Evaluation of agronomic, seed quality traits, and transcript abundance in conventional high oleic soybeans with mutant FAD2-1A and FAD2-1B genes

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

Lauren Kathryn Richardson
May 2016

Copyright © 2016 by Lauren K. Richardson
All rights reserved.

ACKNOWLEDGEMENTS

The completion of this thesis project and graduate program would not have been possible without the constant encouragement and guidance of others. I would like to thank my advising professor, Dr. Vince Pantalone, for giving me the incredible opportunity to see into the world of plant breeding despite having no previous experience. Dr. Pantalone has allowed me to not only learn in the classroom, but to learn through hands on experience in the field and through numerous trips to professional meetings and workshops. He has a particular gift for enabling his students to succeed by showing confidence in their abilities and through his care for professional and personal well-being. I would also like to thank Dr. Tarek Hewezi for allowing me to work in his lab to gain new experience in molecular biology. I would also like to acknowledge Dr. Hem Bhandari and Dr. Feng Chen for serving as advising committee members and exceptional teachers.

Next, I would like to thank all members of the 'Bean Team' for helping me with this project every step of the way and for making graduate school such a positive experience. I would like to acknowledge Chris Smallwood and Rachel Fulton for their never ending supply of pep-talks and great patience. Also, I would like to thank the farm crew and fellow graduate students for being constant sources of help and encouragement.

Finally, I would like to thank my family, dear friends, and fiancé for being a never ending supply of love and support over the past two years.

ABSTRACT

Two soybean omega-6 fatty acid desaturase genes [FAD2-1A and FAD2-1B] are responsible for converting oleic acid into linoleic acid (Okuley *et al.*, 1994). Plant introductions [PI] 603452 and 283327 have naturally occurring mutations in FAD2-1A and FAD2-1B, respectively, which hinder the conversion of oleic acid to polyunsaturated fatty acids. The end result is more functional soybean oil with exceptionally high levels of oleic acid. The homozygous alleles of FAD2-1A and FAD2-1B were combined in a $BC_3F_{2:4}$ [backcross three, second filial generation selected, advanced to the fourth filial generation] population of forty-eight lines and were evaluated in Tennessee. Each line belonged to a double homozygous genotypic class: i) double wild-type [WT] (AABB), ii) mutant FAD2-1A with WT FAD2-1B (aaBB), iii) the converse (AAbb), and iv) double mutant (aabb). All genotypic groups yielded similarly to each other and to three high yielding checks ($p<0.05$). When one mutant allele was present the high oleic content was increased significantly (aaBB-34.4% and AAbb-26.2%) compared to AABB-21.3%. When both mutations were combined, aabb, the mean oleic content was significantly greater ($p<0.05$) than all other genotypic groups (80.4%). The range of oleic acid content across all twelve double mutant groups was 77.1% - 81.8%. It appears that the high oleic genotype has a significant effect on total oil, crude protein, and the concentration of five amino acids. The aabb group had significantly more oil (240 ± 1 g kg^{-1}) and crude protein (400 ± 7 g kg^{-1}) than all other groups. This increase in crude protein is consistent with the significantly lowered amino acid concentration ($p<0.05$) of lysine, methionine,

tryptophan, and threonine in the aabb group. Conversely, the high oleic group had the greatest concentration of cysteine ($p<0.05$). It appears that mutant FAD2-1 genes from PI 603452 and PI 283327 are a useful source for developing high oleic soybean lines in Tennessee with no yield drag. It will be beneficial to explore the effect of these mutant alleles on total oil and protein in the future.

TABLE OF CONTENTS

INTRODUCTION 1
Breeding High Oleic Soybeans 3
Objectives 7
References 8

CHAPTER I Evaluation of agronomic traits of 48 $BC_3F_{2:4}$ derived lines belonging to four double homozygous genotypic classes for the FAD2-1A and FAD2-1B loci 13
Abstract 14
Introduction 15
Materials and Methods 15
Plant Materials 15
DNA Extraction and SNP Genotyping 19
Agronomic Traits 20
Data Analysis 21
Results and Discussion 22
References 25
Appendix 27

CHAPTER II Evaluation of concentration and quality of seed oil and protein of 48 field tested $BC_3F_{2:4}$ soybeanLines 33
Abstract 34
Introduction 34
Materials and Methods 37
Gas Chromatography 37
Near Infrared Reflectance 38
Results and Discussion 38
Fatty Acid Concentrations 38
Oil, Protein, and Amino Acids 39
References 43
Appendix 46

CHAPTER III Evaluation of transcript abundance of FAD2-1A and FAD2-1B with different allelic combinations at different growth stages of soybean 56
Abstract 57
Introduction 57
Materials and Methods 59
Plant Growth and Sample Collection 59
RNA Extraction and Sample Preparation 60
Reverse Transcription Quantitative PCR 62
Results and Discussion 64
References 68

Appendix 70

CONCLUSION 72
References 73
VITA 74

LIST OF TABLES

Table 1.1 Yield (kg ha-1), maturity (DAP), height (cm), and lodging of 48 experimental lines, three high yielding checks, and recurrent parent, TN10-4037, averaged over three replications and three locations 31
Table 2.1 Fatty acid mean concentrations of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes 46
Table 2.2 Protein, oil, and amino acid concentration means of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes4..... 8
Table 2.3 Oil, protein, and fatty acid concentrations of 48 experimental lines, three high yielding checks, and recurrent parent, TN10-4037, averaged over three replications and three locations..... 49
Table 2.4 Ratio of mean grams of amino acids per mean grams of crude protein of 48 experimental lines, three high yielding checks, and recurrent parent, TN10-4037, averaged over three replications and three locations 53
Table 3.1 FAD2-1A and FAD2-1B genotype for experimental and control lines 70
Table 3.2 FAD2-1A mRNA fold change relative to FAD2-1B 70

LIST OF FIGURES

Figure 1.1 Seed yield means of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes 27
Figure 1.2 Height (cm) means of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes 28
Figure 1.3 Lodging means of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes 29
Figure 1.4 Days to maturity means of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes 30
Figure 2.1 Oleic concentration means of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes 46
Figure 2.2 Oleic concentration means of 12 $BC_3F_{2:4}$ lines with mutant FAD2-1A and FAD2-1B alleles 47
Figure 2.3 Ratio of mean grams of amino acids per mean grams of crude protein of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes 49
Figure 3.1 FAD2-1A mRNA fold change between embryo growth stages 71
Figure 3.2 FAD2-1B mRNA fold change between embryo growth stages 71

INTRODUCTION

In 2014 the United States produced and exported more soybeans than any other country at a value of $40.3 billion (Soy Stats, 2015). Soybean's global status as a leading cash crop is a consequence of its many uses. Humans directly consume a modest amount of the harvested seeds, but they are primarily used as a source of meal protein and oil. Soybean meal is a common component of animal feed, and soybean oil is used in cooking oils, processed foods, biodiesel, and industrial products such as lubricants and paints (North Carolina Soybean Producers Association, Inc., 2007).

Soybean oil is affordably used in a wide variety of products and accounts for 56% of world oil-seed production (Soy Stats, 2014). The typical fatty acid composition of soybean oil is 11% palmitic acid (16:0), 4% stearic acid (18:0), 24% oleic acid (18:1), 54% linoleic acid (18:2), and 7% linolenic acid (18:3) (Cober et al., 2009). This fatty acid profile touts some beneficial qualities. For example, soybean oil has a relatively low percentage of saturated fatty acids and a relatively high percentage of polyunsaturated fatty acids (PUFAs). Mozaffarian et al. (2010) reported that diets with reduced saturated fats replaced by PUFAs are associated with less coronary heart disease related events. However, a relatively high level of PUFAs does not optimize both health benefits and functional properties for many end users.

PUFAs show oxidative instability. Oxidative instability is the cause of oil degradation or rancidity. Oxidative instability increases with the number of double

bonds present in the fatty acid carbon chain. Double bonds react more readily with free radical oxygen and increase the rate of decomposition. Food manufacturers have used hydrogenation to reduce the amount of PUFAs and increase shelf life, but the hydrogenation process produces *trans* fats (Chaiyasit et al., 2007). For at least two decades, *trans* fats have been associated with increased risk of heart attack (Ascherio et al., 1997). In June 2015 the FDA made a final determination that partially hydrogenated oils (PHOs) are not “generally recognized as safe” (GRAS) for human consumption. Manufacturers have 3 years to comply with the ruling for total elimination of PHOs from food products (FDA, 2015).

As food manufacturers reformulate products to exclude PHOs there will remain a demand for oils that are relatively stable at room temperature. The monounsaturated fatty acid in soybean is oleic acid. The single double bond in oleic acid makes it less susceptible to degradation than PUFAs while preventing it from being categorized with health jeopardizing saturates. For producers to avoid unhealthy saturates, maintain oxidative stability, and eliminate partial hydrogenation, oils high in oleic acid content may become a superior option. The identified health benefits of consuming oleic fatty acids are numerous. Oleic acid is associated with lowered bad cholesterol in the blood stream, reduced risk of coronary heart disease, decreased blood pressure, suppression of tumor growth, and decreased insulin resistance (Kris-Etherton, 1999, Terés et al., 2008, Yamaki et al., 2005, Sala-Vila et al., 2011). The oxidative stability of oleic acid

has also demonstrated usefulness in food and industrial applications. In a test of frying stability and sensory acceptance, combination high oleic, low linolenic soybean oil produced non-fishy flavor and longer lasting potato chips (Warner and Gupta, 2005). In consideration of U.S. biodiesel production, Fallen *et al*., (2012) suggested that increased oleic acid concentration improved the performance of soybean based biofuels. Improvement of biodiesel production remains important after the renewal of the Farm Bill in 2014 which allocates $880 million dollars for energy programs for the next 5 years (USDA, 2014). U.S. biodiesel production in 2013 was approximately 1.36 trillion gallons compared to 14 million gallons in 2003 (U.S. EIA, 2014). The forecast for biodiesel in 2016 suggests production will hold relatively stable at 1.29 trillion gallons (U.S. EIA, 2015).

Breeding High Oleic Soybeans

Plant breeders recognized the importance of fatty acid composition in soybean oil over 50 years ago (Burton, *et al.,* 1994; Fehr, 2007). But, breeding for oil quality traits has been challenging. In 2003, Cahoon suggested that dramatic modification of oil composition is nearly impossible through conventional breeding methods. Nevertheless, continued breeding efforts are proving otherwise with the use of marker assisted selection.

Omega-6 fatty acid desaturase (FAD2) is an enzyme that converts omega-9 oleic acid into omega-6 linoleic precursors (Okuley *et al.*, 1994). There are at least four expressed oleate desaturase genes designated FAD2. Two of

these genes, FAD2-2 and FAD2-3, are constitutively active and referred to as “house-keeping” genes (Li *et al*., 2007). FAD2-1A and FAD2-1B are embryo specific and responsible for the majority of oleic acid accumulated in the mature seed (Heppard *et al.*, 1996, Tang *et al.,* 2005).

Several sources of FAD2-1A and FAD2-1B mutants associated with increased oleic acid content have been identified and characterized. An x-ray irradiated line, M23, has a complete deletion of FAD2-1A and produces a mean oleic content of 52.4% (Sandhu *et al.* 2007; Takagi and Rahman, 1996). Line KK21 is an x-ray irradiated line with a single nucleotide deletion and frame shift mutation in the FAD2-1A locus and a mean oleic content of 47.2% (Anai, *et al.,* 2008). Line 17D is an ethyl methanesulfonate (EMS) mutagenized line with a novel S117N allele in FAD2-1A and mean oleic content of approximately 33.5% (Dierking and Bilyeu, 2009). FAD2-1B mutant alleles were generated through reverse genetics in lines E11 and B12 to produce mean oleic content of 43.2% and 29.2%, respectively (Hoshino *et al.,* 2010). The difference in oleic acid accumulation between mutant lines demonstrates that some mutations are more effective at disrupting the desaturase pathway. In order to achieve stable high oleic accumulation in the field, the best combination of mutant FAD2-1A and FAD2-1B alleles is desired. Sources of natural mutants for FAD2-1A and FAD2-1B have been discovered housed within the USDA National Plant Germplasm System. PI 603452 has a single deletion in the FAD2-1A gene. This deletion results in a frame shift and premature termination of peptide translation (Pham *et*

al., 2011). PI 283327 has a missense mutation that causes a highly conserved proline residue to be replaced by an arginine residue in the amino acid sequence and thus inhibits the proper functioning of the omega-6 fatty acid desaturase enzyme. This allele is designated P137R (Pham *et al.,* 2010).

All characterized sources of mutant FAD2-1A and FAD2-1B soybeans show increased accumulation in oleic acid when compared to wild-type controls. However, none of these sources have exceeded mid-oleic concentrations with a single gene mutation. In order to achieve high oleic content (>70%), FAD2-1A and FAD2-1B mutants have been combined using a candidate gene molecular breeding approach. Hoshino *et al.,* (2010) combined FAD2-1B mutant lines Ell and B12 with the FAD2-1A mutant sources KK21 and M23 to generate multiple lines with >70% oleic acid. In 2010, Pham *et al.* working out of the University of Missouri, observed transgressive segregation for oleic acid accumulation in two populations with double mutants. The double homozygous mutant lines developed from a cross between M23 and PI 283327 had an average oleic content of 82.2% with a standard deviation of 1.2% across three environments. The average oleic content was 2-4% lower in Columbia, MO, the coldest growing environment. While this suggests some environmental instability, the phenotypic variation within each location was narrow. Pham *et al. (*2011) generated three lines of soybeans homozygous for the mutant alleles from PI 603452 and PI 283327 with oleic content greater than 82%. The report by Pham *et al.* (2011)

suggests that the oleic acid trait still needs to be tested for agronomic traits across various environments.

A recent analysis by La *et al.,* 2014 of seed oil, protein concentration, and yield in high oleic soybeans containing FAD2 missense mutations from 17D and PI 283327 showed no significant decrease in yield and an increase in both seed oil and protein concentration. The increase in both total oil and protein was a surprising result considering these characters are usually negatively correlated (La *et al.,* 2014). If high oleic soybeans are capable of yielding equivalently with top commodity cultivars and have the added advantage of increased total oil and protein, they may become more valuable to farmers and processors.

In order for soybean oil to remain competitive with oils high in monounsaturated fat (canola, sunflower, olive, etc.) in the new age of zero hydrogenation, high oleic soybeans could become the new commodity standard. The United Soybean Board has the long-term increase in soybean profitability as the driving force for improving oil quality. And, it appears that end users are also looking forward to high oleic soybeans making a permanent mark on the industry. The USB reported that soybean processor, Perdue Agribusiness, in Salisbury, Maryland is constructing multi-million dollar storage tanks specifically for high oleic soybeans (Soy Innovation, 2014).

Soybeans account for 30.7% of total crop value in Tennessee and were valued at $785 million in 2014 (Soy Stats, 2015). As soybean growers in other parts of the U.S. move forward in production of high oleic soybeans, it is

important that Tennessee farmers have access to high oleic cultivars with strong yield potential and appropriate adapted maturity. In fact, the USB claims that the demand for high oleic soybeans is there for large companies, but the limiting factor is having a consistent, abundant supply (USB, 2015). This research aims to further Tennessee farmers' role in producing high oleic soybeans through several objectives:

Objectives

1. Evaluate agronomic traits of 48 $BC_3F_{2:4}$ lines belonging to four double homozygous genotypic classes for the FAD2-1A and FAD2-1B loci grown at 3 locations in Tennessee. Agronomic factors include yield, height, lodging, and days to maturity.
2. Evaluate the fatty acid profiles and the total oil content of 48 field tested lines using gas chromatography and near infrared spectroscopy, respectively.
3. Evaluate the protein of 48 field tested lines by measuring crude protein and amino acid content using near infrared spectroscopy.
4. Measure transcript abundance of FAD2-1A and FAD2-1B genes with four double homozygous allelic combinations in three embryo growth stages using reverse transcription quantitative PCR (RT-qPCR).

References

Anai, T., T. Yamada, R. Hideshima, T. Kinoshita, S.M. Rahman, Y. Takagi. 2008. Two high-oleic-acid soybean mutants, M23 and KK21, have disrupted microsomal omega-6 fatty acid desaturase, encoded by GmFAD2-1a. Breeding Sci 58:447-452.

Ascherio, A. and W. C. Willett. 1997. Health effects of trans fatty acids. Am. J. Clin. Nutr. 66:10065-10105.

Burton, J.W., R.F. Wilson, and C.A. Brim. 1994. Registration of N79-2077-12 and N87-2122-4, two soybean germplasm lines with reduced palmitic acid in seed oil. Crop Sci. 34:313.

Cahoon, E.B. 2003. Genetic enhancement of soybean oil for industrial uses: prospects and challenges. AgBioForum. 6:11-13.

Chaiyasit, W., R. Elias, J.D. Mcclements, E. Decker. 2007. Role of physical structure in bulk oils on lipid oxidation. Crit. Rev. Food Sci. 47:299-317.

Cober, E.R., S.R. Cianzio, V.R. Pantalone, and I. Rajcan. 2009. Soybean. In: J. Vollman and I. Rajcan, editors, Oil crops: Handbook of plant breeding, volume 4. Springer Science + Business Media LLC. p. 57-90.

Dierking, E.C. and K.D., Bilyeu. 2009. New sources of soybean seed meal and oil composition traits identified through TILLING. BMC Plant Biol. 9:89

Fallen, B. K., Rainey, C.E. Sams, D.A. Kopsell, and V.R. Pantalone. 2012. Evaluation of agronomic and seed characteristics in elevated oleic acid soybean lines in the south-eastern US. J. Am. Oil. Chem. Soc. 89:1333-1343.

FDA (Food and Drug Administration). Department of Health and Human Services. 2015. FDAhttps://www.federalregister.gov/articles/2015/06/17/2015-14883/final-determination-regarding-partially-hydrogenated-oils

Fehr, W.R. 2007. Breeding for modified fatty acid composition in soybean. Crop Sci. 47:572-587.

Kinney, A.J. and G.H. Miao. 1996. Developmental and growth temperature regulation of two different microsomal [omega]-6 desaturase genes in soybeans. Plant Physiol. 110:311-319.

Hoshino, T. 2010. Novel GmFAD2-1b mutant alleles created by reverse genetics induce marked elevation of oleic acid content in soybean seeds in combination with GmFAD2-1a mutant alleles. Breeding Sci. 60:419-425.

Kris-Etherton, P. 1999. AHA science advisory. Monounsaturated fatty acids and risk of cardiovascular disease. American Heart Association. Nutrition Committee. J. Nutr. 129: 2280-2284.

La, T.C., S.M. Pathan, T. Vuong, J. Lee, A.M. Scaboo, J.R. Smith, A.M. Gillen, J. Gillman, M.R. Ellersieck, H.T. Nguyen, and J.G. Shannon. 2014. Effect of high-oleic acid soybean on seed oil, protein concentration, and yield. Crop Sci. 54:2054-2062.

Li, L., X. Wang, J. Gai, and D. Yu. 2007. Molecular cloning and characterization of a novel microsomal oleate desaturase gene from soybean. Plant Physiol. 164:1516-1526.

Mozaffarian, D., R. Micha, and S. Wallace. 2010. Effects on coronary heart disease of increasing polyunsaturated fat in place of saturated fat: a systematic review and meta-analysis of randomized controlled trials. PLoS Medicine 7: E1000252.

North Carolina Soybean Producers Association, Inc. 2007. How soybeans are used. <http://www.ncsoy.org/ABOUT-SOYBEANS/Uses-of-Soybeans.aspx>

Okuley, J., J. Lightner, K. Feldmann, N. Yadav, E. Lark, J. Browse. 1994. Arabidopsis FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. The Plant Cell. 6:147-158.

Pham, A., J. Lee, J.G. Shannon, and K.D. Bilyeu. 2010. Mutant alleles of FAD2-1a and FAD2-1b combine to produce soybeans with the high oleic acid seed oil trait. BMC Plant Biol. 10:195.

Pham, A., J. Lee, J. Shannon, K. Bilyeu. 2011. A novel FAD2-1A allele in a soybean plant introduction offers an alternate means to produce soybean seed oil with 85% oleic acid content. Theor. Appl. Genet. 123:793-802.

Sala-Vila, A., M. Cofan, R. Mateo-Gallego, A. Cenarro, F. Civeira, E. Ortega, E. Ros. 2011. Inverse association between serum phospholipid oleic acid and insulin resistance in subjects with primary dyslipidaemia. Clin Nutr. 30:590-593

Sandhu, D., J.L. Alt, C.W. Scherder, W.R. Fehr, M.K. Bhattacharyya. 2007. Enhanced oleic acid content in the soybean mutant M23 is associated with

the deletion in the Fad2-1a gene encoding a fatty acid desaturase. J Am Oil Chem Soc. 84:229-235.

Soy Innovation. 2014. Processor shows long-term commitment to high oleic. United Soybean Board. http://soyinnovation.com/processor-shows-long-term-commitment-high-oleic/.

Soy Stats. 2014. A Reference Guide to Important Soybean Facts and Figures. The American Soybean Association. http://soystats.com/wp-content/uploads/SoyStats_2014.pdf.

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

Takagi, Y. and S.M. Rahman. 1996. Inheritance of high oleic acid content in the seed oil of soybean mutant M23. Theor Appl Genet. 92:179-182.

Tang, G.Q., W.P. Novitzky, H.C. Griffin, R.E. Huber SC Dewey. 2005. Oleate desaturase enzymes of soybean: evidence of regulation through differential stability and phosphorylation. Plant J. 44:433-446.

Terés, S., G. Barceló-Coblijn, M. Benet, R. Álvarez, R. Bressani, J.E. Halver, P.V. Escribá. 2008. Oleic acid content is responsible for the reduction in blood pressure induced by olive oil. Proc. Natl. Acad. Sci. U.S.A. 105:13811-13816.

United Soybean Board. 2015. High oleic soybeans. http://unitedsoybean.org/wp-content/uploads/50780-1_HO_NatlReachPromo_Advertorial_FNL_HR_March15update_HR2.pdf

USDA (United States Department of Agriculture). 2014. 2014 farm bill highlights. http://www.usda.gov/wps/portal/usda/usdahome?navid=farmbill.

USEIA (United States Energy Information Administration). 2014. Short-term energy outlook: renewables and CO_2 emissions. http://www.eia.gov/forecasts/steo/report/renew_co2.cfm.

USEIA (United States Energy Information Administration). 2015. Frequently Asked Questions: How much biodiesel is produced, imported, exported, and consumed in the United States? http://www.eia.gov/tools/faqs/faq.cfm?id=927&t=4.

Warner, K and M. Gupta. 2005. Potato chip quality and frying oil stability of high oleic acid soybean oil. J. Food Sci. 70:6

Yamaki, T., I. Nagamine, K. Fukumoto, T. Yano, M. Miyahari, H. Sakurai. 2005. High oleic peanut oil modulates promotion stage in lung tumorigenesis of mice treated with methyl nitrosourea. Food Sci. Technol. Res. 11:231-235.

CHAPTER I
EVALUATION OF AGRONOMIC TRAITS OF 48 $BC_3F_{2:4}$ DERIVED LINES BELONGING TO FOUR DOUBLE HOMOZYGOUS GENOTYPIC CLASSES FOR THE FAD2-1A AND FAD2-1B LOCI

Abstract

Natural mutations in the FAD2-1A and FAD2-1B loci were combined to produce high oleic soybean lines with a genetic background well adapted to Tennessee in order to evaluate the trait source effect on seed yield and other agronomic traits. A backcross population was developed with twelve lines per each of four genotypic classes based on allele type at the FAD2-1A and FAD2-1B loci. These classes were AABB, aaBB, AAbb, and aabb, where wild-type alleles were represented by capital letters and mutant alleles represented by lower-case letters. Forty-eight experimental lines, the recurrent parent, and three high-yielding maturity checks were planted in three replications at three locations across Tennessee. There were no significant seed yield differences between genotypic groups nor the recurrent parent or check group ($p < 0.05$). The mean yields for the high oleic group (aabb), wild-type group (AABB), check group, and recurrent parent were: 3508 kg^{-1} ha, 3513 kg^{-1} ha, 3536 kg^{-1} ha, and 3671 kg^{-1} ha respectively. This does not suggest that every line in the population yielded similarly. There were significant differences within each genotypic group for measured seed yield, height, lodging, and days to maturity. The check cultivars were less lodged than all experimental genotypic groups suggesting that the recurrent parent used in the experimental population contributed to increased lodging. All experimental groups matured within 115 and 116 days after planting which was similar to two check lines LD06-7620 and LD00-2817 with relative maturities of 4.2 and 4.5, respectively. Based on overall agronomic performance, it appears that the mutant FAD2-1A and FAD2-1B sources PI 603452 and PI

283327, respectively, are useful for developing high oleic soybean cultivars for Tennessee production.

Introduction

The high oleic trait will only be valuable to farmers if it comes in a package of desirable agronomic traits. Seed yield and days to maturity are two major factors growers use to select cultivars that will best suit their unique operations. Plant height and lodging are also factors that can influence seed yield performance. All of these factors come together in every cultivar of soybean, and choosing which cultivars to plant each year is no small task to farmers. Cultivar selection can be the management factor responsible for the greatest yield difference potential (Mueller and Elmore, 2014). Palle Pederson with Iowa State University Extension suggests that high-yield and yield stability should be the first and foremost factors considered by farmers because cultivar selection is ultimately about risk management. This experiment tests conventional high oleic soybeans for the potential to become marketable cultivars.

Materials and Methods

Plant Materials

The experimental population was developed from a cross between TN10-4037 (female) and breeding material from North Carolina State University, 'NC-Raleigh' x ('NC-Burton' x (PI 603452 x PI 283327)). Line TN10-4037 is a reselection made within the University of Tennessee soybean breeding program from TN05-3027, a line derived from a high-diversity cross made by USDA-ARS

at the University of Illinois at Urbana-Champaign. NC-Burton is a determinant maturity group V cultivar released in 2012. It is resistant to stem canker, frogeye leaf spot, and bacterial pustule (Cardinal *et al.*, 2012). NC-Raleigh, registered in 2006, is a determinant maturity group VII cultivar with resistance to soybean mosaic virus, stem canker, frogeye leaf spot, and bacterial pustule (Burton *et al.,* 2006). PI 603452 and PI 283327 are plant introductions containing the FAD2-1A and FAD2-1B alleles, respectively, responsible for increased oleic acid content. PI 603452 has a single nucleotide deletion in FAD2-1A identified as a natural source of increased oleic accumulation (Pham *et al.,* 2011). PI 283327 has an identified missense mutation in FAD2-1B responsible for increased oleic content (Pham *et al.,* 2010).

The original cross between TN10-4037 and NC-Raleigh x (NC-Burton x (PI 603452 x PI 283327)) was made at the East Tennessee Research and Education Center (ETREC) in Knoxville, TN. A backcross strategy was employed in attempt to capture the unique FAD2-1 alleles from both plant introductions in the better adapted genetic background of TN10-4037. The F_1 progeny were backcrossed to TN10-4037 to produce BC_1F_1 seeds. BC_1F_1 individuals were subsequently backcrossed to TN10-4037 two more times to form BC_3F_1 seeds with approximately 93.75% expected genome commonality with recurrent parent TN10-4037. BC_3F_1 seeds were grown at a nursery in Isabela, Puerto Rico in January 2014. The harvested BC_3F_2 seeds were hand sown at ETREC on June

27, 2014 into four, 6 m rows at a rate of 1 seed per 7.6 cm to total 80 seeds per row.

After flowering, each plant was given a DNA tag number from 1 to 426, and a leaf sample was collected. DNA from each leaf sample was preserved at room temperature by pressing leaf samples onto FTA Plantsaver Cards (GE Healthcare, Little Chalfont, Buckinghamshire, United Kingdom). The genotype of each individual plant at the FAD2-1A and FAD2-1B loci was determined using a single nucleotide polymorphism (SNP) genotyping assay (Shi, *et al.,* 2015). Details on SNP genotyping follow below. All individual plants were harvested and classified in 1 of 9 genotypic classes based on FAD2-1A and FAD2-1B SNP melting curve data. The assumption of independent assortment at the FAD2-1A and FAD2-1B loci was used to test the observed ratio of genotypes against the expected number. The observed ratio of genotypes passed the chi-square test at 0.05 level probability and confirmed there is no difference between the number of observed and expected genotypes. The plant height was recorded from the soil to the tip of the plant for twelve representative individuals in the four genotypic classes homozygous at each locus. From this time forward each of these double homozygous classes will be referred to as AABB, aaBB, AAbb, and aabb, where 'A' represents FAD2-1A, 'B' represents FAD2-1B, uppercase represents the wild-type alleles, and lowercase represents the mutant alleles of interest. BC_3F_3 seeds were harvested from the representative individual BC_3F_2 plants.

The 12 representative individuals in each of the four double homozygous classes form a population of 48 lines. Forty seeds from each line were planted December 2014 in the lighted winter nursery at the USDA Tropical Agriculture Research Station in Isabela, Puerto Rico. Two seeds were planted on each of 20 hill plots spaced 15.2 cm apart with 71.0 cm fallow space separating each line. A wooden stake was placed at the front of each line with designations LR-1, LR-2…LR-48. The purpose of planting $BC_3F_{2:3}$ seeds in winter nursery was to increase the seed stock needed for a 2015 multi-location yield trial. Each line was bulk harvested and sent back to Knoxville, TN.

Bulk harvested $BC_3F_{2:4}$ seeds were planted in 3 locations for field evaluation in the spring of 2015. Each of the 48 lines, the recurrent parent, TN10-4037, and 3 high yielding checks with expected similar relative days to maturity, LD06-7620, LD00-2817, and LD07-3309, were grown in 2 row plots, with 3 replications at each of three locations in Tennessee: Knoxville, Springfield, and Milan. These locations represent East, Middle, and West Tennessee soybean growing areas. A randomized complete block design was used at all locations to control for potential field variation. In Knoxville and Springfield, seeds were planted in 6.1 m plots on May 13, 2015 and May 21, 2015, respectively. In Milan, seeds were planted in 9.1 m plots on June 8, 2015. The longer plot lengths used at the Milan, Tennessee location were taken into account when calculating yield data for individual plots.

All 2-row plots were combine harvested at Knoxville, Springfield, and Milan on September 23, October 7, and October 15, 2015, respectively. The combines used for harvest recorded weight (lbs.) and relative moisture (%) of each plot. This data was used with plot length to calculate yield at 13% moisture basis of each plot in bu/acre and converted to kg/ha.

DNA Extraction and SNP Genotyping

DNA from 422 BC_3F_2 single plant tissue samples was eluted using the Whatman Sample Purification Protocol (GE Healthcare, Little Chalfont, Buckinghamshire, United Kingdom). A circle 1.20 mm in diameter was punched from each tissue press and placed in an individual well of a 96-well Roche Light Cycler (Basel, Switzerland) PCR plate. Using a multi-channel electronic pipet, 120 µl of Whatman FTA purification reagent was added to each well and allowed to sit for 5 minutes. The reagent was removed from the top of the punches using new sterile pipet tips. This process was repeated once more with the purification reagent and twice more with 120 µl of TE buffer. Following wash, samples were prepared for SNP genotyping following a protocol developed by the lab of Dr. Kristin Bilyeu at USDA-ARS Columbia, MO. The SNP procedure was performed twice for each sample in order to determine the genotype at both the FAD2-1A (PI 603452) and FAD2-1B (PI 283327) loci. Primer mixes for FAD2-1A and FAD2-1B were prepared with 86 µL water, 4.0 µl forward primer, and 10.0 µL reverse primer for each gene. The master mix for both FAD2-1A and FAD2-1B, per reaction, included 7.8 µl water, 10.0 µl 2x Titanium Taq, 2.0 µl primer mix

(described above), and 0.2 µl locus specific probe. Each reaction cell underwent 40 cycles of amplification before genotype characterization using the sequence-specific fluorophore-labeled oligonucleotide Simple Probe. If there was a SNP in the binding region of the probe-target hybrid, the probe became destabilized and denatured at a lower temperature than a hybrid and the perfectly matching sequence. Each plant sample was thus categorized as homozygous wild-type, heterozygous, or homozygous mutant at each of the FAD2-1A and FAD2-1B loci. This classification was the source of the 9 genotypic classes that all individual plants were placed into.

Agronomic Traits

During flowering, all plots were scored for white or purple flowers and off types were pulled for the purpose of line purification. Because the recurrent parent, TN10-4037, has white flowers it was expected that every experimental plot would also have white flowers. This was true; however, purple flowering check cultivars in other plots allowed confirmation of correct planting order.

When plots started to show signs of senescence, all of the plots were walked through every 2-3 days to record maturity notes. When a plot appeared to be 95% mature, based on pod and stem color, the date was recorded. This date marked how many days from planting it took an individual plot to reach full maturity. Full maturity marks readiness for harvest. When a maturity date was recorded, the plot was also scored for lodging and height. Each plot was given a lodging score of whole numbers from 1 to 5 where 1 is completely upright and 5

is completely lodged on the ground. The height was taken by measuring an average looking individual with a metal yard stick to the nearest inch and later converted to centimeters. To minimize data variation, lodging and height measurements for all locations were taken by the graduate student investigator. The graduate student could not be in 3 locations in the state for several weeks taking maturity notes on all plots. However, at each location only one individual recorded the date of full maturity to minimize variation at that location.

Data Analysis

Agronomic and seed quality data were analyzed using a mixed model analysis of variance (ANOVA) via the GLIMMIX procedure of SAS/STAT® software version 9.4. (SAS Institute Inc., 2013, Cary, NC, USA). The fixed effects were the genotypic classes (AABB, aaBB, AAbb, aabb) and lines nested within genotypic classes. Each line was given LR- as a prefix followed by a number designation (LR-1…LR-48). The random effects were the locations, G x E or location by line(genotypic class) interaction, and replications nested within location. Fisher's Least Significant Difference (LSD) method was used to separate means for genotypic classes and lines nested within genotypic classes at the 0.05 significance level. Pearson correlations among traits were determined using "Hmisc: Harrell Miscellaneous" package of R software (Harrell, 2015 and R Core Team, 2015).

Results and Discussion

Seed yield was averaged across all lines within each of the four genotypic classes. The seed yield for three elite checks were averaged together and the mean recurrent parent yield was determined. There were no mean yield differences between any genotypic classes including elite checks ($p<0.05$) (Figure 1.1). This finding suggests that there is no yield drag associated with the high oleic trait sourced from PI 603452 and PI 283324. However, within each genotypic class there were significant ($p<0.05$) yield differences among individual lines. The high oleic group (aabb) had a yield range of 2809-4055 kg ha^{-1} and the normal oleic group (AABB) ranged from 3248-3971 kg ha^{-1}. The check group ranged from 3362-3638 kg ha^{-1} and the recurrent parent, TN10-4037, had a mean yield of 3671 kg ha^{-1}.

The ability of high oleic soybeans to have high yield is dependent on the source of mutations in the FAD2-1 loci and the genetic background. This is true for transgenic and conventional soybeans. Lines containing transgenic event 335-13 with >85% oleic acid showed no reduction in yield under irrigated and non-irrigated production (Graef *et al.,* 2009). Conversely, HO lines from several populations containing transgene DP-305423-1 yielded lower than their normal oleic counterparts (Spear *et al.,* 2013). In a test of conventional HO lines, La *et al.* (2014) confirmed that FAD2-1A mutant source M23 is associated with significant yield reduction, while FAD2-1A mutant source 17D showed no effect on yield in two populations. It is likely that the large deletion in M23 has an effect on surrounding genes, while point mutations like that in 17D do not (Hoshino *et*

al., 2010). The mutations in FAD2-1A and FAD2-1B genes from PI 603452 and PI 283327 used in the experimental population are also point mutations. It is likely that the nature of these mutations make it possible to generate high yielding high oleic soybeans. On average, all lines yielded significantly more in Knoxville compared to Springfield and Milan. However, there was no GxE interaction.

Height differences were not dependent upon location, but the mid-oleic, aaBB, group was taller than any other experimental group (Figure 1.2). The check group was significantly shorter than all groups, and the recurrent parent, TN10-4037, was significantly taller than all experimental and the check groups (Figure 1.2). There was no correlation between height and seed yield ($r=0.14$, $p > 0.05$) nor lodging and seed yield ($r = -0.07$, $p > 0.05$). However, height and lodging had a strong and significant positive relationship ($r = 0.90$, $p < 0.05$). Figure 1.3 is a bar graph representing the mean lodging scores for each genotypic group. For both height and lodging the check group had the smallest value. Genotypic groups AAbb and aabb were more lodged than the check group, but less lodged than groups aaBB and AABB. It is important to consider lodging and yield for selection. In the 2015 Soybean Variety Performance Tests in Tennessee twenty MG IV lines were tested in five environments and the average lodging score was 1.8 (Allen *et al.,* 2015). Of these twenty lines, seven of the eight lines with lodging scores greater than or equal to two were also ranked in the bottom ten in terms of mean yield.

Days to maturity were different at all three locations. On average all lines matured in Milan 5 days before all lines in Springfield and 10 days before all lines in Knoxville. Days to maturity was the only tested agronomic trait with significant G x E interaction ($p<0.0001$). In Tennessee, early to mid-maturity group IV cultivars are increasingly being grown. Three maturity checks were planted along with the recurrent parent for comparison. The maturity checks and designated relative maturities are: LD00-3309 (4.0), LD06-7620 (4.2), and LD00-2817 (4.5). These checks, on average, matured in 114, 115, and 116 days after planting (DAP), respectively. The recurrent parent, TN10-4037 matured by 116 DAP on average. There were significant differences detected between genotypic groups ($p<0.05$) (Figure 1.4). When comparing genotypic groups, the AABB and AAbb groups matured earlier than the aaBB and aabb groups. The presence of the mutant FAD2-1A allele may prolong days to maturity. Each experimental genotypic class matured by 115 or 116 DAP which is similar to the recurrent parent and check lines LD06-7620 (4.2) and LD00-2817 (4.5). There was a range of 6 days for individual lines within each genotypic class to reach full maturity.

References

Allen, F.L., V.R. Sykes, V.G. Benelli, A.J. Ashworth, R.H. Blair, A.T. McClure, H. Kelly. 2015. Soybean variety performance tests in Tennessee. UT Agronomic Variety Trials. http://varietytrials.tennessee.edu/soybean.htm

Burton, J.W., T.E. Carter, M.O. Fountain, D.T. Bowman. 2006. Registration of 'NC-Raleigh' Soybean. Crop Sci. 46:2710-2711.

Cardinal, A.J., S. Wang, D.T. Bowman, V.R. Pantalone. 2012. Registration of 'NC-Burton' soybean. J. Plant Regist. 6:146-149.

Frank E Harrell Jr, with contributions from Charles Dupont and many others. (2015). Hmisc: Harrell Miscellaneous. R package version 3.17-1. https://CRAN.R-project.org/package=Hmisc

Graef, G., B.J. Lavallee, P. Tenopir, M. Tat, B. Schweiger, A.J. Kinney, J.H. Van Gerpen, T.E. Clemente. 2009. A high-oleic-acid and low-palmitic-acid soybean: agronomic performance and evaluation as a feedstock for biodiesel. Plant Biotech. J. 7:411-421.

Hoshino, T. 2010. Novel GmFAD2-1b mutant alleles created by reverse genetics induce marked elevation of oleic acid content in soybean seeds in combination with GmFAD2-1a mutant alleles. Breeding Sci. 60:419-425.

La, T.C., S.M. Pathan, T. Vuong, J. Lee, A.M. Scaboo, J.R. Smith, A.M. Gillen, J. Gillman, M.R. Ellersieck, H.T. Nguyen, and J.G. Shannon. 2014. Effect of high-oleic acid soybean on seed oil, protein concentration, and yield. Crop Sci. 54:2054-2062.

Mueller, N. and R. Elmore. 2014. Why soybean variety is critical to improving farm yield. University of Nebraska-Lincoln Extension Cropwatch. http://cropwatch.unl.edu/why-soybean-variety-critical-improving-farm-yield

Pederson, P. Variety selection. Iowa State University Extension. http://extension.agron.iastate.edu/soybean/documents/VarietySelection_001.pdf

Pham, A., J. Lee, J.G. Shannon, and K.D. Bilyeu. 2010. Mutant alleles of FAD2-1a and FAD2-1b combine to produce soybeans with the high oleic acid seed oil trait. BMC Plant Biol. 10:195.

Pham, A., J. Lee, J. Shannon, K. Bilyeu. 2011. A novel FAD2-1A allele in a soybean plant introduction offers an alternate means to produce soybean seed oil with 85% oleic acid content. Theor. Appl. Genet. 123:793-802.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Shi, Z., N. Bachleda, A.T. Pham, K. Bilyeu, G. Shannon, H. Nguyen, and Z. Li. 2015. High-throughput and functional SNP detection assays for oleic and linolenic acids in soybean. Mol. Breeding. 35:1-10.

Spear, J.D., W.R. Fehr, and S.R. Schnebly. 2013. Agronomic and seed traits of soybean lines containing the high-oleate transgene DP-305423-1. Crop Sci. 53:888-893.

Appendix

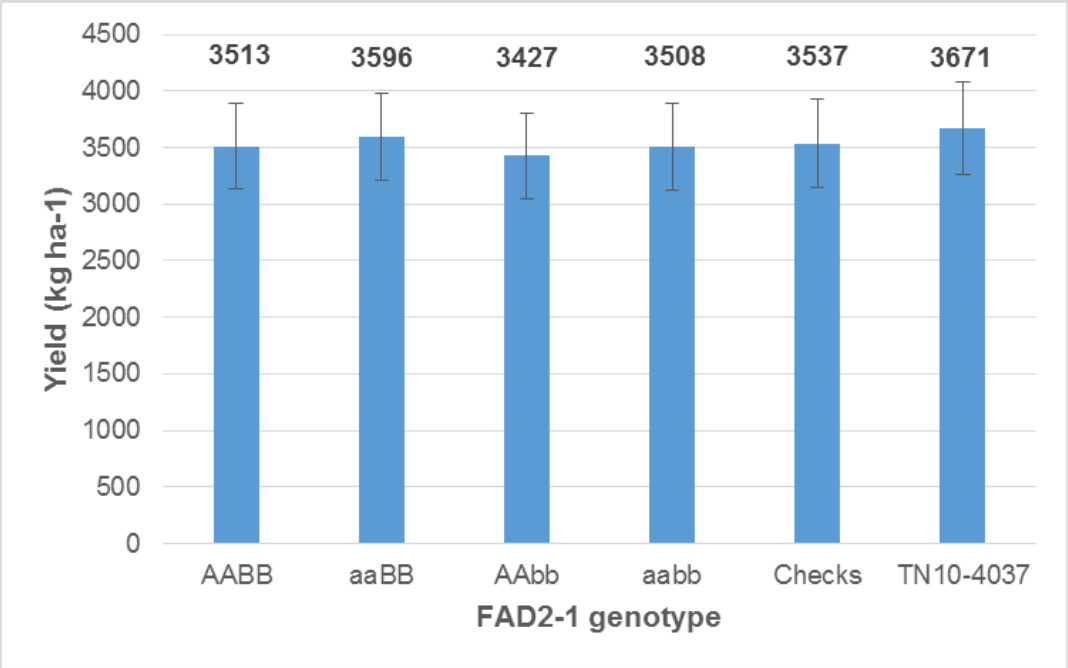


Figure 1.1 Seed yield means of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes. The error bars represent the standard error for each mean value. There were no significant differences in yield between the experimental groups nor between the experimental groups and the checks and recurrent parent ($p < 0.05$).

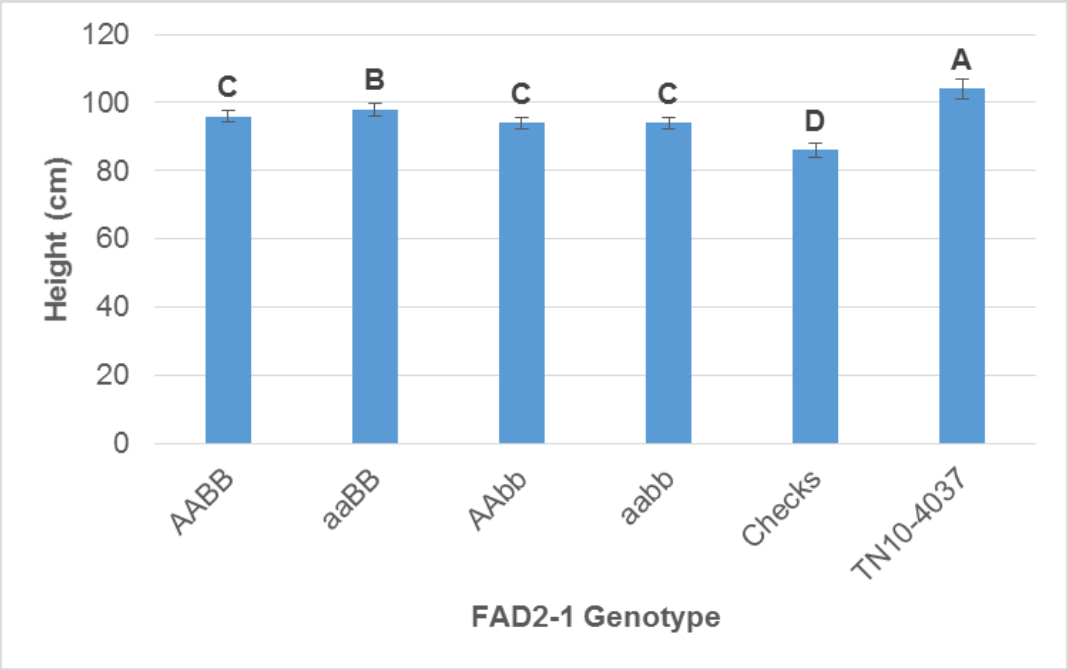


Figure 1.2 Height (cm) means of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes. The error bars represent the standard error for each mean value. Values with the same letter designation are not statistically different ($p < 0.05$)

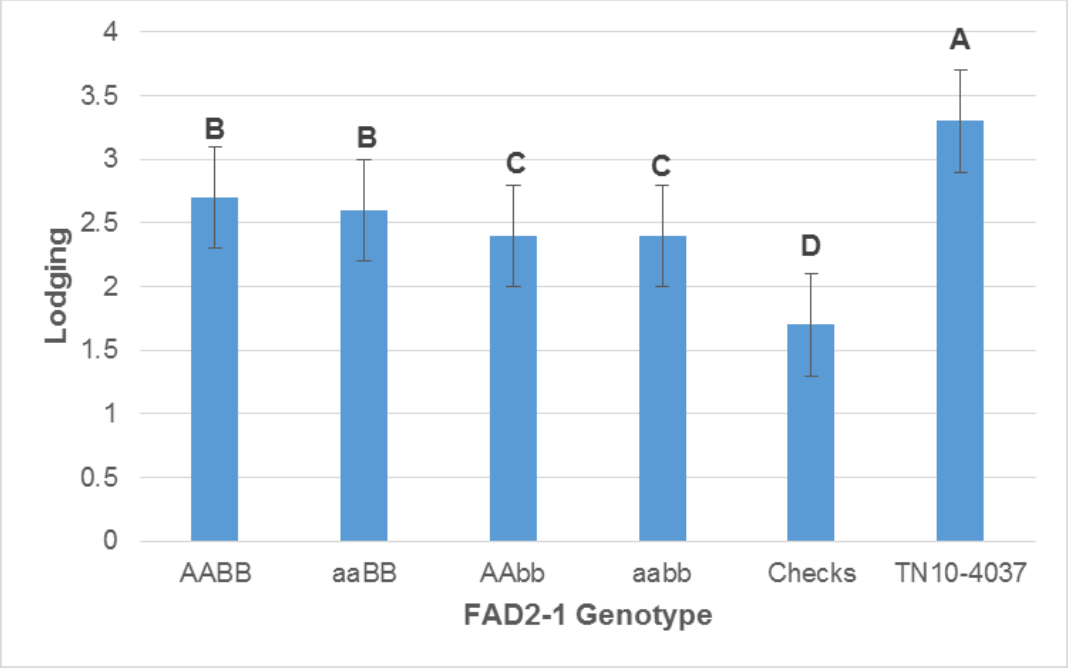


Figure 1.3 Lodging means of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes. The lodging was scored visually between 1 and 5 where 1 is completely upright and 5 is completely prostrate. The error bars represent the standard error for each mean value. Values with the same letter designation are not statistically different ($p < 0.05$).

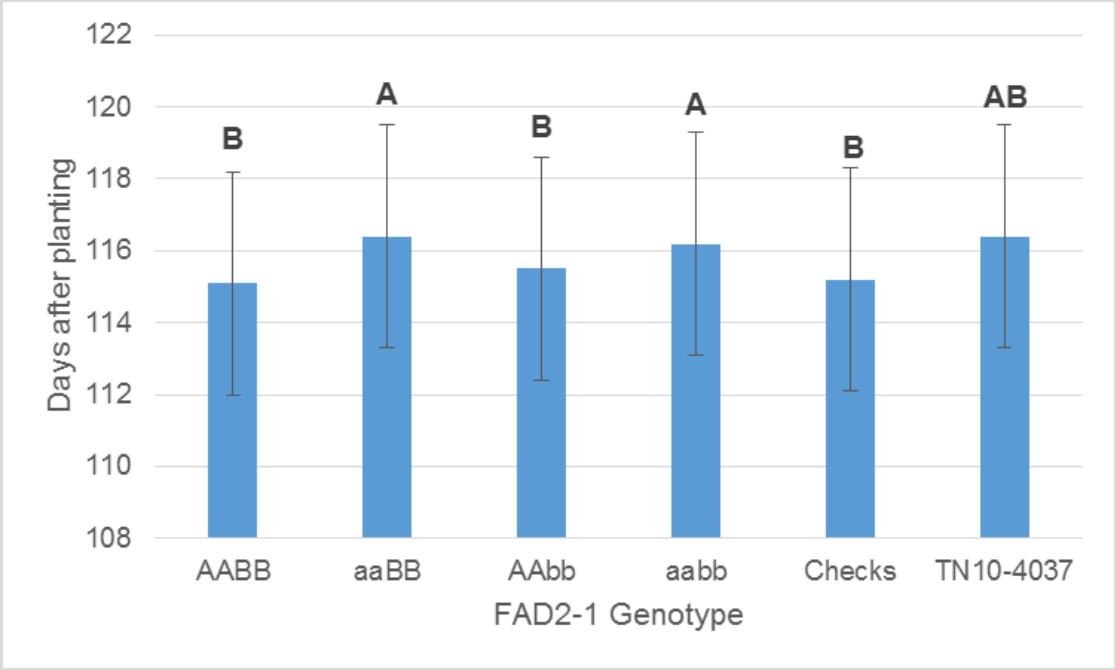


Figure 1.4 Days to maturity means of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes. The days to maturity are reported as maturity in days after planting. The error bars represent the standard error for each mean value. Values with the same letter designation are not statistically different ($p < 0.05$).

Table 1.1 Yield (kg ha-1), maturity (DAP), height (cm), and lodging of 48 experimental lines, three high yielding checks, and recurrent parent, TN10-4037, averaged over three replications and three locations

Line	Type	Yield (kg ha-1)	Maturity (DAP)	Height (cm)	Lodging
TN10-4037	Recurrent Parent	3671.2 ± 410.5	116.4 ± 3.2	104.4 ± 3.1	3.3 ± 0.4
LD00-3309	Check 4.0	3361.9 ± 410.5	114.1 ± 3.2	84.4 ± 3.1	1.3 ± 0.4
LD06-7620	Check 4.2	3609.4 ± 410.5	115.1 ± 3.2	82.8 ± 3.1	1.8 ± 0.4
LD00-2817	Check 4.5	3638.4 ± 410.5	116.4 ± 3.2	90.6 ± 3.1	1.7 ± 0.4
LR-01	aabb	3963.4 ± 410.5	118.3 ± 3.2	101.4 ± 3.1	3.2 ± 0.4
LR-02	aabb	3805.5 ± 410.5	118.2 ± 3.2	99.5 ± 3.1	3.3 ± 0.4
LR-03	aabb	2809.2 ± 410.5	113.7 ± 3.2	93.3 ± 3.1	2.0 ± 0.4
LR-04	aabb	3875.7 ± 410.5	118.4 ± 3.2	100.2 ± 3.1	2.8 ± 0.4
LR-05	aabb	3187.1 ± 410.5	114 ± 3.2	87.0 ± 3.1	1.4 ± 0.4
LR-06	aabb	3674.2 ± 410.5	118.2 ± 3.2	91.5 ± 3.1	2.4 ± 0.4
LR-07	aabb	4055.0 ± 410.5	114.8 ± 3.2	98.7 ± 3.1	2.8 ± 0.4
LR-08	aabb	3254.8 ± 410.5	115.2 ± 3.2	89.5 ± 3.1	2.4 ± 0.4
LR-09	aabb	3186.7 ± 410.5	113.7 ± 3.2	92.1 ± 3.1	1.8 ± 0.4
LR-10	aabb	3074.6 ± 410.5	114.6 ± 3.2	91.5 ± 3.1	1.6 ± 0.4
LR-11	aabb	3605.1 ± 410.5	118.7 ± 3.2	97.3 ± 3.1	2.7 ± 0.4
LR-12	aabb	3605.3 ± 410.5	116.4 ± 3.2	90.1 ± 3.1	2.6 ± 0.4
LR-13	aaBB	3208.4 ± 410.5	114.8 ± 3.2	93.4 ± 3.1	2.2 ± 0.4
LR-14	aaBB	3678.6 ± 410.5	115 ± 3.2	94.7 ± 3.1	2.5 ± 0.4
LR-15	aaBB	3986.0 ± 410.5	117.7 ± 3.2	105.7 ± 3.1	3.0 ± 0.4
LR-16	aaBB	2995.4 ± 410.5	114.1 ± 3.2	86.8 ± 3.1	2.0 ± 0.4
LR-17	aaBB	3680.9 ± 410.5	117.7 ± 3.2	98.2 ± 3.1	2.6 ± 0.4
LR-18	aaBB	3752.4 ± 410.5	116.7 ± 3.2	95.7 ± 3.1	3.0 ± 0.4
LR-19	aaBB	3803.3 ± 410.5	116.8 ± 3.2	102.2 ± 3.1	3.1 ± 0.4
LR-20	aaBB	3139.1 ± 410.5	115.3 ± 3.2	91.2 ± 3.1	2.0 ± 0.4
LR-21	aaBB	3157.0 ± 410.5	116.6 ± 3.2	98.6 ± 3.1	2.5 ± 0.4
LR-22	aaBB	3924.8 ± 410.5	116.2 ± 3.2	103.1 ± 3.1	2.8 ± 0.4
LR-23	aaBB	3890.7 ± 410.5	116.8 ± 3.2	105.3 ± 3.1	2.6 ± 0.4
LR-24	aaBB	3933.5 ± 410.5	119 ± 3.2	106.2 ± 3.1	3.1 ± 0.4
LR-25	AAbb	3315.6 ± 410.5	115.4 ± 3.2	86.4 ± 3.1	2.6 ± 0.4
LR-26	AAbb	3669.7 ± 410.5	118.6 ± 3.2	99.6 ± 3.1	2.4 ± 0.4
LR-27	AAbb	3380.4 ± 410.5	114.4 ± 3.2	95.1 ± 3.1	2.7 ± 0.4
LR-28	AAbb	3869.9 ± 410.5	115.5 ± 3.2	107.2 ± 3.1	3.0 ± 0.4
LR-29	AAbb	3016.9 ± 410.5	113.6 ± 3.2	85 ± 3.1	2.0 ± 0.4

Table 1.1 Continued

Line	Type	Yield (kg ha-1)	Maturity (DAP)	Height (cm)	Lodging
LR-30	AAbb	3609.1 ± 410.5	114.6 ± 3.2	89.2 ± 3.1	1.8 ± 0.4
LR-32	AAbb	3549.2 ± 410.5	118.6 ± 3.2	103.6 ± 3.1	2.7 ± 0.4
LR-33	AAbb	2997.5 ± 410.5	114.3 ± 3.2	88.8 ± 3.1	2.0 ± 0.4
LR-34	AAbb	3512.7 ± 410.5	114.1 ± 3.2	92.8 ± 3.1	2.2 ± 0.4
LR-35	AAbb	3629.0 ± 410.5	113.5 ± 3.2	96.4 ± 3.1	2.3 ± 0.4
LR-36	AAbb	3151.4 ± 410.5	117.8 ± 3.2	93.2 ± 3.1	2.0 ± 0.4
LR-37	AABB	3687.3 ± 410.5	117 ± 3.2	99 ± 3.1	3.4 ± 0.4
LR-38	AABB	3315.5 ± 410.5	114.4 ± 3.2	99.7 ± 3.1	2.6 ± 0.4
LR-39	AABB	3971.0 ± 410.5	118 ± 3.2	104.1 ± 3.1	3.1 ± 0.4
LR-40	AABB	3465.0 ± 410.5	116.8 ± 3.2	98.7 ± 3.1	3.0 ± 0.4
LR-41	AABB	3248.5 ± 410.5	113.7 ± 3.2	92 ± 3.1	1.7 ± 0.4
LR-42	AABB	3336.3 ± 410.5	114.6 ± 3.2	91 ± 3.1	2.7 ± 0.4
LR-43	AABB	3737.6 ± 410.5	114.8 ± 3.2	101.3 ± 3.1	3.2 ± 0.4
LR-44	AABB	3453.7 ± 410.5	114 ± 3.2	90.5 ± 3.1	2.5 ± 0.4
LR-45	AABB	3425.8 ± 410.5	113.8 ± 3.2	92.3 ± 3.1	2.1 ± 0.4
LR-46	AABB	3453.9 ± 410.5	113.5 ± 3.2	97.3 ± 3.1	2.5 ± 0.4
LR-47	AABB	3440.6 ± 410.5	114.6 ± 3.2	95.3 ± 3.1	2.6 ± 0.4
LR-48	AABB	3616.5 ± 410.5	115.4 ± 3.2	98.6 ± 3.1	2.5 ± 0.4

CHAPTER II
EVALUATION OF CONCENTRATION AND QUALITY OF SEED OIL AND PROTEIN OF 48 FIELD TESTED $BC_3F_{2:4}$ SOYBEAN LINES

Abstract

A backcross population of soybean was developed to evaluate the effect of mutant FAD2-1A and FAD2-1B alleles separately and combined on fatty acids and other seed composition traits. In combination, the mutant FAD2-1A and FAD2-1B alleles from PI 603452 and PI 283327 produced a mean of 80.4 ± 1.1% oleic acid. This is about 49 percentage points more oleic acid compared to soybean lines with only one mutant FAD2-1 allele and on average 59 percentage points more than lines containing the wild-type alleles. The high oleic trait was stable in all twelve lines with double mutant FAD2-1 alleles. The high oleic, double mutant group had significantly more total oil (240 ± 1g kg^{-1}) and protein (400 ± 1g kg^{-1}) compared to all other genotypic groups ($p < 0.05$). The concentration of lysine, methionine, tryptophan, and threonine was significantly lower in the double mutant group compared to all other groups, whereas the concentration of cysteine was significantly higher in the double mutant group. The high oleic trait appears to be stable across Tennessee locations. It would be advantageous to verify the apparent increase in both oil and crude protein in order to better understand positive or negative effects high oleic soybeans may have on soybean meal.

Introduction

The high oleic trait source may not only impact the fatty acid profile, but also have an effect on total oil and protein accumulation and important amino acids. Different sources of high oleic cultivars already on the market demonstrate the importance of how the causative mutation source inhibits the fatty acid

desaturase biosynthetic pathway. Dupont-Pioneer and Monsanto have both released transgenic high oleic varieties. Plenish, by Dupont-Pioneer claims to have 75% oleic acid, while Monsanto's Vistive Gold claims 72% (DuPont Pioneer, 2015; Monsanto Company, 2015). The conventional combination of PI 603452 and PI 283327 by Pham *et al.,* (2011) produced non-GMO soybeans with greater than 82% oleic content. This is important because the higher the oleic acid concentration, the higher the oxidative stability, and processors may need to keep different sources of high oleic beans separated.

The disruption of the fatty acid biosynthetic pathway can impact the accumulation of upstream and downstream fatty acids. It is expected that there will be a decrease in polyunsaturated fatty acids (linoleic and linolenic) when the enzyme responsible for making polyunsaturated precursors is disrupted. However, the combination of mutations at both FAD2-1A and FAD2-1B have also resulted in significantly lower saturated fatty acids (Pham *et al.,* 2011).

In 2014 U.S. commodity soybean cultivars had mean protein and oil values of 34.4% and 18.6%, respectively, on a 13% moisture basis (Miller-Garvin *et al.,* 2014). The protein and oil concentrations of soybeans form the basis for their market value. A recent estimation suggests that increased protein content could raise the estimated processed value (EPV) by $7.70-12.96 per acre, depending on the state (Beyondtheelevator.com, 2015). It is difficult to capture both high protein and oil concentration because they are negatively correlated, and both traits can be affected by environmental conditions (Wilson, 2004).

The effect of the high-oleic acid trait on protein and oil concentrations was recently evaluated by researchers at the University of Missouri (La *et al.,* 2014). Their study included mutant FAD2-1B source PI 283327, the same featured in this study, combined with FAD2-1A mutant sources 17D and M23, respectively. Surprisingly, lines derived from PI 283327 and 17D yielded higher protein and higher oil compared to normal oleic soybeans. This combination of improved oil and protein serves as a unique backdrop for the evaluation of protein and oil in mutant lines derived from PI 283327 and PI 603452.

After soybean oil is extracted, the solid by-product usually becomes soybean meal. Soybean meal is a widely used source of protein for livestock and the driving force for the soybean market (Wilson, 2004). Globally, 69% of all protein sources used in animal feeds come from soybean meal (Cromwell, 2012). Soybean meal has the benefits of highly digestible protein and an amino acid profile that nearly meets the nutritional requirements of swine and poultry (Cromwell, 2012). Between swine and poultry there are 5 limiting amino acids: lysine, threonine, tryptophan, methionine, and cysteine (Thakur and Hurburgh, 2007). It is common practice to supplement animal diets with purchased synthetic amino acids. However, the closer soybean meal can come to providing ideal nutrition to livestock, the more livestock producers will benefit from reduced costs from synthetic inputs. It has been demonstrated that meal produced from high-protein soybeans has a greater amount of digestible amino acids compared to conventional soybean meal (Baker and Stein, 2009). In this experiment the

relative accumulation of total oil and crude protein, as well as the relative concentrations of lysine, threonine, tryptophan, methionine, and cysteine were investigated among the four genotypic classes with different combinations of FAD2-1A and FAD2-1B alleles.

Materials and Methods

Gas Chromatography

Five seeds were subsampled from every harvested plot from all 3 locations in fall 2015. Each sample was analyzed using gas chromatography to determine the relative amount of five fatty acids found in soybean oil: palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1), linoleic acid (18:2), and linolenic acid (18:3). Each sample was crushed and transferred to a test tube with 1.0 mL of chloroform: hexane: methanol (8:5:1, by vol) for 4h. Next, 100 µL of oil extract was transferred to a 1.5 mL autosampler vial with 0.75 mL hexane and 75 µL methylation reagent (5 mL sodium methoxide, 10 mL ethyl ether, 20 mL petroleum ether). A crimper was used to secure caps to the vials. A Hewlett Packard HP 6890 series gas chromatograph was used to analyze the fatty acid methyl esters (FAME) (Agilent Technologies, Santa Clara, CA). This chromatograph includes a 7683 series auto sampler, flame ionization detector, and a 30 m x 0.53 mm i.d. J&W (Folsom, CA) 125-2332 capillary column with a 0.5 µm fused stationary phase.

Near Infrared Reflectance

A 50 g sample was taken from every harvested plot. Each plot sample was ground for 20 seconds. Total protein, oil, and amino acid profile all on a dry weight basis were determined using a Perten DA 7250 Near Infrared Reflectance (NIR) analyzer (Perten Instruments, Hägersten, Sweden).

Results and Discussion

Fatty Acid Concentrations

The FAD2-1 genotype was responsible for significantly different accumulations of all five measured fatty acids (Table 2.1). The oleic acid accumulation averaged for genotypic groups followed the trend of other reports where lines with one mutant allele and one wild-type allele produced more oleic acid than the normal lines but 46-54.2 percentage points less than the group with mutant alleles at both loci (Pham *et al.,* 2010 and Pham *et al.,* 2011) (Figure 2.1). Within the high oleic genotypic class there were few significant differences between all 12 lines in oleic acid content (Figure 2.2). Line LR-01 had less oleic acid accumulation than lines LR-03, LR-07, LR-10, and LR-12, and all of these lines were similar to the other 7 tested lines (Figure 2.2). The range of oleic acid accumulation was 77.1-81.8%.

Oleic acid accumulation occurs when polyunsaturated fatty acid production is curbed. This reduction in linoleic and linolenic acid contributes to the improved stability of high oleic soybean oil. High oleic lines averaged 4.2% linoleic acid and 3.9% linolenic acid compared to 54.8% and 7.6% in the wild-type group, respectively. It was reported that oil with 4-6% linolenic acid could

still have instability (Pham e*t al.,* 2012). As a result of that finding, the high oleic trait has been preliminarily combined with a low-linolenic trait linked to the FAD3 gene by some researchers (Pham *et al.,* 2012). There was a significant G x E interaction. Linolenic acid accumulation may not be reliably below 4% in high oleic normal linolenic soybeans depending on the environment. Hence, it would be advisable for breeders to add one or possibly two mutant FAD3 alleles to ensure <3% linolenic acid.

Total saturated fatty acid content in the AABB group was reduced by 30% in the aabb group (Table 1.1). There have been breeding efforts targeting <5% palmitic acid for the purpose of improving health quality and biodiesel performance (Cherrak *et al.,* 2002, Graef e*t al.,* 2009). The aabb lines had 7.7% palmitic acid compared to 12.0% in the AABB group. While not <5% palmitic acid, this decrease may still be useful to overall performance of high oleic soybean oil.

Oil, Protein, and Amino Acids

The effect of FAD2-1A and FAD2-1B genotype on oil, protein, and amino acid accumulation was evaluated using NIR results. The crude protein and oil are reported in g kg^{-1}, and amino acids are reported per unit of crude protein (g amino acids g^{-1} crude protein). The relative amount of amino acid of interest was divided by the relative amount of crude protein for each plot before ANOVA was performed. The amino acids of interest include lysine, methionine, tryptophan, threonine, and cysteine. The mean value and standard deviation for all measured

components within FAD2-1 genotypic classes are found in Table 2.2 Mean differences between genotypic classes were detected for all tested seed components.

The high oleic genotypic class (aabb) had significantly more total oil (240 g kg^{-1} ± 1.0) than all other genotypic classes including the recurrent parent and the check group. The range of total oil for genotypic classes was: 236-240 g kg^{-1}. On average the entire population had almost 5 percentage points higher oil accumulation than the average commodity soybean (186 g kg^{-1}) based on the estimate of Miller-Garvin *et al.*, 2014. This increased oil accumulation is consistent with an evaluation of 51 years of data sets from the Northern and Southern Uniform Soybean Tests which reported higher mean oil concentration in the Southern region compared to the Northern, especially in MG IV lines (Yaklich *et al.,* 2002). Traditionally, oil concentration is positively correlated with temperature (Wilson, 2004). Warm growing season temperatures in Tennessee probably contribute to the oil concentration in combination with genotype. There was a significant G x E interaction for total oil and crude protein.

Crude protein accumulation showed the same pattern as oleic acid accumulation. The genotypic class with wild-type alleles had the lowest mean protein (388 g kg^{-1} ± 7.0), the genotypic groups with one wild-type and one mutant allele had intermediate protein (391 g kg^{-1} ± 7.0 and 390 g kg^{-1} ± 7.0), and the high-oleic double mutant genotypic group had the highest total protein (400 g kg^{-1} ± 7.0). The significantly higher total oil and protein in the high oleic genotypic

group is interesting in light of the usual negative correlation between oil and protein (Wilson, 2004). However, in high oleic soybean lines, this phenomenon has been recently reported. In a recent study the unusual combination of elevated oil and protein was investigated by measuring carbohydrate components and verifying NIR determined oil and protein content via nuclear magnetic resonance and the Dumas method, respectively. There was an inconsistent trend for carbohydrate rankings, but some high oleic groups showed non-significant lower sugar accumulation than the normal groups. The oil and protein verification correlated highly with the original NIR measurements. (La *et al.,* 2014). The current experiment would benefit from further investigation into the increased oil and protein content in the future. However, the previous report by La *et al.,* (2014) suggests that mutations in FAD2-1A and FAD2-1B genes may contribute to altered protein synthesis as well as oil.

The effect on protein is also demonstrated by differences in amino acid concentrations between genotypic classes. None of the amino acids demonstrated significant G x E interaction, and genotype played a different role depending on the amino acid tested (Table 2.2). Interestingly, the high oleic group, genotype aabb, ranked highest or lowest for concentration of amino acid per crude protein for all five tested amino acids. Lysine, methionine, tryptophan, and threonine were all significantly less concentrated in the high oleic group compared to all other genotypic classes. Conversely, cysteine had the highest concentration in the high oleic group (Figure 2.3). The differences in oil, protein,

and amino acid quantity between genotypic groups suggest that mutations at the FAD2-1A and FAD2-1B loci can have ranging effects on soybean seed composition. None of these effects appear to be highly detrimental to breeders interested in selecting for the high oleic acid trait. It is possible that selections for the high oleic acid trait may be coupled with high protein if the simultaneous increase of oil and protein found in this study can be further verified. However, it may be important for breeders to consider the effect of the high oleic trait on important amino acid concentrations because of the potential ramifications on nutritional quality of soybean meal.

References

Baker, K.M. and H.H. Stein. 2009. Amino acid digestibility and concentration of digestible and metabolizable energy in soybean meal produced from conventional, high-protein, or low-oligosaccharide varieties of soybeans and fed to growing pigs. J. Anim. Aci. 87:2282-2290.

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/.

Cherrak, C.M., V.R. Pantalone, E.J. Meyer, D.L Ellis, S.L. Melton, D.R. West, and J.R. Mount. Low-Palmitic, Low-Linolenic Soybean Development. J. Am.Oil Chem. Soc. 80:539-543.

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

DuPont Pioneer. 2015. Plenish high oleic soybean oil. DuPont Pioneer. http://webcache.googleusercontent.com/search?q=cache:ryD5YnaCJIgJ:www.plenish.com/Libraries/Resources/Plenish_High_Oleic_Soybean_Oil_Technical_Data.sflb.ashx+&cd=1&hl=en&ct=clnk&gl=us

Graef, G., B.J. LaVallee, P. Tenopir, M.Tat, B. Schweiger, A.J. Kinney, J.H. Van Gerpen, and T.E. Clemente. 2009. A high-oleic-acid and low-palmitic-acid soybean: agronomic performance and evaluation as a feedstock for biodiesel. Plant Biotech. J. 7:411-421.

Kwanyuen, P., R.F. Wilson, and J.W. Burton. 1998. Soybean protein quality. p. 284-289. In S.S. Koseoglu *et al.* Oilseed and edible oils processing. Vol 1. Am. Oil Chem. Soc. Press, Champaign, IL.

La, T.C., S.M. Pathan, T. Vuong, J. Lee, A.M. Scaboo, J.R. Smith, A.M. Gillen, J. Gillman, M.R. Ellersieck, H.T. Nguyen, and J.G. Shannon. 2014. Effect of high-oleic acid soybean on seed oil, protein concentration, and yield. Crop Sci. 54:2054-2062.

Miller-Garvin, J., J.H. Orf, and S.L. Naeve. 2014. United States Soybean Quality Anuual Report 2014. U.S. Soybean Export Council. http://ussec.org/wp-content/uploads/2015/10/2014-Final-US-Commodity-Soybean-Quality-Report.pdf

Pham, A., J. Lee, J. Grove, and K.D. Bilyeu. 2010. Mutant alleles of FAD2-1a and FAD2-1b combine to produce soybeans with the high oleic acid seed oil trait. BMC Plant Biol. 10:195.

Pham, A., J. Lee, J. Shannon, K. Bilyeu. 2011. A novel FAD2-1A allele in a soybean plant introduction offers an alternate means to produce soybean seed oil with 85% oleic acid content. Theor. Appl. Genet. 123:793-802.

Pham, A..J. G. Shannon, and K.D. Bilyeu. 2012. Combinations of mutant FAD2 and FAD3 genes to produce high oleic acid and low linolenic acid soybean oil. Theor. Appl. Gen. 125:503-515.

Monsanto Company. 2015. About Vistive Gold: Vistive Gold composition vs. other oils. Monsanto Company. http://www.vistivegold.com/About.

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.

Wilson, R.F. 2004. Seed Composition. In: Soybeans: improvement, production, and uses. ASA-CSA-SSSA, Madison, WI. P. 630-631.

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.

Appendix

Table 2.1 Fatty acid mean concentrations of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes

	Fatty acid concentrations (%of total fatty acids)				
FAD2-1 Genotypic Class	Palmitic	Stearic	Oleic	Linoleic	Linolenic
AABB	11.9 ± 0.2a	4.3 ± 0.3bc	21.3 ± 1.1e	54.8 ± 1.0a	7.5 ± 0.2a
aaBB	10.9 ± 0.2d	4.4 ± 0.3b	34.4 ± 1.1b	42.7 ± 1.0c	7.5 ± 0.2a
AAbb	11.6 ± 0.2b	4.3 ± 0.3cd	26.2 ± 1.1c	50.3 ± 1.0b	7.6 ± 0.2a
aabb	7.7 ± 0.2e	3.7 ± 0.3d	80.4 ± 1.1a	4.2 ± 1.0d	3.9 ± 0.2c
Checks	11.2 ± 0.2c	4.5 ± 0.3a	23.8 ± 1.3d	53.3 ± 1.2a	7.1 ± 0.2b
TN10-4037	11.7 ± 0.2ab	4.1± 0.3a	20.9 ± 1.8de	55.4 ± 1.6a	7.7 ± 0.2a

*Letters of significance: two values with the same letter within a single column are not statistically different at a significance value of 0.05

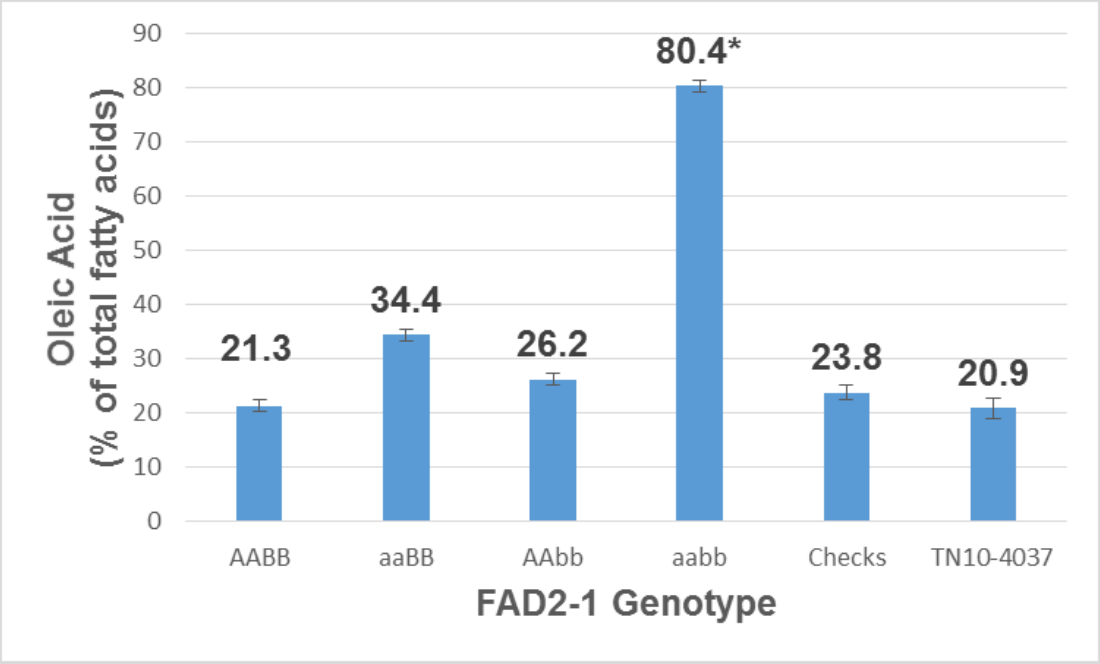


Figure 2.1 Oleic concentration means of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes. The error bars represent the standard error for each mean value. The star indicates that genotypic group, aabb, has statistically more oleic acid compared to all other groups ($p < 0.05$).

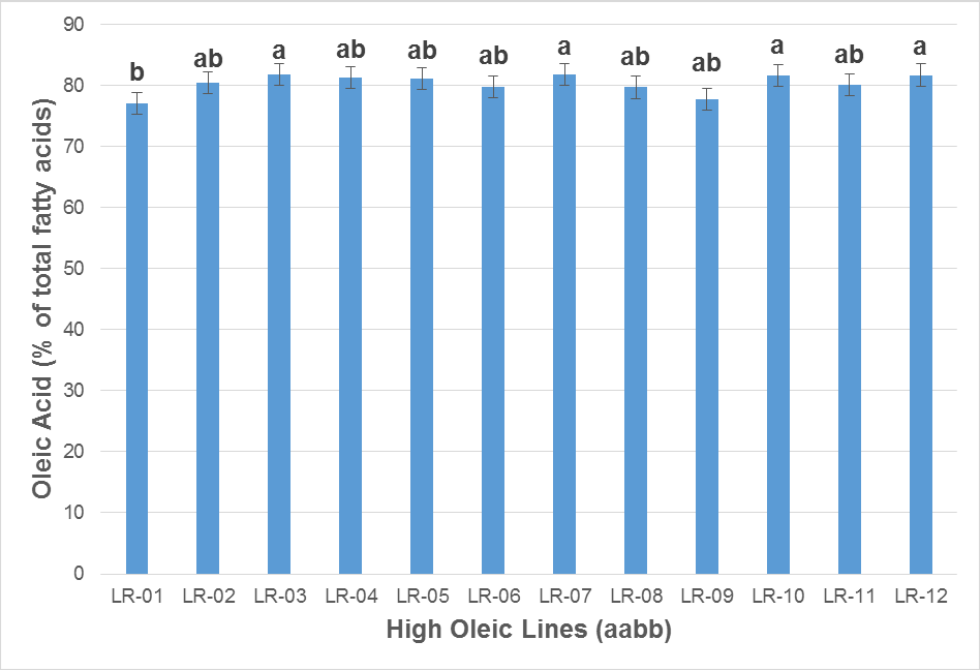


Figure 2.2 Oleic concentration means of 12 $BC_3F_{2:4}$ lines with mutant FAD2-1A and FAD2-1B alleles. The values were averaged across three replications and three locations. The error bars represent the standard error for each mean value. Values with the same letter designation are not statistically different ($p < 0.05$).

Table 2.2 Protein, oil, and amino acid concentration means of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes

Protein (g kg^{-1}), oil (g kg^{-1}), and amino acids (g amino acid g^{-1} crude protein) on a dry weight basis							
FAD2-1 Genotypic Class	Oil	Protein	Lysine	Methionine	Tryptophan	Threonine	Cysteine
AABB	237 ± 1b	388 ± 7c	0.0649 ± 0.0003a	0.0137 ± 0.0001a	0.0117 ± 0.0002a	0.0377 ± 0.0002a	0.0141 ± 0.0002b
aaBB	236 ± 1b	391 ± 7b	0.0644 ± 0.0003b	0.0136 ± 0.0001a	0.0114 ± 0.0002c	0.0375 ± 0.0002b	0.0142 ± 0.0002b
AAbb	236 ± 1b	390 ± 7b	0.0645 ± 0.0003b	0.0136 ± 0.0001a	0.0116 ± 0.0002b	0.0376 ± 0.0002ab	0.0141 ± 0.0002b
aabb	240 ± 1a	400 ± 7a	0.0634 ± 0.0003c	0.0132 ± 0.0001b	0.0105 ± 0.0002d	0.0364 ± 0.0002c	0.0143 ± 0.0002a
Recurrent Parent/Checks	237 ± 1b	389 ± 7bc	0.0646 ± 0.0003b	0.0136 ± 0.0001a	0.0117 ± 0.0002ab	0.0375 ± 0.0002b	0.0138 ± 0.0002c

*Values with the same letter within a single column are not statistically different at the significance value of $p < 0.05$

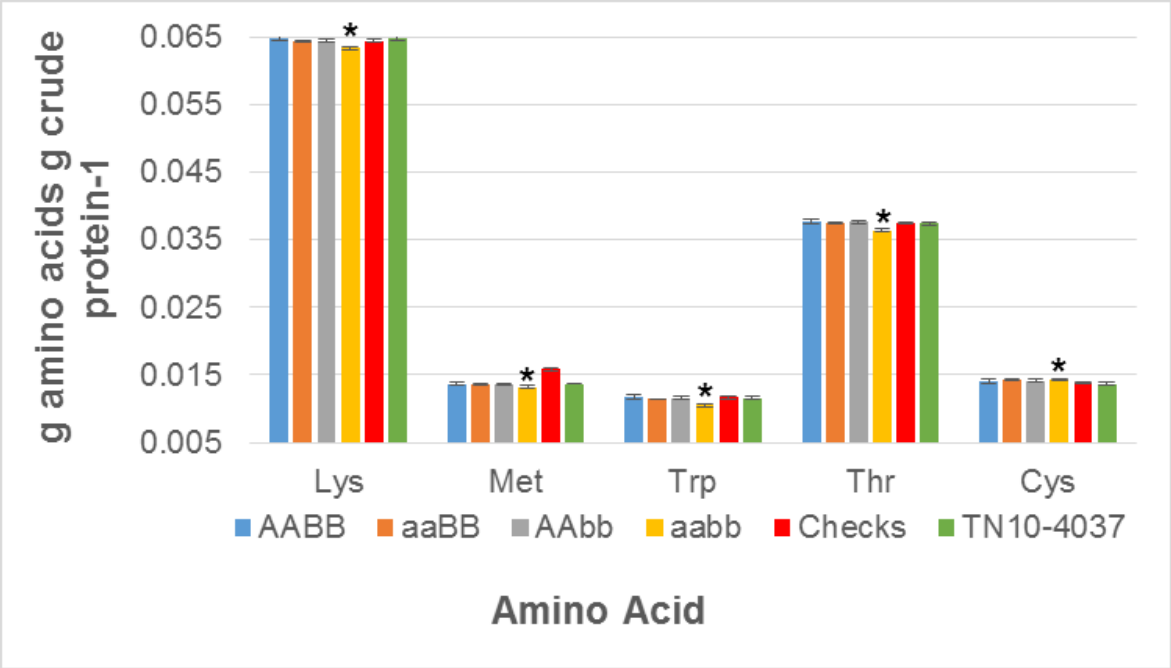


Figure 2.3 Ratio of mean grams of amino acids per mean grams of crude protein of 12 $BC_3F_{2:4}$ lines in each of four experimental groups with different FAD2-1A and FAD2-1B genotypes. Measured amino acids include lysine, methionine, tryptophan, threonine, and cysteine. The star indicates that the genotypic group, aabb, has statistically different amino acid concentration ($p < 0.05$).

Table 2.3 Oil, protein, and fatty acid concentrations of 48 experimental lines, three high yielding checks, and recurrent parent, TN10-4037, averaged over three replications and three locations

Line	Type	Crude Protein (g kg-1)	Oil (%)	Palmitic (% of FAs)	Stearic (% of FAs)	Oleic (% of FAs)	Linoleic (% of FAs)	Linolenic (% of FAs)
TN10-4037	Recurrent Parent	39.6 ± 0.7	23.3 ± 0.2	11.7 ± 0.2	4.1 ± 0.3	20.9 ± 1.8	55.4 ± 1.6	7.7 ± 0.2
LD00-3309	Check 4.0	39.1 ± 0.7	23.4 ± 0.2	11.4 ± 0.2	4.6 ± 0.3	24.3 ± 1.8	52.6 ± 1.6	7 ± 0.2
LD06-7620	Check 4.2	39.4 ± 0.7	23.3 ± 0.2	11.4 ± 0.2	4.5 ± 0.3	22.5 ± 1.8	54.3 ± 1.6	7.2 ± 0.2
LD00-2817	Check 4.5	37.3 ± 0.7	24.6 ± 0.2	10.7 ± 0.2	4.5 ± 0.3	24.6 ± 1.8	53.1 ± 1.6	6.9 ± 0.2
LR-01	aabb	39.9 ± 0.7	23.9 ± 0.2	7.9 ± 0.2	3.7 ± 0.3	77.1 ± 1.8	6.9 ± 1.6	4.2 ± 0.2
LR-02	aabb	40.8 ± 0.7	23.6 ± 0.2	7.5 ± 0.2	3.5 ± 0.3	80.3 ± 1.8	4.5 ± 1.6	4.0 ± 0.2
LR-03	aabb	40.1 ± 0.7	23.9 ± 0.2	7.8 ± 0.2	3.3 ± 0.3	81.7 ± 1.8	3.1 ± 1.6	3.9 ± 0.2
LR-04	aabb	39.7 ± 0.7	23.9 ± 0.2	7.6 ± 0.2	3.9 ± 0.3	81.3 ± 1.8	3.1 ± 1.6	4.1 ± 0.2
LR-05	aabb	40.2 ± 0.7	23.9 ± 0.2	7.8 ± 0.2	3.7 ± 0.3	81.1 ± 1.8	3.3 ± 1.6	3.9 ± 0.2
LR-06	aabb	40.4 ± 0.7	24.0 ± 0.2	7.9 ± 0.2	3.7 ± 0.3	79.8 ± 1.8	4.4 ± 1.6	4.0 ± 0.2
LR-07	aabb	40.2 ± 0.7	23.9 ± 0.2	7.5 ± 0.2	3.7 ± 0.3	81.8 ± 1.8	3.0 ± 1.6	3.9 ± 0.2
LR-08	aabb	39.9 ± 0.7	24.2 ± 0.2	7.5 ± 0.2	3.8 ± 0.3	79.7 ± 1.8	4.7 ± 1.6	4.1 ± 0.2
LR-09	aabb	40.1 ± 0.7	24.1 ± 0.2	8.0 ± 0.2	3.4 ± 0.3	77.7 ± 1.8	6.8 ± 1.6	3.9 ± 0.2
LR-10	aabb	40.1 ± 0.7	23.8 ± 0.2	7.3 ± 0.2	3.6 ± 0.3	81.6 ± 1.8	3.4 ± 1.6	3.9 ± 0.2
LR-11	aabb	39.4 ± 0.7	24.1 ± 0.2	7.5 ± 0.2	3.9 ± 0.3	80.1 ± 1.8	4.3 ± 1.6	4.0 ± 0.2
LR-12	aabb	39.4 ± 0.7	24.2 ± 0.2	7.7 ± 0.2	3.8 ± 0.3	81.7 ± 1.8	2.8 ± 1.6	3.9 ± 0.2
LR-13	aaBB	38.9 ± 0.7	23.6 ± 0.2	11.2 ± 0.2	4.5 ± 0.3	36.4 ± 1.8	41.0 ± 1.6	7.3 ± 0.2
LR-14	aaBB	38.7 ± 0.7	23.6 ± 0.2	10.6 ± 0.2	4.3 ± 0.3	32.9 ± 1.8	43.5 ± 1.6	7.9 ± 0.2
LR-15	aaBB	39.1 ± 0.7	23.5 ± 0.2	11.1 ± 0.2	4.3 ± 0.3	33.1 ± 1.8	44.0 ± 1.6	7.4 ± 0.2
LR-16	aaBB	39.5 ± 0.7	23.3 ± 0.2	10.8 ± 0.2	4.1 ± 0.3	35.1 ± 1.8	42.3 ± 1.6	7.5 ± 0.2
LR-17	aaBB	39.5 ± 0.7	23.2 ± 0.2	11.1 ± 0.2	4.4 ± 0.3	35.0 ± 1.8	41.8 ± 1.6	7.6 ± 0.2

Table 2.3. Continued

Line	Type	Crude Protein (g kg-1)	Oil (%)	Palmitic (% of FAs)	Stearic (% of FAs)	Oleic (% of FAs)	Linoleic (% of FAs)	Linolenic (% of FAs)
LR-18	aaBB	38.6 ± 0.7	23.7 ± 0.2	11.0 ± 0.2	4.2 ± 0.3	34.0 ± 1.8	43.0 ± 1.6	7.7 ± 0.2
LR-19	aaBB	39.1 ± 0.7	23.6 ± 0.2	11.2 ± 0.2	4.2 ± 0.3	32.7 ± 1.8	44.2 ± 1.6	7.6 ± 0.2
LR-20	aaBB	39.0 ± 0.7	23.8 ± 0.2	11.0 ± 0.2	4.4 ± 0.3	34.4 ± 1.8	43.0 ± 1.6	7.1 ± 0.2
LR-21	aaBB	39.0 ± 0.7	23.5 ± 0.2	11.1 ± 0.2	4.5 ± 0.3	32.1 ± 1.8	44.6 ± 1.6	7.6 ± 0.2
LR-22	aaBB	38.7 ± 0.7	23.8 ± 0.2	10.8 ± 0.2	4.6 ± 0.3	34.7 ± 1.8	42.1 ± 1.6	7.6 ± 0.2
LR-23	aaBB	39.4 ± 0.7	23.4 ± 0.2	10.9 ± 0.2	4.3 ± 0.3	35.6 ± 1.8	41.7 ± 1.6	7.4 ± 0.2
LR-24	aaBB	39.5 ± 0.7	23.5 ± 0.2	10.5 ± 0.2	4.3 ± 0.3	36.2 ± 1.8	41.6 ± 1.6	7.2 ± 0.2
LR-25	AAbb	38.6 ± 0.7	23.2 ± 0.2	11.7 ± 0.2	4.5 ± 0.3	24.3 ± 1.8	51.2 ± 1.6	8.2 ± 0.2
LR-26	AAbb	39.4 ± 0.7	23.2 ± 0.2	11.6 ± 0.2	4.5 ± 0.3	25.6 ± 1.8	50.9 ± 1.6	7.3 ± 0.2
LR-27	AAbb	39.0 ± 0.7	23.6 ± 0.2	11.3 ± 0.2	4.1 ± 0.3	26.6 ± 1.8	50.4 ± 1.6	7.5 ± 0.2
LR-28	AAbb	38.8 ± 0.7	23.7 ± 0.2	11.1 ± 0.2	4.0 ± 0.3	28.3 ± 1.8	49.1 ± 1.6	7.4 ± 0.2
LR-29	AAbb	38.9 ± 0.7	23.7 ± 0.2	11.9 ± 0.2	4.2 ± 0.3	26.2 ± 1.8	50.1 ± 1.6	7.5 ± 0.2
LR-30	AAbb	38.9 ± 0.7	23.8 ± 0.2	11.3 ± 0.2	4.1 ± 0.3	27.0 ± 1.8	50.1 ± 1.6	7.4 ± 0.2
LR-32	AAbb	38.9 ± 0.7	23.5 ± 0.2	11.6 ± 0.2	4.5 ± 0.3	25.4 ± 1.8	50.4 ± 1.6	8.0 ± 0.2
LR-33	AAbb	38.7 ± 0.7	23.9 ± 0.2	11.5 ± 0.2	4.4 ± 0.3	26.2 ± 1.8	50.2 ± 1.6	7.6 ± 0.2
LR-34	AAbb	38.8 ± 0.7	23.6 ± 0.2	11.9 ± 0.2	4.1 ± 0.3	25.9 ± 1.8	50.3 ± 1.6	7.7 ± 0.2
LR-35	AAbb	39.5 ± 0.7	23.7 ± 0.2	11.8 ± 0.2	4.3 ± 0.3	25.9 ± 1.8	50.1 ± 1.6	7.8 ± 0.2
LR-36	AAbb	39.4 ± 0.7	23.5 ± 0.2	11.5 ± 0.2	4.2 ± 0.3	26.2 ± 1.8	50.5 ± 1.6	7.5 ± 0.2
LR-37	AABB	38.8 ± 0.7	23.7 ± 0.2	11.6 ± 0.2	4.2 ± 0.3	20.0 ± 1.8	56.6 ± 1.6	7.5 ± 0.2
LR-38	AABB	38.3 ± 0.7	23.6 ± 0.2	11.5 ± 0.2	4.3 ± 0.3	22.3 ± 1.8	54.4 ± 1.6	7.4 ± 0.2
LR-39	AABB	38.9 ± 0.7	23.5 ± 0.2	11.9 ± 0.2	4.9 ± 0.3	20.9 ± 1.8	54.8 ± 1.6	7.4 ± 0.2
LR-40	AABB	39.1 ± 0.7	23.9 ± 0.2	12.2 ± 0.2	4.0 ± 0.3	21.0 ± 1.8	55.1 ± 1.6	7.5 ± 0.2

Table 2.3. Continued

Line	Type	Crude Protein (g kg-1)	Oil (%)	Palmitic (% of FAs)	Stearic (% of FAs)	Oleic (% of FAs)	Linoleic (% of FAs)	Linolenic (% of FAs)
LR-41	AABB	39.1 ± 0.7	23.7 ± 0.2	12.3 ± 0.2	4.0 ± 0.3	22.4 ± 1.8	53.8 ± 1.6	7.4 ± 0.2
LR-42	AABB	39.2 ± 0.7	23.2 ± 0.2	12.1 ± 0.2	4.1 ± 0.3	20.1 ± 1.8	55.8 ± 1.6	7.8 ± 0.2
LR-43	AABB	39.4 ± 0.7	23.3 ± 0.2	11.9 ± 0.2	4.1 ± 0.3	19.9 ± 1.8	55.8 ± 1.6	8.1 ± 0.2
LR-44	AABB	39.1 ± 0.7	23.5 ± 0.2	12.5 ± 0.2	4.5 ± 0.3	21.7 ± 1.8	53.8 ± 1.6	7.3 ± 0.2
LR-45	AABB	38.1 ± 0.7	24.0 ± 0.2	11.9 ± 0.2	4.4 ± 0.3	20.8 ± 1.8	55.2 ± 1.6	7.5 ± 0.2
LR-46	AABB	38.0 ± 0.7	24.0 ± 0.2	12.1 ± 0.2	4.4 ± 0.3	21.9 ± 1.8	54.0 ± 1.6	7.5 ± 0.2
LR-47	AABB	39.0 ± 0.7	23.7 ± 0.2	11.9 ± 0.2	4.4 ± 0.3	22.2 ± 1.8	54.1 ± 1.6	7.3 ± 0.2
LR-48	AABB	37.7 ± 0.7	24.3 ± 0.2	11.5 ± 0.2	4.6 ± 0.3	22.0 ± 1.8	54.2 ± 1.6	7.6 ± 0.2

Table 2.4. Ratio of mean grams of amino acids per mean grams of crude protein of 48 experimental lines, three high yielding checks, and recurrent parent, TN10-4037, averaged over three replications and three locations

Line	Type	Lys	Met	Trp	Thr	Cys
		(g AA g-1 CP)				
TN10-4037	Recurrent Parent	0.0646 ± 0.0003	0.0137 ± 0.0001	0.0116 ± 0.0002	0.0374 ± 0.0002	0.0137 ± 0.0002
LD00-3309	Check 4.0	0.0646 ± 0.0003	0.0136 ± 0.0001	0.0117 ± 0.0002	0.0374 ± 0.0002	0.0138 ± 0.0002
LD06-7620	Check 4.2	0.0640 ± 0.0003	0.0134 ± 0.0001	0.0117 ± 0.0002	0.0374 ± 0.0002	0.0138 ± 0.0002
LD00-2817	Check 4.5	0.0648 ± 0.0003	0.0136 ± 0.0001	0.0117 ± 0.0002	0.0375 ± 0.0002	0.0138 ± 0.0002
LR-01	aabb	0.0635 ± 0.0003	0.0131 ± 0.0001	0.0107 ± 0.0002	0.0366 ± 0.0002	0.0143 ± 0.0002
LR-02	aabb	0.0630 ± 0.0003	0.0130 ± 0.0001	0.0106 ± 0.0002	0.0361 ± 0.0002	0.0141 ± 0.0002
LR-03	aabb	0.0634 ± 0.0003	0.0131 ± 0.0001	0.0104 ± 0.0002	0.0366 ± 0.0002	0.0143 ± 0.0002
LR-04	aabb	0.0635 ± 0.0003	0.0131 ± 0.0001	0.0106 ± 0.0002	0.0368 ± 0.0002	0.0142 ± 0.0002
LR-05	aabb	0.0634 ± 0.0003	0.0131 ± 0.0001	0.0104 ± 0.0002	0.0364 ± 0.0002	0.0143 ± 0.0002
LR-06	aabb	0.0629 ± 0.0003	0.0130 ± 0.0001	0.0105 ± 0.0002	0.0361 ± 0.0002	0.0142 ± 0.0002
LR-07	aabb	0.0632 ± 0.0003	0.0131 ± 0.0001	0.0104 ± 0.0002	0.0363 ± 0.0002	0.0142 ± 0.0002
LR-08	aabb	0.0632 ± 0.0003	0.0130 ± 0.0001	0.0104 ± 0.0002	0.0365 ± 0.0002	0.0142 ± 0.0002
LR-09	aabb	0.0634 ± 0.0003	0.0131 ± 0.0001	0.0103 ± 0.0002	0.0363 ± 0.0002	0.0141 ± 0.0002
LR-10	aabb	0.0635 ± 0.0003	0.0134 ± 0.0001	0.0104 ± 0.0002	0.0364 ± 0.0002	0.0146 ± 0.0002
LR-11	aabb	0.0635 ± 0.0003	0.0133 ± 0.0001	0.0108 ± 0.0002	0.0365 ± 0.0002	0.0145 ± 0.0002
LR-12	aabb	0.0635 ± 0.0003	0.0132 ± 0.0001	0.0105 ± 0.0002	0.0363 ± 0.0002	0.0145 ± 0.0002
LR-13	aaBB	0.0646 ± 0.0003	0.0137 ± 0.0001	0.0115 ± 0.0002	0.0374 ± 0.0002	0.0141 ± 0.0002
LR-14	aaBB	0.0645 ± 0.0003	0.0137 ± 0.0001	0.0114 ± 0.0002	0.0377 ± 0.0002	0.0140 ± 0.0002
LR-15	aaBB	0.0643 ± 0.0003	0.0136 ± 0.0001	0.0116 ± 0.0002	0.0375 ± 0.0002	0.0142 ± 0.0002
LR-16	aaBB	0.0643 ± 0.0003	0.0135 ± 0.0001	0.0114 ± 0.0002	0.0372 ± 0.0002	0.0142 ± 0.0002

Table 2.4 Continued

Line	Type	Lys	Met	Trp	Thr	Cys
LR-17	aaBB	0.0641 ± 0.0003	0.0134 ± 0.0001	0.0114 ± 0.0002	0.0374 ± 0.0002	0.0141 ± 0.0002
LR-18	aaBB	0.0647 ± 0.0003	0.0137 ± 0.0001	0.0114 ± 0.0002	0.0375 ± 0.0002	0.0143 ± 0.0002
LR-19	aaBB	0.0643 ± 0.0003	0.0135 ± 0.0001	0.0115 ± 0.0002	0.0373 ± 0.0002	0.0141 ± 0.0002
LR-20	aaBB	0.0642 ± 0.0003	0.0135 ± 0.0001	0.0115 ± 0.0002	0.0375 ± 0.0002	0.0142 ± 0.0002
LR-21	aaBB	0.0645 ± 0.0003	0.0135 ± 0.0001	0.0113 ± 0.0002	0.0374 ± 0.0002	0.0140 ± 0.0002
LR-22	aaBB	0.0646 ± 0.0003	0.0137 ± 0.0001	0.0113 ± 0.0002	0.0374 ± 0.0002	0.0142 ± 0.0002
LR-23	aaBB	0.0644 ± 0.0003	0.0133 ± 0.0001	0.0112 ± 0.0002	0.0373 ± 0.0002	0.0138 ± 0.0002
LR-24	aaBB	0.0642 ± 0.0003	0.0136 ± 0.0001	0.0115 ± 0.0002	0.0376 ± 0.0002	0.0141 ± 0.0002
LR-25	AAbb	0.0647 ± 0.0003	0.0136 ± 0.0001	0.0116 ± 0.0002	0.0377 ± 0.0002	0.0141 ± 0.0002
LR-26	AAbb	0.0640 ± 0.0003	0.0134 ± 0.0001	0.0114 ± 0.0002	0.0374 ± 0.0002	0.0139 ± 0.0002
LR-27	AAbb	0.0645 ± 0.0003	0.0138 ± 0.0001	0.0117 ± 0.0002	0.0375 ± 0.0002	0.0142 ± 0.0002
LR-28	AAbb	0.0643 ± 0.0003	0.0135 ± 0.0001	0.0116 ± 0.0002	0.0370 ± 0.0002	0.0140 ± 0.0002
LR-29	AAbb	0.0648 ± 0.0003	0.0135 ± 0.0001	0.0116 ± 0.0002	0.0377 ± 0.0002	0.0140 ± 0.0002
LR-30	AAbb	0.0646 ± 0.0003	0.0136 ± 0.0001	0.0114 ± 0.0002	0.0377 ± 0.0002	0.0142 ± 0.0002
LR-32	AAbb	0.0645 ± 0.0003	0.0136 ± 0.0001	0.0117 ± 0.0002	0.0379 ± 0.0002	0.0142 ± 0.0002
LR-33	AAbb	0.0648 ± 0.0003	0.0136 ± 0.0001	0.0116 ± 0.0002	0.0377 ± 0.0002	0.0143 ± 0.0002
LR-34	AAbb	0.0646 ± 0.0003	0.0136 ± 0.0001	0.0116 ± 0.0002	0.0377 ± 0.0002	0.0140 ± 0.0002
LR-35	AAbb	0.0642 ± 0.0003	0.0136 ± 0.0001	0.0114 ± 0.0002	0.0374 ± 0.0002	0.0141 ± 0.0002
LR-36	AAbb	0.0643 ± 0.0003	0.0135 ± 0.0001	0.0116 ± 0.0002	0.0374 ± 0.0002	0.0139 ± 0.0002
LR-37	AABB	0.0648 ± 0.0003	0.0137 ± 0.0001	0.0116 ± 0.0002	0.0377 ± 0.0002	0.0142 ± 0.0002
LR-38	AABB	0.0649 ± 0.0003	0.0138 ± 0.0001	0.0119 ± 0.0002	0.0378 ± 0.0002	0.0144 ± 0.0002
LR-39	AABB	0.0644 ± 0.0003	0.0134 ± 0.0001	0.0116 ± 0.0002	0.0376 ± 0.0002	0.0138 ± 0.0002
LR-40	AABB	0.0646 ± 0.0003	0.0136 ± 0.0001	0.0116 ± 0.0002	0.0376 ± 0.0002	0.0138 ± 0.0002

Table 2.4 Continued

Line	Type	Lys (g AA g-1 CP)	Met (g AA g-1 CP)	Trp (g AA g-1 CP)	Thr (g AA g-1 CP)	Cys (g AA g-1 CP)
LR-41	AABB	0.0646 ± 0.0003	0.0136 ± 0.0001	0.0116 ± 0.0002	0.0377 ± 0.0002	0.0140 ± 0.0002
LR-42	AABB	0.0648 ± 0.0003	0.0136 ± 0.0001	0.0116 ± 0.0002	0.0376 ± 0.0002	0.0139 ± 0.0002
LR-43	AABB	0.0648 ± 0.0003	0.0135 ± 0.0001	0.0116 ± 0.0002	0.0375 ± 0.0002	0.0141 ± 0.0002
LR-44	AABB	0.0647 ± 0.0003	0.0137 ± 0.0001	0.0117 ± 0.0002	0.0375 ± 0.0002	0.0143 ± 0.0002
LR-45	AABB	0.0650 ± 0.0003	0.0137 ± 0.0001	0.0118 ± 0.0002	0.0375 ± 0.0002	0.0141 ± 0.0002
LR-46	AABB	0.0652 ± 0.0003	0.0137 ± 0.0001	0.0116 ± 0.0002	0.0377 ± 0.0002	0.0143 ± 0.0002
LR-47	AABB	0.0645 ± 0.0003	0.0135 ± 0.0001	0.0117 ± 0.0002	0.0376 ± 0.0002	0.0140 ± 0.0002
LR-48	AABB	0.0653 ± 0.0003	0.0138 ± 0.0001	0.0117 ± 0.0002	0.0379 ± 0.0002	0.0142 ± 0.0002

CHAPTER III
EVALUATION OF TRANSCRIPT ABUNDANCE OF FAD2-1A AND FAD2-1B WITH DIFFERENT ALLELIC COMBINATIONS AT DIFFERENT GROWTH STAGES OF SOYBEAN

Abstract

Transcript abundance for FAD2-1A and FAD2-1B were measured using quantitative RT-qPCR (qPCR). Four experimental lines from the field population were chosen to represent the four genotypic groups for oleic acid (AABB, aaBB, AAbb, and aabb). Genotype-compliment lines were planted as controls. Embryos were collected at three major seed development stages (R4, R5, and R6). Statistical analysis was hindered by the failure to maintain biological replicates. A few trends in samples that had consistently produced CT values are noticeable. In general, FAD2-1A was more highly expressed than FAD2-1B. This is consistent with the greater effect that mutant FAD2-1A has on increasing oleic acid content. The expression changes during seed development of FAD2-1A and FAD2-1B may be affected by genotype. If this were true, it would be possible that control of mutant FAD2-1A and FAD2-1B from PI 603452 and PI 283327, respectively, begins before the translation level.

Introduction

Before FAD1-1A and FAD2-1B were identified as individual isoforms, they were characterized as one ω-6 fatty acid desaturase gene (FAD2-1) along with less similar FAD2-2 (Kinney and Miao, 1996). Steady state transcript abundance measured via northern blot analysis showed FAD2-1 transcript increased and peaked during mid-maturation of seed development in concurrence with fatty acid biosynthesis, while FAD2-2 showed a stable expression level across embryo growth stages (Kinney and Miao, 1996). In 2005, Tang *et al.* reported that the soybean genome expresses two isoforms of FAD2-1 (FAD2-1A and FAD2-1B).

Tang also reported that while these isoforms differ by 24 amino acids only, they show different activity and temperature stability. Tang points to post-translational modification as a potential regulatory factor of these closely related enzymes.

Human attempts to regulate FAD2-1A and FAD2-1B function have been achieved using mutagenesis and genetic engineering. (Dierking and Bilyeu, 2009, Park *et al.,* 2014, Sandhu *et al.,* 2007, Takagi and Rahman, 1996). For plant introductions with increased oleic content, candidate gene approaches were used to search for sequence polymorphisms associated with the extreme phenotype. Pham *et al.,* (2010) proposed a causative sequence polymorphism in the FAD2-1B sequence of PI 283327. This missense mutation is the source of an amino acid replacement fatal to proper protein folding. PI 603452 has a sequence polymorphism in FAD2-1A positively identified to be the source of premature termination of translation (Pham e*t al.,* 2011).

The combination of these two mutant alleles has been successfully used to produce soybeans with >80% oleic content (Pham e*t al.,* 2011). In this experiment, reverse transcription quantitative polymerase chain reaction (RT-qPCR) was used to measure the relative transcript abundance of FAD2-1A and FAD2-1B in four genotypic classes homozygous at both loci and the focus of this entire study. To review, ‘A’ refers to the FAD2-1A locus and ‘B’ refers to the FAD2-1B locus. Capital letters indicate the presence of wild-type alleles that are common to most commodity soybeans. Lower case letters indicate the presence of the mutant allele from either PI 603452 or PI 283327. A single line chosen for

each genotypic class is therefore categorized as AABB, aaBB, AAbb, or aabb. If the mutant alleles are primarily controlled at the translation level, then it is expected that all genotypic classes will have similar, if not the same, transcript abundance for FAD2-1A and FAD2-1B. However, differences in transcript abundance may point to upstream control such as by a transcription factor.

While FAD2-1 expression has been characterized previously at different growth stages, a comparison between FAD2-1A and FAD2-1B growth stage expression via qPCR has not been performed (Kinney and Miao, 1996). To further understand the separate expression of the FAD2-1 isoforms within their respective genotypic classes and across different seed development stages, seed samples were collected at three growth stages (R4, R5, and R6) based on relative seed size. Transcript abundance of FAD2-1A, FAD2-1B, and both mature FAD2-1 genes were measured using RT-qPCR.

Materials and Methods

Plant Growth and Sample Collection

In December 2014, 2 seeds were planted into 80 pots in a greenhouse at the University of Tennessee Institute of Agriculture campus in Knoxville, TN. There were 10 pots (20 seeds) for each of 8 lines, 4 experimental and 4 corresponding controls. Each of the experimental lines represent a double homozygous genotypic class and were chosen from excess seeds of lines selected to be evaluated in 2015 yield trials. The control lines have the same combination of alleles at the FAD2-1A and FAD2-1B loci as their corresponding

experimental lines. The genotype, the experimental line designations, and the corresponding controls are listed in Table 3.1. The plants were grown on a 12 hour light and dark cycle. An insufficient number of seeds germinated in the HO control line KB12-1#70, sourced from Dr. Kristin Bilyeu from USDA-ARS in Columbia, Missouri. A second set of seeds from this line as well as the corresponding experimental line HO-78 were replanted in January 2015. Seeds were collected continually as the embryos reached R4, R5, and R6 growth stages. Seeds were considered in R4 when the pod was approximately 2 cm and the embryo was less than 0.3 cm in length, R5 categorized seeds filled approximately half of the pod, and R6 seeds completely filled the pod (Fehr, *et al.,* 1971). These sizes were used to visually group seeds together rather than collect from all pods on plants on a given day after flowering. Embryos were removed from their pods and placed in labeled envelopes that were immediately placed in liquid nitrogen to be transported back to -80° C storage.

RNA Extraction and Sample Preparation

Samples collected from different days within the same line-growth stage were pooled together. Twenty-four mortars, pestles, and spatulas were autoclaved. Each sample was ground to a semi-fine powder by hand using mortar and pestle sitting in a container of liquid nitrogen with a continual addition of liquid nitrogen directly onto the seeds. All samples were kept in -80° C storage until handled and when handled, they were placed directly into liquid nitrogen.

RNA extraction protocol was sourced from Verwoerd *et al.,* (1989). TLES extraction buffer was prepared from 10ml Tris.Cl pH 8.0, 2.5 ml 4M $LiCl_2$, 2 ml 0.5M EDTA pH 8.0, 10 ml SDS 10%, and RNase free water to 100 ml. A warm water bath was prepared at 70°C. For each 500 mg sample of ground tissue, 750 µl of TLES extraction buffer and 750 µl phenol (saturated pH 8.0) were combined and incubated at 70°C for 20 minutes. After incubation 500 mg of powdered tissue was added to the mixture and mixed for 30 seconds with a vortex. Next, 750 µl chloroform-isoamyl alcohol (24:1) was added and vortexed for 30 seconds. Samples were centrifuged for 20 minutes at 10,000 rpm at 4°C and the supernatants were recovered in new tubes. An equal volume of $LiCl_2$ 4M was added to these tubes, the tubes vortexed for 10 seconds and incubated for 4°C overnight. The next morning, samples were centrifuged for 30 minutes at 14,000 rpm at 4°C. The pellets were recovered and dissolved in 250 µl nuclease free water. Next, 25 µl of 3M NaOCH pH 5.6 was added to each sample followed by 500µl of 100% ethanol and mixed well. Samples were incubated 1hr at -80°C and then centrifuged for 30 minutes at 14,000 rpm. Samples were washed again with 500 µl 70% ethanol and centrifuged for 5 minutes at 4°C followed by the addition of 50µl RNase free water. Samples were finally subjected to DNase treatment. 3 µl of RNA sample was combined with 8 µl water and 1 µl amplification grade DNase to sit at room temperature for 15 minutes. Then, 1 µl 25mM EDTA was added and samples heated at 65°C for 10 minutes. RNA purity and concentration was determined spectrophotometrically using the NanoDrop 1000

Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, U.S.). Working samples of 500 µl and RNA concentration of 25 ng/µl were prepared for direct pipetting into 96-well plates.

Reverse Transcription Quantitative PCR

Primers were designed to bind specifically to FAD2-1A, FAD2-1B, and mature FAD2-1 transcripts. It is primarily of interest to compare transcript abundance of isoforms FAD2-1A and FAD2-1B and between genotypes and growth stages within each gene. However, a primer was designed to bind both FAD2-1A and FAD2-1B mature RNA in order to look for a sign of post-transcriptional regulation.

Five parameters were used to design primers: melting temperature of at least 60°C, guanine-cytosine content of at least 50%, 21-23 base pairs in length, total RNA sequence length between primers of 160-220 base pairs, and at least three guanine or cytosine nucleotide bases. All potential primers were tested for these qualities using OligoAnalyzer3.1 (Integrated DNA Technologies, Coralville, IA). Sequences were aligned to ensure no overlapping primer binding would occur between FAD2-1A and FAD2-1B primers using MultAlin software (Corpet, 1988). Finally, the inverse complement sequence of each chosen primer was determined using SeWeR Tools (Basu, 2004). The FAD2-1A primer set consists of F-5‘ATGCCGCAGTAGAGGACGATCAC3’ and R-5‘GAGGTGGAAGTAGGTG-GTGGCAA3’. The FAD2-1B primer set consists of F-5’CAGCGTTGGCTTTCCC-TGTAGCT3’ and R-5’GGGAACGCTGAAAGCAGTGCGG3.’ The mature FAD2-1

primer set consists of F-5’CGCCGCCATCACTCCAACACAGG3’ and R-5’GGGA-GCATAAGGGTGGTAGC3.’

For each RNA sample tested with each primer pair for the FAD2-1 genes of interest in quadruplicate, a quadruplicate sample using a reference gene was simultaneously run. The reference gene is also referred to as the endogenous control. The endogenous control gene chosen in this experiment was that of soybean ubiquitin protein (Glyma20G141600). Ubiquitin is a highly conserved protein in most eukaryotes (Callis and Vierstra, 1989). Endogenous control genes are expected to be constitutively active and not change expression between samples because the two samples being compared are normalized using them. (Taylor, 2010). The ubiquitin primer set consists of F-5’GTGTAATGTTGGATGTGTTCCC-3’ and R-5’ACACAATTGAGTTCAACACAAACCG-3’.

RT-qPCR was performed with the 7900 HT Fast Real-Time PCR instrument using Verso SYBR Green 1-Step qRT-PCR ROX Mix assay kit (Thermofisher, Waltham, MA). Each reaction, done in replicates of 4 contained 13 µl of Verso SYBR Green PCR master mix and 2 µl of 25 ng/µl RNA samples. The comparative CT method was employed to make deductions about FAD2-1A and FAD2-1B gene expression. In this method, the output value of interest for each sample is the CT value or cycle threshold value. SYBR Green is a dye that preferentially binds to double stranded DNA. With each cycle of PCR the target gene copy number is being increased exponentially, and as the SYBR Green dye

binds to the double stranded replicated gene of interest it emits a fluorescence. Its signal value is recorded when the fluorescence is large enough to be detected above the background signal as the CT value. The CT value is inversely proportional to the original amount of RNA present in a sample. Therefore, samples that have lower CT values have greater relative transcript abundance than samples with higher CT values. To determine the relative expression difference between two samples the CT value of the reference gene, ubiquitin, was subtracted from the CT values of the two samples of interest. This provided two ΔCT values. Then, the difference of the ΔCT value for each sample of interest was determined and raised to the second power to obtain the relative fold change for pairwise comparisons.

Results and Discussion

The results for this experiment are first addressed with an acknowledgement of concerns about experimental design and execution. While samples for each line-growth stage combination were collected multiple times, all samples belonging to the same line-growth stage were pooled together upon sample preparation. The experimenter failed to see the need to maintain biological replicates for each sample and therefore only had experimental quadruplicates. While mean cycle threshold (CT) values were only calculated if all technical replicates in a tested sample were within 0.9 values of each other, there is no way to verify the accuracy of measured transcript abundance for any given sample nor test for significance. Furthermore, several samples had to be

run a second, and possibly third, time due to large replicate variation. The variation between initial and repeat samples was very large in all instances. It is unsure whether this is due to RNA sample degradation or variation in the internal control efficiency when more stock solution of ubiquitin primer was diluted. The fold change results for repeat samples will not be reported due to their extreme values. All transcript abundance comparisons reported here are viewed as a preliminary look at gene expression differences between FAD2-1A and FAD2-1B. Possible trends will be acknowledged, however, no significance will be attributed to any mRNA fold change between samples.

The expression of FAD2-1A and FAD2-1B was compared in all lines in the R4, R5, and R6 growth stage. For experimental lines LR-102 (AABB), LR-76 (aaBB), and LR-16 (AAbb) the expression of FAD2-1A mRNA was increased over the mRNA of FAD2-1B in all three growth stages (Table 3.2). There was a trend in these lines between growth stages as well. At R4 the upregulation of FAD2-1A compared was largest compared to FAD2-1B. The fold change increase was intermediate during R5 and smallest during R6. The genotype compliment control line for LR-75, PI 603452 (aaBB) did not show this same trend (Table 3.2). The genotype compliment control lines for LR-102 and LR-16 are not reported due to experimental error. Line LR-78 with double mutant alleles is also not reported for experimental error. However, the genotype compliment control line for LR-78, KB12-1#70 demonstrated the same increased expression trend for FAD2-1A and FAD2-1B (Table 3.2). The greater transcript abundance in

FAD2-1A is consistent with the influence of the mutant FAD2-1A allele on the oleic content. When the mutant FAD2-1A allele was combined with the wild-type FAD2-1B allele, the mean oleic content of 12 field tested lines was 47.7 ± 1.3%. This is significantly greater than the mid-oleic content of lines with a mutant FAD2-1B allele combined with a wild-type FAD2-1A allele (38.4 ± 1.3%).

Previously it was reported that FAD2-1 expression was enhanced during seed development, which is consistent with seed oil deposition (Kinney and Miao, 1996). For three lines (LR-102, LR-76, and LR-16), with three FAD2-1A and FAD2-1B allele combinations, the expression difference between R4-R5 and R5-R6 for FAD2-1A and FAD2-1B are represented in Figure 3.1 and Figure 3.2. Line LR-102 with wild-type alleles had an increase in expression in FAD2-1A from R4 to R5 and a reduction in expression from R5 to R6. Line LR-76 (aaBB) showed no expression change in FAD2-1A between growth stages, and line LR-16 (AAbb) showed a reduction in expression from R4 to R5 and from R5 to R6. The three different trends in transcript abundance between growth stages for FAD2-1A were not the same for FAD2-1B. Lines LR-102, LR-76, and LR-16 showed a general decrease in expression from R4 to R5 and R5 to R6.

Based on transcript abundance, it appears that the FAD2-1A isoform is more heavily expressed than FAD2-1B. This is consistent with the differing effect of mutant FAD2-1A and FAD2-1B alleles at increasing oleic acid content. There are differing trends in FAD2-1A and FAD2-1B expression between embryo development growth stages between lines that have different genotypes. This

may be an indicator that the mutations in FAD2-1A and FAD2-1B from PI 603452 and PI 283327 may prevent the conversion of oleic acid to linoleic precursors starting with the control of transcription. Overall, this experiment should be repeated to verify the possible effect of genotype on transcript abundance in FAD2-1A and FAD2-1B.

References

Basu, M.K. 2004. SeWeR: Tools for sequence analysis. Scilico, LLC. http://www.bioinformatics.org/sewer/

Callis, J. and R. Vierstra. 1989. Ubiquitin and ubiquitin genes in higher plants. Oxf. Surv. Plant Mol. Cell Biol. 6:1-30.

Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. Nucl. Acids. Res. 16: 10881-10890.

Dierking, E.C. and K.D., Bilyeu. 2009. New sources of soybean seed meal and oil composition traits identified through TILLING. BMC Plant Biol. 9:89

Fehr, W.R., C.E. Caviness, D.T. Burmood, and J.S. Pennington. 1971. Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. Crop Sci. 11:929-931.

Kinney, K.L. and G.H. Miao. 1996. Developmental and growth temperature regulation of two different microsomal [omega]-6 desaturase genes in soybeans. Plant Physiol. 110:311-319.

Integrated DNA Technologies. 2016. OligoAnalyzer 3.1. Coralville, IA, U.S. https://www.idtdna.com/calc/analyzer

Park, H., G. Graef, Y. Xu, P. Tenopir, T.E. Clemente. 2014. Stacking of a stearoyl-ACP thioesterase with a dual-silenced palmitoyl-ACP thioesterase and ∆12 fatty acid desaturase in transgenic soybean. Plant Biotech. J. 12:1035-1043.

Pham, A., J. Lee, J. Grove, and K.D. Bilyeu. 2010. Mutant alleles of FAD2-1a and FAD2-1b combine to produce soybeans with the high oleic acid seed oil trait. BMC Plant Biol. 10:195.

Pham, A., J. Lee, J. Shannon, K. Bilyeu. 2011. A novel FAD2-1A allele in a soybean plant introduction offers an alternate means to produce soybean seed oil with 85% oleic acid content. Theor. Appl. Genet. 123:793-802.

Sandhu, D., J.L. Alt, C.W. Scherder, W.R. Fehr, M.K. Bhattacharyya. 2007. Enhanced oleic acid content in the soybean mutant M23 is associated with the deletion in the Fad2-1a gene encoding a fatty acid desaturase. J. Am. Oil Chem. Soc. 84:229-235.

Takagi, Y. and S.M. Rahman. 1996. Inheritance of high oleic acid content in the seed oil of soybean mutant M23. Theor. Appl. Genet. 92:179-182.

Taylor, Sean. 2010 A practical guide to publishing RT-qPCR data that conform to the MIQE guidelines. Am. Biotechnol. Lab. 8:14-19.

Tang, G.Q., W.P. Novitzky, H.C. Griffin, R.E. Huber SC Dewey. 2005. Oleate desaturase enzymes of soybean: evidence of regulation through differential stability and phosphorylation. Plant J. 44:433-446.

Verwoerd, T.C., B.M. Dekker, A. Hoekema. 1989. A small-scale procedure for the rapid isolation of plant RNAs. Nucleic Acids Res. 17:2362.

Appendix

Table 3.1 FAD2-1A and FAD2-1B genotype for experimental and control lines

FAD2-1 Genotype	Experimental Line	Corresponding Control
AABB	LR-102	TN10-4037 (Recurrent Parent)
aaBB	LR-76	PI 603452
AAbb	LR-16	PI 283327
aabb	LR-78	KB12-1#70

Table 3.2 FAD2-1A mRNA fold change relative to FAD2-1B

Line	Genotype	Fold change (R4)	Fold change (R5)	Fold change (R6)
LR-102	AABB	35	19	9
LR-76	aaBB	28	11	2
LR-16	AAbb	75	32	12
PI603452	aaBB	23	32	2
KB12-1#70	aabb	25	9	3

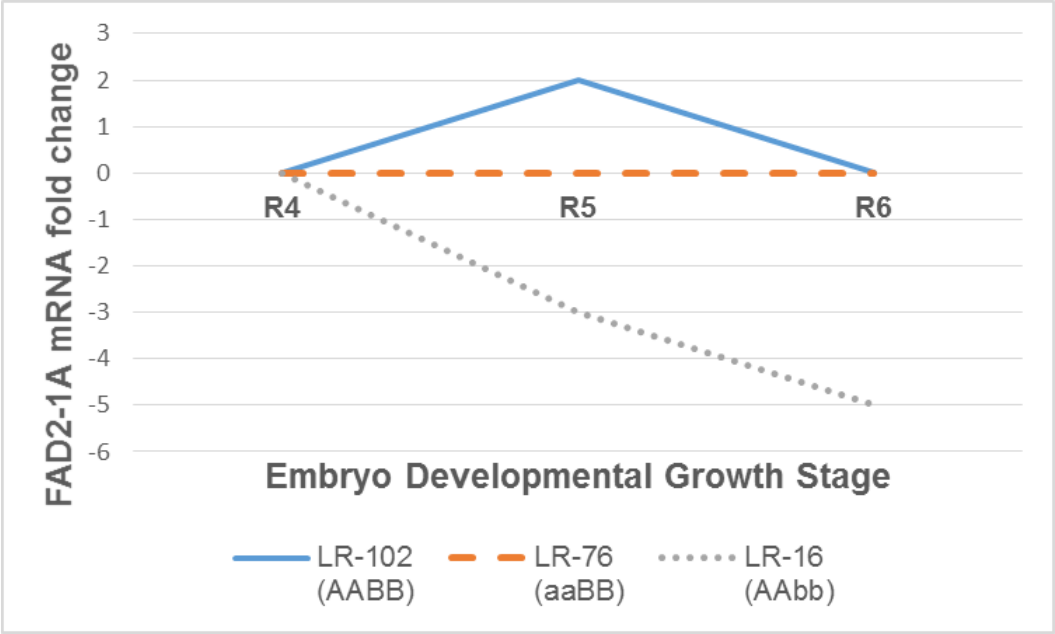


Figure 3.1 FAD2-1A mRNA fold change between embryo growth stages

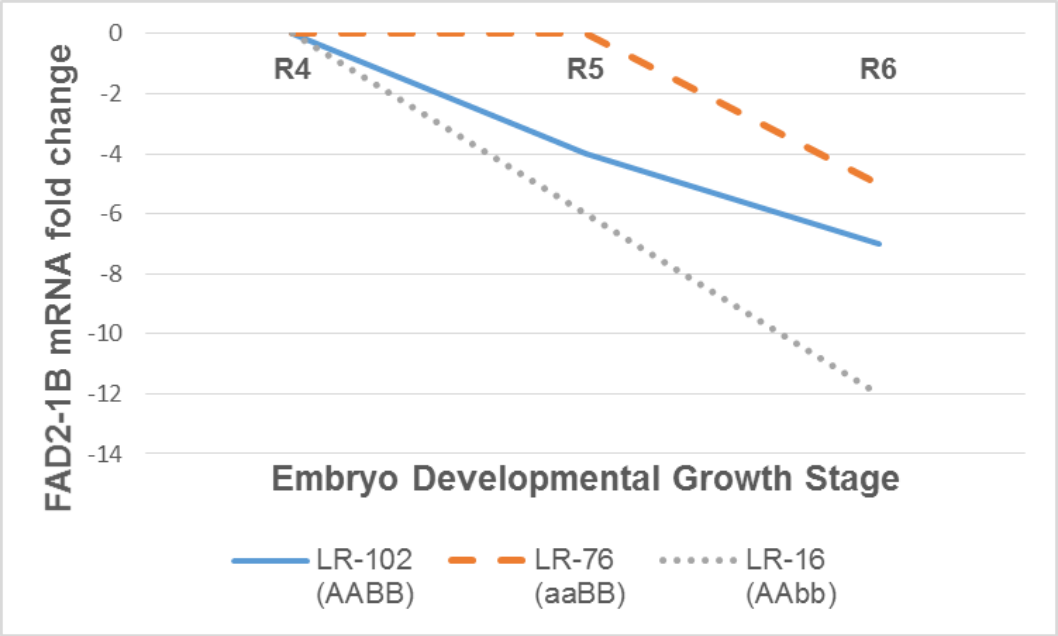


Figure 3.2 FAD2-1B mRNA fold change between embryo growth stages

CONCLUSION

In 2015, farmers in nine states in the Eastern Corn Belt received premiums between 40¢ and 80¢ per bushel to grow high oleic soybeans (Gullickson, 2015). However, soybeans are an important commodity crop in other regions as well, including the state of Tennessee. This experiment confirms that the high oleic trait sourced from two plant introductions with naturally occurring mutations can be successfully integrated into lines appropriate for Tennessee growers. The conventional high oleic trait tested showed no apparent yield drag nor negative effect on the days to maturity. Furthermore, it appears that high oleic soybean populations developed for Tennessee may have the potential to yield greater total oil and protein. This phenomena should be verified in the future. While increased protein is commonly a goal for soybean breeders, in this experiment it appeared that the concentration of four of five nutritionally important amino acids were significantly reduced when expressed as grams of amino acids per gram of crude protein. It would be advantageous for high oleic soybean meal to be tested in feeding trials with hogs and chickens to look for any advantages or drawbacks compared to normal oleic soybean meal. Overall, high oleic soybeans developed in Tennessee have no glaring drawbacks in terms of agronomic performance and seed composition, and they should be a feasible option for farmers in the future.

References

Gullickson, G. 2015. High-oleic beans spawn healthy oil for booming market. Agriculture.com. http://www.agriculture.com/crops/soybeans/technology/higholeic-bes-spawn-healthy-oil-f_143-ar51491

VITA

Lauren K. Richardson was born in Bakersfield, CA. She received a B.S. in Biochemistry from Carson-Newman University and hopes to continue working in the field of plant breeding after she graduates from the University of Tennessee with a M.S. in Plant Sciences.