein and yield are negatively genetically correlated. Significant differences between Dan and WT experimental groups were observed for all amino acids tested as well ($p < 0.05$). It appears that protein concentrations of Dan NILs were raised with significant reductions in oil and yield. It would be beneficial to further explore the effect of the Dan high protein allele on agronomic and seed quality traits in other soybean growing regions across the United States to test the effect of the genotype by environment interaction.

TABLE OF CONTENTS

INTRODUCTION AND LITERATURE REVIEW 1
Protein Quality 2
Breeding Efforts with Danbaekkong 4
Breeding Efforts for High Protein Soybeans 4
Objectives 8
References 9
CHAPTER I Evaluation of agronomic traits of 24 $F_{8:10}$ near isogenic lines for the Danbaekkong high protein allele 14
Abstract 15
Introduction 17
Materials and Methods 18
Plant Materials 18
DNA Extraction and SNP Genotyping 22
Agronomic Traits 24
Data Analysis 25
Results and Discussion 26
References 29
Appendix 32
CHAPTER II Evaluation of seed protein and oil of 24 $F_{8:10}$ field tested near-isogenic soybean lines 38
Abstract 39
Introduction 40
Materials and Methods 43
Plant Materials 43
Near Infrared Reflectance 44
Gas Chromatography 45
Results and Discussion 46
Near Infrared Reflectance 46
Gas Chromatography 48
References 50
Appendix 54
CHAPTER III Evaluation of agronomic and seed quality traits of near isogenic lines of soybeans that differ at the Dt1 stem termination locus on chromosome 19 61
Abstract 62
Introduction 64
Materials and Methods 66
Plant Materials 66
DNA Extraction and SNP Genotyping 67
Agronomic Traits 67
Near Infrared Reflectance 69
Gas Chromatography 69
Data Analysis 70
Results and Discussion 71

Agronomic Traits 71
Protein, Oil, and Amino Acids 76
Fatty Acids 79
References 81
Appendix 85
CONCLUSION 99
References 100
VITA 101

LIST OF TABLES

Table 1.1 Depiction of population development from inception of cross between G03-3101 and LD00-2817P to summer 2016 field trial. 32
Table 1.2 Chi-Square test results on 72 plants with genotypes: Danbaekkong (Dan), heterozygous (Dan/WT), and wild type (WT) in the MC5 population. The Chi-Squared table value of 5.99 at 2 degrees of freedom with a p-value of 0.05 exceeds the calculated value of 4.75; thus, this population of $F_{7:8}$ plants shows expected 1:2:1 Mendelian inheritance at the chromosome 20 Danbaekkong locus. 33
Table 1.3 Mean yield, maturity, lodging, and height for 24 $F_{8:10}$ lines in two experimental groups [wild type (WT) and Danbaekkong (Dan)] with differences at the Danbaekkong loci. A third experimental group consisting of two check lines (TN12-4100 and Ellis) was included in the study for reference. 34
Table 1.4 Means of agronomic traits for 24 $F_{8:10}$ lines in two experimental groups with differences at the Danbaekkong loci 35
Table 1.5 Mean yield, maturity, lodging, and height of 24 $F_{8:10}$ lines and two checks averaged over three replications and three locations 36
Table 2.1 Mean concentrations of protein, oil, and amino acids for 24 $F_{8:10}$ lines in two experimental groups with differences at the Danbaekkong loci. 54
Table 2.2 Means of protein and oil (g kg^{-1}) for 24 $F_{8:10}$ lines in two experimental groups within a location. 55
Table 2.3 Means of protein and oil (g kg^{-1}) for 12 $F_{8:10}$ Wild Type (WT) lines and 12 $F_{8:10}$ Danbaekkong (Dan) type lines within a location. 56
Table 2.4 Mean concentrations of fatty acids for 24 $F_{8:10}$ lines in two experimental groups with differences at the Danbaekkong loci. 57
Table 2.5 Mean protein, oil, and amino acid concentrations of 24 $F_{8:10}$ lines and two checks averaged over three replications and three locations. 58
Table 2.6 Mean fatty acid concentrations of 24 $F_{8:10}$ lines and two checks averaged over three replications and three locations. 59
Table 3.1 Depiction of population development of conventional line TN14-4001 from inception of the cross of G03-3101× LD00-2817P in 2010 to 2016 test. 85
Table 3.2 Depiction of population development of conventional line TN14-4003 from inception of the cross of LD00-2817P and 17D × S08-14788 #3 in 2010 to 2016 test. 86
Table 3.3 Depiction of population development of RR1 line TN15-5802 from inception of the cross of TN09-029 and USG 74T59 in 2010 to 2016 test. 87
Table 3.4 Single nucleotide polymorphisms (SNPs) used to determine the DNA for genes of interest in the stem termination experiment. 88
Table 3.5 Statistical values for 17 traits of interest for 12 $F_{8:10}$ conventional lines in five experimental groups: i. TN14-4001 indeterminate, ii. TN14-4001 determinate, iii. TN14-4003 indeterminate, iv. TN14-4003 determinate, and v. TN14-4003 segregating, with differences at the Dt_1 loci for stem termination. Experimental group values were pulled from estimate statements run in SAS. LSD values were generated based on a line by line comparison for overall traits of interest. Checks used in this study were Ellis and LD00-2817P. 89

Table 3.6 Statistical values for 17 traits of interest for 12 $F_{7:9}$ RR1 TN15-5802 lines in three experimental groups: indeterminate, determinate, and segregating, with differences at the Dt_1 loci for stem termination. Experimental group values were pulled from estimate statements run in SAS. LSD values were generated based on a line by line comparison for overall traits of interest. Checks used in this study were TN13-5537RR1 and TN13-5538RR1. 90
Table 3.7 Mean concentrations of protein, oil, and amino acids for 10 conventional $F_{8:10}$ near isogenic lines (NIL) in five experimental groups with differences at the Dt_1 loci. Checks used in this study were Ellis and LD00-2817P. 91
Table 3.8 Mean concentrations of protein, oil, and amino acids for 10 RR1 $F_{7:9}$ near isogenic lines (NIL) in three experimental groups with differences at the Dt_1 loci. Checks used in this study were TN13-5537RR1 and TN13-5538RR1. 92
Table 3.9 Mean concentrations of fatty acids for 10 conventional $F_{8:10}$ near isogenic lines (NIL) in five experimental groups with differences at the Dt_1 loci. Check included in this study were Ellis and LD00-2817P. 93
Table 3.10 Mean concentrations of fatty acids for 10 RR1 $F_{7:9}$ near isogenic lines (NIL) in three experimental groups with differences at the Dt_1 loci. Checks used in this study were TN13-5537RR1 and TN13-5538RR1. 94
Table 3.11 Mean agronomic trait values and protein, oil, amino acid, and fatty acid concentrations of 10 conventional $F_{8:10}$ experimental lines and two checks averaged over three replications and three locations. Checks included in this study were Ellis and LD00-2817P. 95
Table 3.12 Mean agronomic trait values and protein, oil, amino acid, and fatty acid concentrations of 10 RR1 $F_{7:9}$ experimental lines and two checks averaged over three replications and three locations. Checks used in this study were TN13-5537RR1 and TN13-5538RR1. 96

LIST OF FIGURES

Figure 1.1 Means of tested agronomic traits for 24 $F_{8:10}$ lines in two experimental groups [wild type (WT) and Danbaekkong (Dan)] across three replications and three locations. The mean of two checks (TN12-4100 and Ellis) was also averaged across three replications and three locations. 37

Figure 2.1 Means of amino acids [g aa kg^{-1} crude protein (cp)] for 24 $F_{8:10}$ lines in two experimental groups [wild-type (WT) and Danbaekkong (Dan) type] across three locations.....60

Figure 3.1 Mean agronomic trait values for 10 conventional $F_{8:10}$ lines in five experimental groups: i. TN14-4001 indeterminate (ind), ii. TN14-4001 determinate (det), iii. TN14-4003 ind, iv. TN14-4003 det, and v. TN14-4003 segregating (seg), with differences at the Dt_1 loci. The TN14-4001 experimental groups were composed of two ind and two det $F_{8:10}$ lines, while the TN14-4003 group was composed of four ind, one det, and one seg $F_{8:10}$ lines. The same letters within a graph denote the groups are not significantly different at $p < 0.05$. 97

Figure 3.2 Mean agronomic trait values for three RR1 experimental groups: indeterminate, determinate, and segregating, from line TN15-5802, with differences at the Dt_1 loci. The TN15-5802 experimental groups were composed of three indeterminate, five determinate, and two segregating $F_{7:9}$ lines. The same letters within a graph denote the groups are not significantly different at $p < 0.05$. 98

INTRODUCTION AND LITERATURE REVIEW

Soybean [*Glycine max* (L.) Merr.] is a domesticated annual plant species native to Eastern Asia that belongs to the *Glycine* genus and *Fabaceae* family (Hymowitz, 2004). Soybean is the world's leading source of vegetable oil and high quality protein meal (Warrington et al., 2015). It was first domesticated in China around 1100 BC before being introduced to countries such as Japan, Philippines, Thailand, and Nepal, where landraces were developed (Aoyagi and Shurtleff, 2004; Hymowitz, 2004). Trade routes on land and sea allowed for the expansion of soybean to other countries (Hymowitz, 2004). For centuries, soybean uses were developed to include items such as tofu and soy sauce, and soybean became a cornerstone of nutrition in Eastern Asia (Hymowitz, 2004).

Soybean was first introduced to North America in 1765, and later became a forage crop in America in the 1920s (Aoyagi and Shurtleff, 2004; Hymowitz, 2004; North Carolina Soybean Producers Association, Inc., 2014). Thanks to breeding efforts and improvements in pest management, irrigation, fertilizer, production, and equipment, soybean became a valuable agricultural crop worldwide. Currently, soybean is grown for its oil and high protein meal (Aoyagi and Shurtleff, 2004). In the United States (U.S.) in 2014, 31% of the crop area planted was occupied by soybeans (Soy Stats, 2015). In 2015, there were 33.4 million hectares of soybean harvested in the U.S. (Miller-Garvin and Naeve, 2015).

Soybean growth has expanded across North America and into South America to countries such as Brazil. Expansion of the crop is necessary because of the increasing high demand for soybeans. Competitive expansion will encourage continued

improvement for better quality protein and oil through breeding and molecular techniques (Aoyagi and Shurtleff, 2004), which will benefit consumers of soy products. Expansion of soybean is also necessary to help meet the growing world demand for protein. However, expansion into vulnerable areas such as rainforests is shortsighted and devastates the rainforest ecosystem.

The soybean is a legume that is able to fix nitrogen due to a symbiotic relationship with *Bradyrhizobia japonicum,* which establishes nodules on the soybean root (Casteel, 2011). The rhizobia live in root nodules on the soybean plant and fix nitrogen for the plant in exchange for sugars. The crop provides benefits to farmers by reducing the need for chemical fertilizer (Araújo et al., 2015). This is one reason why soybeans are utilized in crop rotation (North Carolina Soybean Producers Association, Inc., 2014). The nitrogen fixation performed by the root nodule bacteria plays a large role in creating the soy proteins present in these legumes (Friedman and Brandon, 2001). Soybeans are the world's greatest source of vegetable protein for livestock feed and human diets (Araújo et al., 2015; Friedman and Brandon, 2001).

Protein Quality

Soybean is one of the world's leading sources of vegetable oil and high quality protein meal (Diers et al., 1992; Lin et al., 2011; Miller-Garvin and Naeve, 2015; Warrington et al., 2015; Wilcox and Shibles, 2001). Traditionally, the main determining factor for the value of soybeans has been protein and oil content (Charron et al., 2005). The quality of soy protein can be influenced by a multitude of factors, including but not limited to composition of amino acids, natural toxic factors in the seed, digestibility, or

processing (Edwards 3rd et al., 2000; Lin et al., 2011). Processing the meal improves digestibility, which makes it easier on the stomachs of consumers (Friedman and Brandon, 2001). Prior to 1950, the economic importance of soybean was in its oil (Chung et al., 2003). However commercial interest has shifted toward a major emphasis on the meal protein, which is a by-product of soybean oil extraction (Chung et al., 2003). When soybeans are processed, they are first cleaned then cracked, dehulled, and rolled into flakes causing rupturing of oil cells. The oil and meal of the soybean is separated in this process allowing for extraction of the oil (Soy Stats, 2015). Most of the flakes are used for animal feed and the remaining soybean meal is processed into products containing soy protein (Soy Stats, 2015; Wilcox and Shibles, 2001). A small percentage of the soybean meal is consumed by humans in soy milk, flours, soy protein, tofu, and many other products (Friedman and Brandon, 2001; North Carolina Soybean Producers Association, Inc., 2014).

Protein in soybeans is measured on a dry-weight basis by scientists and on a 13% moisture basis by grain elevators, grain transporters, and processors. The seed proteins in legumes such as soybeans are categorized based on solubility pattern. In soybean seed, the majority of the storage proteins are globulins, further subdivided into 7S vicilin-type (β-conglycinin) and 11S leguminin-type (glycinin) (Friedman and Brandon, 2001; Kim and Wicker, 2005; Panthee et al., 2004; Warrington et al., 2015; Wilcox and Shibles, 2001). These two types of storage proteins have differing gel-forming abilities. The ratios between the two types also have an influence on soy product quality (Kim and Wicker, 2005).

The interaction of genotype and environment has a large effect on the protein quality of soybeans (Lee et al., 2010). Chung et al. (2003), reported that when temperatures are high as soybean seeds develop, there is an elevation in seed oil, and severe drought can result in lower seed protein. It is also worthy of mention that in the United States, northwestern states produce soybean seed with lower protein than southeastern states (Chung et al., 2003), and this regional discrepancy is currently generating major economic concerns. Krishnan et al. notes that, although the average protein content of soybean in the United States is 40%, there is a current trend towards breeding for yield rather than protein quality; this could have a negative effect on the concentration of protein.

Breeding Efforts with Danbaekkong

The Danbaekkong (Dan) protein allele on Chromosome 20 (Gm 20) has been linked to higher protein concentration in soybeans. Inheritance for the Dan allele is dominant, therefore DNA or a progeny row test must be looked at to determine whether a soybean plant is homozygous dominant or heterozygous for the allele. A protocol for single nucleotide polymorphisms (SNP) for seed protein concentration on Gm 20 was created by the Boerma Lab at the University of Georgia in Athens, GA. This protocol has recently been revised by the current leader, Dr. Zenglu Li, who is using the Dan allele for molecular marker assisted selection (MAS) (Warrington, 2015).

Breeding Efforts for High Protein Soybeans

Soybeans epitomize a low-cost high-quality protein source for the world (Aoyagi and Shurtleff, 2004; Joshi et al., 2013; Wilcox, 1998). A greater value of soybean per

hectare can be achieved with an increase of protein concentration by just 1 percent (Beyond the Elevator, 2015). Goals in soybean breeding have been to increase protein concentration within the soybean, while maintaining a high yield and oil concentration, and to develop stable cultivars with good performance throughout different locations (Brim and Burton, 1979; Diers et al., 1992; Lee et al., 2010; Warrington et al., 2015; Warrington et al., 2014, Wilcox, 1998), however this is a major challenge.

There is a negative genetic correlation between protein and yield and between protein and oil concentration that has resulted in many lines not being released (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). There have been successful attempts at increasing protein concentration in soybean cultivars, for example the creation of Provar (Weber and Fehr, 1970) and Prolina (Burton et al., 1999) high protein soybean cultivars, however selection for this trait is hindered by differences in environment (Warrington et al. 2014). For example, severe drought during soybean seed development can result in lower seed protein, while soybean seed development during high temperatures can cause an increase in seed oil (Chung et al., 2003). It is also worthy of mention that protein is affected more by genotypic variation than by the environment (Lee et al., 2010; Shorter et al., 1977).

The balance of amino acids which make up protein is important in determining the nutritional value of the soybean (Friedman and Brandon, 2001; Thakur and Hurburgh, 2007; Warrington et al., 2015; Wilcox and Shibles, 2001). Protein is necessary in both human and animal diets in order to supply essential amino acids to the body (Thakur and

Hurburgh, 2007; Warrington et al., 2015). Soybeans contain eight essential amino acids not produced naturally in the human body (Soy Stats, 2015). Soybeans are deficient in four essential amino acids for chickens: methionine (Met), cysteine (Cys), threonine (Thr) and lysine (Lys) (Friedman and Brandon, 2001; Warrington et al., 2015; Wilcox and Shibles, 2001). Since monogastric animals cannot produce these amino acids naturally and cannot obtain them by consuming soybean meal, these amino acids must be obtained through diets supplemented with synthetic amino acids (Warrington et al, 2015; Wilcox and Shibles, 2001) at a large expense for animal producers. Baker et al. (2011) suggests that there is a greater concentration of amino acids in high protein soybean meal than in conventional soybean meal, and therefore there are more digestible amino acids available for growing swine and chickens fed high protein soybean meal rather than conventional. Higher quality protein could help to solve some of these amino acid deficiency problems and lower excess costs for farmers caused by supplementing animal feed.

Soybean reproduces through self-fertilization. For this reason, the goal in soybean breeding is to bring crops to genetic uniformity before being released as a cultivar or being used in genetic studies. The interest of this experiment lies primarily in one gene for improving seed protein concentration that resides on Gm 20 in the genome of cultivar Danbaekkong (Dan). This research aims to create sets of near isogenic lines (NIL) with differences at that one locus of interest, allowing for the effect of that locus to be tested in multiple environments. We are interested in detecting any enhancing or detrimental effects that the locus may have on agronomic or seed quality traits. The

literature is conflicting regarding this trait: In Illinois the elevated protein allele (Dan) reduces seed yield, whereas in Georgia there is no effect on yield (Diers et al, 2010; Warrington et al., 2015). With Tennessee geographically situated between Illinois and Georgia, we wish to determine the effect of the allele in Tennessee environments before we embark on incorporating the trait to many breeding lines.

Objectives

There were five objectives of this study:

1. To determine if marker assisted selection (MAS) for Danbaekkong protein allele on chromosome 20 influences seed quality.
2. To evaluate the protein concentration and quality of 24 field-tested near isogenic lines (NIL) of soybean grown in three locations across Tennessee. Near-infrared reflectance (NIR) spectroscopy was used to measure total protein and amino acid content.
3. To evaluate agronomic and seed quality traits of 24 field-tested NILs of soybean grown in three locations across Tennessee. Traits included days to maturity, plant height, lodging, and yield.
4. To determine if stem termination has an effect on yield of soybeans grown in Tennessee.
5. To evaluate the effect of stem termination on increased protein concentration, oleic acid, and other seed quality traits.

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

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.

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.

Beyond the Elevator. 2015. More protein can bring more value to soybeans. Our Soy Checkoff. http://beyondtheelevator.com/2015/more-protein-can-bring-more-value-to-soybeans/

Burton, J.W., T.E. JR. Carter, and R.F. Wilson. 1999. Registration of 'Prolina' soybean. Crop Sci. 39: 294.

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.

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.

Diers, B.W., T. Carey, D. Thomas, A. Colgrove, and T. Niblack. 2010. Registration of LD00-2817P soybean germplasm line with resistance to soybean cyst nematode from PI 437654. J. Plant Reg. 4:141-144.

Edwards 3rd, H.M., M.W. Douglas, C.M. Parsons, and D.H. Baker. 2000. Protein and energy evaluation of soybean meals processed from genetically modified high-protein soybeans. Poult. Sci. 79: 525-527.

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.

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, Y and L. Wicker. 2005. Soybean cultivars impact quality and function of soymilk and tofu. J. Sci. Food Agric. 85:2514-2518

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

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

Panthee, D.R., P. Kwanyuen, C.E. Sams, D.R. West, A.M. Saxton, and V.R. Pantalone. 2004. Quantitative trait loci for β-conglycinin (7S) and glycinin (11S) fractions of soybean storage protein. JAOCS. 81: 1005-1012.

Shorter, R., D.E. Byth, and V.E. Mungomery. 1977. Estimates of selection parameters associated with protein and oil content of soybean seeds (*Glycine max* (L.) Merr.). Australian Journal of Agricultrual Research. 28: 211-222.

Soy Stats. 2015. A Reference Guide to Important Soybean Facts and Figures. The American Soybean Association. http://soystats.com/wp-content/uploads/SoyStats-2015.pdf

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.

Weber, C.R. and W.R. Fehr. 1970. Registration of provar soybeans. Crop Sci. 10: 728.

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.

CHAPTER I
EVALUATION OF AGRONOMIC TRAITS OF 24 $F_{8:10}$ NEAR ISOGENIC LINES FOR THE DANBAEKKONG HIGH PROTEIN ALLELE

Abstract

Soybean [*Glycine max* (L.) Merr.] is the world's leading source of vegetable oil and high quality protein meal. Increasing soybean protein concentration through selection while maintaining oil concentration and yield has been a challenging goal for plant breeders, as there is a negative correlation between protein and oil and protein and yield. The objective of this study was to determine if marker assisted selection (MAS) for the Danbaekkong (Dan) protein allele influences agronomic traits. Agronomic traits evaluated include yield, days to maturity, plant height, and lodging. The Dan protein allele on Chromosome 20 (Gm 20) has been linked to higher protein concentration in soybeans. A population of 12 wild type (WT) and 12 mutant Dan $F_{8:10}$ near isogenic lines (NIL) of soybean was created from a cross between G03-3101 and LD00-2817P. These 24 NILs were grown in a replicated three location field trial across Tennessee. Notes on agronomic traits were taken on all plots at all locations. Yield and moisture were recorded for all plots at the time of combine harvest. There were significant differences in yield between the two experimental groups and the check group ($p < 0.05$). The yield ranges for the Dan, WT, and check groups were: 2713-3183 kg ha^{-1}, 3106-3600 kg ha^{-1}, and 3341-3939 kg ha^{-1}, respectively. On average, the checks matured prior to the WT and Dan lines with a range of maturation from 124-126 days after planting (DAP) on average. The range of maturation for the WT experimental group was 132-137 DAP and for the Dan experimental group, 134-136 DAP on average. While there were no significant differences between the heights of WT and Dan experimental groups, there were significant differences observed between the heights of the check group and each experimental group ($p < 0.05$) as the mean height of the check group was 4 cm shorter

than the WT group and 5 cm shorter than the Dan group. There were significant differences between lodging in the two experimental groups and between the check and each experimental group ($p < 0.05$). As expected, the check group displayed the least lodging with a mean score of 1.7, while the WT group had a mean lodging score of 2.1 and the Dan group a 2.3.

Introduction

The agronomic traits such as yield, maturity, height, and lodging, associated with a particular line of soybean are of high importance when introgressing the Danbaekkong (Dan) gene for increased protein. Plant breeders focus their attention on improving the agronomic traits of soybeans to increase soybean value and to satisfy farmer preferences. Without high values for agronomic traits, lines will not be selected to further develop into cultivars. Seed yield is a major factor that influences the decision of farmers to grow a specific cultivar. Yield paired with other agronomic traits such as days to maturity, plant height, and lodging provide an inclusive profile for cultivar selection in soybean. Good values in all these categories provide cultivars likely to be adapted by farmers.

The characteristics of soybeans can be divided into two genetic categories, qualitative and quantitative traits. Qualitative traits are controlled by a single or a few genes and often follow Mendelian inheritance (Bernardo, 2014a; Bernardo, 2014b). Examples of qualitative traits are flower color and pubescence color. These traits are not typically affected much by environmental conditions. Quantitative traits are controlled by many genes with varying levels of additive, dominance, and epistatic effects (Bernardo, 2014a; Bernardo, 2014b). These traits typically have larger non-genetic effects, meaning they are more likely to be influenced by environmental conditions than qualitative traits (Bernardo, 2014a; Bernardo, 2014b). Examples of quantitative traits are seed protein and oil concentrations, which are both economically important (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, stating that in particular, the content of protein in soybean seeds is the main determinant of the crop's value.

There is not much information on the Dan high protein allele, therefore there is a need to identify and test the effects its presence ensues. This experiment analyzes differences between near isogenic lines containing the Gm 20 Dan allelic form versus the wild type allele for the possibility of developing commercialized high protein, high yielding cultivars.

Materials and Methods

Plant Materials

The experimental $F_{8:10}$ lines grown in a replicated multi location Tennessee field trial in the summer of 2016 were created from an $F_{7:8}$ line made from a cross between G03-3101 and LD00-2817P (Table 1.1). Highly homozygous (>99%) near isogenic lines (NILs) were created by harvesting individual F_8 plants that differed in DNA at soybean Gm 20 at the Dan locus. NILs are created several generations after a plant with a gene of interest (e.g. Dan) is crossed with a different standard (wild type) line. The F_1 plant is selfed and produces an F_2 generation that is 50% inbred. This F_2 generation exhibits the most genetic variation. As selections are made and the generations are advanced, the genomes among selected plants become more uniform. In our experiment, the F_8 generation plant was heterozygous for Dan. This allowed NILs to be created and to be tested for the true effect of the elevated protein allele. The NILs created differ at one to a few loci (Bernardo, 2014c). Thus, NILs provide a precise way of testing the effect of one gene on important agronomic and seed quality traits. Choosing to study a set of F_8 derived NILs is important because the genome should be highly genetically uniform (99.22% identical) among lines, with the exception that one line would carry the Dan

allele and another line would carry the wild type allele. An F_8 derived NIL would have very few genetic differences barring this allele.

Line G03-3101 was derived from a cross of Benning × Danbaekkong. G03-3101 is a determinate, unreleased line developed by the University of Georgia. The line G03-3101 belongs to the maturity group VI and has gray pubescence and purple flower color. It was selected as a parent because the Dan allele was in the homozygous state in 2010, when the initial cross was made. LD00-2817P is an F_4 derived line from a cross of Ina × Dwight, developed at the University of Illinois (Diers et al., 2010). LD00-2817P is an indeterminate germplasm line with resistance to soybean cyst nematode (USDA GRIN, 2010). The line has a relative maturity of 4.5, gray pubescence, and purple flower color (USDA GRIN, 2010). One of the purposes of crossing G03-3103 × LD00-2817P was to create a population whose progeny lines exhibited a wide range in maturity for selection among those well adapted to Tennessee.

The relative maturity call of the plants produced from the cross G03-3101 × LD00-2817P in 2014 row 70,184 was 4.9. Therefore, all NILs created from this cross were expected to have a similar relative maturity. For this reason, two maturity group IV-late checks, TN12-4100 and Ellis, were included. These checks were included in the randomization and planted in the field trial. Inclusion of the checks in the field trial allowed for agronomic traits of the experimental NILs to be compared with known lines of high performance.

The population development of G03-3101 × LD00-2817P began in the summer of 2010 when the two parents were crossed at ETREC. In winter 2010-2011, the F_1 seeds

produced from the initial cross were sent to a winter nursery in Isabela, Puerto Rico, where they were grown and harvested as individual F_1 plants. The F_2 seed was planted in summer 2011 at ETREC, one pod from each F_2 plant was harvested, and the seed was sent to a winter nursery in Homestead, Florida in 2011-2012. The F_3 seed was planted and one pod was picked from each F_3 plant. The F_4 seed was grown at ETREC in summer 2012, and approximately 500 single plants were harvested. The plants were threshed, capturing the $F_{4:5}$ seed into coin envelopes, and then were grown at ETREC in the summer of 2013. One single plant was pulled from 2013 row 50,449 because it was identified as heterozygous, and its $F_{5:6}$ seed was increased at a winter nursery in Homestead, Florida, in the winter of 2013-2014. The $F_{5:6}$ plants were bulk harvested and their $F_{5:7}$ seed was grown at ETREC in the summer of 2014. A single F_7 plant with SNP #3242 was harvested and threshed, and its $F_{7:8}$ seed was grown in row MC5 in the summer of 2015.

The $F_{7:8}$ seeds produced from the cross G03-3101 × LD00-2817P were hand planted at ETREC. Plant row MC5 was created using approximately150 seeds from one $F_{7:8}$ derived line of known DAN/WT heterozygous genotype of cross G03-3101 × LD00-2817P and was planted in June 2015, in one 6.09 meter row. Each of the 72 individual plants from row MC5 were tagged with a specific number. One leaf from every plant was collected, and the DNA was tested using a single nucleotide polymorphism (SNP) genotyping assay to determine the genotype of the plant. When running a genetic experiment, it is possible to test whether the genetics are as expected. A Chi-Square test was run on the MC5 population (Table 1.2) based on the SNP genotypes determined by

the Light Cycler. There are 20 homologous pairs of chromosomes in soybeans. At any locus, there will typically be two associated allele types, however in this population all NILs should have the same allele type with the exception of the target Dan allele. Each pair of alleles forms the genotype at the Gm 20 loci. The MC5 population represented random individual plants that are the offspring formed from a DAN/WT plant. In this experiment, the three possible genotype combinations were DAN/DAN (homozygous dominant), DAN/WT (heterozygous), and WT/WT (homozygous recessive). A Chi-Squared test was run to test if the data fit the Mendelian expected genetic ratio of 1:2:1. For this population of 72 plants, the expected ratio was 18 DAN/DAN: 36 DAN/WT: 18 WT/WT, and the observed values were 24 DAN/DAN: 27 DAN/WT: 21 WT/WT. Since there were 72 total genotypes in the MC5 population, this yielded a Chi-Square calculation of 4.75, and the Chi-Squared from the table was 5.99 at 2 degrees of freedom with a p-value of 0.05. The population passed the test, meaning that it is representative of one descending from a heterozygote after one generation, per Chi-Square.

Using the SNP genotype information, 12 homozygous dominant and 12 homozygous recessive F_8 near isogenic single plants from row MC5 were harvested in November 2015. By single plant harvesting within the $F_{7:8}$ plant row, F_8 derived NILs were created. Only WT and mutant Dan single plants were pulled. No heterozygous plants were pulled because our interest in soybean is as a pure-line crop. Of the plants pulled, 12 WT and 12 mutant Dan plants were chosen based on representative pod density, appearance, and lodging. These 24 plants were single plant threshed in November 2015. Following threshing, the single plants' seeds were weighed and

recorded in grams. The seed was then packaged and sent to the winter nursery in Isabela, Puerto Rico (PR). In PR, twenty hills were planted in every row. Two seeds were planted in every hill to ensure germination of at least one plant. Each hill was spaced one foot apart, creating 6.09 meter rows. The seed of the 24 NILs was increased in PR over the winter, bulk harvested, and the seed stock was used to grow 2016 field trials over three environments in Tennessee.

The bulk harvested seeds were planted in Knoxville, Springfield, and Milan, Tennessee in the spring of 2016, to represent East, Middle, and West Tennessee growing environments. Seeds were planted in Knoxville on May 25, 2016, in Springfield on June 8, 2016, and in Milan on May 23, 2016. The 24 lines and two checks were planted in 6.1 m, two row plots with three replications per location. To control for variation within each field, a randomized complete block design was used at all three locations.

Plots were combine harvested on October 14, 2016, October 18, 2016, and October 25, 2016, at Knoxville, Springfield, and Milan, respectively. Weight (lbs) and relative moisture (%) of the plots harvested were recorded by the combine. Weight and moisture in combination with plot length were used in calculating yield in bu/acre on a 13% moisture basis. This data was later converted to kg ha^{-1}.

DNA Extraction and SNP Genotyping

Using DNA markers such as SNP genotyping to find chromosomal locations controlling important plant traits is valuable in plant breeding (Fasoula et al., 2004; Song et al., 2004). In this experiment, 72 single plants from row MC5 were individually DNA tagged with numbers 001-072 in August 2015. After tagging the plants, one leaf tissue

sample was collected from each plant and was rubbed onto pre-labeled Whatman FTA Elute cards (GE Healthcare Life Sciences, Buckinghamshire, England). A small punch measuring 1.20 mm was taken out of each smeared leaf DNA sample and placed into an individual well in a 96-well Roche Light Cycler PCR plate. A template of the plate was made using Excel (Mircrosoft Corporation, Seattle, WA) to ensure each DNA punch corresponded with the correct well. When all 72 samples were punched, they were washed individually with Whatman FTA purification reagent. This was done using a multichannel pipette set to 100 µl. The reagent sat in the wells for five minutes then was removed using the multichannel pipette, this time set to 200 µl and changing to new sterile tips between each row of wells. This was repeated in an identical step. Following this, the plate was washed with Whatman TE-1 buffer, using identical steps. After washing the plate twice with TE-1 buffer, the plate was set to air dry.

After air-drying, the plates were ready for SNP genotyping. The mixture for the Light Cycler PCR was then prepared following a protocol for SNP Assay for Protein on Linkage Group (LG)-I from the University of Georgia in Athens, GA. In a clean 2 ml tube the following were mixed together: 300 µl SNP61899 F2 (forward primer), 150µl SNP61899 R1 (reverse primer), 60µl SNP61899-SP[C] (simple probe), 200µl Genotype Master Mix, 150µl $MgCl_2$, and 640µl water. The mixture was pipetted up and down to mix evenly. Two checks, G03-3101 and LD00-2817P, were added to the bottom right two wells so melting curve data from the 72 samples in the plate could be compared to these two melting curves. A 20 µl sample of the mixture from the 2 ml tube was then pipetted into each of the 96 wells. Following this, all wells were covered with a plastic

film that sealed the plate, and a scraper was used to push down the plastic creating a tight seal over the top of each well. The plate was transferred to a plate centrifuge to spin down all the solution in the plate for 60 seconds. The plate was then moved into the Light Cycler, and a new experiment was started from an existing SNP Assay for Protein on LG-1 template. The pre-incubation was then set at 95˚C for ten minutes. The PCR amplification was set at 95˚C for ten s, 55˚C for 15 s, and 72˚C for 20 s for 35 cycles; and the melting curve was set between 40˚-75˚C. The test was run and the results were analyzed by comparing the melting temperature of each cell on the plate to that of the two checks. Each sample was then categorized as either homozygous Dan, heterozygous, or homozygous WT.

Agronomic Traits

Each plot was walked through during peak flowering, and flower color was recorded when approximately 90% of the plants in a specific plot were in bloom. This differed slightly from the recommended 95%. Both parents G03-3101 and LD00-2817P had purple flowers; therefore, it was expected that all plots would have purple flowers. Any soybean off type with white flowers was pulled during this time to ensure the purity of the lines. A white flowering check, Ellis, was also planted to confirm proper planting order.

Each plot was again walked through at the first sign of senescence. From this point on, each plot was walked every 3-4 days, and maturity notes were taken until every plot was fully mature. A plot was considered fully mature when 90% or more of the pods appeared dried down and attained their mature color. This again differed slightly from

the recommended 95%. During this time days to maturity, lodging, height, and pubescence color were recorded. Days to maturity was calculated as the number of days after planting the plot took to mature. Plots were given a lodging score from 1 to 5, where 1 was completely upright and 5 was completely prostrate or flat on the ground. Height was recorded to the nearest inch by measuring one average looking single plant per plot. Height was later converted to centimeters. All lodging and height scores were taken by the same individual to minimize any variation in data. Days to maturity was recorded by the same individual within the same location to minimize differences in calls. However, different individuals recorded maturity at different locations.

Data Analysis

The experimental design was a randomized complete block model. Data on agronomic 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 24 NILs grown in 2016. For the purpose of analysis, fixed effects were individual lines. To keep track of lines, each was given the prefix MC- followed by consecutive numbers ranging from 1-24. Random effects were replication nested within location, location, and GxE or location by line interaction. Differences between the two experimental NIL groups were tested through use of contrast statements. Estimate statements for check, WT, and Dan genotypes were included in this analysis. Fisher's protected LSD was used to compare experimental groups within each test.

Results and Discussion

Seed yield was averaged over three replications and three locations for all NIL lines within two experimental groups, WT and Dan. Seed yield for a third class (checks) was also averaged. The two checks included in the study were TN12-4100 and Ellis. There were significant differences in yield between the two experimental groups and the check group ($p < 0.05$) (Figure 1.1). The high protein Dan experimental group was numerically the lowest for yield with a mean of 2979.2 kg ha^{-1} and a range from 2713-3183 kg ha^{-1} for the 12 NILs, and the WT experimental group fell between the check and Dan groups with a yield mean of 3260 kg ha^{-1} and a yield range of 3106-3600 kg ha^{-1} for the 12 NILs (Table 1.3). The check group averaged the highest yields with a mean of 3640 kg ha^{-1} and a range from 3341-3939 kg ha^{-1}, as would be expected for high yielding checks (Table 1.3). This result supports previous findings and further confirms the fact that protein and yield are negatively correlated (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).

Maturity of the lines varied throughout all locations. On average, the checks matured prior to the WT and Dan lines. Maturity groups grown in Tennessee typically range from early group IV to group V. Two late maturity group IV checks, TN12-4100 and Ellis, were planted in this field trial. On average, these checks matured 124 and 126 days after planting (DAP), respectively. The range of maturation for the WT experimental group was 132-137 DAP and for the Dan experimental group, 134-136 DAP on average (Table 1.3). When comparing check to WT and check to Dan maturation across locations, there were significant differences between each location (p <

0.05) (Figure 1.1). However, when comparing WT to Dan maturation across locations, there was only a significant difference at the Springfield, TN location ($p < 0.05$). This could be indicative of G × E interaction.

Plant height (cm) was averaged over three replications and locations for all lines within WT and Dan experimental groups. Plant height for an additional check group was averaged as well. While there were no significant differences between the heights of WT and Dan experimental groups, there were significant differences observed between the heights of the check group and each experimental group ($p < 0.05$) (Table 1.4). The Dan experimental group consistently had the tallest average heights with a mean of 82.2 cm and a range from 79-87 cm, and the WT experimental group fell between the check and Dan groups with average heights of 80.5 cm, ranging from 73-85 cm (Table 1.3). The check group averaged the shortest heights of 76.8 cm, ranging from 72-80 cm (Table 1.3).

Lodging and height typically have a positive relationship since taller plants lacking in lodging resistance often have an increased susceptibility to lodging. When averaged across replications and locations, there were significant differences between lodging in the two experimental groups and between the check and each experimental group ($p < 0.05$) (Figure 1.1). The Dan experimental group was the most lodged, followed closely by the WT group. As expected, the check group was the least lodged. One of the checks (Ellis) is a well-established cultivar, notable for its resistance to lodging. The other check (TN12-4100) is also known for its low lodging scores.

Height, lodging, and yield of soybeans are all traits taken into account when selecting cultivars for production. Variations in height often come into play when stem termination varies. Wilcox and Sediyama (1981) reported that lodging of indeterminate, taller plants increased as density of plants grown increased; however, determinate, shorter plants did not lodge under similar conditions. It was also worthy to note that all 73 determinate lines in that study were resistant to lodging, while the 93 indeterminate lines ranged in lodging (Wilcox and Sediyama, 1981). Additionally, severely lodged soybeans often have lower yields than upright soybeans (Wilcox and Sediyama, 1981).

References

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

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

Bernardo, R. 2014c. Modes of reproduction and types of cultivars, p. 71-93. Essentials of plant breeding. Stemma Press, Woodbury, MN.

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.

Diers, B.W., T. Carey, D. Thomas, A. Colgrove, and T. Niblack. 2010. Registration of LD00-2817P soybean germplasm line with resistance to soybean cyst nematode from PI 437654. J. Plant Reg. 4:141-144.

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

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.

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.

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

USDA Germplasm Resources Information Network (GRIN). 2010. PI 658519: Glycine max (L.) Merr. https://npgsweb.ars-grin.gov/gringlobal/accessiondetail.aspx?id=1828074

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

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

Appendix

Table 1.1 Depiction of population development from inception of cross between G03-3101 and LD00-2817P to summer 2016 field trial.

Year	Season	Location	Generation grown	Procedure	Row Designation
2010	Summer	ETREC[a]	Parentals	Cross P1 x P2	Cross 10-19
2010-2011	Winter	TARS[b]	F1	Harvest individual F_1 plants	VP-195 to VP-200
2011	Summer	ETREC	F2	Pick one pod per plant	20133-20148
2011-2012	Winter	27 Farms[c]	F3	Pick one pod per plant	VP12-053 to VP12-060
2012	Summer	ETREC	F4	Harvest ~500 single plants	40146-40161
2013	Summer	ETREC	F4:5	Single plants pulled from row 50449	50432-50493
2013-2014	Winter	27 Farms	F5:6	Bulk harvest individual row	VH14-3274
2014	Summer	ETREC	F5:7	Single plant harvest DNA tag #3242 F_7 derived HET plant	70184
2015	Summer	ETREC	F7:8	DNA tag; single plant harvest homozygous dominant and recessive	MC5
2015-2016	Winter	TARS	F8:9	Bulk harvest each row	MC1-MC24
2016	Summer	ETREC, HRREC[d], RECM[e]	F8:10	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 Chi-Square test results on 72 plants with genotypes: Danbaekkong (Dan), heterozygous (Dan/WT), and wild type (WT) in the MC5 population. The Chi-Squared table value of 5.99 at 2 degrees of freedom with a p-value of 0.05 exceeds the calculated value of 4.75; thus, this population of $F_{7:8}$ plants shows expected 1:2:1 Mendelian inheritance at the chromosome 20 Danbaekkong locus.

Genotype	Observed	Expected	(O-E)2/E
Dan	24	18	2.0
Dan/WT	27	36	2.25
WT	21	18	0.50
			4.75

Table 1.3 Mean yield, maturity, lodging, and height for 24 $F_{8:10}$ lines in two experimental groups [wild type (WT) and Danbaekkong (Dan)] with differences at the Danbaekkong loci. A third experimental group consisting of two check lines (TN12-4100 and Ellis) was included in the study for reference.

				Experimental Groups										
Trait	**GO3-3101[a]**	**LD00-2817P[b]**	**P value**		**Check**			**WT**			**Dan**		**LSD value**	**CV**
				Min	**Mean**	**Max**	**Min**	**Mean**	**Max**	**Min**	**Mean**	**Max**		
Yield (Kg Ha^{-1})	N/A[c]	N/A	0.0037	3341.5	3640.4	3939.3	3105.9	3260.6	3600.1	2713.9	2979.2	3183.9	230.5	15.4
Maturity (Days)	N/A	N/A	<.0001	124.1	124.8	125.1	132.2	136.1	137.6	134.8	135.7	136.7	1.6	2.2
Lodging (1-5scale)	1	1	0.0255	1.6	1.6	1.6	1.6	2.1	2.4	2.1	2.3	2.6	0.3	26.2
Height (cm)	81.28	66.04	0.0014	72.5	76.7	80.9	73.6	80.4	85.7	79.0	82.1	87.1	2.9	7.0

[a]GO3-3010 is a parent and was not included in yield trial. It was grown adjacent to the trial at the ETREC location only.
[b]LD00-2817P is a parent and was not included in yield trial. It was grown adjacent to the trial at the ETREC location only.
[c]N/A, not available. The parental lines were not included in this study.

Table 1.4 Means of agronomic traits for 24 $F_{8:10}$ lines in two experimental groups with differences at the Danbaekkong loci

	Agronomic traits means			
Experimental Group	Yield	Maturity	Lodging	Height
	kg ha^{-1}	Days	1-5 scale	cm
Check	3640.5	124.8	1.66670	76.8
WT	3260.7	136.1	2.12960	80.5
Dan	2979.2	135.8	2.34260	82.2
LSD (0.05)	230.5	1.6	0.26186	2.9

†Checks included in field trial were: TN12-4100 and Ellis
‡WT: Wild type allele (Gm 20)
§Dan: Danbaekkong type allele (Gm 20)

Table 1.5 Mean yield, maturity, lodging, and height of 24 $F_{8:10}$ lines and two checks averaged over three replications and three locations

Line	Class	Yield	Maturity	Lodging	Height
		kg ha^{-1}	DAP	1-5 scale	cm
TN12-4100	Check†	3341.6	123.8	1.7	81.0
Ellis	Check	3939.4	125.9	1.7	72.5
MC-1	WT‡	3105.9	137.3	2.1	84.1
MC-2	WT	3148.1	137.1	2.2	85.8
MC-3	WT	3148.8	136.8	2.3	83.0
MC-4	WT	3149.5	137.8	2.2	83.5
MC-5	WT	3241.5	136.9	2.2	80.4
MC-6	WT	3196.6	137.7	2.2	82.7
MC-7	WT	3461.1	132.2	2.2	77.0
MC-8	WT	3600.1	133.2	2.0	73.7
MC-9	WT	3409.6	133.4	1.7	76.5
MC-10	WT	3337.1	136.6	2.2	78.7
MC-11	WT	3223.5	137.4	2.4	81.8
MC-12	WT	3106.2	137.1	1.7	78.2
MC-13	Dan§	3012.8	135.7	2.3	83.3
MC-14	Dan	2985.2	136.0	2.7	82.4
MC-15	Dan	2795.4	135.4	2.6	81.0
MC-16	Dan	3183.9	135.2	2.2	82.7
MC-17	Dan	2929.1	136.8	2.3	82.1
MC-18	Dan	2962.7	135.8	2.1	81.3
MC-19	Dan	3177.2	135.3	2.4	80.4
MC-20	Dan	2810.3	134.9	2.2	79.0
MC-21	Dan	2993.4	135.2	2.2	80.2
MC-22	Dan	2713.9	136.0	2.3	84.1
MC-23	Dan	3100.2	136.1	2.2	82.4
MC-24	Dan	3086.8	136.6	2.4	87.2

†Checks included in field trial were: TN12-4100 and Ellis
‡WT: Wild type allele (Gm 20)
§Dan: Danbaekkong type allele (Gm 20)

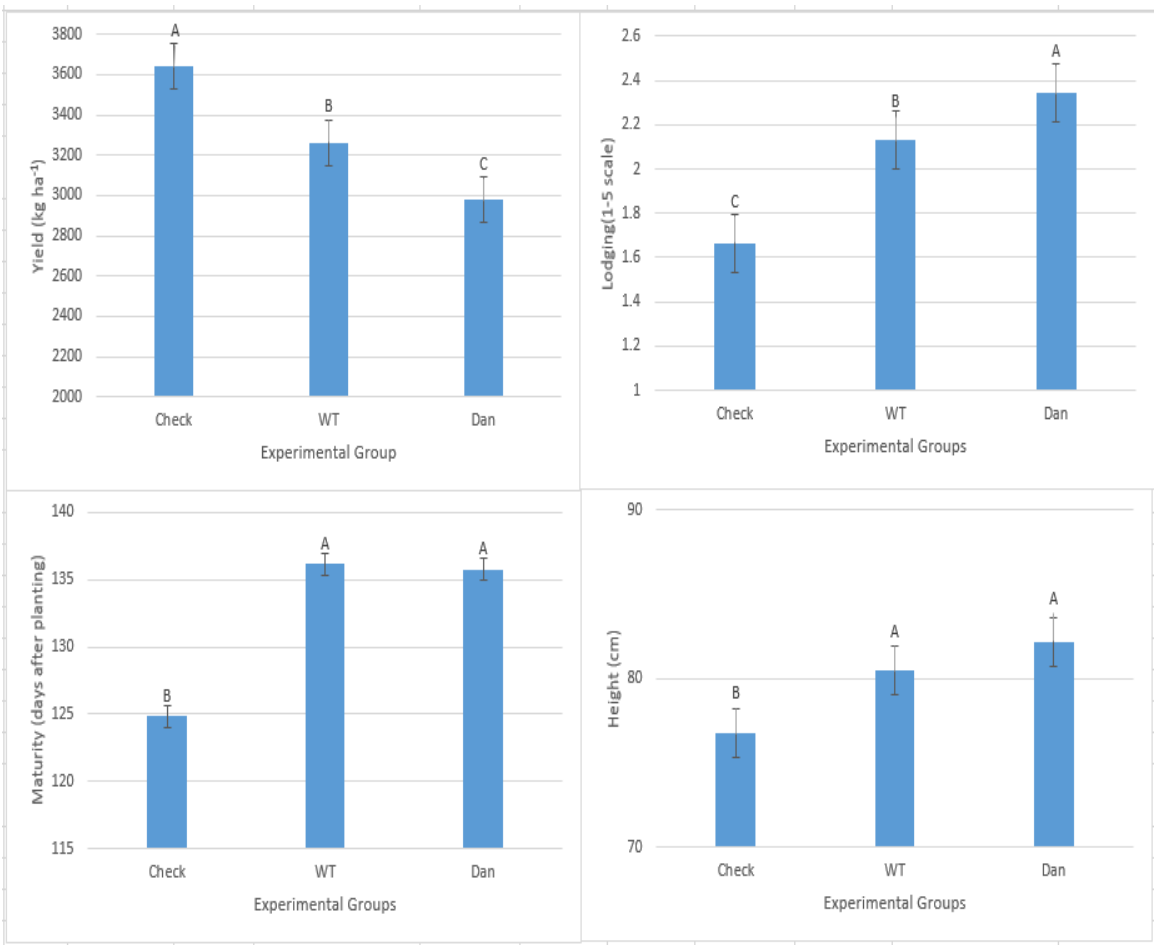

Figure 1.1 Means of tested agronomic traits for 24 $F_{8:10}$ lines in two experimental groups [wild type (WT) and Danbaekkong (Dan)] across three replications and three locations. The mean of two checks (TN12-4100 and Ellis) was also averaged across three replications and three locations.

CHAPTER II
EVALUATION OF SEED PROTEIN AND OIL OF 24 $F_{8:10}$ FIELD TESTED NEAR-ISOGENIC SOYBEAN LINES

Abstract

A population of 24 $F_{8:10}$ near isogenic lines (NIL) of soybean was created from a cross between G03-3101 and LD00-2817P. Of these 24 NILs, 12 were wild type (WT) and 12 were mutant Danbaekkong (Dan). These NILs were grown in a replicated three location field trial across Tennessee. The objectives of this study were to evaluate the seed protein, oil, and amino acid content of the NILs using near-infrared reflectance (NIR) spectroscopy and to analyze fatty acids using gas chromatography (GC). In total, five amino acids: cysteine, methionine, tryptophan, lysine, and threonine and five fatty acids: palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1), linoleic acid (18:2), and linolenic acid (18:3) were analyzed. Protein, oil, and all fatty acids are reported in g kg^{-1}. Amino acids are reported as g amino acid (aa) kg^{-1} crude protein (cp). There were significant differences detected between the means of WT and Dan experimental groups for all tested traits with the exception of stearic acid ($p < 0.05$). Significant differences between Dan and WT experimental groups were observed for all amino acids tested ($p < 0.05$). Dan means were consistently lower than both check and WT means for the five tested amino acids. The Dan experimental group had significantly more protein (414 g kg^{-1}) and significantly less oil (206.9 g kg^{-1}) than WT and check groups ($p < 0.05$). Therefore, the negative correlation between protein and oil in soybean is further confirmed by this result.

Introduction

The Danbaekkong (Dan) gene for high protein may have an impact on amino acids and total oil accrual as well as having an effect on total protein accumulation. The Dan protein allele is located on Chromosome 20 (Gm20) and has been linked to higher protein concentration in soybeans. Protein in soybeans comes from the meal, which is a by-product of soybean oil extraction (Chung et al., 2003; Wilcox, 1998; Wilson, 2004). There are two types of soybean meal available, conventional (43-44% protein) and dehulled high protein (47-49% protein) soybean meal (Cromwell, 2012). In conventional soybean meal, the hulls are blended back into the meal to standardize crude protein (cp) to approximately 44%, and in dehulled soybean meal the hulls are left out, resulting in higher protein content (Cromwell, 2012). When soybeans are processed, they are rolled into flakes causing rupturing of oil cells. The oil and protein meal is separated, allowing for extraction of the oil (Soy Stats, 2015). Most of the flakes are used in animal feed, and the remaining meal is processed into products containing soy protein (Soy Stats, 2015). A small percentage of the soybean meal is consumed by humans in flours, soy protein, soy milk, tofu, and many other products (Friedman and Brandon, 2001; North Carolina Soybean Producers Association, Inc., 2014).

Prior to 1950, soybean oil was of the utmost importance (Chung et al., 2003). Commercial interest has since shifted toward an emphasis on protein meal (Chung et al., 2003). Traditionally, the main determining factors for the value of soybeans have been protein and oil content (Diers et al., 1992; Charron et al., 2005; Wilson, 2004). In 2015, U.S. commodity soybean cultivars had a mean protein value of 34.3% and a mean oil value of 19.7% on a 13% moisture basis (Miller-Garvin and Naeve, 2015). This shows

an increase in oil concentration from previous years but does not allow farmers to produce seed with adequate protein concentration to make a 48% meal while simultaneously having high soybean oil yields, which is a common difficulty (Miller-Garvin and Naeve, 2015; Wilson, 2004). Moreover, the long term trend for higher seed yield at the expense of lower protein is becoming a cause for increased concern, with fewer and fewer cultivars able to meet demands for 48% meal protein. An increase in the protein concentration of soybean by as little as 1% can achieve greater estimated process value for the crop of anywhere between $7.70 and $12.96 per acre (Beyond the Elevator, 2015). A major challenge with increasing protein concentration is maintaining high yield and oil concentration (Lee et al., 2010; Warrington et al., 2015; Warrington et al., 2014; Wilson, 2004). This is because there is a negative genetic correlation between protein and yield and between protein and oil concentration (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). For example, Brim and Burton (1979), reported a linear increase in mean seed protein with successive cycles of selection, while there was a linear decrease in the seed oil concentration. Wilcox (1998), reported an increase in mean seed protein concentration from 438 to 484 g kg^{-1} during eight cycles of recurrent selection, and a decrease in seed oil of 2.3 g kg^{-1} per cycle during this same time.

Genotype by environment (G x E) interaction has a large effect on the protein quality of soybeans (Lee et al., 2010). Although protein is affected more by genotypic variation than by the environment (Lee et al., 2010; Shorter et al., 1977), selection for

high protein is hindered by differences in environment (Warrington et al. 2014). Chung et al. (2003) reported that with high temperatures as soybean seeds develop, there is notable elevation in seed oil and reduction in protein. Severe drought can result in lower seed protein as well. In the United States, northwestern states produce soybean seed with lower protein than southeastern states (Chung et al., 2003). Additionally, Yaklich et al. (2002) reports that soybean mean protein and oil concentrations were higher in the Southern Region uniform tests than in the Northern Region uniform tests.

Amino acids are the building blocks which form proteins (Cromwell, 2012). The balance of amino acids in protein is important in determining the nutritional value of the soybean (Cromwell, 2012; Friedman and Brandon, 2001; Warrington et al., 2015). Essential amino acids are supplied to the human body through consumption of animal and plant based protein (Miller-Garvin and Naeve, 2015; Thakur and Hurburgh, 2007; Warrington et al., 2015; Wilcox and Shibles, 2001). Soybean amino acid balance provides nearly all essential amino acid nutritional requirements of both swine and poultry (Wilson, 2004). Animals get the amino acids needed through feed proteins such as soybean meal, which is highly digestible (Cromwell, 2012; Miller-Garvin and Naeve, 2015; Wilcox and Shibles, 2001). Soybean deficiency in essential amino acids for chickens and swine can result in increased costs for farmers (Friedman and Brandon, 2001; Warrington et al., 2015). Since monogastric animals cannot produce the amino acids methionine, cysteine, threonine, and lysine naturally or obtain them through consumption of soybean meal, these amino acids must be obtained through diets supplemented with synthetic amino acids (Warrington et al, 2015; Wilcox and Shibles,

2001). Animal producers incur large expenses resulting from supplementation of amino acids in the food of their animals. Baker et al. (2011) suggests that there is a greater concentration of amino acids in high protein soybean meal than in conventional soybean meal. Therefore, more digestible amino acids are available for growing swine and chickens fed high protein soybean meal rather than conventional. Higher quality protein could help to solve some of these amino acid deficiency problems and lower excess supplementation costs for poultry and swine producers.

Materials and Methods

Plant Materials

The experimental population of 24 $F_{8:10}$ lines grown in a replicated multi location Tennessee field trial in the summer of 2016 were created from an $F_{7:8}$ line made from a cross between G03-3101 and LD00-2817P (Table 1.1). A set of 24 highly homozygous (>99%) near isogenic lines (NILs) were created by harvesting individual F_8 plants that differed in DNA at Gm 20 Dan locus. The NILs consisted of 12 that were wild type (WT) and 12 that were Danbaekkong (Dan) type. The NILs were denoted MC-1 to MC-24, where MC-1 to MC-12 were WT and MC-13 to MC-24 were Dan type. NILs are created several generations after a plant with a gene of interest (e.g. Dan) is crossed with a different standard (wild type) line. The F_1 plant is selfed and produces an F_2 generation that is 50% inbred. As selections are made and the generations are advanced, the genomes among selected plants become more uniform. In our experiment, the F_8 generation plant was heterozygous for Dan. This allowed NILs to be created and to be tested for the true effect of the elevated protein allele. The NILs created differ at one to a

few loci (Bernardo, 2014). Thus, NILs provide a precise way of testing the effect of one gene on important agronomic and seed quality traits. Choosing to study a set of F_8 derived NILs is important because the genome should be highly genetically uniform (99.22% identical) among lines, with the exception that one line would carry the Dan allele and another line would carry the WT allele. Any other genetic differences that an F_8 derived NIL would have would be very few.

The relative maturity of plants produced from the cross G03-3101 × LD00-2817P in 2014 row 70,184 was 4.9. Therefore, all NILs created from this cross would be expected to have a similar relative maturity. For this reason, two maturity group IV-late checks, TN12-4100 and Ellis, were included in the randomization and planted in the field trial. This allowed for agronomic traits of experimental NILs to be compared with known lines of high performance.

Near Infrared Reflectance

Total protein, oil, and amino acid concentrations were determined by near-infrared reflectance spectroscopy (NIR) using a Perten DA 7250 analyzer. University of Minnesota, in cooperation with Perten, developed calibration equations to analyze soybean seed quality traits. The equation was developed on the HPLC measurements made in the University of Missouri analytical laboratory (Warrington et al., 2015). The machine was calibrated using 900 North American soybean established cultivars and breeding lines (Warrington et al., 2015). A sample of approximately 20g of seed was taken from each replication in every location for each of the 24 lines grown in a 2016 replicated field trial in three locations across Tennessee. Thus, there were 24 lines × 3

replications × 3 locations for a total of 216 samples run on the NIR. The seed samples were ground using a Foss water-cooled Knifetec 1095 Sample Mill (Tecator, Hoganas, Sweden) to create uniform particle size. The values for protein, oil, and amino acid concentrations were determined based on 13% seed moisture and were converted to g kg^{-1} dry weight. The amino acid values were further converted by dividing by the protein concentration to report as g amino acid kg^{-1} of cp. Analysis of variance was conducted using SAS version 9.4 (SAS Institute Inc., Cary, NC, USA). Entries were considered fixed effects and replications and environments as random effects. Pre-planned single degree of freedom contrasts were utilized to statistically test the effect of the Dan vs. WT allele for all seed quality traits. Fisher's protected LSD was used to compare Dan, WT, and checks.

Gas Chromatography

Components of the soybean oil were analyzed using gas chromatography (GC). A soybean oil profile consists of varying combinations of five predominant fatty acids: palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1), linoleic acid (18:2), and linolenic acid (18:3) (Min et al., 2015). A five seed sample was taken from each replication in every location for each of the 24 lines grown in a 2016 replicated field trial in three locations across Tennessee. Thus, there were 24 lines × 3 replications × 3 locations for a total of 216 samples run on the GC as well. Each five seed sample was crushed with a hammer to expose sufficient surface area from which to extract soybean oil. The crushed seeds were transferred into individual test tubes with a 3 mL volume of extraction solvent consisting of 2,000 mL chloroform, 1250 mL hexanes, and 500 mL

methanol. The test tubes were then capped with plastic stoppers and sat overnight. The next day, a 100 µL sample of the oil extract was transferred to a 1.5 mL vial. A 0.75 mL volume of hexanes and a 75 µL volume of methylation reagent were then added to each vial. The methylation reagent consisted of 5 mL 0.5M sodium methoxide solution in methanol, 10 mL ethyl ether, and 20 mL petroleum ether. Caps were then crimped to fasten to each vial. The fatty acid methyl esters were analyzed using a Hewlett Packard HP 6890 series gas chromatograph (Agilent Technologies, Santa Clara, CA) equipped with a flame ionization detector, 30 m x 0.53 mm 125-2332 capillary column, and a 7683 series auto sampler.

Results and Discussion

Near Infrared Reflectance

There is a recent trend towards improving the quality of protein and oil in the grain market (Hurburgh et al., 1990). Wilson (2004) reported that among accessions of the USDA Soybean Germplasm Collection, the genetic variation in protein and oil varied greatly, stating that protein concentrations ranged from 34-56% of seed dry weight, and oil concentrations ranged from 8-27%, with means of 42.1% protein and 19.5% oil. A negative correlation between protein and yield makes commercial production of above average protein soybeans a difficult sell. Wilson (2004) explained that soybeans with above average protein may not be profitable to grow due to an economic plateau which occurs around 48% protein in soymeal.

NIR was used to evaluate the effect of the Dan gene on soybean seed protein and oil concentrations. Protein and oil are reported in g kg^{-1}. Mean concentrations of protein,

oil, and amino acids for 24 $F_{8:10}$ NILs with differences at the Danbaekkong loci are found in Table 2.1. There were significant differences between the means of WT and Dan experimental groups for all traits tested on the NIR. As predicted, the Dan experimental group (414 g kg^{-1}) had significantly more protein than WT and check groups ($p < 0.05$). However, the negative genetic correlation between protein and oil was further confirmed as the Dan experimental group (206.9 g kg^{-1}) had significantly less oil than the WT and check groups (227.3 and 231.3 g kg^{-1}, respectively) ($p < 0.05$). The range of mean protein and oil for experimental groups across locations was 378-418 g kg^{-1} and 205-232 g kg^{-1}, respectively.

While higher protein and lower oil concentrations in the Dan group were consistent across locations, plants grown in Springfield had the highest protein means and lowest oil means for all three experimental groups (Table 2.2). When looking at individual locations, all Dan lines (MC-13 – MC-24) at the Springfield location had mean protein concentration values above and mean oil concentration values below the total mean concentration values for three locations (Table 2.3). Along with a negative genetic correlation for protein and oil, environmental factors could have come into play here, as different locations showed different ranges for protein and oil. Due to the negative genetic correlation realized between protein and oil concentrations, studies focusing on increasing the quality of either one trait or both, are important.

The G x E interaction has an effect on protein and oil concentrations. Hurburgh and Brumm (1990) found yearly variation in protein and oil concentrations within the same location in their three year study. They observed standard deviations of 1.0% for

protein and 0.5% for oil of soybean deliveries to grain elevators. Over the course of their study, they found approximately 15% of the samples to be above and 15% to be below average in both constituents. The remainder were above in one constituent and below in the other, following typical trends of negatively correlated traits.

NIR was used to evaluate the effect of the Dan gene on soybean seed amino acid concentrations as well. Amino acid values were divided by the protein concentrations to report as g amino acid kg^{-1} cp. Amino acids of interest were: cysteine (Cys), methionine (Met), tryptophan (Try), lysine (Lys), and threonine (Thr), due to their importance for poultry and swine nutrition. Significant differences between Dan and WT experimental groups were observed for all amino acids tested ($p < 0.05$) (Figure 2.1). Dan means were consistently significantly lower than both check and WT means for each of the individual amino acids (Table 2.1). The balance and constitution of amino acids that comprise soybean protein are what mark a quality protein. The decrease in amino acids, although slight, in the Dan experimental group leads to a decrease in protein quality compared to the WT and check groups. This is an important consideration before breeders embark on introgressing the Dan gene to elite germplasm.

Gas Chromatography

The effect of the Dan gene on soybean seed oil and fatty acid accumulation was evaluated using GC results. Fatty acids are reported in g kg^{-1}. The GC measured sample concentrations of palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2), and linolenic (18:3) acids. Mean concentrations of fatty acids for $F_{8:10}$ NILs with differences at the Danbaekkong locus are found in Table 2.4. Significant differences were detected

between WT and Dan experimental groups in all fatty acids except 18:0 ($p < 0.05$). For 18:3, significant differences were detected between all experimental groups ($p < 0.05$). Commercial soybean oil is approximately: 100 g kg^{-1} 16:0, 40 g kg^{-1} 18:0, 220 g kg^{-1} 18:1, 540 g kg^{-1} 18:2, and 100 g kg^{-1} 18:3, on average (Wilson, 2004). Mean values for fatty acids averaged over all three locations for the WT group were: 110 g kg^{-1} 16:0, 40 g kg^{-1} 18:0, 200 g kg^{-1} 18:1, 570 g kg^{-1}18:2, and 80 g kg^{-1} 18:3. Mean values averaged across locations for the Dan group were: 120 g kg^{-1} 16:0, 40 g kg^{-1} 18:0, 190 g kg^{-1} 18:1, 570 g kg^{-1} 18:2, and 80 g kg^{-1} 18:3. The small variation in concentrations of fatty acids between the WT and Dan experimental groups shows that the Dan allele does not substantially negatively impact any fatty acid accumulation.

References

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. 2014. Modes of reproduction and types of cultivars, p. 71-93. Essentials of plant breeding. Stemma Press, Woodbury, MN.

Beyond the Elevator. 2015. More Protein Can Bring More Value to Soybeans. Our Soy Checkoff. http://beyondtheelevator.com/2015/more-protein-can-bring-more-value-to-soybeans/.

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.

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.

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.

Hurburgh, C.R. Jr., and T.J. Brumm. 1990. Protein and oil content of soybeans received at county elevators. Appl. Eng. Agr. 6: 65-68.

Hurburgh, C.R. Jr., T.J. Brumm, J.M. Guinn, and R.A. Hartwig. 1990. Protein and oil patterns in U.S. and world soybean markets. J. Am. Oil Chem. Soc. 67: 966-973.

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.

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

Min, C.W., R. Gupta, S.W. Kim, S.E. Lee, Y.C. Kim, D.W. Bae, W.Y. Han, B.W. Lee, J.M. Ko, G.K. Agrawal, R. Rakwal, and S.T. Kim. 2015. Comparative biochemical and proteomic analyses of soybean seed cultivars differing in protein and oil content. J. Agric. Food Chem. 63: 7134-7142.

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

Shorter, R., D.E. Byth, and V.E. Mungomery. 1977. Estimates of selection parameters associated with protein and oil content of soybean seeds (*Glycine max* (L.) Merr.). Aust. J. Agric. Res. 28: 211-222.

Soy Stats. 2015. A Reference Guide to Important Soybean Facts and Figures. The American Soybean Association. http://soystats.com/wp-content/uploads/SoyStats-2015.pdf

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.

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., B. Vinyard, M. Camp, and S. Douglass. 2002. Analysis of seed protein and oil from soybean northern and southern region uniform tests. Crop Sci. 42:1504-1515.

Yaklich, R.W. 2001. β-Conglycinin and glycinin in high-protein soybean seeds. J. Agric. Food Chem. 49:729-73.

Appendix

Table 2.1 Mean concentrations of protein, oil, and amino acids for 24 $F_{8:10}$ lines in two experimental groups with differences at the Danbaekkong loci.

	Protein, oil, and amino acid concentrations						
Experimental Group	Protein	Oil	Cysteine	Methionine	Tryptophan	Lysine	Threonine
	g kg^{-1}		g aa kg^{-1} cp				
Check†	389.0	231.3	15.2	14.1	11.4	65.1	37.1
WT‡	382.1	227.3	14.7	13.9	11.5	65.1	37.5
Dan§	414.4	206.9	14.4	13.6	10.9	64.0	36.9
LSD (0.05)	3.4	1.9	0.3	0.1	0.2	0.3	0.2
	1.7064	0.951	0.13093	0.071265	0.079025	0.13855	0.112325

† Checks included in field trial were: TN12-4100 and Ellis
‡ WT: Wild type allele (Gm 20)
§ Dan: Danbaekkong type allele (Gm 20)

Table 2.2 Means of protein and oil (g kg^{-1}) for 24 $F_{8:10}$ lines in two experimental groups within a location.

	Knoxville, TN		Springfield, TN		Milan, TN	
Experimental Groups	Protein	Oil	Protein	Oil	Protein	Oil
	g kg^{-1}		g kg^{-1}		g kg^{-1}	
Check†	371.1	244.1	400.4	220.5	396.3	228.8
WT‡	366.8	237.1	402.1	213.9	377.5	230.8
Dan§	397.4	216.3	432.1	195.2	413.8	209.3
LSD (0.05)	5.6	4.0	4.0	2.8	7.3	3.2

† Checks included in field trial were: TN12-4100 and Ellis
‡ WT: Wild type allele (Gm 20)
§ Dan: Danbaekkong type allele (Gm 20)

Table 2.3 Means of protein and oil (g kg^{-1}) for 12 $F_{8:10}$ Wild Type (WT) lines and 12 $F_{8:10}$ Danbaekkong (Dan) type lines within a location.

		Knoxville, TN		Springfield, TN		Milan, TN	
	Type	Protein	Oil	Protein	Oil	Protein	Oil
		g kg^{-1}		g kg^{-1}		g kg^{-1}	
MC-1	WT	362.0	237.3	430.4	197.2	375.7	231.8
MC-2	WT	361.6	237.7	431.4	194.2	375.8	228.8
MC-3	WT	375.2	231.9	432.7	193.1	387.0	225.0
MC-4	WT	366.0	236.0	431.7	194.8	373.6	231.2
MC-5	WT	367.1	238.2	426.8	198.4	371.1	232.3
MC-6	WT	376.1	232.9	429.1	195.6	388.9	225.4
MC-7	WT	361.9	242.3	430.8	196.6	382.0	227.6
MC-8	WT	368.0	236.3	435.8	195.1	377.7	232.5
MC-9	WT	364.4	241.5	439.0	192.0	373.3	235.9
MC-10	WT	363.2	239.3	429.9	196.6	379.0	233.4
MC-11	WT	367.6	236.2	434.1	192.3	372.1	233.3
MC-12	WT	368.5	235.4	433.2	196.5	374.0	232.7
MC-13	Dan	412.3	208.4	430.4	197.2	430.4	197.2
MC-14	Dan	413.4	206.7	431.4	194.2	431.4	194.2
MC-15	Dan	415.3	205.5	432.7	193.1	432.7	193.1
MC-16	Dan	413.9	207.5	431.7	194.8	431.7	194.8
MC-17	Dan	409.5	206.9	426.8	198.4	426.8	198.4
MC-18	Dan	414.5	206.1	429.1	195.6	429.1	195.6
MC-19	Dan	418.4	206.6	430.8	196.6	430.8	196.6
MC-20	Dan	418.0	206.4	435.8	195.1	435.8	195.1
MC-21	Dan	414.1	207.6	439.0	192.0	439.0	192.0
MC-22	Dan	413.0	208.0	429.9	196.6	429.9	196.6
MC-23	Dan	415.8	206.5	434.1	192.3	434.1	192.3
MC-24	Dan	414.7	206.8	433.2	196.5	433.2	196.5
LSD (0.05)		5.6	4.0	4.0	2.8	7.3	3.2

Table 2.4 Mean concentrations of fatty acids for 24 $F_{8:10}$ lines in two experimental groups with differences at the Danbaekkong loci.

	Fatty acid concentrations				
Experimental Group	Palmitic	Stearic	Oleic	Linoleic	Linolenic
	g kg^{-1}				
Check†	115.8	40.6	199.4	569.7	74.4
WT‡	114.4	39.5	196.3	568.7	81.1
Dan§	116.4	39.5	185.9	573.7	84.7
LSD (0.05)	1.3	1.2	4.8	4.1	2.0
	0.647815	0.595155	2.41123	2.03922	1.00103

†Checks included in field trial were: TN12-4100 and Ellis
‡WT: Wild type allele (Gm 20)
§Dan: Danbaekkong type allele (Gm 20)

Table 2.5 Mean protein, oil, and amino acid concentrations of 24 $F_{8:10}$ lines and two checks averaged over three replications and three locations.

Line	Class	Crude Protein	Oil	Cys	Met	Trp	Lys	Thr
		g kg^{-1}		g aa kg^{-1} cp				
TN12-4100	Check†	384.0	232.8	15.1	14.0	11.4	65.2	37.3
Ellis	Check	393.9	229.8	15.2	14.1	11.4	64.9	37.0
MC-1	WT‡	379.9	227.4	14.5	13.9	11.6	65.2	37.4
MC-2	WT	379.9	225.8	14.5	13.8	11.4	65.0	37.3
MC-3	WT	390.8	221.8	14.6	13.7	11.2	65.0	37.3
MC-4	WT	381.6	226.6	14.8	13.9	11.4	64.9	37.5
MC-5	WT	379.9	228.4	14.7	14.0	11.6	65.2	37.5
MC-6	WT	391.6	221.7	14.7	13.9	11.5	64.9	37.4
MC-7	WT	380.9	229.7	14.8	14.0	11.4	65.2	37.7
MC-8	WT	380.9	228.7	14.7	13.9	11.4	65.4	37.5
MC-9	WT	379.6	231.7	14.9	14.1	11.7	65.2	37.4
MC-10	WT	379.6	229.0	14.8	14.1	11.5	65.3	37.6
MC-11	WT	378.5	228.9	14.6	13.9	11.6	65.3	37.6
MC-12	WT	382.6	227.3	14.6	13.7	11.5	64.7	37.3
MC-13	Dan§	412.3	208.4	14.4	13.6	11.0	64.0	37.0
MC-14	Dan	413.4	206.7	14.6	13.6	11.0	64.0	37.0
MC-15	Dan	415.3	205.5	14.5	13.5	11.0	63.8	36.8
MC-16	Dan	413.9	207.5	14.1	13.4	11.0	63.9	36.8
MC-17	Dan	409.5	206.9	14.6	13.6	11.1	64.2	37.0
MC-18	Dan	414.5	206.1	14.4	13.6	10.8	64.1	36.9
MC-19	Dan	418.4	206.6	14.3	13.4	10.9	63.8	36.8
MC-20	Dan	418.0	206.4	14.5	13.6	10.8	63.9	37.0
MC-21	Dan	414.1	207.6	14.2	13.5	10.9	63.8	36.9
MC-22	Dan	413.0	208.0	14.6	13.6	11.0	64.2	36.9
MC-23	Dan	415.8	206.5	14.4	13.6	11.0	64.0	36.9
MC-24	Dan	414.7	206.8	14.2	13.6	10.9	64.0	37.0

†Checks included in field trial were: TN12-4100 and Ellis
‡WT: Wild type allele (Gm 20)
§Dan: Danbaekkong type allele (Gm 20)

Table 2.6 Mean fatty acid concentrations of 24 $F_{8:10}$ lines and two checks averaged over three replications and three locations.

Line	Class	Palmitic	Stearic	Oleic	Linoleic	Linolenic
		g kg^{-1}				
TN12-4100	Check†	118.6	39.3	198.4	567.5	76.3
Ellis	Check	112.9	41.9	200.5	572.0	72.5
MC-1	WT‡	114.1	39.8	196.9	567.1	82.1
MC-2	WT	114.4	39.1	196.0	567.8	82.7
MC-3	WT	115.3	39.5	192.9	570.4	81.8
MC-4	WT	115.6	40.0	203.3	560.8	80.3
MC-5	WT	114.0	40.0	201.1	564.1	80.7
MC-6	WT	116.6	39.1	191.1	569.3	83.8
MC-7	WT	113.2	38.9	195.6	574.3	77.9
MC-8	WT	112.5	39.8	191.1	574.8	82.0
MC-9	WT	114.5	40.6	199.1	567.6	78.1
MC-10	WT	113.6	38.3	192.5	573.2	82.4
MC-11	WT	113.9	39.6	198.5	567.8	80.2
MC-12	WT	114.4	39.6	197.5	567.7	80.8
MC-13	Dan§	119.1	38.6	185.7	572.9	83.8
MC-14	Dan	116.6	39.4	185.5	573.0	85.5
MC-15	Dan	117.0	39.3	181.7	576.9	85.1
MC-16	Dan	116.3	39.6	183.1	572.9	88.3
MC-17	Dan	116.6	39.6	183.8	575.3	84.6
MC-18	Dan	115.3	39.8	188.0	573.2	83.8
MC-19	Dan	116.0	39.0	182.8	576.7	85.4
MC-20	Dan	113.8	40.0	189.0	572.3	84.8
MC-21	Dan	116.5	39.5	187.4	573.7	82.8
MC-22	Dan	116.1	39.6	185.6	573.7	85.2
MC-23	Dan	116.4	39.5	190.0	571.7	82.4
MC-24	Dan	116.4	39.7	188.3	571.4	84.2

†Checks included in field trial were: TN12-4100 and Ellis
‡WT: Wild type allele (Gm 20)
§Dan: Danbaekkong type allele (Gm 20)

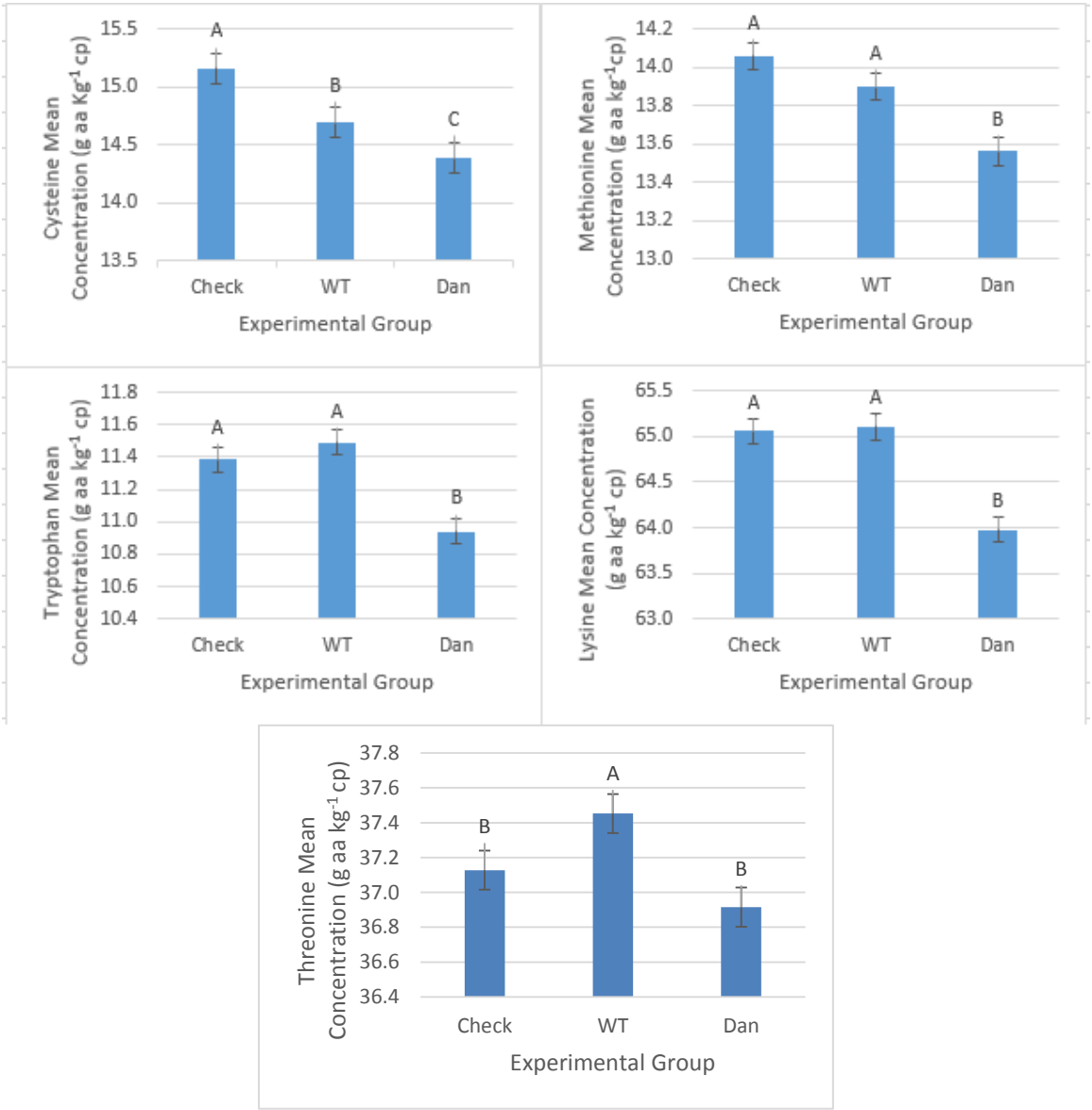

Figure 2.1 Means of amino acids [g aa kg^{-1} crude protein (cp)] for 24 $F_{8:10}$ lines in two experimental groups [wild-type (WT) and Danbaekkong (Dan) type] across three locations.

CHAPTER III
EVALUATION OF AGRONOMIC AND SEED QUALITY TRAITS OF NEAR ISOGENIC LINES OF SOYBEANS THAT DIFFER AT THE DT1 STEM TERMINATION LOCUS ON CHROMOSOME 19

Abstract

Height variation in soybeans can be due to the Dt1 stem termination gene located on chromosome 19 (Gm 19). Two distinct stem termination phenotypes, determinate (*det*) and indeterminate (*ind*), result from this gene. Determinate soybeans are typically shorter than their *ind* counterparts. Two populations, one conventional and one containing the Roundup Ready® (RR1) trait for resistance to glyphosate herbicide, were developed to conduct replicated field trials in two Tennessee locations. Conventional lines used in the first test were TN14-4001 and TN14-4003. The RR1 line used in the second test was TN15-5802. The objectives of this study were to determine if stem termination has an effect on yield of soybean grown in Tennessee and to evaluate the effect of stem termination on protein concentration, oleic acid, and other agronomic and seed quality traits. Interest lies in testing the effect on seed quality and agronomics of near isogenic lines (NILs) that differ at the Gm 19 Dt_1 locus to guide southern breeders in selections, since there is typically a divide between preference of stem termination soybean grown in the northern and southern United States. Experimental groups for both tests were *det*, *ind*, and segregating (*seg*). There were no significant differences in yield between conventional experimental group near isogenic lines (NIL) of TN14-4001 or TN14-4003 ($p < 0.05$). Additionally, there were no significant differences when RR1 *ind* and *det* experimental NIL groups were compared to the *seg* group; however, there was a significant difference in yield between *ind* and *det* experimental groups ($p < 0.05$). No significant differences in protein were identified within the conventional test; however, there was a significant difference in oil between the *ind* and *det* experimental group NILs

of TN14-4003 ($p < 0.05$). Additionally, there were no significant differences identified in protein or oil between any experimental groups in the RR1 test.

Introduction

There are two distinct phenotypes for stem termination in soybean: determinate (*det*), with the recessive genotype dt_1/dt_1, and indeterminate (*ind*), with the dominant genotype Dt_1/Dt_1 (Beaver and Johnson, 1981b; Bernard, 1972; Carlson and Lersten, 2004; Curtis et al., 2000; Heatherly and Smith, 2004; Kilgore-Norquest and Sneller, 2000) on Gm 19 (SoyBase and the Soybean Breeder's Toolbox, 2016). The difference between the two is the timing of the stem termination (Ablett et al., 1994; Heatherly and Smith, 2004). First reported in 1933, the gene pair *Dt1dt1* was based on segregation of a cross between *ind* (Manchu) and *det* (Ebony) soybean lines (Thompson et al., 1997). The stem termination alleles determine when seed development occurs (Rennie and Tanner, 1991). Other differences that have been found between different stem termination types were length of flowering, length of reproductive periods, height at maturity, and number of nodes (Ablett et al., 1994).

Determinate soybeans are generally maturity group V or later (Curtis et al., 2000; Heatherly and Smith, 2004). They are better adapted to longer growth season areas, such as the southern U.S., and are generally shorter in height and have fewer main stem nodes than *ind* soybeans (Beaver and Johnson, 1981a; Bernard, 1972; Kilgore-Norquest and Sneller, 2000; Lin and Nelson, 1988; Rennie and Tanner, 1991; Thompson et al., 1997). Determinate plants terminate apical stem growth abruptly (Heatherly and Smith, 2004) and stop growing after flowering occurs (Beaver and Johnson, 1981b; Rennie and Tanner, 1991). This results in a reduction of stem length due to a reduction in node number (Heatherly and Smith, 2004). Plants found to be *det* also have a decreased problem with lodging, due to their shorter stature (Rennie and Tanner, 1991).

Indeterminate plants are generally maturity group IV and earlier (Heatherly and Smith, 2004). They are better adapted to short growing season areas, such as the northern U.S. (Beaver and Johnson, 1981b), and plants continue to grow during their reproductive phase resulting in longer stems and more internodes than *det* plants (Carlson and Lersten, 2004; Heatherly and Smith, 2004; Kilgore-Norquest and Sneller, 2000; Rennie and Tanner, 1991; Thompson et al., 1997). Indeterminate soybeans have an even distribution of pods along the stem at maturity, while *det* soybeans have a dense cluster of pods at the ends of the branches (Carlson and Lersten, 2004). There is an increase in likelihood of lodging for *ind* soybeans due to their tall stature (Ablett et al., 1989). In highly productive environments, a reduction of yield up to 23% has been reported due to lodging (Cooper, 1981).

There is a trend between the type of stem termination planted and geographic location. Generally, in the mid-west and northern U.S., *ind* soybeans are grown (Curtis et al., 2000; Kilgore-Norquest and Sneller, 2000). In the southern U.S., planting of *det* soybeans is preferred (Kilgore-Norquest and Sneller, 2000). However, in recent years many southern soybean producers, especially in Arkansas and Mississippi, have shifted to the early soybean production system (Heatherly and Smith, 2004; Mengistu and Heatherly, 2006), which requires earlier maturing cultivars that are *ind* in growth habit for planting in late March to early April. Due to its lodging resistance, there is also interest in growing *det* soybeans in the northern U.S. (Lin and Nelson, 1988). We are interested in testing the effect on seed quality and agronomics of NILs that differ at the chromosome 19 Dt1 locus to guide southern breeders in selections.

Materials and Methods

Plant Materials

Single plants were harvested from within each of experimental lines **TN14-4001** (G03-3101 × LD00-2817P) (Table 3.1), **TN14-4003** (LD00-2817P × (17D × S08-14788 #3)) (Table 3.2), and **TN15-5802** (TN09-029 × USG 74T59) (Table 3.3). These lines were chosen because when grown in 2015, their phenotypes appeared to segregate for stem termination. Each plant per line was single plant threshed, and the seeds were weighed. Based on the number of seeds per plant, one, two, or four rows were grown in a winter nursery in Homestead, Florida to increase seed. These rows were bulk harvested in Florida, and the seed was used to form two location, two replication field tests.

Phenotypes were confirmed when grown in the 2016 field test, by SNP genotyping and additional phenotypic selection. Lines TN14-4001 and TN14-4003 are both conventional herbicide technologies and were therefore grown in the same two replication, two location field test. From line TN14-4001, there were two *det* and two *ind* lines, creating four individual genotypes for comparison. From line TN14-4003, there was one *det*, four *ind*, and one segregating (*seg*) line, creating six individual genotypes for comparison. Line TN15-5802 is a Roundup Ready® (RR1) herbicide technology line and was therefore grown in a separate two replication, two location field test. From line TN15-5802, there were five *det*, three *ind*, and two *seg* lines, creating ten individual genotypes for comparison. The seeds in both tests were planted in a randomized complete block design in Knoxville, TN on May 25, 2016, and in Springfield, TN on June 8, 2016.

DNA Extraction and SNP Genotyping

To confirm the genotypes of the plants growing in the field trial, four plants from each row in the first replication of the Knoxville location were tagged, and leaf samples were taken. DNA was stabilized onto pre-labeled Whattman FTA Elute cards (GE Healthcare Life Sciences, Buckinghamshire, England). SNP analysis was run on these samples for various genes (Table 3.4). SNP analysis was run on all three lines for the Dt_1 gene using the SNP assay for Dt_1 developed by Dr. Kristin Bilyeu, USDA-ARS, Columbia, MO (personal communication). Additionally, SNP analysis for the Danbaekkong (Dan) locus for seed protein concentration was run on all plants derived from line TN14-4001. The analysis was performed using the SNP assay for protein on chromosome 20, developed by Dr. Zenglu Li, the University of Georgia in Athens, GA (personal communication). SNP analysis was run on all plants derived from the line TN14-4003 for the High Oleic FAD2-1A (17D- version) using a SNP assay for high oleic at GmFAD2-1a (17D), and FAD2-1B using SNP assay for high oleic at GmFAD2-1b, both developed by Dr. Kristin Bilyeu, USDA-ARS, Columbia, MO (personal communication). The genotype results from SNP analysis confirmed the phenotypic selections made the previous year in the field, and allowed for the addition of the segregating genotypic class.

Agronomic Traits

Each plot was walked through during peak flowering. Flower color was recorded during this time. In the conventional test, both parents of line TN14-4001 had purple flowers, therefore it was expected that all plots would have purple flowers. Also in the conventional test, TN14-4003, had one parent with purple flowers (LD00-2817P) and one

parent with white flowers (17D × S08-14788 #3). The expectation here was that there should be segregation in flower color, which was observed. Two checks, a white flowering check (Ellis) (Pantalone et al., 2017) and a purple flowering check (LD00-2817P) were planted to confirm proper planting order and serve as yield checks. In the RR1 test, both parents (TN09-029 and USG 74T59) had purple flowers, therefore it was expected that all plots would have purple flowers. Two white flowering checks, TN13-5538RR1 and TN13-5537RR1, were also planted to confirm proper planting order and serve as yield checks. These two lines together constitute the new glyphosate herbicide cultivar Go Soy 54G16 marketed by Genetics Optimized (G0) (Dickson, TN). Any off-type soybean with incorrect flower color was pulled during this time to ensure purity of the lines.

Each plot was again walked through at the first sign of senescence. From this point on, plots were walked every 3-4 days and maturity notes were taken until each plot reached full maturity. A plot was considered fully mature when 90% or more of the pods appeared dried down. During this time days to maturity, lodging, height and pubescence color were recorded. Days to maturity was calculated based on number of days after planting the plot took to mature. Plots were given a lodging score from 1 to 5, where 1 was completely upright and 5 was completely lodged or flat on the ground. Height was recorded to the nearest inch by measuring an average looking single plant at random. Height was later converted to centimeters. In order to minimize variation in data, lodging and height scores were taken by the same individual. Days to maturity were recorded by the same individual to minimize differences in calls.

Near Infrared Reflectance

Near-infrared reflectance spectroscopy (NIRS) was used to determine concentrations of total protein and oil using a Perten DA 7250 analyzer. Calibration equations developed by the University of Minnesota, in cooperation with Perten, were used to analyze soybean seed quality traits. The equations were developed using HPLC measurements made at the University of Missouri analytical laboratory (Warrington et al., 2015). The Perten NIRS machine was calibrated using wet chemistry data from 900 North American soybean cultivars and breeding lines (Warrington et al., 2015). We used the Perten NIRS to measure protein and oil concentration and amino acid composition. A seed sample weighing approximately 20 g was sampled from each plot of the conventional and RR1 studies. Both studies consisted of 10 lines grown in a 2016 replicated field trial in two Tennessee locations. Therefore, 2 tests × 10 lines × 2 replications × 2 locations totaled 80 samples to run on the NIR. Each seed sample was ground using a Foss water-cooled Knifetec 1095 Sample Mill (Tecator, Hoganas, Sweden) to create uniform particle size. The values for protein and oil concentrations were converted to g kg^{-1} dry weight. Amino acid values were converted to g amino acid per kg crude protein (cp).

Gas Chromatography

Soybean oil fatty acids were analyzed using gas chromatography (GC). The five fatty acids tested were palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1), linoleic acid (18:2), and linolenic acid (18:3). A soybean oil profile consists of varying combinations of these fatty acids. Both studies consisted of 10 lines grown in a 2016 replicated field trial in two Tennessee locations. Therefore, 2 tests × 10 lines × 2

replications × 2 locations totaled 80 samples to run on the GC. A five seed sample was taken from each plot in both Tennessee locations. Each five seed sample was crushed, exposing more surface to extract soybean oil. Crushed seeds were then transferred to test tubes with a 3 mL volume of extraction solvent (2000 mL chloroform, 1250 mL hexanes, and 500 mL methanol). The test tubes were then capped with plastic stoppers to sit overnight. The next day, a 100 µL sample of the oil extract was transferred to a 1.5 mL vial. A 0.75 mL volume of hexanes and a 75 µL volume of methylation reagent (5 mL 0.5M sodium methoxide solution in methanol, 10 mL ethyl ether, and 20 mL petroleum ether) were then added to each vial. Caps were securely fastened to each vial. The fatty acid methyl esters were analyzed using a Hewlett Packard HP 6890 series gas chromatograph (Agilent Technologies, Santa Clara, CA) equipped with a flame ionization detector, 30 m x 0.53 mm 125-2332 capillary column, and a 7683 series auto sampler.

Data Analysis

Both conventional and RR1 studies were grown in a randomized complete block experimental design. 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). Least squares means were estimated in this model for the 10 lines and two checks grown in each of the conventional and RR1 studies in 2016. Fixed effects were individual lines. For simplicity, the original line name followed by a dash and number were used to identify lines, for example TN14-4001-3, where the last number corresponds to a plant ID that can be traced back to growth in a

winter nursery. Random effects were replication nested within location, location, and GxE or location by line. Differences within experimental groups were tested through use of contrast statements. Estimate statements for check, *det*, *ind*, and *seg* genotypes for both conventional and RR1 studies were run. Fisher's protected LSD was used to compare experimental groups within each test.

Results and Discussion

Agronomic Traits

The purpose of this study was to determine if agronomic and seed quality traits varied depending on whether a line was *det*, *ind*, or *seg*. Statistical values for 17 traits of interest for 12 $F_{8:10}$ conventional lines in five experimental groups with differences at the Dt_1 loci for stem termination are provided in Table 3.5. Conventional seed yield was separately averaged over two replications and locations for all lines. Across environments, an understanding of phenotypic stability is important to optimize yield performance (Wang et al., 2016), a trait farmers exhibit much economic interest in. For the conventional seed yields, all lines within five experimental groups (4001 *ind*, 4001 *det*, 4003 *ind*, 4003 *det*, and 4003 *seg*) were averaged. Seed yield for a sixth group of checks was additionally averaged. The two checks included in the study were Ellis and LD00-2817P. Studies have shown soybean architecture traits are involved in allocating energy for seed yield (Wang et al., 2016). It does not appear that the Dt_1 gene for the stem termination trait affects yield in this way. The checks ranged from an average of 2188 kg ha^{-1} for LD00-2817P to 4290 kg ha^{-1} for Ellis (Table 3.5). It is not surprising that Ellis out yielded LD00-2817P, because Ellis was developed in Tennessee and is well

adapted to the region, frequently taking the top place in yield trials (Pantalone et al., 2017), while LD00-2817P was developed and is well adapted for Illinois environments (Diers et al., 2010). There were no significant differences in yield between experimental groups derived from lines TN14-4001 or TN14-4003 ($p < 0.05$) (Figure 3.1). The TN14-4001 *ind* averaged a yield of 3422 kg ha^{-1} (Table 3.5). The TN14-4001 *det* yield averaged 3326 kg ha^{-1} (Table 3.5). The TN14-4003 *ind* yield averaged 2778 kg ha^{-1} (Table 3.5). The TN14-4003 *det* yield averaged 2470 kg ha^{-1} (Table 3.5). The TN14-4003 *seg* yield averaged 2440 kg ha^{-1} (Table 3.5).

Statistical values for 17 traits of interest for 12 $F_{7:9}$ RR1 lines in three experimental groups with differences at the Dt_1 loci for stem termination are provided in Table 3.6. For the RR1 seed yields, all lines within three experimental groups (*ind*, *det*, and *seg*) derived from line TN15-5802 were averaged by group. A fourth class of checks was averaged as well. The two checks included in the study were TN13-5538 RR1 and TN13-5537 RR1. There was a significant difference in yield between *ind* and *det* experimental groups ($p < 0.05$) (Figure 3.2). There were no significant differences when *ind* and *det* experimental groups were compared to the *seg* group ($p < 0.05$) (Table 3.6). The check group was the highest yielder with a mean of 3753 kg ha^{-1}. The RR1 *ind* experimental group was the next highest yielder, with a mean of 3504 kg ha^{-1} (Table 3.6) and ranging from 3444-3535 kg ha^{-1}. The RR1 *seg* experimental group was the third highest for yield, with a mean of 3318 kg ha^{-1} and ranging in yield from 3022-3613 kg ha^{-1}. The RR1 *det* experimental group was the lowest yielder, with a mean of 3189 kg ha^{-1} (Table 3.6) and ranging from 3031-3367 kg ha^{-1}.

Maturity of the lines varied across both locations. On average, the lines planted in Knoxville matured prior to lines planted in Springfield, TN for both conventional and RR1 tests. The conventional lines, TN14-4001 and TN14-4003, had relative maturities of 4.7 and 4.9, respectively. Two maturity group checks with similar relative maturities, LD00-2817P (4.5) and Ellis (4.9), were planted in this field trial. On average, these checks matured 113 and 124 days after planting (DAP), respectively. The average maturation for the TN14-4001 *det* group was 117 DAP and for the *ind* group, 118 DAP (Table 3.5). Average maturation for the TN14-4003 *det* group was 123 DAP while the *ind* group averaged 114 DAP (Table 3.5). The TN14-4003 *seg* group averaged maturity of 115 DAP (Table 3.5). Thus, the maturity of the checks fully bracketed the range of maturities of the experimental groups. There were no significant differences in maturity between experimental groups derived from line TN14-4001 ($p < 0.05$) (Table 3.5). Additionally, there were no significant differences in maturity between *ind* and *seg* experimental groups derived from line TN14-4003 ($p < 0.05$) (Figure 3.1). Previous studies have looked into the effect of maturity genes in combination with different growth habits (Curtis et al., 2000). There were significant differences in maturity between *ind* and *det* and between *det* and *seg* experimental groups derived from line TN14-4003 ($p < 0.05$) (Figure 3.1). Curtis et al. (2000) concluded that maturity genes in combination with growth habit genes have an effect on agronomic traits in soybean NILs.

The RR1 line TN15-5802 had a relative maturity of 5.3. Two maturity group checks with similar maturities, TN13-5537 RR1 and TN13-5538 RR1, were planted in this field trial. On average, the checks matured 123 and 124 DAP, respectively.

Determinate plants in the RR1 test averaged maturation of 124 DAP, while *ind* and *seg* averaged 128 DAP (Table 3.6). It appears that the *ind* growth habit delayed maturity. There were significant differences in maturation when comparing *ind* to *det* and *det* to *seg* experimental groups ($p < 0.05$) (Figure 3.2). When comparing *ind* to *det* growth habits in previous studies, research has shown the *ind* growth habit increased time to maturity over the *det* growth habit (Curtis et al., 2000). There was no significant difference between maturation of *ind* to *seg* experimental groups ($p < 0.05$) (Figure 3.2).

Variations in plant height can often be explained by differences at the Dt_1 locus. Conventional and RR1 plant height was separately averaged over two replications and locations for all lines. For conventional plant heights, all lines within five experimental groups (4001 *ind*, 4001 *det*, 4003 *ind*, 4003 *det*, and 4003 *seg*) were averaged by group. There was a significant difference in heights between *ind* and *det* plants within line TN14-4001 ($p < 0.05$) (Figure 3.1). Typically, shorter *det* plants exhibit less lodging and allocate more energy to reproductive organs (Wang et al., 2016). There were additional significant differences in heights between *ind* and *det* and between the *det* and *seg* experimental groups within line TN14-4003 ($p < 0.05$) (Figure 3.1). There was no significant difference between heights of *ind* and *seg* experimental groups within line TN14-4003 ($P < 0.05$) (Figure 3.1).

For the RR1 plant heights, all lines within three experimental groups (*ind*, *det*, and *seg*) were averaged by group. There were significant differences when comparing heights of all experimental groups ($p < 0.05$) (Table 3.6). As expected, the *det* group had the shortest average height (73.2 cm), while the *ind* group had the tallest average height

(125.5 cm). The average height for the *seg* group (111.7 cm) fell between the aforementioned groups (Table 3.6). Additionally, plant height is positively correlated with lodging (Wang et al., 2016).

Height and lodging have a direct effect on yield of soybeans (Curtis et al., 2000; Wang et al., 2016). Therefore, breeders commonly consider these traits when developing cultivars (Wilcox and Guodong, 1997). Lodging and height typically have a positive relationship since taller plants lack in lodging resistance. Lodging resistance is a desirable trait for soybeans, as yield typically increases and loss of seed typically decreases when present (Wang et al., 2016). Severely lodged soybeans often have lower yields than upright soybeans (Wilcox and Sediyama, 1981). Lodging can be influenced by a number of factors such as plant height, pod density, and shade. For example, in a study conducted by Wicox and Sediyama in 1981, lodging of tall, *ind* plants increases as density of plants grown is increased; however, under similar conditions relatively short, *det* plants did not lodge. In that experiment, all 73 *det* lines were resistant to lodging, while the 93 *ind* lines ranged in lodging (Wilcox and Sediyama, 1981).

Lodging was averaged separately for conventional and RR1 tests across locations and replications. For the conventional test there was a significant difference in lodging between experimental groups within line TN14-4001 ($p < 0.05$) (Figure 3.1). Additionally, there was a significant difference in lodging between *ind* and *det* experimental groups within line TN14-4003 ($p < 0.05$) (Figure 3.1). There were no significant differences in lodging between *ind* and *seg* and between *det* and *seg* experimental groups within line TN14-4003 ($p < 0.05$) (Figure 3.1). The TN14-4001 *ind*

lines were the most lodged with an average score of 3.1. The TN14-4001 *det* and TN14-4003 *ind* were the most upright with average lodging scores of 2 for both. Recent progress has been made in improving lodging resistance in *ind* plants (Ablett et al., 1989), and TN14-4003 *ind* may be another example of good lodging resistance in an *ind* line. Studies focusing on main stem strength of soybeans have resulted in an increased understanding of lodging resistant cultivars (Liu et al., 2016).

For the RR1 test, there were significant differences in lodging when comparing both *ind* to *det* and *det* to *seg* experimental groups ($p < 0.05$) (Figure 3.2). There was no significant difference in lodging between *ind* and *seg* experimental groups ($p < 0.05$) (Figure 3.2). The *seg* group (lodging score of 2.8) and the *ind* group (lodging score 2.6) were the most lodged. The *det* group was the most upright with a 2.1 lodging score (Figure 3.2).

Protein, Oil, and Amino Acids

Near infrared reflectance was used to evaluate the effect of the Dt_1 gene on soybean seed protein, oil, and amino acid (aa) concentration. Protein and oil are reported in g kg^{-1}. Amino acid values were divided by the protein concentrations and are reported as g aa kg^{-1} crude protein (cp). Five aa were analyzed: cysteine (Cys), methionine (Met), tryptophan (Trp), lysine (Lys), and threonine (Thr). These five aa are the most nutritionally limiting in the diets of poultry and swine (Abrams, 2015), and the majority of soybean meal produced in the U.S. goes to chicken and swine feed. Mean concentrations of protein, oil, and the five aa of interest for conventional $F_{8:10}$ lines are found in Table 3.7. For both *det* and *ind* soybean lines, an inverse relationship between

protein and yield exists, making it difficult to develop cultivars with both an increase in protein and yield (Wilcox and Guodong, 1997). Line TN14-4001 has the presence of high protein Danbaekkong (Dan) allele; therefore, both *ind* (454.8 g kg^{-1}) and *det* (452.7 g kg^{-1}) progeny of this line had significantly higher protein than the checks (397.8 g kg^{-1}) (Table 3.7). It is also worthy of mention that these groups had significantly lower oil concentrations than the checks, further confirming the negative genetic correlation between protein and oil. There were no significant differences in protein observed in the conventional test among the *det*, *ind*, or *seg* groups from either line (Table 3.7). This could be due to the fact that NILs were grown, which were very similar for protein concentration in this study (Wilcox and Guodong, 1997). Many studies have been conducted on the concentrations of protein in *det* and *ind* isolines (Wilcox and Guodong, 1997). For example, Escalante and Wilcox (1993) studied protein concentration differences between 10 *ind* and *det* NILs of soybean and found no differences among lines. This lack of difference in protein concentration suggests that it is feasible to grow both *ind* and *det* soybeans in Tennessee environments, without detrimental effects on protein. There was, however, a significant difference in oil in the conventional test, between the TN14-4003 *ind* and *det* lines ($p < 0.05$) (Table 3.7). Indeterminate NILs from line TN14-4003 had an overall mean oil concentration higher than that of the *det* NILs in the same line (Table 3.7). This finding supports that of Wilcox and Guodong (1997), who concluded that seed oil concentration in *det* plants averaged lower than *ind* progenies. Since there was an increase in oil and no significant difference in protein or yield within this line, it may be beneficial for Tennessee farmers to grow *ind* soybeans of

this specific line. There were no significant differences in Cys among groups in the conventional test. For Met, only the check group was significantly different, producing 13.8 g Met kg^{-1} cp, significantly higher than values for all other groups. This suggests that increases in seed protein concentration dilute the concentration of methionine per unit of crude protein. There was no significant difference in Trp between the TN14-4001 *det* (10.7 g Trp kg^{-1} cp) and *ind* (10.7g Trp kg^{-1} cp), but the difference was significant for TN14-4003 *det* (10.7 g Trp kg^{-1} cp) compared to *ind* (10.0 g Trp kg^{-1} cp) (Table 3.7). Similarly there was no significant difference in Lys for the TN14-4001 *det* (63.1 g Lys kg^{-1} cp) compared to *ind* (62.8 g Lys kg^{-1} cp), but the difference was significant for TN14-4003 *det* (63.9 g Lys kg^{-1} cp) compared to *ind* (63.3 g Lys kg^{-1} cp). The same trend was observed for Thr, where there was no significant difference for the TN14-4001 *det* (36.0 g Thr kg^{-1} cp) compared to *ind* (35.8 g Thr kg^{-1} cp), but the difference was significant for TN14-4003 *det* (36.9 g Thr kg^{-1} cp) compared to *ind* (35.9 g Thr kg^{-1} cp) (Table 3.7).

Mean concentrations for RR1 $F_{7:9}$ lines are found in Table 3.8. There were no significant differences observed in protein, oil, or any of the amino acids of interest among groups in the RR1 test ($p < 0.05$) with the following exceptions: The check group had significantly lower protein concentration (389.5 g kg^{-1}) than all other groups, the check group had significantly higher Lys (64.8 g Lys kg^{-1} cp) than all other groups, and the check group had significantly higher Thr (37.4 g Thr kg^{-1} cp) than all other groups (Table 3.8). These findings suggest that it is feasible for Tennessee farmers to grow either *ind* or *det* RR1 soybeans without negative effects on seed quality traits.

Fatty Acids

The effect of the Dt_1 gene on soybean seed fatty acid composition was evaluated using GC results. Fatty acids are reported in g kg^{-1}. We report the GC measured sample concentrations of palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2), and linolenic acids (18:3). Mean concentrations of fatty acids for conventional $F_{8:10}$ lines are found in Table 3.9. Both palmitic and linolenic fatty acids had significant differences between *ind* and *det* and *det* and *seg* experimental groups in line TN14-4003 ($p < 0.05$) (Table 3.9). Palmitic acid means were 65.4 g kg^{-1} for the seg lines and 110.2 g kg^{-1} for the *det* lines, showing significant variation between experimental groups. Additionally, significant differences in both oleic and linoleic acid were observed between all experimental groups derived from line TN14-4003 ($p < 0.05$). Line TN14-4003 is a high oleic line; therefore, the oleic acid concentration is higher than average in its NIL progeny. The mean values for oleic acid concentration were 727.8 g kg^{-1} *ind*, 268.1 g kg^{-1} *det*, and 815.7 g kg^{-1} *seg*. The *det* experimental group lost the high oleic trait, while the *ind* lines retained it. There is no linkage to the Dt_1 gene on Gm 19 associated with the loss of this trait, as FAD2-1A and FAD2-1B are located on Gm 10 and 20, respectively. Additional studies should be conducted to determine the cause of the loss of the high oleic trait from the *det* experimental group. Additionally, the mean values for the linoleic acid concentration were 116.8 g kg^{-1} *ind*, 516.3 g kg^{-1} *det*, and 43.1 g kg^{-1} *seg*. No significant difference in stearic acid between experimental groups was observed in the conventional test.

Similarly, mean concentrations of fatty acids for RR1 $F_{7:9}$ lines are found in Table 3.10. There was a significant difference in linolenic acid between the *ind* and *det* RR1 lines ($p < 0.05$) (Table 3.10). There was a significant difference in oleic and linoleic acid

between both the *ind* and *det* and the *ind* and *seg* RR1 lines ($p < 0.05$). There was no significant difference in the palmitic acid between any RR1 experimental groups. Commercial soybean oil is approximately: 100 g kg^{-1} palmitic, 40 g kg^{-1} stearic, 220 g kg^{-1} oleic, 540 g kg^{-1} linoleic, and 100 g kg^{-1} linolenic acid, on average (Wilson, 2004). Mean values for fatty acids averaged over both locations for the *ind* group were: 116 g kg^{-1} palmitic, 45 g kg^{-1} stearic, 239 g kg^{-1} oleic, 531 g kg^{-1} linoleic, and 68 g kg^{-1} linolenic acid. Mean values averaged across locations for the *det* group were: 118 g kg^{-1} palmitic, 43 g kg^{-1} stearic, 217 g kg^{-1} oleic, 551 g kg^{-1} linoleic, and 72 g kg^{-1} linolenic acid. Mean values averaged across locations for the *seg* group were: 119 g kg^{-1} palmitic, 44 g kg^{-1} stearic, 219 g kg^{-1} oleic, 547 g kg^{-1} linoleic, and 71 g kg^{-1} linolenic acid. The small variation in concentrations of fatty acids between the *ind*, *det*, and *seg* experimental groups shows that the Dt_1 allele does not substantially negatively impact any fatty acid accumulation.

References

Ablett, G.R., R.I. Buzzell, W.D. Beversdorf, and O.B. Allen. 1994. Comparative stability of 40 indeterminate and semideterminate soybean lines. Crop Sci. 34: 347-351.

Ablett, G.R., W.D. Beversdorf, and V.A. Dirks. 1989. Performance and stability of indeterminate and determinate soybeans in short season environments. Corp Sci. 29:1428-1433.

Abrams, J. 2015. Identification and confirmation of quantitative trait loci for protein concentration and improved amino acid composition in *Glycine max*. Ph. D. Dissertation. University of Tennessee.

Beaver, J.S. and R.R. Johnson. 1981a. Response of determinate and indeterminate soybeans to varying cultural practices in the northern USA. Agron. J. 73:833-838.

Beaver, J.S. and R.R. Johnson. 1981b. Yield stability of determinate and indeterminate soybeans adapted to the northern United States. Corp Sci. 21:449-454.

Bernard, R.L. 1972. Two genes affecting stem termination in soybeans. Crop Sci. 12:235-239.

Carlson, J.B., and N.R. Lersten. 2004. Reproductive morphology, p. 59-95. In: H.R. Boerma and J.E. Specht (eds.). Soybeans: improvement, production, and uses. 3rd ed. ASA Inc., CSSA Inc., SSSA Inc., Madison, WI.

Cooper, R.L. 1981. Development of short statured soybean cultivars. Crop Sci. 21:238-131.

Curtis, D.F., J.W. Tanner, B.M. Luzzi, and D.J. Hume. 2000. Agronomic and phenological differences of soybean isolines differing in maturity and growth habit. Crop Sci. 40: 1624-1629.

Diers, B.W., T. Carey, D. Thomas, A. Colgrove, and T. Niblack. 2010. Registration of LD00-2817P soybean germplasm line with resistance to soybean cyst nematode from PI 437654. J. Plant Reg. 4:141-144.

Escalante, E.E., and J.R. Wilcox. 1993. Variation in seed protein among nodes of determinate and indeterminate soybean nearisolines. Crop Sci. 33:1855-1858.

Heatherly, L.G., and J.R. Smith. 2004. Effect of soybean stem growth habit on height and node number after beginning bloom in the midsouthern USA. Crop Sci. 44:1855-1858.

Kilgore-Norquest, L. and C.H. Sneller. 2000. Effect of stem termination on soybean traits in southern U.S. production systems. Crop Sci. 40: 83-90.

Lin, M.S. and R.L. Nelson. 1988. Relationship between plant height and flowering date in determinate soybean. Crop Sci. 28:27-30.

Liu, W., Y. Deng, S. Hussain, J. Zou, J. Yuan, L. Luo, C. Yang, X. Yuan, and W. Yang. 2016. Relationship between cellulose accumulation and lodging resistance in the stem of relay intercropped soybean [*Glycine max* (L.) Merr.]. Field Crops Research. 196: 261-267.

Mengistu, A., and L.G. Heatherly. 2006. Planting date, irrigation, maturity group, year, and environment effects on *Phomopsis longicolla*, seed germination, and seed

health rating of soybean in the early soybean production system of the midsouthern USA. Crop Protection. 25:310-317.

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: (in press).

Rennie, B.D. and J.W. Tanner. 1991. Interactions between stem termination type and linolenic acid content in soybean. Can. J. Plant Sci. 71: 649-653.

SoyBase and the Soybean Breeder's Toolbox. 2016. Report for Sequence Feature Glyma.19g194300. USDA and Iowa State University.

Thompson, J.A., R.L Bernard, and R.L Nelson. 1997. A third allele at the soybean *dt1* locus. Crop Sci. 37:757-762.

Wang, C., T. Wu, S. Sun, R. Xu, J. Ren, C. Wu, B. Jiang, W. Hou, and T. Han. 2016. Seventy-five years of improvement of yield and agronomic traits of soybean cultivars released in the Yellow-Huai-Hai river valley. Crop Sci. 56:2354-2364.

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

Wilcox, J.R. and Z. Guodong. 1997. Relationships between seed yield and seed protein in determinate and indeterminate soybean populations. Crop Sci. 37:361-.

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.

Appendix

Table 3.1 Depiction of population development of conventional line TN14-4001 from inception of the cross of G03-3101× LD00-2817P in 2010 to 2016 test.

Year	Season	Location	Generation grown	Procedure	Row Designation
2010	Summer	ETREC[a]	Parentals	Cross P1 x P2	Cross 10-19
2010-2011	Winter	TARS[b]	F_1	Harvest individual F_1 plants	VP-195 to VP-200
2011	Summer	ETREC	F_2	Pick one pod per plant	20133 - 20148
2011-2012	Winter	27 Farms[c]	F_3	Pick one pod per plant	VH12-055
2012	Summer	ETREC	F_4	Single plants	40150
2013	Summer	ETREC	$F_{4:5}$	Bulk harvest	50447
2013-2014	Winter	27 Farms	$F_{4:6}$	Bulk harvest	VH14-3287 to VH14-3294
2014	Summer	ETREC	$F_{4:7}$	Bulk harvest	PYT4e1CON
2015	Summer	ETREC	$F_{4:8}$	SingleF_8 plants	IYT4LCON
2015-2016	Winter	27 Farms	$F_{8:9}$	Bulk harvest	VH16-001 to VH16-041
2016	Summer	ETREC, HRREC[d]	$F_{8:10}$	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

Table 3.2 Depiction of population development of conventional line TN14-4003 from inception of the cross of LD00-2817P and 17D × S08-14788 #3 in 2010 to 2016 test.

Year	Season	Location	Generation grown	Procedure	Row Designation
2010	Summer	ETREC[a]	Parentals	Cross P1 x P2	Cross 10-02
2010-2011	Winter	TARS[b]	F_1	Harvest individual F_1 plants	VP-011 to VP-025
2011	Summer	ETREC	F_2	Pick one pod per plant	20003 - 20018
2011-2012	Winter	27 Farms[c]	F_3	Pick one pod per plant	VH12-001 to VH12-020
2012	Summer	ETREC	F_4	Single plants	40003-40041
2013	Summer	ETREC	$F_{4:5}$	Bulked or single plant	50010
2013-2014	Winter	27 Farms	$F_{4:6}$or $F_{5:6}$	Bulk harvest	VH14-3299 to VH14-3302
2014	Summer	ETREC	$F_{4:7}$or $F_{5:7}$	Bulk harvest	PYT4e1Con
2015	Summer	ETREC	$F_{4:8}$or $F_{5:8}$	SingleF_8 plants	IYT4eCON
2015-2016	Winter	27 Farms	$F_{8:9}$	Bulk harvest	VH16-042 to VH16-074
2016	Summer	ETREC, HRREC[d]	$F_{8:10}$	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

Table 3.3 Depiction of population development of RR1 line TN15-5802 from inception of the cross of TN09-029 and USG 74T59 in 2010 to 2016 test.

Year	Season	Location	Generation grown	Procedure	Row Designation
2010	Summer	ETREC[a]	Parentals	Cross P1 x P2	Cross 10-20 RR1
March2011	Spring	UTIA[b] Greenhouse	F_1	Harvest individual F_1 plants	10-20-1 to 10-20-8
2011	Summer	ETREC	F_2	Pick one pod per plant	25011 - 25018
2011-2012	Winter	27 Farms[c]	F_3	Pick one pod per plant	VH12-501 to VH12-506
2012	Summer	ETREC	F_4	Single plants	45002-45011
2013	Summer	ETREC	$F_{4:5}$	Single plants	57105
2014	Summer	ETREC	$F_{5:6}$	Bulk harvest	67038
2015	Summer	ETREC	$F_{5:7}$	Single F_7 plants	PYT5RR1
2015-2016	Winter	27 Farms	$F_{7:8}$	Bulk harvest	VH16-075 to VH16-104
2016	Summer	ETREC, HRREC[d]	$F_{7:9}$	Combine harvest individual plots	

[a]ETREC, East Tennessee Research and Education Center
[b] UTIA, University of Tennessee Institute of Agriculture
[c]27 Farms, 27 Farms of Homestead, Inc., Homestead, FL
[d] HRREC, Highland Rim Research and Education Center

Table 3.4 Single nucleotide polymorphisms (SNPs) used to determine the DNA for genes of interest in the stem termination experiment.

		Gene of Interest for SNP			
Line	**Herbicide**	**Dan†**	**Dt_1‡**	**FAD2-1a (17D)§**	**FAD2-1b¶**
TN14-4001	Conventional	1#	1	0††	0
TN14-4003	Conventional	0	1	1	1
TN15-5802	RR1‡‡	0	1	0	0

† Dan= Danbaekkong allele for increased protein
‡ Dt_1= Stem Termination type
§ FAD2-1a (17D)= High Oleic acid
¶ FAD2-1b= High Oleic acid
# 1= Indicates the population was tested
†† 0= Indicates the population was not tested
‡‡ RR1= Roundup Ready 1 herbicide technology

Table 3.5 Statistical values for 17 traits of interest for 12 $F_{8:10}$ conventional lines in five experimental groups: i. TN14-4001 indeterminate, ii. TN14-4001 determinate, iii. TN14-4003 indeterminate, iv. TN14-4003 determinate, and v. TN14-4003 segregating, with differences at the Dt_1 loci for stem termination. Experimental group values were pulled from estimate statements run in SAS. LSD values were generated based on a line by line comparison for overall traits of interest. Checks used in this study were Ellis and LD00-2817P.

		Experimental Groups										
Trait	**P value**	**Check**	**4001 Ind†**	**4001 Det‡**	**4003 Ind§**	**4003 Det¶**	**4003 Seg#**	**Min**	**Mean**	**Max**	**LSD value**	**CV**
Yield (Kg Ha^{-1})	0.086	3239.0	3421.9	3325.6	2777.5	2470.0	2440.0	2187.8	2999.4	4290.1	1019.4	19.3
Protein (g Kg^{-1})	<.0001	397.8	454.8	452.7	422.5	421.3	423.4	389.4	428.8	456.2	10.2	1.8
Oil (g Kg^{-1})	0.0002	231.0	216.4	212.3	229.0	214.3	225.5	212.3	222.9	238.9	9.5	1.5
Palmitic (g Kg^{-1})	<.0001	107.4	109.1	111.6	72.4	110.2	65.4	65.0	93.4	112.9	6.9	4.0
Stearic (g Kg^{-1})	0.0006	46.2	38.1	36.5	38.5	37.8	36.9	35.6	39.2	48.5	4.2	6.8
Oleic (g Kg^{-1})	<.0001	270.3	273.5	225.2	727.8	268.1	815.7	222.5	461.1	815.7	78.5	8.8
Linoleic (g Kg^{-1})	<.0001	510.5	513.6	555.3	116.8	516.3	43.1	39.4	348.8	556.7	69.8	10.5
Linolenic (g Kg^{-1})	<.0001	65.7	65.7	71.5	44.6	67.7	39.0	39.0	57.5	72.3	6.4	4.9
Cysteine (g Kg^{-1} cp)	0.9354	14.8	14.5	14.4	14.6	14.8	14.6	14.2	14.6	14.9	0.7	3.0
Methionine (g Kg^{-1} cp)	0.0748	13.8	13.3	13.3	13.3	13.5	13.2	13.0	13.4	13.9	0.4	1.7
Tryptophan (g Kg^{-1} cp)	<.0001	11.4	10.7	10.7	10.0	10.7	9.7	9.6	10.5	11.5	0.4	2.4
Lysine (g Kg^{-1} cp)	<.0001	64.4	62.8	63.1	63.3	63.9	62.9	62.7	63.4	64.4	0.5	0.6
Threonine (g Kg^{-1} cp)	0.0028	36.9	35.8	36.0	35.9	36.9	36.1	35.8	36.2	36.9	0.5	1.1
Maturity (Days)	<.0001	118.9	117.5	117.4	114.3	122.8	115.0	113.8	116.9	124.0	2.0	1.1
Lodging (1-5 scale)	0.0484	2.5	3.1	2.0	2.0	3.0	2.3	1.8	2.4	3.3	0.8	22.7
Height (cm)	0.0001	75.9	114.9	69.5	101.9	80.0	102.9	67.9	92.6	118.7	17.7	5.0
Meal (g Kg^{-1})	<.0001	470.9	529.8	525.1	498.9	489.8	498.3	464.8	502.9	530.5	10.7	1.7

†4001 Ind= TN14-4001 indeterminate near isogenic line (NIL)
‡4001 Det= TN14-4001 determinate NIL
§4003 Ind=TN14-4003 indeterminate NIL
¶4003 Det=TN14-4003 determinate NIL
#4003 Seg=TN14-4003 NIL segregating for stem termination

Table 3.6 Statistical values for 17 traits of interest for 12 $F_{7:9}$ RR1 TN15-5802 lines in three experimental groups: indeterminate, determinate, and segregating, with differences at the Dt_1 loci for stem termination. Experimental group values were pulled from estimate statements run in SAS. LSD values were generated based on a line by line comparison for overall traits of interest. Checks used in this study were TN13-5537RR1 and TN13-5538RR1.

		Experimental Groups					
Trait	**P value**	**Check**	**Ind†**	**Det‡**	**Seg§**	**LSD value**	**CV**
Yield (Kg Ha^{-1})	0.0751	3753.4	3503.7	3189.3	3318.0	353.9	11.8
Protein (g Kg^{-1})	0.001	389.5	400.5	403.2	405.4	7.3	2.1
Oil (g Kg^{-1})	0.4318	226.5	229.2	225.5	225.2	5.2	1.6
Palmitic (g Kg^{-1})	0.0332	113.3	116.3	117.7	118.9	3.2	3.1
Stearic (g Kg^{-1})	0.0172	41.3	45.4	43.2	44.3	2.3	5.9
Oleic (g Kg^{-1})	0.0023	241.4	238.8	216.5	219.4	14.9	5.4
Linoleic (g Kg^{-1})	0.0026	533.1	531.4	550.7	546.8	12.1	1.7
Linolenic (g Kg^{-1})	0.1668	70.9	68.1	72.0	70.7	4.3	5.2
Cysteine (g Kg^{-1} cp)	0.6829	15.3	15.3	15.4	15.2	0.3	2.5
Methionine (g Kg^{-1} cp)	0.2881	14.1	14.0	14.0	13.9	0.2	1.6
Tryptophan (g Kg^{-1} cp)	0.1671	11.3	11.3	11.2	11.3	0.3	2.8
Lysine (g Kg^{-1} cp)	0.0648	64.8	64.2	64.3	64.1	0.4	0.6
Threonine (g Kg^{-1} cp)	0.1151	37.4	37.0	37.1	36.9	0.3	1.1
Maturity (Days)	<.0001	123.9	128.4	123.8	127.9	1.3	1.2
Lodging (1-5scale)	0.0097	1.5	2.7	2.1	2.9	0.6	24.6
Height (cm)	<.0001	76.5	125.5	73.2	111.7	11.6	7.0
Meal (g Kg^{-1})	0.0005	458.6	473.1	474.4	476.6	7.5	1.8

†Indeterminate growth habit
‡Determinate growth habit
§Segregating for stem termination

Table 3.7 Mean concentrations of protein, oil, and amino acids for 10 conventional $F_{8:10}$ near isogenic lines (NIL) in five experimental groups with differences at the Dt_1 loci. Checks used in this study were Ellis and LD00-2817P.

	Protein, oil, and amino acid concentrations						
	Protein	Oil	Cysteine	Methionine	Tryptophan	Lysine	Threonine
Line and Genotypic Class	g kg^{-1}		g aa kg^{-1} cp				
Check	397.8	231.0	14.8	13.8	11.4	64.4	36.9
TN14-4001 ind†	454.8	216.4	14.5	13.3	10.7	62.8	35.8
TN14-4001 det‡	452.7	212.3	14.4	13.3	10.7	63.1	36.0
TN14-4003 ind§	422.5	229.0	14.6	13.3	10.0	63.3	35.9
TN14-4003 det¶	421.3	214.3	14.8	13.5	10.7	63.9	36.9
TN14-4003 seg#	423.4	225.5	14.6	13.2	9.7	62.9	36.1
LSD (0.05)	9.0	8.4	0.6	0.4	0.3	0.5	0.5

† Indeterminate stem termination in conventional NILs of TN14-4001
‡ Determinate stem termination in conventional NILs of TN14-4001
§ Indeterminate stem termination in conventional NILs of TN14-4003
¶ Determinate stem termination in conventional NILs of TN14-4003
# Segregating for stem termination in conventional NILs of TN14-4003

Table 3.8 Mean concentrations of protein, oil, and amino acids for 10 RR1 $F_{7:9}$ near isogenic lines (NIL) in three experimental groups with differences at the Dt_1 loci. Checks used in this study were TN13-5537RR1 and TN13-5538RR1.

	Protein, oil, and amino acid concentrations						
	Protein	Oil	Cysteine	Methionine	Tryptophan	Lysine	Threonine
Line and Genotypic Class	g kg^{-1}		g aa kg^{-1} cp				
Check	389.5	226.5	15.3	14.1	11.3	64.8	37.4
TN15-5802 ind†	400.5	229.2	15.3	14.0	11.3	64.2	37.0
TN15-5802 det‡	403.2	225.5	15.4	14.0	11.2	64.3	37.1
TN15-5802 seg§	405.4	225.2	15.2	13.9	11.3	64.1	36.9
LSD (0.05)	7.3	5.2	0.3	0.2	0.3	0.4	0.3

† Indeterminate stem termination in Roundup Ready 1 NILs of TN15-5802
‡ Determinate stem termination in Roundup Ready 1 NILs of TN15-5802
§ Segregating stem termination in Roundup Ready 1 NILs of TN15-5802

Table 3.9 Mean concentrations of fatty acids for 10 conventional $F_{8:10}$ near isogenic lines (NIL) in five experimental groups with differences at the Dt_1 loci. Check included in this study were Ellis and LD00-2817P.

	Fatty acid concentrations				
	Palmitic	Stearic	Oleic	Linoleic	Linolenic
Line and Genotypic Class	g kg^{-1}				
Check	107.4	46.2	270.3	510.5	65.7
TN14-4001 ind†	109.1	38.1	273.5	513.6	65.7
TN14-4001 det‡	111.6	36.5	225.2	555.3	71.5
TN14-4003 ind§	72.4	38.5	727.8	116.8	44.6
TN14-4003 det¶	110.2	37.8	268.1	516.3	67.7
TN14-4003 seg#	65.4	36.9	815.7	43.1	39.0
LSD (0.05)	6.1	3.7	69.7	62.0	5.7

† Indeterminate stem termination in conventional NILs of TN14-4001
‡ Determinate stem termination in conventional NILs of TN14-4001
§ Indeterminate stem termination in conventional NILs of TN14-4003
¶ Determinate stem termination in conventional NILs of TN14-4003
# Segregating for stem termination in conventional NILs of TN14-4003

Table 3.10 Mean concentrations of fatty acids for 10 RR1 $F_{7:9}$ near isogenic lines (NIL) in three experimental groups with differences at the Dt_1 loci. Checks used in this study were TN13-5537RR1 and TN13-5538RR1.

	Fatty acid concentrations				
	Palmitic	Stearic	Oleic	Linoleic	Linolenic
Line and Genotypic Class	g kg^{-1}				
Check	113.3	41.3	241.4	533.1	70.9
TN15-5802 ind†	116.3	45.4	238.8	531.4	68.1
TN15-5802 det‡	117.7	43.2	216.5	550.7	72.0
TN15-5802 seg§	118.9	44.3	219.4	546.8	70.7
LSD (0.05)	3.2	2.3	14.9	12.1	4.3

† Indeterminate stem termination in Roundup Ready 1 NILs of TN15-5802
‡ Determinate stem termination in Roundup Ready 1 NILs of TN15-5802
§ Segregating stem termination in Roundup Ready 1 NILs of TN15-5802

Table 3.11 Mean agronomic trait values and protein, oil, amino acid, and fatty acid concentrations of 10 conventional $F_{8:10}$ experimental lines and two checks averaged over three replications and three locations. Checks included in this study were Ellis and LD00-2817P.

Line	Class	Yield	Maturity	Lodging	Height	Crude Protein	Oil	Cys	Met	Trp	Lys	Thr	Palmitic	Stearic	Oleic	Linoleic	Linolenic
		kg ha^{-1}	DAP		cm	g kg^{-1}		g aa kg^{-1} cp					g kg^{-1}				
LD00-2817P	Check	2187.8	113.8	2.8	79.4	389.4	238.9	14.8	13.8	11.5	64.4	36.9	110.4	48.5	247.1	531.0	63.2
Ellis	Check	4290.1	124.0	2.3	72.4	406.3	223.2	14.8	13.9	11.3	64.4	36.9	104.4	43.9	293.6	490.1	68.2
TN 14-4001-3	Det	3570.6	117.3	2.0	67.9	456.2	212.4	14.2	13.2	10.7	63.1	35.9	112.9	35.8	222.5	556.7	72.3
TN 14-4001-5	Det	3080.6	117.5	2.0	71.1	449.1	212.3	14.6	13.3	10.7	63.2	36.2	110.4	37.3	227.8	553.8	70.7
TN 14-4001-9	Ind	3753.0	117.0	3.3	111.1	456.2	214.8	14.4	13.3	10.7	62.7	35.8	109.0	36.3	281.1	507.4	66.3
TN 14-4001-11	Ind	3090.7	118.0	3.0	118.7	453.4	217.9	14.6	13.3	10.7	62.9	35.9	109.3	39.9	265.9	519.8	65.1
TN 14-4003-18	Det	2470.0	122.8	3.0	80.0	421.3	214.3	14.8	13.5	10.7	63.9	36.9	110.2	37.8	268.1	516.3	67.7
TN 14-4003-17	Ind	2881.1	114.3	1.8	95.3	418.1	238.8	14.4	13.4	10.8	63.3	35.8	91.2	37.6	484.3	331.1	55.9
TN 14-4003-19	Ind	2543.8	114.8	2.0	98.4	426.6	226.5	14.4	13.0	9.7	63.1	35.9	65.0	40.0	815.7	39.4	40.0
TN 14-4003-20	Ind	2623.2	113.8	2.0	108.6	419.7	225.1	14.6	13.3	9.8	63.4	35.8	66.1	35.6	810.9	45.6	41.9
TN 14-4003-22	Ind	3062.0	114.3	2.3	105.4	425.7	225.6	14.9	13.4	9.6	63.2	36.1	67.3	40.9	800.2	51.1	40.5
TN 14-4003-15	Seg	2440.0	115.0	2.3	102.9	423.4	225.5	14.6	13.2	9.7	62.9	36.1	65.4	36.9	815.7	43.1	39.0

† Determinate stem termination
‡ Indeterminate stem termination
§ Segregating for stem termination

Table 3.12 Mean agronomic trait values and protein, oil, amino acid, and fatty acid concentrations of 10 RR1 $F_{7:9}$ experimental lines and two checks averaged over three replications and three locations. Checks used in this study were TN13-5537RR1 and TN13-5538RR1.

Line	Class	Yield	Maturity	Lodging	Height	Crude Protein	Oil	Cys	Met	Trp	Lys	Thr	Palmitic	Stearic	Oleic	Linoleic	Linolenic
		kg ha^{-1}	DAP		cm	g kg^{-1}		g aa kg^{-1} cp					g kg^{-1}				
TN13-5537RR1	Check	3582.7	123.0	1.5	71.8	384.1	228.5	15.1	14.0	11.5	64.9	37.6	113.7	42.4	238.1	535.3	70.5
TN13-5538RR1	Check	3924.0	124.8	1.5	81.3	394.8	224.6	15.5	14.2	11.0	64.7	37.3	112.9	40.3	244.7	530.9	71.3
TN 15-5802-25	Det	3031.3	123.5	1.8	78.1	402.7	222.7	15.6	14.1	11.4	64.1	37.1	118.0	43.5	216.2	550.0	72.3
TN 15-5802-26	Det	3367.6	124.0	2.0	67.9	402.8	228.6	15.4	14.1	11.2	64.5	37.3	119.0	40.3	196.7	568.1	76.0
TN 15-5802-27	Det	3107.0	122.3	2.0	71.1	409.3	224.8	15.4	13.9	11.1	63.9	36.7	117.8	45.0	231.7	538.8	66.7
TN 15-5802-28	Det	3137.2	123.3	2.5	67.9	393.0	227.3	15.3	14.2	11.2	64.9	37.3	117.2	43.8	222.1	546.0	71.0
TN 15-5802-29	Det	3303.7	125.8	2.3	80.6	408.2	224.3	15.3	13.9	10.9	64.0	37.1	116.7	43.2	215.9	550.5	73.8
TN 15-5802-31	Ind	3530.6	129.8	3.0	126.4	408.8	224.4	15.5	14.1	11.1	64.4	36.9	114.7	47.5	253.3	518.4	66.0
TN 15-5802-32	Ind	3535.7	128.0	2.5	125.7	402.1	229.9	15.1	13.8	11.1	64.1	37.0	113.6	46.0	241.5	531.3	67.6
TN 15-5802-36	Ind	3444.9	127.5	2.5	124.5	390.7	233.2	15.2	14.1	11.6	64.2	37.2	120.6	42.7	221.7	544.3	70.7
TN 15-5802-33	Seg	3022.9	127.8	2.5	114.0	407.5	225.2	15.1	13.9	11.4	64.0	36.7	121.4	43.7	217.8	544.4	72.7
TN 15-5802-34	Seg	3613.0	128.0	3.3	109.5	403.4	225.2	15.3	13.9	11.2	64.2	37.1	116.4	44.9	221.0	549.2	68.6

† Indeterminate stem termination in Roundup Ready 1 NILs of TN15-5802
‡ Determinate stem termination in Roundup Ready 1 NILs of TN15-5802
§ Segregating stem termination in Roundup Ready 1 NILs of TN15-5802

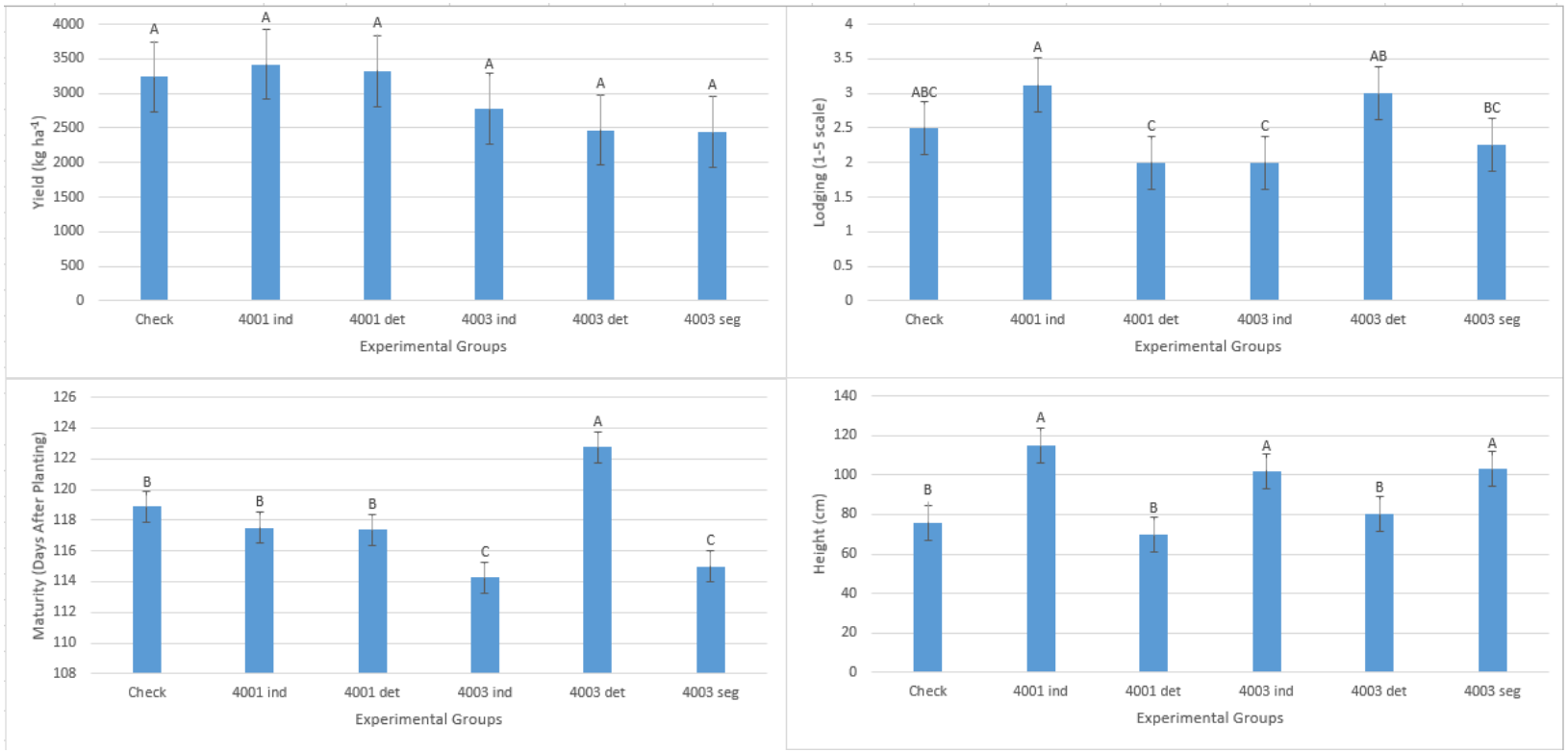

Figure 3.1 Mean agronomic trait values for 10 conventional $F_{8:10}$ lines in five experimental groups: i. TN14-4001 indeterminate (ind), ii. TN14-4001 determinate (det), iii. TN14-4003 ind, iv. TN14-4003 det, and v. TN14-4003 segregating (seg), with differences at the Dt_1 loci. The TN14-4001 experimental groups were composed of two ind and two det $F_{8:10}$ lines, while the TN14-4003 group was composed of four ind, one det, and one seg $F_{8:10}$ lines. The same letters within a graph denote the groups are not significantly different at $p < 0.05$.

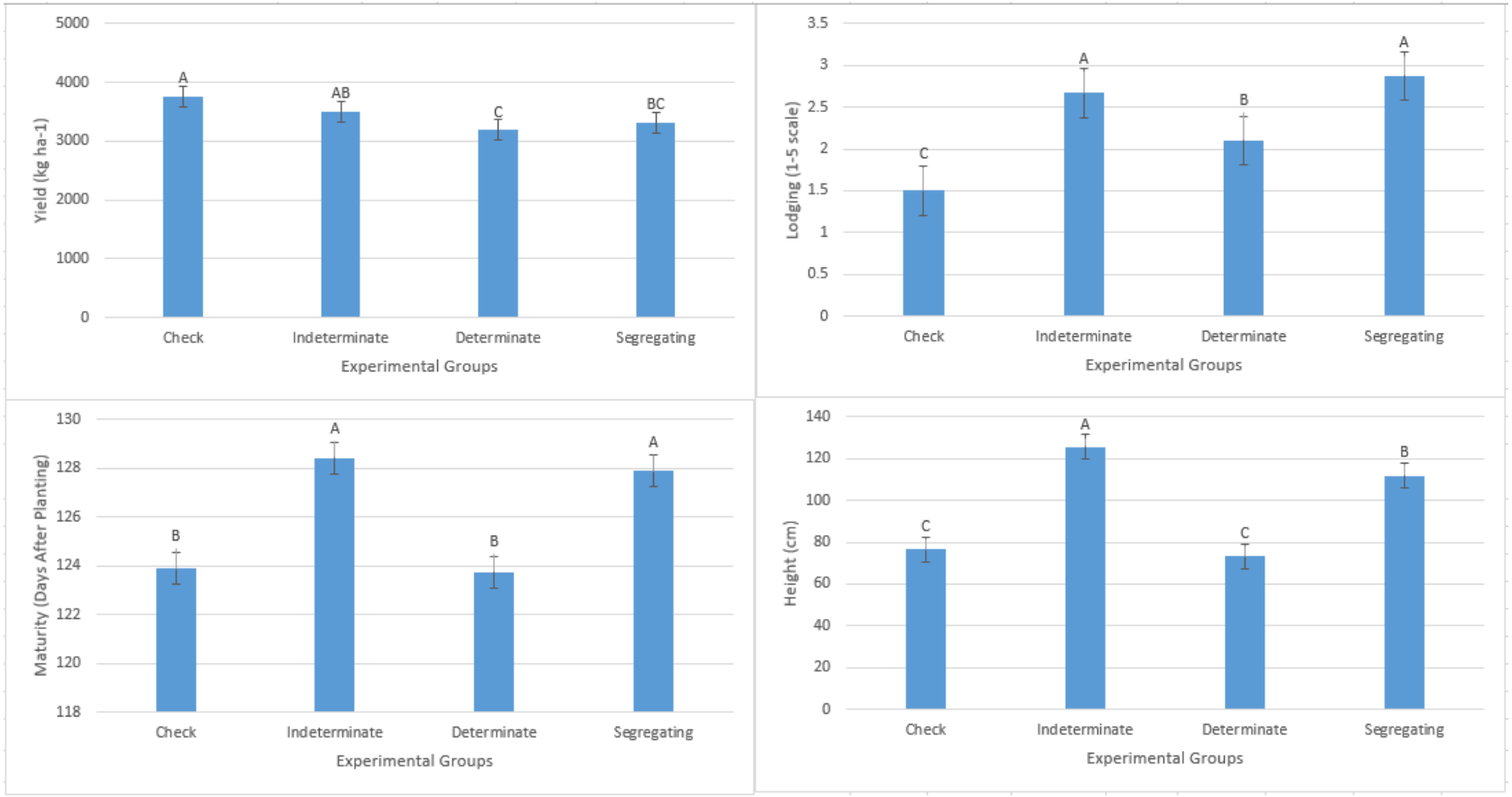

Figure 3.2 Mean agronomic trait values for three RR1 experimental groups: indeterminate, determinate, and segregating, from line TN15-5802, with differences at the Dt_1 loci. The TN15-5802 experimental groups were composed of three indeterminate, five determinate, and two segregating $F_{7:9}$ lines. The same letters within a graph denote the groups are not significantly different at $p < 0.05$.

CONCLUSION

A recent trend towards improving the quality of protein and oil in the grain market (Hurburgh et al., 1990) drives breeders to seek material with natural high protein alleles such as Danbaekkong (Dan). Wilson (2004) explains that soybeans with above average protein may not be profitable to grow due to an economic plateau which occurs around 48% protein in soymeal. Commercial production of soybeans with above average protein concentration is a difficult sell to farmers, due to the negative genetic correlation between protein and yield. In this study, yield was significantly lower in the Dan lines than in the wild type (WT) counterparts, supporting the aforementioned correlation. Farmers may not be willing to overlook the decrease in yield of Dan soybeans despite the increase of protein. While protein was successfully increased in the Dan lines, the quality of the protein was significantly reduced as all Dan lines averaged lower concentrations of the five tested amino acids when converted to g of amino acid per g of crude protein. An additional characteristic of stem termination was considered on lines from TN14-4001 with the same pedigree. All lines were homozygous Dan, however they segregated for stem termination and results showed significant differences between indeterminate and determinate lines only in height in lodging, as would be expected. No significant differences were observed between indeterminate and determinate groups in yield, protein, oil, or any of the five tested amino acids and fatty acids. Overall, the increase in protein concentration due to the high protein allele Dan in Tennessee environments, was not able to overcome the negative genetic correlation between protein and oil and protein and yield.

References

Hurburgh, C.R. Jr., T.J. Brumm, J.M. Guinn, and R.A. Hartwig. 1990. Protein and oil patterns in U.S. and world soybean markets. J. Am. Oil Chem. Soc. 67: 966-973.

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.

VITA

Mia J. Cunicelli was born in Pittsburgh, PA. She received a B.S. in Biology and a minor in Chemistry from the Virginia Military Institute. She hopes to pursue a career in plant breeding upon earning a M.S. degree in Plant Science and graduating from the University of Tennessee.