MOLECULAR BREEDING FOR IMPROVEMENT OF FATTY ACIDS IN SOYBEAN

by

NICOLE MELISSA BACHLEDA

(Under the Direction of Zenglu Li)

ABSTRACT

Soybean is a major global oil crop. Development of soybean cultivars with enhanced fatty acids is an important breeding goal that will produce healthier oil for human consumption. In this study, we developed near-isolines for high oleic and low linolenic traits with various combinations of *FAD2-1A* and *FAD2-1B*, or different *FAD3A* alleles. The lines with mutant *FAD2-1A* and *FAD2-1B* alleles produced an average of 806 g kg^{-1} oleic acid and with mutant *FAD3A* alleles reduced the linolenic acid to 37 g kg^{-1}. Replicated yield tests indicated that these traits didn't cause yield reduction. Modified fatty acid composition did not have a major impact on seed germination. The *FATB1a* gene is responsible for reduced palmitic acid in soybean line N87-2122-4. We identified the *FATB1a* gene deletion region in the soybean genome and successfully developed a TaqMan marker assay to detect the deletion. The marker can be used effectively in marker-assisted breeding of the reduced palmitic acid trait in soybean.

INDEX WORDS: Soybean, DNA marker, Molecular Breeding, Oleic Acid, Linolenic Acid, Palmitic Acid, and Yield

MOLECULAR BREEDING FOR IMPROVEMENT OF FATTY ACIDS IN SOYBEAN

by

NICOLE MELISSA BACHLEDA

BS, University of Illinois Urbana-Champaign, 2011

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment

of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2015

MOLECULAR BREEDING FOR IMPROVEMENT OF FATTY ACIDS IN SOYBEAN

by

NICOLE MELISSA BACHLEDA

Major Professor:	Zenglu Li
Committee:	Roger Boerma
	Ali Missaoui

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2015

DEDICATION

This thesis is dedicated to the most important people in my life. To my partner, Andrew Sticha, for being my sunshine through the toughest parts of graduate school and loving me every step of the way. To my parents, Paul and Gloria Bachleda, for their continued love, guidance, and support as I grow in life and education. To my brother and sister, Steven and Melissa, who can always put things into perspective and make me laugh about anything.

ACKNOWLEDGEMENTS

Above all, I am deeply grateful to Dr. Zenglu Li, who has taught me far more than any class in the ways of professionalism and scientific integrity. I would like to thank my committee members, Dr. Roger Boerma and Dr. Ali Missaoui for bringing well rounded guidance to my research. Dr. Anh Pham deserves credit for being an essential mentor in lab work and my first friend in Athens. I thank Dr. Timothy Grey for his technical guidance and assistance with the seed germination test. I would also like to thank Dr. Kristen Bilyeu from USDA-ARS, Columbia, MO for fatty acid analyses of soybean lines. I am sincerely thankful to Dale Wood, Dr. Donna Harris, Gina Bishop, Earl Baxter, Brice Wilson, Chris Collin, Dr. Rebecca Tashiro, Dr. Justin Vaughn, Tatyana Nienow, Collen Wu, and everyone from the soybean breeding and molecular genetics lab for their technical assistance in the lab, field, and greenhouse. This research was funded by the United Soybean Board and the USDA Agriculture and Food Research Initiative grant.

TABLE OF CONTENTS

Page

CHAPTER

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Introduction and History of Soybean Production and Commercial Value

Soybean [*Glycine max* (L.) Merr.] is an annual legume originating from East Asia (Hymowitz, 1970). It has a trifoliolate leaf system and self-fertilizing flowers in the axillary nodes. Flowers are white or purple. Soybean plants have a fine pubescence of gray or tawny covering the whole plant, including the seed pods which hold 2 to 4 seeds. Pod walls range in color from tan, brown, or black. The seed hilum can be black, gray, brown, buff, or yellow. Like other legumes, soybean has the ability to “fix” nitrogen from a symbiotic interaction with rhizobacteria in the roots that convert atmospheric nitrogen into ammonia (Wacek and Brill, 1976).

Soybean was domesticated from wild soybean (*Glycine soja*) in China. Multiple domestications have been suggested in the north-eastern region, the Yellow River Valley, and the southern region in China (Li et al., 2010). Soybean was brought to the USA in 1765. Samuel Bowen first brought it to Georgia (GA), where it was used to produce soy sauce and vermicelli noodles for export to England (Hymowitz and Harlan, 1983). For about 100 years, soybean production was slow to gain popularity, and it was grown primarily as a forage crop. As early as 1904, soybean was known to be a good source of oil from the work of George Washington Carver (Shurtleff and Aoyagi, 2004). While less than 2,000 hectares were grown in 1907, soybean hectarage were increased very rapidly between 1907 and 1930. Starting in 1920, soybean quickly began to spread through the Corn Belt including Iowa, Illinois, Indiana, southern Michigan, western Ohio, eastern Nebraska, eastern Kansas, southern Minnesota and parts of Missouri (Hart, 1986). From 1920 to 1924, the top five producers of soybean went from North Carolina, Virginia, Alabama, Missouri, and Kentucky to Illinois, North Carolina, Missouri, Indiana, and Ohio (Shurtleff and Aoyagi, 2004). It was during World War II that soybean really took off as a seed crop, with major applications to oil and protein production (Gibson and Benson, 2005). By 1942, the USA became the world's leading soybean producing nation, and was producing over 76% of the world's soybeans by 1969 (Gibson and Benson, 2005; Hart, 1986; Hymowitz and Harlan, 1983; Hymowitz and Shurtleff, 2005). While soybean is a major crop in the Corn Belt, it still represents 3% of the total crop value grown in GA (Soystats, 2015).

Currently in the USA, soybeans are valued at over $40 billion, with production over 108 million metric tons from 33.9 million hectares (Soystats, 2015). The USA exported 17,949 million metric tons of soybean in 2014, accounting for the majority of world soybean production (Soystats, 2015). While a lot of soybean goes towards human and animal consumption, biodiesel is a growing industry for soybean (Soystats, 2015). The global trade of soybeans and soybean products has surpassed global trade in either wheat or total coarse grains (corn, barley, sorghum, rye, oats, millet, and mixed grains) (Westcott and Hansen, 2015).

Soybean Biology and Value as an Oil Crop

Soybean is valued in the USA for its uses as meal and oil. One metric ton of soybeans produced 0.8 metric tons of meal and 209 liters of biodiesel (Soystats, 2015). Soybean meal is a very important protein source for livestock. The combined region of Africa and the Middle East was projected to increase import of soybean meal by almost 20% by 2024. China was expected to increase its world import share of soybean from 65 to 71% by 2024 (Westcott and Hansen, 2015). Soybean oil is used for human consumption and commercial applications. In 1999, the USA produced 1.9 million liters of biodiesel and then grew to 6,615 million liters in 2014. Soybean oil accounted for 55% of the U.S. vegetable oil consumption in 2014 (Soystats, 2015).

While popular, soybean oil cannot always meet optimal physical, chemical, and nutritional properties, since these properties are not necessarily mutually compatible (Gunstone, 2006). Soybean oil is often hydrogenated to improve oxidative stability for both food and industrial applications, however, this increases saturated and *trans* fatty acids. Oil blending and interesterification are two other methods to reach desired physical, chemical, and/or nutritional properties (United Soybean Board, 2014). Genetic changes to soybean can also be utilized to direct the fatty acid development to the desired composition and functionality.

Increased oleic acid content of soybean oil will enhance its applications to the food industry. Soybean competes with a number of other oils for consumption in the USA. In 2014, soybean oil accounted for 55% of the vegetable oil consumed in the USA, with rapeseed oil at 16%, corn oil at 11%, palm oil at 8%, coconut oil at 3%, and the remaining 7% was other oils such as palm kernel, peanut, sunflower, and safflower oils (Soystats, 2015). Each oil has its own unique fatty acid profile (Table 1.1), which lends different physical, chemical, and nutritional qualities. Canola oil is the major competitor of soybean oil in the USA, and has undergone significant genetic changes through genetic modification and breeding efforts to lower erucic acid, increase oleic acid, and decrease linoleic acid with a current profile of 5% palmitic, 1% stearic, 59% oleic, 22% linoleic, and 11% linolenic acids (Abbadi and Leckband, 2011; Scarth and Tang, 2006; USDA, 2015). Corn, a major crop in the USA, has an oil fatty acid profile of

11% palmitic, 2% stearic, 25% oleic, 58% linoleic, and 1% linolenic acids (Duckett et al., 2009; USDA, 2015). Olive oil, valued for its flavor and health benefits, has a fatty acid profile of 7–20% palmitic, 0–3.5 % stearic, 56–83% oleic, 3–20% linoleic, and <1.5% linolenic acids (Gutfinger, 1981; USDA, 2015). Peanut, an important crop to the southeastern USA, has a fatty acid profile of 10% palmitic, 4% stearic, 59% oleic, 20% linoleic, and 1% linolenic acids, with the remaining 4% attributed to small quantities of other fatty acids (Carrin and Carelli, 2010).

Soybean oil also has industrial uses in addition to food applications. However, poor oxidative stability makes soy oil a poor choice for biodiesel. A greater amount of polyunsaturated fatty acids such as linoleic and linolenic acids would improve the oil’s ability to perform in cold weather, preventing it from ‘gelling’ (Fallen et al., 2012). There is a direct inverse relationship between oxidative stability and biodiesel cold performance (Clemente, 2005), so improving soybean oil quality will benefit industrial applications.

Oil Extraction Process

In order to extract the oil from soybean seeds, the seed must first be cleaned of the hulls and any other debris. The cleaned seeds are then coarsely ground to increase surface area before solvents are used for extraction. The most common solvent used is the volatile hydrocarbon hexane, which is later distilled out of the oil (Kim and Yoon, 1990; Sharma et al., 2002; Wahl and Colburn, 2009). The oil then undergoes several refining steps depending on end-use applications; degumming, bleaching (for use in cooking), winterization (for use in refrigerated goods), and heating with alkaline and acidic substances to remove color, odor, and bitterness (Wiedermann, 1981).

Hydrogenation

Once soybean oil is extracted, it undergoes the hydrogenation process to improve shelf life and food applications, such as margarine and baking uses (Ghotra et al., 2002). Hydrogenation improves oxidative stability by removing carbon-carbon double bonds and increasing saturated and solid fat content (Clemente and Kinney, 2005; Ghotra et al., 2002; Gunstone, 2006; Pereira et al., 2013; Veldsink et al., 1997). In simple chemistry terms, oxidation is the loss of electrons. In soybean oil, oxidation is an interaction with oxygen which causes lipid peroxides and is responsible for rancidification. Oxidative stability is the ability to resist rancidification. Certain oils are more resistant to rancidification than others; the more double bonds a fatty acid has, the less stable it is (Liu, 1997). Fully hydrogenated soybean oil has a high C18:0 content (~850 g kg^{-1}), but this saturated fatty acid has none of the adverse effects on cardiovascular diseases reported for other saturated fats, such as palmitic acid (Ribeiro et al., 2009).

The most common method to hydrogenate soybean oil is the Horiuti-Polanyi mechanism (Dijkstra, 2006; Horiuti and Polanyi, 1934; Singh et al., 2011; Singh et al., 2009; Yang et al., 2013). While there may be differences in the catalyst surface used in this method, the Horiuti-Polanyi mechanism works by dissociating the molecular hydrogen from the catalyst surface, and these atomic hydrogens then interact with the fatty acids in soybean oil (Mattson et al., 2013; Singh et al., 2011; Yang et al., 2013). The success of this process is due to the amount of atomic hydrogen present. A high concentration of hydrogen helps complete hydrogenation, however, a lower concentration favors incomplete hydrogenation resulting in the production of *trans* fatty acids (Dijkstra, 2006).

Soybean oil is typically hydrogenated as a slurry in a batch autoclave over nickel based catalyst. Reaction conditions are between 110–190°C, 30–70 psi hydrogen pressure, and with 0.01–0.15 wt. % nickel (Singh et al., 2011; Veldsink et al., 1997). Conventional hydrogenation methods produce upwards of 10 wt. % *trans* fatty acids, but there has been work to improve hydrogenation methods resulting in <4 wt. % *trans* fatty acids. Singh et al. (2009) was able to achieve 3.5 wt.% *trans* fatty acid by using platinum instead of nickel at a much lower temperature, 70°C, and at 65 psi hydrogen pressure. It should be considered that heat treatments are expensive and not entirely effective (Hildebrand and Hymowitz, 1981; Reinprecht et al., 2006).

Trans-fatty Acids

The isomers of molecules are differentiated by carbon-carbon double bonds in two physical configurations: *trans* or *cis* (Thomas, 2000). *Cis* indicates that carbon groups on either side of a carbon-carbon double bond are situated on the same plane of symmetry, while *trans* means they are across from each other on the plane of symmetry of the carbon-carbon double bond (Thomas, 2000). Even with the improved methods to reduce *trans* fatty acids, any amount may still be too much to outweigh the commercial benefits of hydrogenation. In 2006, the U.S. Food and Drug Administration began requiring *trans* fat labeling and in 2013 affirmed that *trans* fats are not generally regarded as safe (Food and Drug Administration, 2013).

Blending and Interesterification

Full-hydrogenation is used in order to eliminate *trans* fat production. Two ways to get around the need to fully hydrogenate soybean oil are blending and interesterification. Blending combines fully-hydrogenated and non-hydrogenated soybean oil. This method avoids the need to fully hydrogenate all the soybean oil and can still create a final product with desired functional properties (United Soybean Board, 2014). Interesterification is an alternative process to hydrogenation, and can be either chemical or enzymatic (Farfán et al., 2015; Ribeiro et al., 2009).

There are several oil modification methods to achieve improved oil performance, but by genetically modifying soybeans to produce oil with more desirable composition, post-harvest modification will become unnecessary. A significant amount of research has already gone into understanding the fatty acid biosynthesis pathway and the genetic control involved in soybean oil production, highlighted in the next section.

Identification of Genes Controlling Fatty Acids

Soybean oil is made up of five major fatty acids: linoleic (52%), oleic (25%), palmitic (11%), linolenic (8%), and stearic (4%) acids (Fehr, 2007). Palmitic acid (16:0), a saturated fat, is synthesized into stearic acid (18:0) by elongase. The Δ9-desaturase then catalyzes the stearic acid into oleic acid (18:1), a monounsaturated fatty acid. Oleic acid is converted by Δ12-desaturase into linoleic acid (18:2), which is further converted by Δ6-desaturase to α-linolenic acid (18:3), both polyunsaturated fatty acids (Putten, 2009) (Figure 1.1).

Two identified genes, *FAD2-1A* and *FAD2-1B*, have been reported to increase oleic acid content in soybean seed oil (Pham et al., 2010). The *FAD2* family of genes control the desaturation of oleic acid and the mutant alleles of these genes prevent the desaturation, thus causing a build-up of oleic acid (Okuley et al., 1994). Cardinal et al. (2007) showed that mutant

FATB1a allele was responsible for low palmitic acid in N87-2122-4. The mutant alleles of *FAD3* loci were reported to control low linolenic acid (Bilyeu et al., 2011).

Palmitic Acid

Palmitic acid is a major saturated fatty acid in soybean oil, and reduction of this fatty acid in soybean seed oil would improve the oil quality for human consumption. A growing body of evidence is challenging whether saturated fat is directly associated with coronary events (Puaschitz et al., 2015), however, the American Heart Association still recommended limiting saturated fat consumption, and replacing it with mono and polyunsaturated fats.

Previous genetic studies identified *fap1* and *fap3* loci to be responsible for decreasing palmitic acid in soybean lines C1726 (85 g kg$^{-1)}$ and A22 (68 g kg^{-1}), respectively (Erickson et al., 1988; Lee et al., 2007; Schnebly et al., 1994). A combination of mutant alleles at these two loci resulted in <40 g kg^{-1} (Fehr, 2007). A mutation at the *fap2* locus identified in high palmitic acid line C1727 (172 g kg^{-1}) (Wilcox and Cavins, 1990) was related to reduced KAS-II activity rather than or in addition to increased 16:0-ACP thioesterase activity (Wilson et al., 2001). Previous research by Wilson et al. (2001) indicated that the fap_{nc} allele was included in a deletion of 16:0-ACP thioesterase enzyme in low palmitic acid line N87-2122-4 (50 g kg^{-1}).

Soybean line N87-2122-4 has palmitic acid of 53 g kg^{-1} and is an F_6-derived line from a cross between N78-2245 and N79-2077 (Burton et al., 1994). Parental line N78-2245 has normal palmitic acid levels (100 g kg^{-1}), but increased oleic acid (383 g kg^{-1}) and decreased linoleic (426 g kg^{-1}) and linolenic acids (55 g kg^{-1}). In contrast, parental line N79-2077 has decreased palmitic acid (60 g kg^{-1}), which was derived from parents PI 90406, PI 92567, and N69-2774 in a fifth cycle of a high-oleic recurrent selection experiment (Burton et al., 1994).

Since its release in 1994, soybean line N87-2122-4 has been a source for the low palmitic trait in soybean for improving nutritional quality of soybean oil. The allele conferring reduced palmitic acid from N79-2077 was designated as fap_{nc} (Wilson et al., 2001). Using F_2 and $F_{2:3}$ populations derived from 'Benning' x N87-2122-4 and SSR markers, one major and one minor QTL mapped on Linkage Group (LG) A1 (Chr 5) and on LG M (Chr 7), respectively, and were reported to be associated with reduced palmitic acid in soybean N87-2122-4 (Li et al., 2002). These QTL have an additive effect and explained 51% of the total phenotypic variation for palmitic acid content in the F_2 generation and 43% of the variation in the $F_{2:3}$ families. By isolating the full length cDNAs from three unique *FATB* genes, *GmFATB1a*, *GmFATB1b*, and *GmFATB2a*, Cardinal et al. (2007) reported that the isoform designated *GmFATB1a* represented the specific gene deleted in the lines carrying the fap_{nc} allele, which was responsible for reduced palmitic acid; however, the location and size of the genomic deletion was unknown.

Stearic Acid

Stearic acid is an 18-carbon chain saturated fatty acid. While being the minor saturated fat of soybean oil, its value should not be overlooked. Stearic acid has health advantages over palmitic and other saturated fatty acids in baking and other food industry applications (Kris-Etherton and Yu, 1997; Thijssen et al., 2005). Stearic acid has a neutral effect on blood cholesterol concentration. When stearic acid is used to replace other saturated fats, there is evidence of positive effects on low-density lipoprotein cholesterol levels (Hunter et al., 2010; Kris-Etherton and Yu, 1997). However, there are conflicting reports to these claims. When high stearic acid diets replaced oleic acid, low density lipoprotein (LDL) cholesterol was increased (Denke and Grundy, 1991; Judd et al., 2002; Kris-Etherton et al., 1993), high density lipoproteins (HDL) were decreased, (Judd et al., 2002), and triglycerides were increased (Denke and Breslow, 1988; Judd et al., 2002). It has also been reported that substituting high stearic acid for linoleic acid increased LDL cholesterol (Kris-Etherton et al., 1993; Zock and Katan, 1992), decreased HDL cholesterol (Zock and Katan, 1992), increased triglycerides (Kris-Etherton et al., 1993; Zock and Katan, 1992), and increased the total cholesterol over HDL ratio (Hunter et al., 2000; Kris-Etherton et al., 1993; Thijssen and Mensink, 2005; Zock and Katan, 1992). Beyond the health aspects to food industry applications, stearic acid has many industrial uses. Stearic acid is used as a lubricant in paper making (Flynn et al., 1996) and in injection molding and pressing of ceramic powders (Tseng et al., 1999). It is also of major use in the production of rubber for the automotive industry (Beniska and Dogadkin, 1959; Dogadkin and Beniska, 1958; Luyt, 1993).

A Δ9-stearoyl-acyl carrier protein desaturase (SACPD) adds a double bond at C9 position, converting stearic acid into oleic acid (18:1) (Pantalone et al., 2002; Wilson, 2004), which has been localized to the *Fas* locus on LG B2 (Chr 14) (Spencer et al., 2003). Three forms of SACPD, have been reported: SACPD-A, SACPD-B (Byfield and Upchurch, 2007), and SACPD-C (Zhang et al., 2008). Increased levels of stearic acid have been associated with molecular changes to SACPD, preventing stearic acid from being converted to oleic acid. A mutation in SACPD-A increased stearic acid in soybean line A6 to 300 g kg^{-1}, a mutation in SACPD-B increased stearic acid in soybean line FA41545 to 150 g kg^{-1}, and a mutation in SACPD-C increased stearic acid in soybean line A81-606085 to 190 g kg^{-1} (Boersma et al., 2012; Graef et al., 1985; Hammond and Fehr, 1983).

Oleic Acid

Oleic acid is a monounsaturated fatty acid (MUFA) and has a naturally occurring *cis* configuration about the 9th carbon-carbon double bond (Thomas, 2000), which can be found in most plant and animal oils. This MUFA occurs in soybean seed and has been of great interest due to its effects on human health when consumed. A diet high in MUFAs resulted in a 6.3% reduction in low density lipoprotein (LDL) when compared to an average American diet (Berglund et al., 2007). The *FAD2* family of genes is related to the quantity of oleic acid by interacting with fatty acid Δ12-desaturase. By preventing the desaturation of oleic acid to linoleic acid, oleic acid is increased which results in a significant decrease in consuming LDL. One *FAD2-1A* (Glyma10g42470) mutation was isolated from an M_2 generation of γ-irradiated soybean and was mapped to the Chr 10 (Bachlava et al., 2008). *FAD2-1B* (Glyma20g24530) is another gene related to oleic acid synthesis that worked in the same way as *FAD2-1A* and was mapped to Chr 20 (Bachlava et al., 2008; Bachlava et al., 2009). It was suggested that both *FAD2-1A* and *FAD2-1B* are duplicated genes due to evidence that soybean is a paleopolyploid crop (Anai et al., 2008; Schlueter et al., 2007).

Multiple mutations in *FAD2-1A* have been reported, including the deletion of the *FAD2-1A* gene in soybean line M23 (Alt et al., 2005) and the missense mutation in *FAD2-1A* from line 17D (*FAD2-1A* S117N) (Dierking and Bilyeu, 2009). Pham et al. (2010) identified the two missense mutant alleles at *FAD2-1B* locus from PI 283327 (*FAD2-1B* P137R) and PI 567189A (*FAD2-1B* I143T) and reported that a combination of mutant alleles at *FAD2-1A* and *FAD2-1B* loci could produce 800 g kg^{-1} oleic acid in soybean lines. This has created the opportunity to develop high oleic soybean cultivars through conventional breeding methods.

Linoleic Acid

Linoleic acid is the major polyunsatured fat (PUFA) in soybean, with a double bond at C9 and another at C12. A Δ12-desaturase catalyzes oleic acid into linoleic acid, which is further transformed into linolenic acid by a Δ15-desaturase. Linoleic acid is hydrogenated to either stearic or *trans* fatty acids, which may increase LDL and decrease HDL cholesterol relative to linoleic acid itself (Zock and Katan, 1992). When 7% of dietary energy of stearic acid was replaced by linoleic acid, a decrease in ex vivo platelet aggregation and therefore an increase in aggregation time was reported in men (Thijssen et al., 2005) which is negatively associated with mortality from coronary heart disease (Hornstra and ten Hoor, 1975). Similarly, Zock and Katan (1992) replaced 8% of dietary energy with either *trans* fatty acid, stearic acid, or linoleic acid and found that men that consumed the linoleic diet had lower LDL cholesterol and triglycerides than those consuming the stearic acid diet, however, there was not an observable difference in women. Many of the same genes utilized to increase oleic acid consequently reduce linoleic acid.

Linolenic Acid

Linolenic acid is a PUFA that makes up 8% of the oil composition of soybean seed and partially responsible for poor stability of soy oil which necessitates hydrogenation. It has three double bonds with *cis* configurations located at the 9th, 12th, and 15th carbons (Thomas, 2000). All members of *FAD3* gene family (*FAD3A*, *FAD3B*, and *FAD3C*) influences the production of linolenic acid, but the *FAD3A* was found to have the greatest impact in soybean seed (Bilyeu et al., 2011). Its function is to block Δ15-desaturase from converting linoleic acid to linolenic acid, resulting in low linolenic content in soybean seed.

The first reported low linolenic acid mutant line was C1640, which had an average linolenic acid content of 34 g kg^{-1} (Wilcox and Cavins, 1986). Chappell and Bilyeu (2006) were able to determine that the low linolenic trait of line C1640 was due to a nonsense mutation resulting in a premature stop codon in the omeg-3 fatty acid desaturase gene. Other low linolenic lines have been developed using chemical mutagenesis. Using seed treatment with chemical mutagens such as ethyl methanesulfonate (EMS), Hammond and Fehr (1983) were able to identify a locus in soybean line A5 responsible for reduced linolenic acid content (35 to 41 g kg^{-1}). Further study determined that the reduced linolenic acid content in this soybean line was conditioned by a *fan1* allele (Rennie and Tanner, 1991). Subsequently, Fehr and Hammond (2000) found a f*an2* allele in soybean line A23. Later, the third locus responsible for linolenic acid content designated as *fan3* was found in A26, which was developed by treating soybean line A89-144003 with EMS (Fehr and Hammond, 2000). By combining the three alleles- *fan1* (A5), *fan2* (A23), and *fan3* (A26) through hybridization and selection, soybean line A29 was developed to have the lowest linolenic acid content to date (<20 g kg^{-1}) (Fehr and Hammond, 2000; Ross et al., 2000). The low linolenic acid soybean line CX1512-44 was also developed using a mutagenesis approach by Dr. James Wilcox at Purdue University (Bilyeu et al., 2005). Bilyeu et al. (2011) found a combination of two mutant genes, *FAD3A* and *FAD3C*, in CX1512-44 reduced linolenic acid content to about 30 g kg^{-1}. In addition, they were able to confirm the association of the linolenic acid phenotypes with the *FAD3* genotypes in the progeny of crosses derived from CX1512-44.

Genotype by Environment Interactions of Fatty Acids

The production and levels of fatty acids in soybean vary depending on the environment where the plants grow. Specifically, the time during pod fill is the most influential stage, with cooler temperatures causing a decrease in oleic acid (Zuil et al., 2012). A 5-year study in Europe indicated that soybean seed oil content was more affected by the yearly conditions than the 15 different locations (Sudaric et al., 2005). Another study in Canada reported similar results, with year accounting for a large proportion of variation in fatty acid content (Primomo et al., 2002). Primomo et al. (2002) also suggested that, of all measured environmental factors, temperature played a significant role in oleic acid concentrations. Wolf et al. (1982) reported a positive correlation between soybean seed oil content and temperature during seed development. They indicated that as temperature increased, linoleic and linolenic acids decreased and oleic acid increased, also noting that palmitic and stearic acids remained essentially the same.

Genetic Improvement of Fatty Acids and Yield

Genetic improvement of fatty acids in soybean is one of important breeding goals. Currently, the main focus on genetic improvement of fatty acids is for oleic, linolenic, and palmitic acids of soybean. La et al. (2014) compared seed yield, fatty acids, and total oil and protein concentrations between 40 $F_{4:7}$ high oleic lines carrying both mutant *FAD2-1A* and *FAD2-1B* alleles and 40 normal oleic acid lines derived from six crosses in six environments. Their results indicated that the yield between high and normal oleic acid lines varied among the populations, but seed yield of the high oleic lines were within 2% of the normal oleic lines. They also found that the high oleic lines on average produced >790 g kg^{-1} oleic acid and were significantly higher in protein and oil contents than normal oleic lines in all six populations (La et al., 2014).

Transgenic high oleic soybean lines have been developed by silencing *FAD2* genes, and these lines have shown to have a consistent oleic acid concentration across multiple environments without affecting yield (Brace et al., 2011). Kinney (1996) studied the yield and oleic acid levels (840 – 880 g kg^{-1}) of lines containing the event 260-05 across three generations over 2 years and concluded that the high oleic phenotype did not negatively affect yield. Kinney (1996) also found that elite conventional lines yielded about 3,564 kg ha^{-1} and the high oleic lines with event 260-05 yielded 3,665 kg ha^{-1}.

DuPont Pioneer developed Plenish® high oleic soybean cultivars using a transgenic approach. These high oleic soybean cultivars could reach 750 g kg^{-1} oleic acid content. The research plot yield data from 346 observations in Indiana and Ohio indicated an average yield on Plenish® high oleic soybean commercial cultivars was similar to the average yield of Pioneer® brand elite soybean checks in 2011–2013 (www.plenish.com). Similarly, Monsanto Company has developed VistiveGold® high oleic soybean cultivars with a fatty acid profile of 720 g kg^{-1} oleic, 30 g kg^{-1} linolenic, and 60 g kg^{-1} palmitic acids. Yield tests indicated that VistiveGold® soybean cultivars were competitive with the proven Monsanto soybean varieties (www.vistivegold.com).

Control of palmitic acid in vegetable oils is under the command of 16:0-ACP thioesterase activity. After comparing the yield between F_4–derived lines with low and normal palmitic acids from two populations of ‘Corsica’ × N97-3681-11 and ‘Brim’ × N97-3708-13, Cardinal et al. (2007) concluded that the negative influence of the fap_{nc} allele on seed yield and plant height had not been eliminated through breeding. Therefore, innovative breeding approaches are needed to develop soybean lines with competitive yield and desired fatty acid profile.

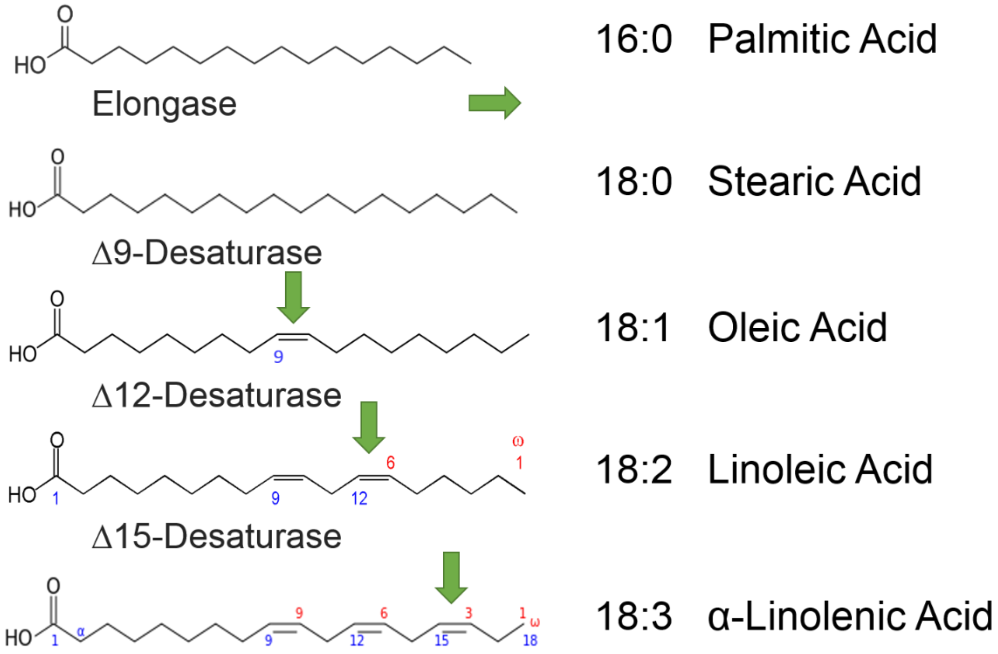

Figure 1.1 Diagram of fatty acid biosynthesis pathway.

Table 1.1 Summary of typical fatty acid composition of major popular oils used in the USA†

Oil type	Palmitic acid (16:0)	Stearic acid (18:0)	Oleic acid (18:1)	Linoleic acid (18:2)	Linolenic acid (18:3)	Others
	g kg^{-1}					
Canola (classic)	40	10	130	140	90	590
Canola (low erucic acid)	50	10	590	220	110	0
Coconut	70	20	70	0	0	840
Corn	110	20	250	580	10	0
Olive	70–200	0–35	560–830	30–200	<15	0
Palm	400	50	410	120	10	20
Peanut	100	40	590	200	10	40
Safflower	20–100	10–100	70–420	550–810	<10	0
Soybean	110	40	250	520	80	0
Soybean (fully-hydrogenated)	120	870	0	0	0	10
Sunflower	70	50	230	640	10	20

† Table was modified from Veldsink et al. (1997).

References

Abbadi, A. and G. Leckband. 2011. Rapeseed breeding for oil content, quality, and sustainability. Eur. J. Lipid Sci. Tech. 113:1198–1206.

Alt, J., W. Fehr, G. Welke, and D. Sandhu. 2005. Phenotypic and molecular analysis of oleate content in the mutant soybean line M23. Crop Sci. 45:1997–2000.

Anai, T., T. Yamada, R. Hideshima, T. Kinoshita, S.M. Rahman, and 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.

Bachlava, E., R.E. Dewey, J. Auclair, S. Wang, J.W. Burton, and A.J. Cardinal. 2008. Mapping genes encoding microsomal ω-6 desaturase enzymes and their cosegregation with QTL affecting oleate content in soybean. Crop Sci. 48:640–650.

Bachlava, E., R.E. Dewey, J.W. Burton, and A.J. Cardinal. 2009. Mapping and comparison of quantitative trait loci for oleic acid seed content in two segregating soybean populations. Crop Sci. 49:433–442.

Beniska, J. and B. Dogadkin. 1959. The influence of activators on the vulcanization process. II. The influence of zinc oxide on the structure of vulcanizates. Rubber Chem. Tech. 32:780–784.

Berglund, L., M. Lefevre, H.N. Ginsberg, P.M. Kris-Etherton, P.J. Elmer, P.W. Stewart, A. Ershow, T.A. Pearson, B.H. Dennis, P.S. Roheim, R. Ramakrishnan, R. Reed, K. Stewart, K.M. Phillips, and D. Investigators. 2007. Comparison of monounsaturated fat with carbohydrates as a replacement for saturated fat in subjects with a high metabolic risk profile: studies in the fasting and postprandial states. Am. J. Clin. Nutr. 86:1611–1620.

Bilyeu, K., J.D. Gillman, and A.R. LeRoy. 2011. Novel mutant allele combinations produce soybeans containing 1% linolenic acid in the seed oil. Crop Sci. 51:259.

Bilyeu, K., L. Palavalli, D. Sleper, and P. Beuselinck. 2005. Mutations in soybean microsomal omega-3 fatty acid desaturase genes reduce linolenic acid concentration in soybean seeds. Crop Sci. 45:1830–1836.

Bilyeu, K.D., J.D. Gillman, and A. LeRoy. 2011. Novel *FAD3* mutant allele combinations produce soybeans containing 1% linolenic acid in the seed oil. Crop Sci. 51:259–264.

Boersma, J.G., J.D. Gillman, K.D. Bilyeu, G.R. Ablett, C. Grainger, and I. Rajcan. 2012. New mutations in a delta-9-stearoyl-acyl carrier protein desaturase gene associated with enhanced stearic acid levels in soybean seed. Crop Sci. 52:1736–1742.

Brace, R.C., W.R. Fehr, and S.R. Schnebly. 2011. Agronomic and seed traits of soybean lines with high oleate concentration. Crop Sci. 51:534–541.

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.

Byfield, G.E. and R.G. Upchurch. 2007. Effect of temperature on delta-9 stearoyl-ACP and microsomal omega-6 desaturase gene expression and fatty acid content in developing soybean seeds. Crop Sci. 47:1698–1704.

Cardinal, A.J., J.W. Burton, A.M. Camacho-Roger, J.H. Yang, R.F. Wilson, and R.E. Dewey. 2007. Molecular analysis of soybean lines with low palmitic acid content in the seed oil. Crop Sci. 47:304.

Carrin, M.E. and A.A. Carelli. 2010. Peanut oil: compositional data. Eur. J. Lipid Sci. Technol. 112:697–707.

Chappell, A.S. and K.D. Bilyeu. 2006. A *GmFAD3A* mutation in the low linolenic acid soybean mutant C1640. Plant Breeding 125:535–536.

Clemente, T.E. and A.J. Kinney. 2005. Modifying soybean oil for enhanced performance in biodiesel blends. Fuel Process Tech. 86:1137–1147.

Denke, M.A. and J.L. Breslow. 1988. Effects of a low fat diet with and without intermittent saturated fat and cholesterol ingestion on plasma lipid, lipoprotein, and apolipoprotein levels in normal volunteers. J. Lipid Res. 29:963–969.

Denke, M.A. and S.M. Grundy. 1991. Effects of fats high in stearic acid on lipid and lipoprotein concentrations in men. Am. J. Clin. Nutr. 54:1036–1040.

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

Dijkstra, A.J. 2006. Revisiting the formation of *trans* isomers during partial hydrogenation of triacylglycerol oils. Eur. J. Lipid Sci. Tech. 108:249–264.

Dogadkin, B. and I. Beniska. 1958. The action of vulcanization activators. Rubber Chem. Technol. 31:329–342.

Duckett, S.K., S.L. Pratt, and E. Pavan. 2009. Corn oil or corn grain supplementation to steers grazing endophyte-free tall fescue. II. Effects on subcutaneous fatty acid content and lipogenic gene expression. J. Anim. Sci. 87:1120–1128.

Erickson, E.A., J.R. Wilcox, and J.F. Cavins. 1988. Inheritance of altered palmitic acid percentages in two soybean mutants. J. Hered. 79:465–468.

Fallen, B.D., 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.

Farfán, M., A. Alvarez, A. Gárate, and P. Bouchon. 2015. Comparison of chemical and enzymatic interesterification of fully hydrogenated soybean oil and walnut oil to produce a fat base with adequate nutritional and physical characteristics. Food Technol. Biotechnol. 53:361–366.

Fehr, W.R. 2007. Breeding for modified fatty acid composition in soybean. Crop Sci. 47:S72-S87.

Fehr, W.R. and E.G. Hammond. 2000. Reduced linolenic acid production in soybean. U.S. Patent 6,133,509. Date issued: 17 October.

Flynn, G.J., J.M. Krasniewski, and J. Wilson. 1996. Lubricant additives for paper coating compositions. US Patent 5,527,383. Date issued: 18 June.

Food and Drug Administration, Department of Health and Human Services. 2013. Tentative determination regarding partially hydrogenated oils; request for comments and scientific data and information. Federal Register 78(127):67,169–67,175

Ghotra, B.S., S.D. Dyal, and S.S. Narine. 2002. Lipid shortenings: a review. Food Res. Int. 35:1015–1048.

Gibson, L. and G. Benson. 2005. Origin, history, and uses of soybean (*Glycine max*). Iowa State University, Department of Agronomy. http://agron-www.agron.iastate.edu.

Graef, G.L., W.R. Fehr, and E.G. Hammond. 1985. Inheritance of three stearic acid mutants of soybean. Crop Sci. 25:1076–1079.

Gunstone, F.D. 2006. Modifying lipids - why and how? In: Modifying lipids for use in food, editor, F.D. Gunstone. CRC Press LLC, Boca Raton, FL, USA. P. 1–8.

Gutfinger, T. 1981. Polyphenols in olive oils. J. Am. Oil Chem. Soc. 58:966–968.

Hammond, E.G. and W.R. Fehr. 1983. Registration of A5 germplasm line of soybean (reg. No. GP44). Crop Sci. 23:192–192.

Hammond, E.G. and W.R. Fehr. 1983. Registration of A6 germplasm line of soybean. Crop Sci. 23:192–193.

Hart, J.F. 1986. Change in the Corn Belt. Geogr. Rev. 58:51–72.

Hildebrand, D.F. and T. Hymowitz. 1981. Two soybean genotypes lacking lipoxygenase-1. J. Am. Oil Chem. Soc. 58:583–586.

Horiuti, I. and M. Polanyi. 1934. Exchange reactions of hydrogen on metallic catalysts. Transactions of the Faraday Society 30:1164–1172.

Hornstra, G. and F. ten Hoor. 1975. The filtragometer: a new device for measuring platelet aggregation in venous blood of man. Thromb. Haemostasis 34:531–544.

Hunter, J.E., J. Zhang, and P.M. Kris-Etherton. 2010. Cardiovascular disease risk of dietary stearic acid compared with *trans*, other saturated, and unsaturated fatty acids: a systematic review. Am. J. Clin. Nutr. 91:46–63.

Hunter, K.A., L.C. Crosbie, A. Weir, G.J. Miller, and A.K. Dutta-Roy. 2000. A residential study comparing the effects of diets rich in stearic acid, oleic acid, and linoleic acid on fasting blood lipids, hemostatic variables and platelets in young healthy men. J. Nutr. Biochem. 11:408–416.

Hymowitz, T. 1970. On the domestication of soybean. Econ. Bot. 24:408–421.

Hymowitz, T. and J.R. Harlan. 1983. Introduction of soybean to North America by Samuel Bowen in 1765. Econ. Bot. 37:371–379.

Hymowitz, T., and W.R. Shurtleff. 2005. Debunking soybean myths and legends in the historical and popular literature. Crop Sci. 45:473–476.

Judd, J.T., D.J. Baer, B.A. Clevidence, P. Kris-Etherton, R.A. Muesing, and M. Iwane. 2002. Dietary *cis* and *trans* monounsaturated and saturated FA and plasma lipids and lipoproteins in men. Lipids. 37:123–131.

Kim, I. and S. Yoon. 1990. Effect of extraction solvents on oxidative stability of crude soybean oil. J. Am. Oil Chem. Soc. 67:165–167.

Kinney, A.J. 1996. Development of genetically engineered soybean oils for food applications. J. Food Lipids 3:273–292.

Kris-Etherton, P.M., J. Derr, D.C. Mitchell, V.A. Mustad, M.E. Russell, E.T. McDonnell, D. Salabsky, and T.A. Pearson. 1993. The role of fatty acid saturation on plasma lipids, lipoproteins, and apolipoproteins: I. Effects of whole food diets high in cocoa butter, olive oil, soybean oil, dairy butter, and milk chocolate on the plasma lipids of young men. Metabolism 42:121–129.

Kris-Etherton, P.M. and S. Yu. 1997. Individual fatty acid effects on plasma lipids and lipoproteins: human studies. Am. J. Clin. Nutr. 65:1628S–1644S.

La, T.C., S.M. Pathan, T. Vuong, J.D. 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.

Lee, J.D., K.D. Bilyeu, and J.G. Shannon. 2007. Genetics and breeding for modified fatty acid profile in soybean seed oil. J. Crop Sci. Biotech. 10:201–210.

Li, Y.H., W. Li, C. Zhang, L. Yang, R.-Z. Chang, B.S. Gaut, and L.J. Qiu. 2010. Genetic diversity in domesticated soybean (*Glycine max*) and its wild progenitor (*Glycine soja*) for simple sequence repeat and single-nucleotide polymorphism loci. New Phytol. 188:242–253.

Li, Z., R.F. Wilson, W.E. Rayford, and H.R. Boerma. 2002. Molecular mapping genes conditioning reduced palmitic acid content in N87-2122-4 soybean. Crop Sci. 42:373–378.

Liu, K., editor. 1997. Soybeans: chemistry, technology, and utilization. International Thomson Publishing Asia, Singapore. doi:10.1007/9781461517634

Luyt, A.S. 1993. A differential scanning calorimetric study of the interaction of zinc benzothiazole-2-thiolate, sulfur, stearic acid, and zinc stearate in the absence and presence of polyisoprene. J. Appl. Polym. Sci. 47:2177–2185.

Mattson, B., W. Foster, J. Greimann, T. Hoette, N. Le, A. Mirich, S. Wankum, A. Cabri, C. Reichenbacher, and E. Schwanke. 2013. Heterogeneous catalysis: the Horiuti–Polanyi mechanism and alkene hydrogenation. J. Chem. Educ. 90:613–619.

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

Pantalone, V.R., R.F. Wilson, W.P. Novitzky, and J.W. Burton. 2002. Genetic regulation of elevated stearic acid concentration in soybean oil. J. Am. Oil Chem. Soc. 79:549–553.

Pereira, G.G., S. Marmesat, D. Barrera-Arellano, and M.C. Dobarganes. 2013. Evolution of oxidation in soybean oil and its biodiesel under the conditions of the oxidation stability test. Int. J. Fats Oils 64:482–488.

Pham, A.T., J.D. 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 Biology 10:195.

Primomo, V.S., D.E. Falk, G.R. Ablett, J.W. Tanner, and I. Rajcan. 2002. Genotype x environment interactions, stability, and agronomic performance of soybean with altered fatty acid profiles. Crop Sci. 2:37–44.

Puaschitz, N.G., E Strand, T.M. Norekval, J. Kierkes, L. Dahl, G.Frodahl, T. Svingen, J. Assmus, H. Schartum-Hansen, J. Oyen, E.K.R. Pedersen, C.A. Drevon, G.S. Tell, and O. Nygard. 2015. Dietary intake of saturated fat is not associated with risk of coronary events or mortality in patients with established coronary artery disease. J. Nutr. 145(2):299–305.

Putten, A.V.V.T.H., T. König, D. Fabritius, T. Kiy, A. Leske, and J. Plote. 2009. Method for producing γ-linolenic acids from a ciliate culture by adding suitable precursor molecules to said culture medium. U.S. Patent 7,476,522. Date issued: 13 January.

Reinprecht, Y., V.W. Poysa, I. Rajcan, G.R. Ablett, and K.P. Pauls. 2006. Agronomic performance of soybean with seed lipoxygenase nulls and low linolenic acid content. Can. J. Plant Sci. 86:379–387.

Rennie, B., D. and J.W. Tanner. 1991. New allele at the *Fan* locus in the soybean line A5. Crop Sci. 31:297–301.

Ribeiro, A.P.B., R.C. Basso, R. Grimaldi, L.A. Gioielli, A.O. dos Santos, L.P. Cardoso, and L.A. Guaraldo Gonçalves. 2009. Influence of chemical interesterification on thermal behavior, microstructure, polymorphism and crystallization properties of canola oil and fully hydrogenated cottonseed oil blends. Food Res. Int. 42:1153–1162.

Ribeiro, A.P.B., R. Grimaldi, L.A. Gioielli, and L.A.G. Gonçalves. 2009. Zero *trans* fats from soybean oil and fully hydrogenated soybean oil: physico-chemical properties and food applications. Food Res. Int. 42:401–410.

Ross, A.J., W.R. Fehr, G.A. Welke, and S.R. Cianzio. 2000. Agronomic and seed traits of 1%-linolenate soybean genotypes. Crop Sci. 40:383–386.

Scarth, R. and J. Tang. 2006. Modification of oil using conventional and transgenic approaches. Crop Sci. 46:1225–1236.

Schlueter, J.A., J.Y. Lin, S.D. Schlueter, I.F. Vasylenko-Sanders, S. Deshpande, J. Yi, M. O'Bleness, B.A. Roe, R.T. Nelson, B.E. Scheffler, S.A. Jackson, and R.C. Shoemaker. 2007. Gene duplication and paleopolyploidy in soybean and the implications for whole genome sequencing. BMC Genomics 8:330.

Schnebly, S.R., W. Fehr, G. Welke, E.G. Hammond, and D.N. Duvick. 1994. Inheritance of reduced and elevated palmitate in mutant lines of soybean. Crop Sci. 34:829–833.

Sharma, A., S.K. Khare, and M.N. Gupta. 2002. Three phase partitioning for extraction of oil from soybean. Bioresource Technol. 85:327–329.

Shurtleff, W. and A. Aoyagi. 2004. History of soybeans and soy foods, 1100 B.C. to the 1980s. Copyright Soyinfo Center, Lafayette, California. www.soyinfocenter.com.

Singh, D., P.H. Pfromm, and M.E. Rezac. 2011. Overcoming mass-transfer limitations in partial hydrogenation of soybean oil using metal-decorated polymeric membranes. AIChE J. 57:2450–2457.

Singh, D., M. Rezac, and P. Pfromm. 2009. Partial hydrogenation of soybean oil with minimal *trans* fat production using a Pt-decorated polymeric membrane reactor. J. Am. Oil Chem. Soc. 86:93–101.

Spencer, M., V. Pantalone, E. Meyer, D. Landau-Ellis, and D. Hyten. 2003. Mapping the *Fas* locus controlling stearic acid content in soybean. Theor. Appl. Genet. 106:615–619.

Sudaric, A., D. Simic, and M. Vrataric. 2005. Characterization of genotype by environment interactions in soybean breeding programmes of southeast Europe. Plant Breeding 125:191–194.

Thijssen, M.A., G. Hornstra, and R.P. Mensink. 2005. Stearic, oleic, and linoleic acids have comparable effects on markers of thrombotic tendency in healthy human subjects. J. Nutr. 135:2805–2811.

Thijssen, M.A. and R.P. Mensink. 2005. Small differences in the effects of stearic acid, oleic acid, and linoleic acid on the serum lipoprotein profile of humans. Am. J. Clin. Nutr. 82:510–516.

Thomas, A. 2000. Fats and fatty oils. In: Ullmann's encyclopedia of industrial chemistry. John Wiley and Sons, Hoboken, NJ. DOI: 10.1002/14356007

Tseng, W.J., D. M. Liu, and C.K. Hsu. 1999. Influence of stearic acid on suspension structure and green microstructure of injection-molded zirconia ceramics. Ceram. Int. 25:191–195.

United Soybean Board. 2014. Soybean oil innovations 5th ed. In: SoyConnection. United Soybean Board. www.SoyConnection.com.

USDA. 2015. USDA National Nutrient Database for Standard Reference, Release 28. In: A.R.S. US Department of Agriculture, Nutrient Data Laboratory, editor www.ars.usda.gov.

Veldsink, J.W., M.J. Bouma, N.H. Schöön, and A.A.C.M. Beenackers. 1997. Heterogeneous hydrogenation of vegetable oils: a literature review. Catal. Rev. 39:253–318.

Wacek, T. and W.J. Brill. 1976. Simple, rapid assay for screening nitrogen-fixing ability in soybean. Crop Sci. 16:519–523.

Wahl, J. and H. Colburn. 2009. Solvent extraction of chemical attribution signature compounds from painted wall board: final report. In: S.a.T.D. Department of Homeland Security, editor Pacific Northwest National Laboratory. www.pnl.gov

Westcott, P. and J. Hansen. 2015. USDA Agricultural Projections to 2024 No. (OCE-151). ERS-USDA. http://www.ers.usda.gov

Wiedermann, L.H. 1981. Degumming, refining and bleaching soybean oil. J. Am. Oil Chem. Soc. 58:159–166.

Wilcox, J.R. and J.F. Cavins. 1986. Registration of C1640 soybean germplasm. Crop Sci. 26:209–210.

Wilcox, J.R. and J.F. Cavins. 1990. Registration of C1726 and C1727 soybean germplasm with altered levels of palmitic acid. Crop Sci. 30:240.

Wilson, R.F. 2004. Seed composition. In: H. R. Boerma and J. Specht, editors, Soybeans: improvement, production and uses. ASA, Madison, WI. p. 621–677.

Wilson, R.F., T.C. Marquardt, W.P. Novitzky, J.W. Burton, J.R. Wilcox, A.J. Kinney, and R.E. Dewey. 2001. Metabolic mechanisms associated with alleles governing the 16:0 concentration of soybean oil. J. Am. Oil Chem. Soc. 78:335–340.

Wolf, R.B., J.F. Cavins, R. Kleiman, and L.T. Black. 1982. Effect of temperature on soybean seed constituents: oil, protein, moisture, fatty acids, amino acids and sugars. J. Am. Oil Chem. Soc. 59:230–232.

Yang, B., X.Q. Gong, H.F. Wang, X.M. Cao, J.J. Rooney, and P. Hu. 2013. Evidence to challenge the universality of the Horiuti-Polanyi mechanism for hydrogenation in heterogeneous catalysis: origin and trend of the preference of a Non-Horiuti-Polanyi mechanism. J. Am. Chem. Soc. 135:15244–15250.

Zhang, P., J.W. Burton, R.G. Upchurch, E. Whittle, J. Shanklin, and R.E. Dewey. 2008. Mutations in a Δ–stearoyl-ACP-desaturase gene are associated with enhanced stearic acid levels in soybean seeds. Crop Sci. 48:2305–2313.

Zock, P.L. and M.B. Katan. 1992. Hydrogenation alternatives: effects of *trans* fatty acids and stearic acid versus linoleic acid on serum lipids and lipoproteins in humans. J. Lipid Res. 33:399–410.

Zuil, S.G., N.G. Izquierdo, J. Luján, M. Cantarero, and L.A.N. Aguirrezábal. 2012. Oil quality of maize and soybean genotypes with increased oleic acid percentage as affected by intercepted solar radiation and temperature. Field Crops Res. 127:203–214.

CHAPTER 2

EFFECTS OF MODIFIED FATTY ACID COMPOSITION ON YIELD, PROTEIN AND OIL CONTENTS, AND SEED GERMINATION IN SOYBEAN[1]

[1] N. Bachleda, D.K. Harris, T. Grey, H.R. Boerma, and Z. Li. To be submitted to *Crop Science.*

Abstract

Typical soybean oil is composed of palmitic, stearic, oleic, linoleic, and linolenic acids. High oleic and low linolenic acid contents in soybean seed are the key compositional traits that improve oxidative stability and increase oil functionality and shelf life. Using a marker-assisted selection method, we have developed near-isolines (NILs) of G00-3213 and G00-3880 for high oleic and low linolenic traits, respectively, through backcrossing. We have yield-tested these NILs that have various combinations of *FAD2-1A* and *FAD2-1B*, or different *FAD3A* alleles that were derived from the same backcross populations. The results indicated that G00-3213 NILs with both homozygous mutant *FAD2-1A* and *FAD2-1B* alleles produced an average of 806 g kg^{-1} oleic acid content, and G00-3213 and G00-3880 NILs with homozygous mutant *FAD3A* alleles reduced the linolenic acid content to 37 and 41 g kg^{-1}, respectively. The results also demonstrated that possessing these mutant alleles did not cause a yield reduction. Furthermore, seed germination tests across 12 temperatures (12.8 - 32.0°C) showed that modified seed composition for oleic or linolenic acids in general did not have a major impact on seed germination. However, there was a possible reduction in seed germination vigor when high oleic seed are planted in cold soil. The observed significant differences ($p<0.05$) of the germination rates between low linolenic and normal linolenic NILs were dependent on the genetic background. The mutant *FAD2-1A* and *FAD2-1B* alleles did not hinder either seed or plant development.

Key words: Soybean, DNA marker, High oleic acid, Low linolenic acid, Yield, Seed germination.

Introduction

Soybean is a major global oil crop, accounting for 55% of U.S. vegetable oil consumption in 2014 (http://soystats.com). This widely consumed oil is made up of five major fatty acids: palmitic (110 g kg^{-1}), stearic (40 g kg^{-1}), oleic (250 g kg^{-1}), linoleic (520 g kg^{-1}), and linolenic (80 g kg^{-1}) acids (Fehr, 2007). Previous research identified mutations of *FAD2-1A* and *FAD2-1B* genes that caused high oleic acid content (Lee et al., 2007). Pham et al. (2010) showed that by combining both *FAD2-1A* and *FAD2-1B* mutant alleles, oleic acid levels of soybean lines can reach >750 g kg^{-1}. A mutation at the *FAD3A* locus is responsible for low linolenic acid (Chappell et al., 2006).

Oleic acid is a monounsaturated fatty acid and has a naturally occurring *cis* configuration at the 9th carbon-carbon double bond (Thomas, 2000), which can be found in most plant and animal oils. This monounsaturated fatty acid occurs in soybean seed and has been of great interest due to its effects on human health when consumed. A diet high in monounsaturated fatty acids resulted in a 6.3% reduction in low density lipoprotein (LDL) when compared to an average American diet (Berglund et al., 2007). By preventing the desaturation of oleic acid to linoleic acid, oleic acid is increased which results in a significant decrease in consuming LDL. The *FAD2* family of genes is related to the quantity of oleic acid by interacting with fatty acid Δ12-desaturase (Anai et al., 2008). One *FAD2-1A* (Glyma10g42470) mutation was isolated from an M_2 generation of γ-irradiated soybean seed and was mapped to the Chromosome (Chr) 10 (Bachlava et al., 2008). *FAD2-1B* (Glyma20g24530) is another gene related to oleic acid synthesis that functions in the same way as *FAD2-1A* and was mapped to Chr 20 (Bachlava et al., 2008; Bachlava et al., 2009). It has been suggested that both *FAD2-1A* and *FAD2-1B* are duplicated genes due to the evidence that soybean is a paleopolyploid crop (Anai et al., 2008; Schlueter et al., 2007).

Multiple versions of mutations at *FAD2-1A* locus have been reported, including the deletion of the *FAD2-1A* gene in soybean line M23 (Alt et al., 2005) and the missense mutation in *FAD2-1A* from line 17D (*FAD2-1A* S117N) (Dierking and Bilyeu, 2009). Pham et al. (2010) identified missense mutant alleles in *FAD2-1B* from PI 283327 and PI 567189A and reported that a combination of mutations at both *FAD2-1A* and *FAD2-1B* loci produced 800 g kg^{-1} oleic acid in soybean lines. This has created an opportunity to develop high oleic soybean cultivars through conventional breeding methods.

Linolenic acid is a polyunsaturated fat (PUFA) that makes up 80 g kg^{-1} of the oil composition of soybean seed (Fehr et al., 2007) and is partially responsible for the poor stability of soybean oil which necessitated hydrogenation to improve the stability of the oil. It has three double bonds with *cis* configurations located at the 9th, 12th, and 15th carbons (Thomas, 2000). All members of the *FAD3* gene family influence the production of linolenic acid, but the *FAD3A* gene was found to have the greatest impact in soybean seed (Bilyeu et al., 2011). Its function is to block Δ15-desaturase from converting linoleic acid to linolenic acid, resulting in low linolenic acid content in soybean seed.

The first reported low linolenic acid mutant line was C1640, which had an average linolenic acid content of 34 g kg^{-1} (Wilcox and Cavins, 1986). Chappell and Bilyeu (2006) determined that the low linolenic trait of line C1640 was due to a nonsense mutation resulting in a premature stop codon in the omega-3 fatty acid desaturase gene. Other low linolenic lines have been developed using chemical mutagenesis. Using seed treatment with chemical mutagens such as ethyl methanesulfonate (EMS), Hammond and Fehr (1983) selected the soybean line A5 with reduced linolenic acid content (35 – 41 g kg^{-1}). The reduced linolenic acid content in this soybean line was determined to be under the control of a *fan1* allele (Rennie and Tanner, 1991). Subsequently, Fehr and Hammond (2000) found a *fan2* allele in soybean line A23. The third locus responsible for linolenic acid content, designated as *fan3,* was found in A26 which was developed by treating soybean line A89-144003 with EMS (Fehr and Hammond, 2000). By combining the three alleles- *fan1* (A5), *fan2* (A23), and *fan3* (A26) through hybridization and selection, soybean line A29 was developed to have the lowest linolenic acid content to date (<20 g kg^{-1}) (Fehr and Hammond, 2000; Ross et al., 2000). The low linolenic acid soybean line CX1512-44 was also developed using a mutagenesis approach by Dr. James Wilcox at Purdue University (Bilyeu et al., 2005). Bilyeu et al. (2011) found that a combination of two mutant alleles at *FAD3A* and *FAD3C* loci in CX1512-44 reduced linolenic content to about 30 g kg^{-1}.

It has been documented that environmental conditions influence crop yield and seed composition, such as seed fatty acids (Cherry et al., 1984; Cherry et al., 1985; Kumar et al., 2006; Maestri et al., 1998; Reddy et al., 2007; Sudaric et al., 2005). Soybean grown in warmer climates produced more oleic acid, and in cooler climates produced a greater percent of linoleic and linolenic acids (Howell and Collins, 1957; Maestri et al., 1998; Wolf et al., 1982). Specifically, there is a positive correlation between high oleic acid in seed and high mean temperature during seed filling (Lee et al., 2009; Oliva et al., 2006). Since Georgia has a warmer climate than the major soybean production region of the USA, the Corn Belt, the variability of linoleic and linolenic acids is lower in soybean seed grown in Georgia.

Utilizing previously discovered genetic resources to achieve the desired fatty acid compositions in soybean production will be an effective approach to quickly develop new soybean cultivars. To facilitate introgressing high oleic and low linolenic traits into elite soybean lines, robust markers are necessary to assist in backcrossing and efficient trait selection. By characterizing the *FAD3A* locus in line A5, Pham et al. (2014) developed a TaqMan detection assay by targeting the deletion junction in A5, which was used to distinguish the homozygotes and heterozygotes of the gene. They also developed TaqMan assays for *FAD3B* and *FAD3C* loci based on single nucleotide polymorphisms (SNPs) identified in A26 and A23, respectively. Based on reported SNPs, Shi et al. (2015) successfully developed TaqMan or Competitive Allele-Specific PCR (KASP) SNP detection assays to select mutant alleles of *FAD2-1A* (17D source) and *FAD2-1B* (PI 283327 source) for the high oleic acid trait as well as those of *FAD3A* (CX1512-44 and C1640 sources) and *FAD3C* (CX1512-44 source) for the low linolenic acid trait in soybean.

Transgenic high oleic soybean lines have been developed by silencing *FAD2* genes, and these lines are shown to have a consistent oleic acid concentration across multiple environments without affecting yield (Brace et al., 2011). Knowlton (1999) transformed the *FAD2-1* gene (event 260-05) into cultivar A2396 (Monsanto Company, St. Louis, MO) using a particle bombardment method, which effectively suppressed the expression of *FAD2-1* gene. Yield tests of lines containing the event 260-05 (oleic acid levels at 840–880 g kg^{-1}) across three generations over 2 years indicated that the high oleic lines yielded 3,665 kg ha^{-1}, while elite conventional lines yielded 3,564 kg ha^{-1} (Kinney, 1996). Kinney concluded that the high oleic trait did not negatively affect yield. La et al. (2014) compared seed yield, fatty acids, total oil and protein concentrations between 40 $F_{4:7}$ high oleic lines carrying both mutant *FAD2-1A* and *FAD2-1B* alleles and 40 normal oleic acid lines derived from same six crosses in six environments. Their results indicated that the yield difference between high and normal oleic acid lines varied among populations, but seed yield of the high oleic lines were within 2% of the normal oleic lines. They also found that the high oleic lines averaged >790 g kg^{-1} oleic acid and had significantly higher protein and oil content than normal oleic lines in all six populations (La et al., 2014). To date, no research has compared the yield, fatty acid profiles, and seed germination rates between near-isolines derived from the same population in southeastern U.S. environments.

To determine if high oleic and low linoleic traits affect soybean seed germination, we used thermal gradient tables for the germination tests in our study. Thermal gradient tables have been used to evaluate weed and crop seeds (Cardina and Hook, 1989; Chatterton and Kadish, 1969; Grey et al., 2011; Mohapatra and Johnson, 1986), which provided a systematic way to look at the seed germination at multiple temperatures. However, to date, no result has been reported for the impacts of high oleic and low linoleic traits on seed germination in soybean.

The objectives of this study were to: i) evaluate effects of high oleic and low linolenic acid traits on yield, fatty acids, protein, and oil contents in southeastern U.S. environments; and ii) determine if these traits affect seed germination in soybean.

Materials and Methods

Population Development

$BC_3F_{2:4}$ populations were developed by backcrossing *FAD2-1A* and *FAD2-1B* mutant alleles for high oleic acid content or the *FAD3A* mutant allele for low linolenic acid content into elite soybean lines, G00-3213 or G00-3880, respectively, using marker-assisted selections. The donor source for high oleic trait was a breeding line derived from 17D (*FAD2-1A*) and PI283327 (*FAD2-1B*) (Pham et al., 2011), courtesy of Dr. Grover Shannon, University of Missouri, Portageville, MO. The donor source for the low linolenic trait was ‘Benning low lin/low palm’ (*FAD3A*) that was developed at the University of Georgia. ‘Benning low lin/low palm’ is an NIL of ‘Benning’ (Boerma et al., 1997) having low linolenic acid trait from C1640 (Wilcox and Cavins, 1986) and low palmitic acid trait from N87-2122-4 (Burton et al., 1994).

Two high-yielding lines, G00-3213 and G00-3880, developed at the University of Georgia were selected as recurrent parents to introgress the high oleic and low linolenic acid traits in parallel. G00-3213 is derived from the cross of ‘N7001’ (Carter et al., 2003) and ‘Boggs’ (Boerma, 2000). G00-3880 is a high yielding line developed from the cross of G93-9201 and ‘Cook’ (Boerma et al., 1992). G93-9201 was developed from the cross G83-559 x [G80-1515(2) x PI 230977].

Each recurrent parent was crossed to the high oleic or low linolenic acid trait donors, respectively, at the University of Georgia Plant Sciences Farm, near Athens, GA. Backcrosses to the recurrent parents were conducted at either the Plant Sciences Farm in GA or the USDA Puerto Rican Winter Nursery using marker-assisted selection at each generation. BC_3F_2 plants were grown at the Plant Sciences Farm in GA during the summer of 2013. A total of 1,101 individual BC_3F_2 plants were genotyped for high oleic and low linolenic acid traits. After fatty acid phenotypes of the seed harvested from selected plants were confirmed using gas chromatography (GC), these seed were grown as plant rows at the USDA Puerto Rican Winter Nursery during the winter of 2013–2014. The seed harvested in bulk from each of these rows were then used for yield trials in 2014.

Genotyping for Marker-Assisted Selection

During the summer of 2013 in Athens, GA, each individual BC_3F_2 plant was tagged one month after planting and then leaf tissue was collected from each individual plant for genotyping. A roughly “penny-sized” (2.8 cm^2) young trifoliolate leaf was collected from each individual plant, placed in a tube, and organized in 96-well racks. After the tissue samples were collected, they were dried overnight in an oven at 55°C. The dried tissues were ground to a fine powder by adding two BBs to each collection tube and shaken using a GenoGrinder (SPEX SamplePrep, Metuchen, NJ). DNA was extracted following a modified CTAB procedure from Keim et al. (1988). From each 96 well plate, DNA concentrations were checked using a Tecan Infinite M100 Pro Reader (TecanGroup Ltd, Morrisville, NC).

To assist backcrossing and selection of target traits, TaqMan markers GSM257 (*FAD2-1A*) and GSM256 (*FAD2-1B*) were used to genotype backcrossed progeny to select alleles controlling the high oleic trait from sources 17D and PI283327, respectively. Similarly, TaqMan marker GSM329 (*FAD3A*) was used for genotyping the low linolenic trait from C1640 (Shi et al., 2015). Two positive and two negative controls for the trait were used in each DNA plate for all marker genotyping. The PCR was conducted in 384-well plates, using the protocol reported by Shi et al. (2015).

Yield Trials

Based on genotyping results of markers GSM257 (*FAD2-1A*) and GSM256 (*FAD2-1B*), four groups of G00-3213 NILs for the oleic trait were created, including NILs with homozygous mutants (referred as high oleic NILs) and homozygous wild types (referred as normal oleic NILs) for both *FAD2-1A* and *FAD2-1B* loci, and NILs with one locus homozygous mutant and other locus homozygous wild type (Table 2.1). Due to the limited number of the NILs, the number of NILs in each group varied from eight to twelve (Table 2.1). Since there was only one G00-3880 NIL in the group of the homozygous mutant for both *FAD2-1A* and *FAD2-1B* loci for oleic trait, the G00-3880 NIL was included in yield trials, but was not used for data analysis.

Similarly for the low linolenic trait, two groups of NILs for either G00-3213 or G00-3880 lines were created based on GSM329 (*FAD3A*) genotyping results, including homozygous mutants (referred as low linolenic NILs) and wild types (referred as normal linolenic NILs) for *FAD3A*. For the G00-3213 NIL pedigree, 10 NILs in each group of low and normal linolenic levels were included, while for the G00-3880 NIL pedigree, 6 NILs in each of the two groups were included. In all yield trials, the recurrent parents G00-3213 and G00-3880 were also included twice in the tests as parental checks (Table 2.1).

These $BC_3F_{2:4}$ NILs were yield tested at two locations in 2014. Within each location, a randomized complete block design with two replications and two rows per plot was used for each set. The plots at both locations had a row length of 4.9 m with 0.76 m between rows and were planted at the density of 34.4 seed m^{-2}. Experiments were planted on 22 May 2014 at the Plant Sciences Farm in Appling coarse sandy loam soil (Fine, kaolinitic, thermic Typic Kanhapludults) and 9 June 2014 at the University of Georgia Southwest Research and Education Center near Plains, GA in Greenville sandy clay loam soil (Fine, kaolinitic, thermic Rhodic Kandiudults). Field notes including maturity, plant height, flower color, pubescence color, pod wall color, and lodging were taken from each plot in Athens and yield data were collected from both locations. The plots at both locations were end-trimmed before harvest, and the final harvest row length were 3.7 m.

Seed Composition Analysis

After harvest, five seeds from each plot at both locations were used for fatty acid analysis using gas chromatography (GC). The GC analyses were conducted with an Agilent Technologies 6890N Network Capillary GC System at Dr. Kristen Bilyeu's lab at the USDA/University of

Missouri, Columbia, MO following the protocols reported by Beuselinck et al. (2006) and Bilyeu et al. (2005). The five seeds per line were crushed to extract oil overnight in chloroform–hexane–methanol. Derivatization was done with a methylating reagent (0.5 M methanolic sodium methoxide–petroleum ether–ethyl ether, 1:5:2, v/v/v) and fatty acid methyl esters were expressed as a percent of the total fatty acids.

Seed from selected NILs underwent additional composition analysis including protein, oil, fiber, and ash contents with a combustion analysis (LECO) using the AOAC Official Method 990.03, in which combustion at high temperature in pure oxygen was measured by thermal conductivity detection and calculated with a factor of 6.25. The analyses were conducted at the Agricultural Experiment Station Chemical Laboratories of the University of Missouri, Columbia, MO.

Seed Germination Test

Due to the limited capacity for the germination tests, entries used for the germination test were selected based on the genotypes of NILs from the yield trials (Table 2.2). Of the G00-3213 NILs, we selected six homozygous mutant NILs and six homozygous wild type NILs for both *FAD2-1A* and *FAD2-1B* loci, and six homozygous mutant NILs and six homozygous wild type NILs for the *FAD3A* locus. Of the G00-3880 NILs, we selected six homozygous mutant NILs and six homozygous wild type NILs for the *FAD3A* locus. One hundred seed for each of the above entries were taken from the first replication at each of Athens and Plains locations and used for the germination tests. The germination rates were tested using the thermal gradient tables at Dr. Timothy Grey’s lab at the University of Georgia, Tifton campus. Each thermal gradient table has nine rows, and accommodated eight entries plus the check.

Twelve temperatures were chosen across a thermal gradient table for the tests. Typically, warm standard germination tests for soybean seed are conducted at 25°C for 7 d. With 25°C being the optimum temperature for soybean germination, we wanted to also challenge the seeds by setting the hottest end of the thermal gradient tables at 32°C. Cold germination tests are typically conducted at 10°C for 7 d and then at 25ºC for additional 5 to 7 d. Since the thermal gradient test was only set up for 7 d (168 h) and little to no germination was expected at 10ºC for 7 d, the coldest end of the thermal gradient tables was set at 12ºC. Each entry was replicated twice at each temperature. Ten seeds per entry at each temperature in each replication were loaded into a 100 x 15 mm petri dish with blotting paper (Anchor Paper Company, St. Paul, MN) and moistened with 10 mL water. Entries were randomized by row on each thermal gradient table. The 12 temperatures ran a gradient from 12°C to 32°C, with ~1.7°C between two temperatures. Starting 24 h after the seeds were plated, germinated seed was counted every day at approximately same time for four days in a row (24 h, 48 h, 72 h, and 96 h). A final count of germinated seed was conducted at 168 h, which is consistent with the standard soybean germination tests. Germination was counted when the seed radicals reached 5 mm (Fig. 2.1) (Grey et al., 2011).

The thermal gradient tables were custom made from aluminum blocks (2.4 m x 0.9 m x 0.76 m) and a 1.0 cm hole was drilled across the side section to allow fluid to be pumped through. Pumps on either end of the table pumped either 12°C or 32°C ethylene glycol plus water (1:10 mixture) at 3.8 L min^{-1} to generate the thermogradient (Grey et al., 2011). Temperature was measured by thermocouples made from duplex insulated PR-T-24 wires (Omega Engineer, Inc. Stamford, CT) at 10-cm intervals along the length of the table. The temperature of each thermocouple was continuously monitored and recorded at 30 min intervals with a Graphtec data logger (MicroDAQ.com Ltd., Contoocook, NH).

Data analysis

Statistical analyses were conducted using SAS Enterprise Guide 4.3 (SAS Institute Inc., Cary, NC). Analysis of variance (ANOVA) using a general linear model (GLM) (PROC GLM) was performed to analyze seed yield, agronomic traits, and fatty acids among or between the NIL groups over two locations. In the GLM, both locations and NIL groups were treated as fixed effects and the error mean squares were used to test the significance of locations, NIL groups and interactions of NIL group by location. Due to the difference of number of NILs in each group, statistical comparisons of the G00-3213 oleic NIL groups were conducted using *Tukey's Studentized Range Test* at the 0.05 probability level. Same test was also used for comparisons between the locations or between low linolenic and normal linolenic NIL groups. If the genotype by environment interactions (NIL group by location) were significant ($p<0.05$) for the oleic and linolenic acids or yield, Least Square Means were used for the comparison of the interactions at the 0.05 probability level. Student's *t*-tests in RStudio (www.rstudio.com) were used to compare the differences of germinations between high oleic and normal oleic NILs, and low linolenic and normal linolenic NILs.

Results and Discussion

Seed Composition

Fatty acid profile between NIL groups

Based on GC analyses of the seed samples from G00-3213 NILs, the high oleic G00-3213 NILs averaged 806 g kg^{-1} oleic acid and 45 g kg^{-1} linolenic acid, whereas the normal oleic G00-3213 NILs averaged 252 g kg^{-1} oleic acid and 67 g kg^{-1} linolenic acid in the seed oil (Table 2.3). This indicates that mutant alleles of both *FAD2-1A* and *FAD2-1B* loci significantly ($p<0.05$) increased oleic acid content and decreased linolenic acid content in the seed oil. In general, average oleic acid content across all oleic NILs in Plains, GA was significantly higher ($p<0.01$) than that in Athens, GA (481 g kg^{-1} vs. 421 g kg^{-1}). Similar trends were observed for high oleic NILs that high oleic NILs produced an average oleic acid content of 831 g kg^{-1} in Plains, GA, while an average oleic acid contents of 785 g kg^{-1} in Athens, GA (Table 2.4). The increase in oleic acid content was likely due to the higher temperature in Plains, GA than in Athens, GA in September and October when the soybean plants were at seed filling stage (Table 2.5). Significant ($p<0.05$) genotype by environment interaction (Location x NIL group) was observed for oleic acid content among oleic NIL group (Table 2.6). The rankings in oleic acid content among the oleic NIL groups were consistent and the interaction were due to the magnitude of oleic acid contents at both locations. We also observed the decrease in linolenic acid content among the oleic G00-3213 NIL groups in Plains, GA. In addition to the increase of oleic acid and decrease of linolenic acid in Plains, the results showed the decrease of palmitic (74 g kg^{-1}), stearic (28 g kg^{-1}) and linoleic (47 g kg^{-1}) acids in the high oleic G00-3213 NIL group.

A study on characterization of the *FAD3A* gene in segregating populations by Pham et al. (2014) reported a correlation coefficient of 0.79 between oleic and linoleic acids in a population consisting of 300 $F_{2:3}$ seeds derived from R07-10231 x IA30234 (high yielding line x low linolenic line), and a correlation coefficient of 0.81 between oleic and linoleic acids in a second population consisting of 300 $F_{2:3}$ seeds derived from R05-1415 x R07-167 (high protein line x high yield/low linolenic line). The reported results coincided with changes observed from this study in palmitic, linoleic, and linolenic acids when oleic acid was increased. The mutant *FAD2-1A* and *FAD2-1B* alleles in this study produced similar oleic acid contents to that reported by La et al. (2014), but higher than those of the transgenic high oleic and low linolenic lines that were developed by Brace et al. (2011).

The linolenic G00-3213 NILs with mutant *FAD3A* allele averaged 37 g kg^{-1} linolenic acid, while the normal linolenic G00-3213 NILs averaged 70 g kg^{-1} linolenic acid (Table 2.7), which was significant different ($p<0.05$). For all linolenic G00-3213 NILs, significant difference ($p<0.05$) of linolenic acids between two locations (64 g kg^{-1} in Athens, GA vs. 53 g kg^{-1} in Plains, GA) were observed. Similar trends were also seen for the low linolenic NILs that the low linolenic NILs had average linolenic acid contents of 38 g kg^{-} and 35 g kg^{-1} in ^{1} in Athens and Plains, GA, respectively (Table 2.4). ANOVA indicated that genotype x environment interactions (Location x NIL group) was significant for the linolenic acid (Table 2.6). The NIL groups had same rankings at both locations and the significance in the interactions was due to the magnitudes of linolenic acids at both locations. In addition, palmitic, stearic, and linoleic acids were significantly increased, while oleic acid content was decreased in the low linolenic G00-3213 NIL group (Table 2.7).

Similar to the low linolenic G00-3213 NILs, low linolenic G00-3880 NILs averaged 41 g kg^{-1} linolenic acid, while the normal linolenic NILs averaged 79 g kg^{-1} linolenic acid. The *Tukey* (0.05) test indicated that the mutant *FAD3A* allele significantly reduced linolenic acid in seed oil in a different genetic background. Both stearic and linoleic acids were higher in the low linolenic G00-3880 NIL group than normal linolenic G00-3880 NIL group (Table 2.8). The G00-3880 NILs did not have significant ($p>0.05$) genotype x environment interaction for linolenic acid. However, the Plains location had significantly ($p<0.05$) lower linolenic acid content than the Athens location, which was consistent with the results of the linolenic G00-3213 NILs. . This study indicated that in the Georgia climate, the mutant *FAD3A* allele could reduce linolenic acid to about 33–38 g kg^{-1}. A follow up study to stack both mutant high oleic alleles at *FAD2-1A* and *FAD2-1B* loci with the mutant low linolenic allele from *FAD3A* is needed to determine the attainable linolenic acid levels in these stacks.

Seed composition between NIL groups

To determine the effects of high oleic or low linolenic traits on protein, oil, fiber, and ash, six NILs from each of the high oleic, normal oleic, low linolenic, and normal linolenic G00-3213 NIL groups were selected from each of two locations for seed composition analyses (Table 2.9). No significant ($p>0.05$) differences were observed between these G00-3213 NIL groups for protein, oil, fiber and ash contents. These findings differed from those reported by La et al. (2014) who compared high and normal oleic lines derived from the same donor parents across six locations. They found that the high oleic lines had higher mean protein content than the normal oleic lines, whereas this study determined that high oleic NILs were not different than normal oleic NILs in protein content. The differing results from the two studies could be due to an environmental effect on protein production.

Similarly, seed of six G00-3213 NILs from each of the low and normal linolenic groups from each location were analyzed for protein, oil, fiber, and ash contents. There was no significant difference ($p>0.05$) in protein, oil, fiber, and ash contents between the low and normal linolenic G00-3213 NIL groups.

Yield and Agronomic Traits

Yield and agronomic traits of NIL groups for oleic acid

Significant differences ($p<0.05$) in yield were observed among the four G00-3213 NIL groups with different *FAD2* allele combinations derived from the G00-3213 $BC_3F_{2:4}$ population (Tables 2.3 and 2.5). High oleic G00-3213 NILs had an average yield of 2,737 kg ha^{-1}, while normal oleic G00-3213 NILs had an average yield of 2,730 kg ha^{-1}. There was no significant yield difference between the high (MUT/MUT) and normal (WT/WT) oleic NIL groups despite significant differences ($p<0.05$) in seed oleic acid content between the two NIL groups. High and normal oleic G00-3213 NILs were similar in maturity (16 Oct. vs. 15 Oct.), which indicated maturity was not a confounding factor in the yield parity results between high and normal oleic G00-3213 NILs. When comparing the yield with the recurrent parent (G00-3213), the average yield of these two G00-3213 NIL groups (high and normal oleic NIL groups) were 222 kg ha^{-1} lower in yield than that of the recurrent parent (Table 2.3). However, one line from the high oleic G00-3213 NIL group yielded 336 kg ha^{-1} greater than the recurrent parent, G00-3213. Although the yield was significantly ($p<0.05$) different between two locations (2,860 kg ha^{-1} for Athens vs. 2,541 kg ha^{-1} for Plains), the genotype by environment interaction (NIL group x location) was not significant ($p>0.05$) (Table 2.4 and 2.5).

Of the four *FAD2* allele combination groups, G00-3213 NILs with a combination of the *FAD2-1A* MUT/*FAD2-1B* WT alleles had the highest average yield of 2,939 kg ha^{-1}, which was significantly higher ($p<0.05$) than the other three G00-3213 NIL groups. Our results are consistent with the report from La et al. (2014), in which the yield difference was not significant between high oleic (750–800 g kg^{-1}, n=35) and normal oleic (200–250 g kg^{-1}, n=35) $F_{4:7}$ lines derived from five different backgrounds, but they were significantly lower ($p<0.05$) than that of the elite parents. Our results were also consistent with the study by Greaf et al. (2009), in which a transgenic high oleic trait tested over eight environments with four replications per location had no impact on yield, with four different transgenic events yielding over 3,508 kg ha^{-1} (863 g kg^{-1} oleic acid).

Although the high oleic G00-3213 NIL group on average yielded below the recurrent parent G00-3213, we observed one NIL from this group that yielded greater than the G00-3213, suggesting that we might be able to breed high oleic lines that yield similar to or higher than the recurrent parent. However, the trait in this study was only tested in one genetic background in one year. To make an informed conclusion, the trait needs to be tested in more genetic backgrounds and multiple years. There were no significant differences ($p>0.05$) observed in seed weight and plant height among the G00-3213 NILs groups except for the NILs with *FAD2-1A* MUT/*FAD2-1B* WT alleles in seed weight (Table 2.3), suggesting that both mutant *FAD2-1A* and *FAD2-1B* alleles did not affect either seed or plant development.

Yield and agronomic traits of NIL groups for linolenic acid

Similar to the oleic NIL groups, the yield of linolenic NIL groups with the MUT or WT *FAD3A* allele derived from the G00-3213 $BC_3F_{2:4}$ population also showed non-significant difference ($p>0.05$) (Table 2.7). The low linolenic G00-3213 NIL group had an average yield of 2,744 kg ha^{-1}, while the normal linolenic G00-3213 NIL group had an average yield of 2,724 kg ha^{-1}. However, they averaged approximately 336 kg ha^{-1} lower in yield than the recurrent parent

G00-3213. The low linolenic G00-3213 NIL group with mutant *FAD3A* alleles matured on average 3 d and 1 d later than the normal linolenic NIL group and the recurrent parent G00-3213, respectively. Unlike the oleic acid G00-3213 NIL groups, no individual low linolenic G00-3213 NIL was found to yield higher than the recurrent parent. There was non-significant ($p>0.05$) genotype by environment interaction (NIL group x Location) for yield detected (Table 2.6).

The low linolenic G00-3880 NILs had an average yield of 2,663 kg ha^{-1}and matured on average on 13 Oct. (Table 2.8). In contrast, the normal linolenic G00-3880 NILs yielded on average 2,724 kg ha^{-1}and matured on 17 Oct. The low linolenic and normal linolenic G00-3880 NIL groups did not have significantly different yield or maturity from each other, but the low linolenic G00-3880 NIL group were 242 kg ha^{-1} lower in yield than the recurrent parent G00-3880. Both low linolenic and normal linolenic G00-3880 NILs were 3-4 d earlier in maturity than the recurrent parent.

The low linolenic G00-3213 NILs had similar seed weight and plant height as the normal G00-3213 NILs. In contrast, low linolenic G00-3880 NILs had significantly ($p<0.05$) higher seed weight and shorter plant height than normal linolenic G00-3880 NILs.

Based on the comparison of yield between low and normal linolenic NILs in the two genetic backgrounds, no significant yield reduction was observed in the low linolenic NIL groups although the average yield of both low and normal linolenic NILs was below their recurrent parents. Our results are also consistent with previous studies (Camacho-Roger, 2006; Reinprecht et al., 2009; Wilcox et al., 1993). After evaluating the offspring derived from a ‘Century’ x C1640 cross, Wilcox et al. (1993) concluded there was no association between the mutant *FAD3A* in low linolenic line C1640 and the agronomic traits. Camacho-Roger (2006) also found no correlation between linolenic acid levels and seed yield or between linolenic levels and seed oil concentration in a test conducted in Tennessee. Similarly, Reinprecht et al. (2006) evaluated the agronomic performance of 307 recombinant inbred lines (RILs) across three locations in two years containing an EMS-induced mutation in *FAD3A* (Reinprecht et al., 2009) and found that there was no yield penalty associated with the low linolenic trait.

Seed Germination

The average germination rates between high and normal oleic NILs derived from the G00-3123 $BC_3F_{2:4}$ population were compared and tested using a Student's *t*-test (Table 2.10). Most significant differences ($p<0.01$) for germination rates between high and normal oleic G00-32313 NILs were observed within the first three days (24 h, 48 h, and 72 h) at some of the lowest temperatures. Significant differences at a probability level of 0.05 were typically observed at the 96 h time-point after seed were plated. The high oleic G00-32313 NILs showed significantly ($p<0.05$) lower germination rates at 14.9°C and 16.7°C at 48 h, 72 h, and 96 h, but were statistically similar to normal oleic G00-32313 NILs at 168 h, which is used as a length of a standard soybean germination test (University of Arkansas Division of Agriculture, Arkansas Soybean Seed Quality Project, 2008). Only at the 18.4°C temperature did the NIL groups show a significant ($p<0.05$) difference in germination at the 168 h time point, with high oleic G00-3213 NILs having lower germination rates when compared to the normal oleic G00-3213 NILs (Fig. 2.2). The germination tests, we can conclude that although differences in germination rates existed during the tests at different temperatures, the high oleic trait in these G00-3213 NILs had minimal impacts on seed germination. Brink et al. (2014) compared transgenic high oleic soybean seed to null seeds using an AOSA standard cool and warm germination test protocol and found no difference in germination or dormancy. Our results are in accordance with the report from Brink et al. (2014).

Likewise, the average germination rates of low and normal linolenic G00-3213 NILs derived from the G00-3213 $BC_3F_{2:4}$ population were compared and tested using a Student's *t*-test (Table 2.11). In general, low linolenic G00-3213 NILs had high germination rates than normal linolenic G00-3213 NILs and most significant differences ($p<0.01$) were observed at 48 h and 72 h after the seed were plated across low to high temperatures. Significant differences ($p<0.05$) were also observed at the 168 h time point. Low linolenic G00-3213 NILs consistently germinated at a higher rate than normal linolenic G00-3213 NILs, except at 32°C, at which temperature the low linolenic NILs germinated at a similar rate. Differences in germination rates still existed at the end of 168 h, with low linolenic G00-3213 NILs germinating at a higher rate at 16.7°C, 18.4°C, 26.3°C, and 30.0°C, and a lower rate at 23.1°C and 32.0°C (Fig. 2.3). It can be concluded that at most temperatures, the low linolenic G00-3213 NILs had higher germination rates than these normal linolenic G00-3213 NILs.

The low and normal linolenic NILs derived from a G00-3880 $BC_3F_{2:4}$ population, however, did not show a similar germination pattern as the G00-3213 NILs (Table 2.12). At 18.4°C, low linolenic G00-3880 NILs showed significantly ($p<0.05$) lower germination rates at the 48 and 72 h time points. At 168 h, there were no significant differences ($p>0.05$) in germination at any temperatures examined (Fig. 2.4). In general, the low linolenic G00-3880 NILs had lower germination rates than these normal linolenic G00-3880 NILs.

Germination of modified fatty acid content soybean seed has not yet been studied in a great detail. Furthermore, the use of a thermal gradient table has not been used to study germination of soybean seeds. However, germination data from the 168 h time point of this study at 30°C could be compared to standard soybean germination tests. A standard germination test would miss differences at lower temperatures and earlier time points. In summary, the high oleic NILs had lower germination rates than these normal oleic lines across all temperatures, but they were not statically significant ($p<0.05$) for most temperatures with a few exceptions. Seed germination vigor of the high oleic lines could be reduced when planted in cold soil. Low linolenic lines performed differently in germination rates at different genetic backgrounds. Due to the limitation in the genetic backgrounds, a follow up study is needed to determine the effects of high oleic and low linolenic traits on the germination rates at different genetic backgrounds.

Figure 2.1. Seed with radicals reaching 5 mm in length

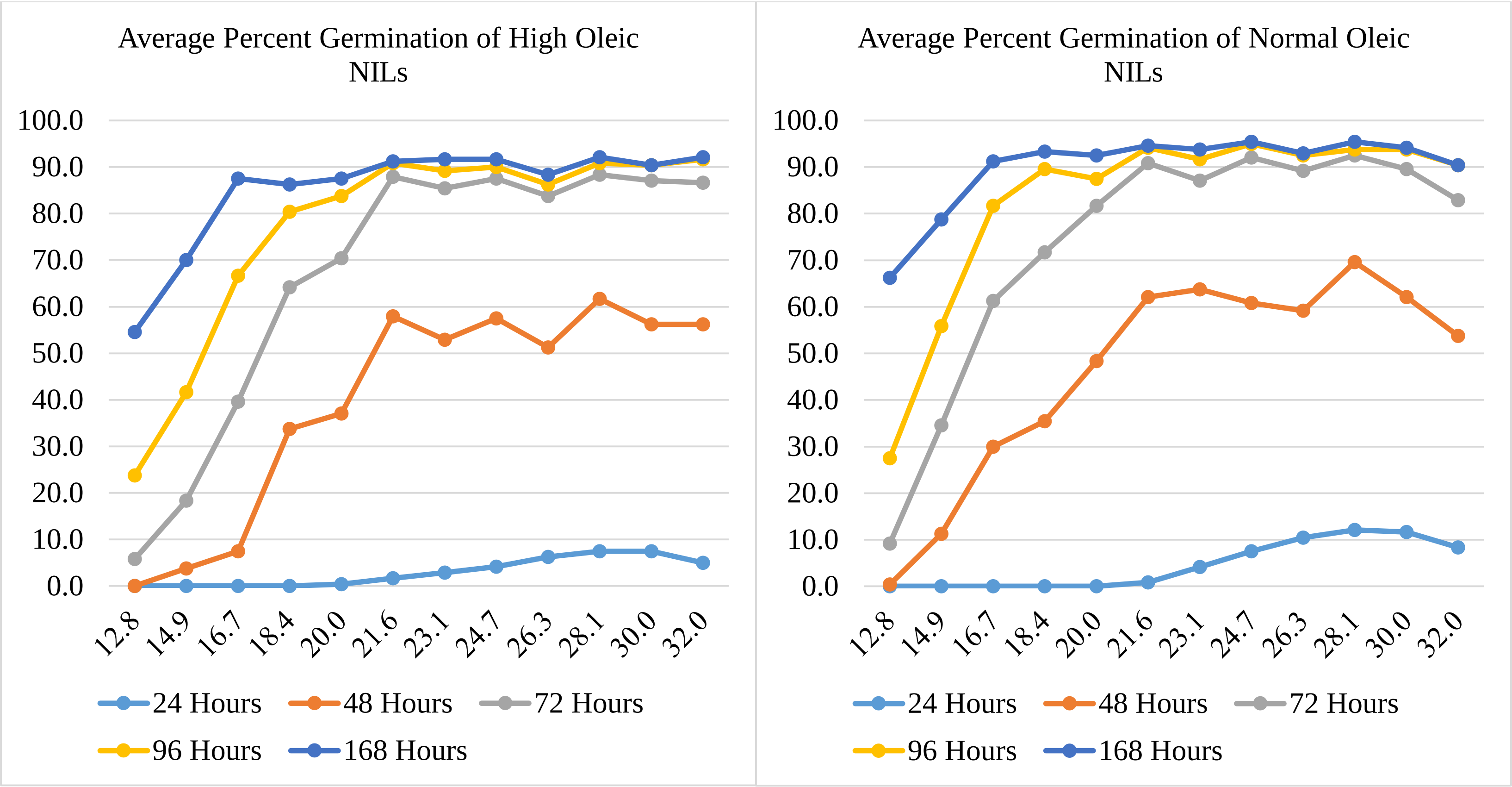

Figure 2.2. Percent germination of high (left) vs. normal (right) oleic acid near-isolines (NILs) derived from a G00-3213 $BC_3F_{2:4}$ population across different temperatures (Y-axis = percent germination; X-axis = temperature in degrees Celsius; colored lines = percent germination obtained from the different time points after the seed were plated).

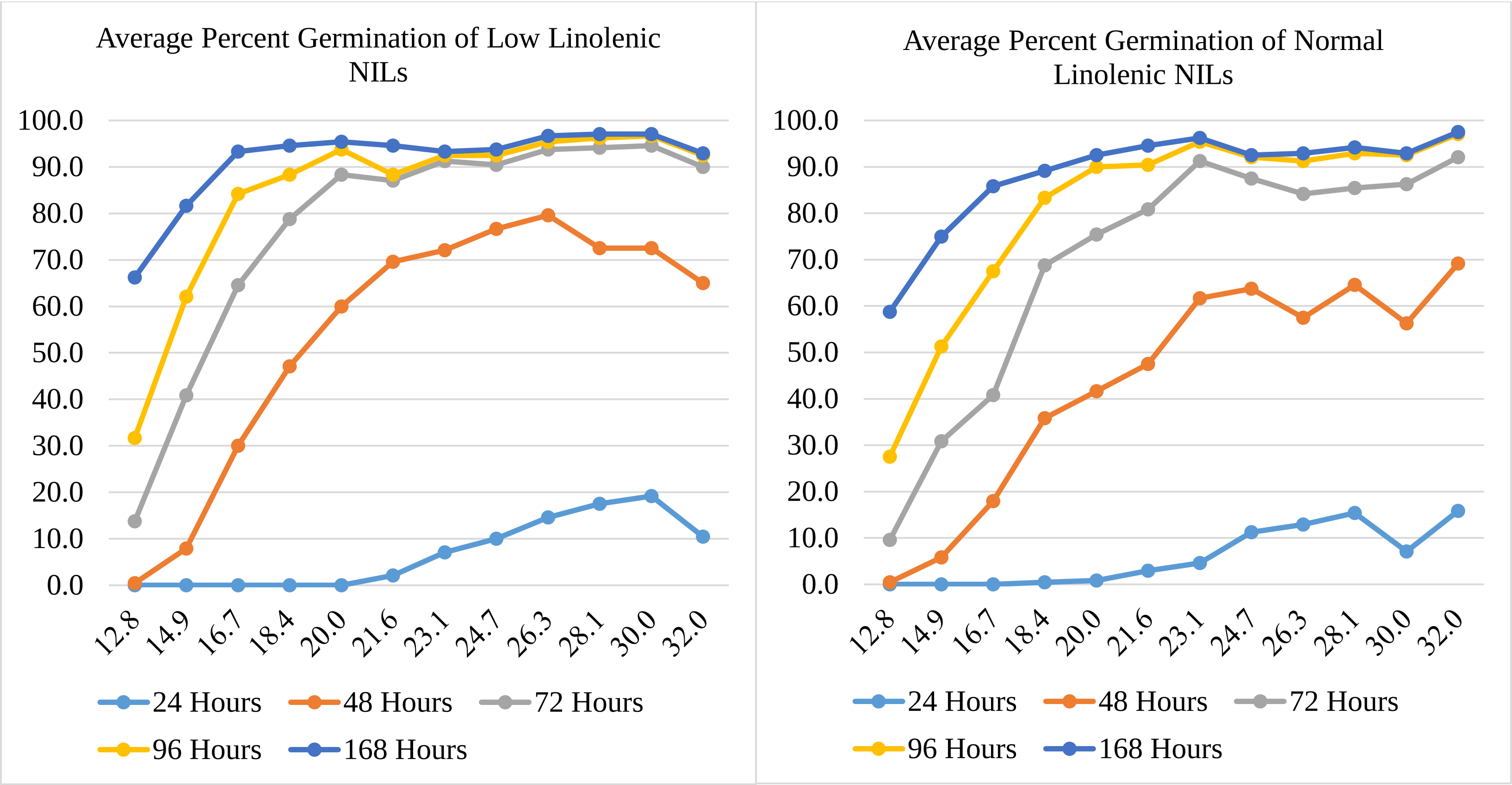

Figure 2.3. Percent germination of low (left) vs. normal (right) linolenic acid near-isolines (NILs) derived from a G00-3213 BC$_3$F$_{2:4}$ population across different temperatures (Y-axis = percent germination; X-axis = temperature in degrees Celsius; colored lines = percent germination obtained from the different time points after the seed were plated).

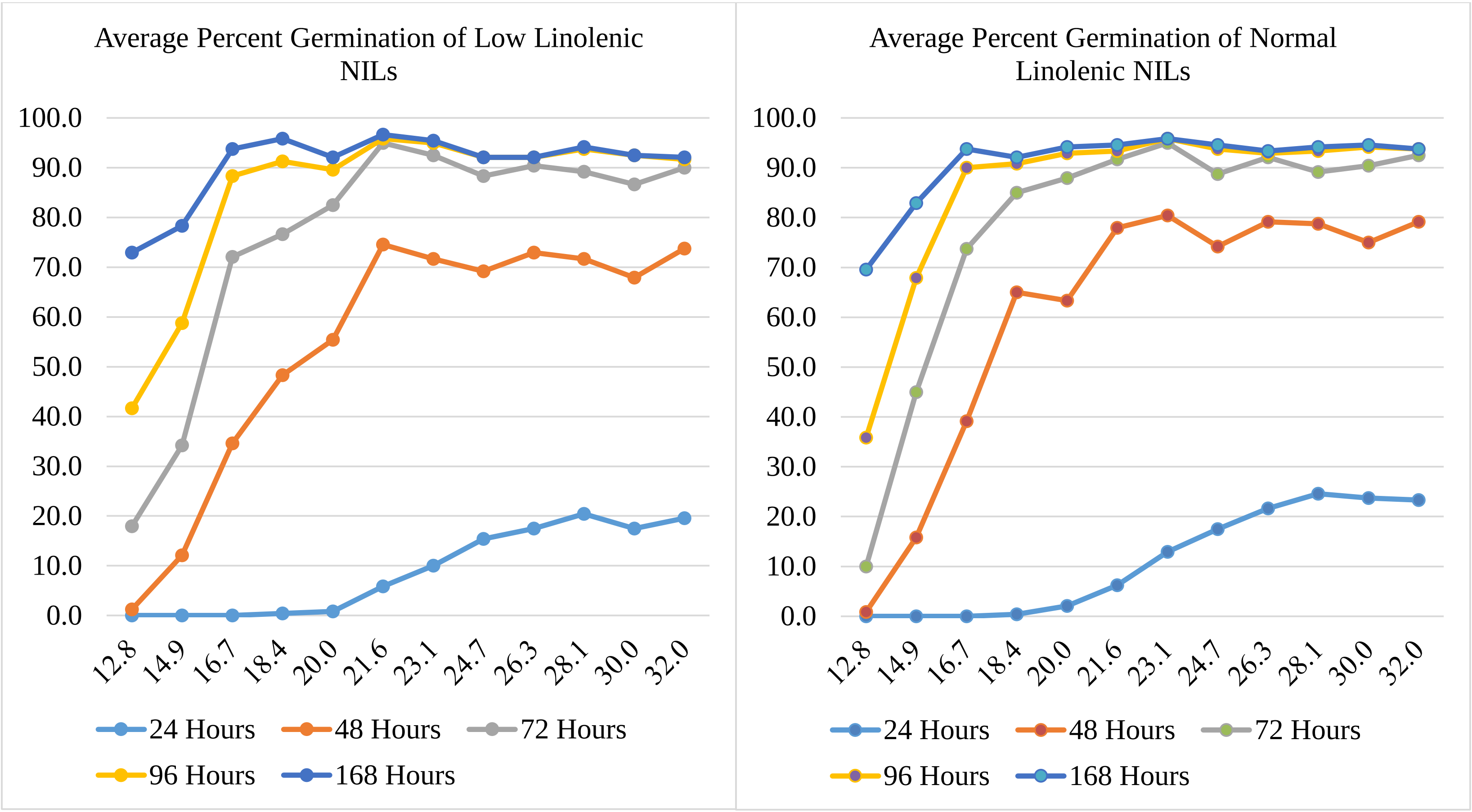

Figure 2.4. Percent germination of low (left) vs. normal (right) linolenic acid near-isolines (NILs) derived from a G00-3880 $BC_3F_{2:4}$ population across different temperatures (Y-axis = percent germination; X-axis = temperature in degrees Celsius; colored lines = percent germination obtained from the different time points after the seed were plated).

Table 2.1. Number of near-isolines (NILs) for oleic and linolenic acids included in the yield trials and their respective genotypes.

Line	NIL Genotype			Number of lines	Source
	FAD2-1A†	*FAD2-1B*‡	*FAD3A*§		
G00-3213 NILs	**MUT**¶	**MUT**	WT #	11	High oleic backcross
	MUT	WT	WT	12	
	WT	**MUT**	WT	8	
	WT	WT	WT	12	
G00-3213 NILs	WT	WT	**MUT**	10	Low linolenic backcross
	WT	WT	WT	10	
G00-3880 NILs	WT	WT	**MUT**	6	Low linolenic backcross
	WT	WT	WT	6	

† Gene responsible for oleic acid.
‡ Gene responsible for oleic acid.
§ Gene responsible for linolenic acid.
¶ MUT= homozygous mutant alleles.
# WT= homozygous wild type alleles.

Table 2.2. Number of near-isolines (NILs) for oleic and linolenic acids included in germination tests using thermogradient tables and their respective genotypes.

Line	**Genotype**			**Number of lines**	**Source**
	***FAD2-1A*†**	***FAD2-1B*‡**	***FAD3A*§**		
G00-3213 NILs	**MUT** ¶	**MUT**	WT#	6	High oleic backcross
	WT	WT	WT	6	
G00-3213 NILs	WT	WT	**MUT**	6	Low linolenic backcross
	WT	WT	WT	6	
G00-3880 NILs	WT	WT	**MUT**	6	Low linolenic backcross
	WT	WT	WT	6	

† Gene responsible for oleic acid.
‡ Gene responsible for oleic acid.
§ Gene responsible for linolenic acid.
¶ MUT = homozygous mutant alleles.
# WT = homozygous wild type alleles.

Table 2.3. Performace of agronomic traits and fatty acid profiles of near-isolines (NILs) with different *FAD2* genotypes derived from the G00-3213 BC$_3$F$_{2:4}$ population.

NIL Genotype		**Yield§**	**Palmitic acid**	**Stearic acid**	**Oleic acid**	**Linoleic acid**	**Linolenic acid**	**Maturity ¶**	**Seed weight**	**Plant height#**
***FAD2-1A*†**	***FAD2-1B*‡**									
		kg ha^{-1}	**g kg^{-1}**					**days**	**mg seed^{-1}**	**cm**
MUT ††	**MUT**	2,737 a§§	74 a	28 a	806 a	47 a	45 a	46 a	164 a	86 a
MUT	**WT‡‡**	2,939 b	109 b	32 b	349 b	440 b	70 b	46 a	172 bc	91 a
WT	**MUT**	2,623 a	108 b	33 b	319 c	471 c	69 b	46 a	168 ac	85 a
WT	**WT**	2,730 a	116 c	33 b	252 d	533 d	67 b	45 a	168 ac	86 a
G00-3213 (Recurrent Parent)		3,060	119	31	247	534	68	44	168	90

† Gene responsible for oleic acid.
‡ Gene responsible for oleic acid.
§ Yield= average yield of the lines in each NIL genotype group from two replications across two locations.
¶ Days after 31 August when 95% of the pods were their mature color. Data was collected from the Athens, GA location.
# Data collected from the Athens, GA location.
†† MUT= homozygous mutant alleles.
‡‡ WT= homozygous wild type alleles.
§§ *Tukey's Studentized Range Test* at the 0.05 probability level was used for comparisons between groups and means with the same letter are not significantly different at the 0.05 probability level.

Table 2.4. Comparsion of yield and oleic and linolenic contents of G00-3213 near-isolines (NILs) derived from the G00-3213 $BC_3F_{2:4}$ population between locations.

Location	NILs for Oleic Acid			NILs for Linolenic Acid		
	Yield	Oleic acid (All NILs)	Oleic Acid (High Oleic NILs only)	Yield	Linolenic acid (All NILs)	Linolenic acid (Low linolenic NILs only)
	kg ha^{-1}	g kg^{-1}		kg ha^{-1}	g kg^{-1}	
Athens	2,860 A	421 A	785 A	2,963 A	64 A	38 a
Plains	2,541 B	481 B	831 B	2,737 B	53 B	35 b

Note: The locations with different lowercase letters indicate significant difference at the 0.05 probability level and locations with different uppercase letters indicate significant difference at the 0.01 probability level.

Table 2.5. Average tempreatures at two test locations, Athens and Plains, GA in August, Septmber and Octorber of 2014†

Location	**August**			**September**			**October**		
	Ave. high‡	**Ave. low§**	**Ave. ¶**	**Ave. high**	**Ave. low**	**Ave.**	**Ave. high**	**Ave. low**	**Ave.**
	˚C								
Athens, GA	32.9	19.6	26.3	28.8	19.0	23.9	25.2	10.6	17.9
Plains, GA	32.9	19.6	26.3	29.8	19.3	24.6	26.6	10.9	18.8

†Data was obtained from http://www.wunderground.com/history
‡ Ave. high = average high temperature within a month.
§ Ave. low = average low temperature within a month.
¶ Ave. = average temperature within a month.

Table 2.6. Analysis of variance (ANOVA) for yield, oleic and linolenic contents among G00-3213 near-isolines (NILs) derived from the G00-3213 $BC_3F_{2:4}$ population and recurrent parent.

Source	**NILs for Oleic Acid**			**NILs for Linolenic Acid**		
	DF†	**Yield**	**Oleic acid**	**DF**	**Yield**	**Linolenic acid**
		MS‡	**MS**		**MS**	**MS**
Location	1	0.97***	1,809.13***	1	0.02 ^{NS}	8.69***
NIL group	3	0.24***	34,451.29***	1	0.002 ^{NS}	215.38***
Location x NIL group	3	0.03^{NS}	50.85*	1	0.03 ^{NS}	8.93**

† DF= Degree of freedom.
‡ MS= Mean square.
* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
*** Significant at the 0.001 probability level.
NS= not significant.

Table 2.7. Performace of agronomic traits and fatty acid profiles of near-isolines (NILs) with different *FAD3A* genotypes derived from the G00-3213 BC$_3$F$_{2:4}$ population.

***FAD3A* NIL Genotype†**	**Yield‡**	**Palmitic acid**	**Stearic acid**	**Oleic acid**	**Linoleic acid**	**Linolenic acid**	**Maturity §**	**Seed weight**	**Plant height¶**
	kg ha^{-1}	**g kg^{-1}**					**days**	**mg seed^{-1}**	**cm**
MUT#	2,744 a‡‡	115 a	37 a	225 a	586 a	37 a	45 a	149 a	89 a
WT††	2,724 a	107 b	31 b	241 b	550 b	70 b	42 b	153 a	88 a
G00-3213 (Recurrent Parent)	3,060	119	31	247	540	68	44	168	90

† Gene responsible for linolenic acid.

‡ Yield= average yield of the lines in each NIL genotype group from two replications across two locations.

§ Days after 31 August when 95% of the pods were their mature color. Data was collected from the Athens, GA location.

¶ Data collected from the Athens, GA location.

MUT= homozygous mutant alleles.

†† WT= homozygous wild type alleles.

‡‡ *Tukey's Studentized Range Test* at the 0.05 probability level was used for comparisons between groups and means with the same letter are not significantly different at the 0.05 probability level.

Table 2.8. Performace of agronomic traits and fatty acid profiles of near-isolines (NILs) with different *FAD3A* genotypes derived from the G00-3880 $BC_3F_{2:4}$ population.

***FAD3A* NIL Genotype†**	**Yield‡**	**Palmitic acid**	**Stearic acid**	**Oleic acid**	**Linoleic acid**	**Linolenic acid**	**Maturity §**	**Seed weight**	**Plant height¶**
	kg ha^{-1}	**g kg^{-1}**					**days**	**mg seed^{-1}**	**cm**
MUT#	2,663 a‡‡	104 a	37 a	225 a	593 a	41 a	43 a	150 a	88 a
WT††	2,724 a	105 a	34 a	224 a	559 b	79 b	44 a	144 b	97 b
G00-3880 (Recurrent Parent)	2,905	116	33	231	543	77	47	150	97

† Gene responsible for linolenic acid.
‡ Yield= average yield of the lines in each NIL genotype group from two replications across two locations.
§ Days after 31 August when 95% of the pods were their mature color. Data was collected from the Athens, GA location.
¶ Data collected from the Athens, GA location.
# MUT= homozygous mutant alleles.
†† WT= homozygous wild type alleles.
‡‡ *Tukey's Studentized Range Test* at the 0.05 probability level was used for comparisons between groups and means with the same letter are not significantly different at the 0.05 probability level.

Table 2.9. Seed composition of selected G00-3213 near-isolines (NILs) derived from the G00-3213 $BC_3F_{2:4}$ population.

NIL Genotype			**N¶**	**Protein**	**Oil**	**Fiber**	**Ash**	**Source**
***FAD2-1A*†**	***FAD2-1B*‡**	***FAD3A*§**						
				g kg^{-1}				
MUT#	**MUT**	**WT††**	6	396 a‡‡	200 a	55 a	49 a	High oleic backcross
WT	**WT**	**WT**	6	390 a	207 a	52 a	51 a	
WT	**WT**	**MUT**	6	383 a	207 a	55 a	51 a	Low linolenic backcross
WT	**WT**	**WT**	6	385 a	207 a	54 a	51 a	
G00-3213 (Recurrent Parent)			2	393	203	51	51	

† Gene responsible for oleic acid.
‡ Gene responsible for oleic acid.
§ Gene responsible for linolenic acid.
¶ Number of lines included in the seed composition analysis.
# MUT= homozygous mutant alleles.
†† WT= homozygous wild type alleles.
‡‡ *Tukey's Studentized Range Test* at the 0.05 probability level was used for comparisons between groups and means with the same letter are not significantly different at the 0.05 probability level.

Table 2.10. Difference of average percent germinations between high and normal oleic G00-3213 near-isolines (NILs) derived from the G00-3213 $BC_3F_{2:4}$ population.

Time point† (h)	**Temperature (°C)**											
	12.8	14.9	16.7	18.4	20.0	21.6	23.1	24.7	26.3	28.1	30.0	32.0
24	0	0	0	0	0	1	-1	-3	-4	-5	-4	-3*
48	-0	-8**	-23**	-2	-11	-4	-11	-3	-8	-8	-6	3
72	-3	-16**	-22**	-8	-11*	-3	-2	-5	-5	-4	-3	4
96	-4	-14*	-15*	-9*	-4	-3	-3	-5*	-6*	-3	-3	1
168	-12	-9	-4	-7*	-5	-3	-2	-4	-5	-3	-4	2

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

†Time point= hours when the germinated seed were counted after seed were plated.

Table 2.11. Difference of average percent germinations between low and normal linolenic G00-3213 near-isolines (NILs) derived from the G00-3213 $BC_3F_{2:4}$ population.

Time point† (h)	**Temperature (°C)**											
	12.8	14.9	16.7	18.4	20.0	21.6	23.1	24.7	26.3	28.1	30.0	32.0
24	0	0	0	-0	-1	-1	3	-1	2	2	12**	-5
48	0	2	12**	11	18**	22**	10	13*	22**	8	16**	-4
72	4	10	24**	10	13**	6	0	3	10**	9	8**	-2
96	4	11	17**	5	4	-2	-3	0	4	3	4*	-5*
168	8	7	8*	5*	3	0	-3	1	4*	3	4*	-5*

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
†Time point= hours when the germinated seed were counted after seed were plated.

Table 2.12. Difference of average percent germinations between low and normal linolenic G00-3880 near-isolines (NILs) derived from the G00-3880 $BC_3F_{2:4}$ population.

Time point† (h)	**Temperature (°C)**											
	12.8	14.9	16.7	18.4	20	21.6	23.1	24.7	26.3	28.1	30	32
24	0	0	0	0	-1	-0	-3	-2	-4	-4	-6	-4
48	0	-4	-5	-17**	-8	-3	-9	-5	-6	-7	-7	-5
72	8	-11	-2	-8*	-5	3	-3	-0	-2	0	-4	-3
96	6	-9	-2	0	-3	3	-1	-2	-1	0	-2	-2
168	3	-5	0	4	-2	2	-0	-3	-1	0	-2	-2

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
†Time point= hours when the germinated seed were counted after seed were plated.

References

Alt, J., W. Fehr, G. Welke, and D. Sandhu. 2005. Phenotypic and molecular analysis of oleate content in the mutant soybean line M23. Crop Sci. 45:1997-2000.

Anai, T., T. Yamada, R. Hideshima, T. Kinoshita, S.M. Rahman, and 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.

Bachlava, E., R.E. Dewey, J. Auclair, S. Wang, J.W. Burton, and A.J. Cardinal. 2008. Mapping genes encoding microsomal ω-6 desaturase enzymes and their cosegregation with QTL affecting oleate content in soybean. Crop Sci. 48:640–650.

Bachlava, E., R.E. Dewey, J.W. Burton, and A.J. Cardinal. 2009. Mapping and comparison of quantitative trait loci for oleic acid seed content in two segregating soybean populations. Crop Sci. 49:433–442.

Berglund, L., M. Lefevre, H.N. Ginsberg, P.M. Kris-Etherton, P.J. Elmer, P.W. Stewart, A. Ershow, T.A. Pearson, B.H. Dennis, P.S. Roheim, R. Ramakrishnan, R. Reed, K. Stewart, K.M. Phillips, and D. Investigators. 2007. Comparison of monounsaturated fat with carbohydrates as a replacement for saturated fat in subjects with a high metabolic risk profile: studies in the fasting and postprandial states. Am. J. Clin. Nutr. 86:1611–1620.

Beuselinck, P.R., D.A. Sleper, and K.D. Bilyeu. 2006. An assessment of phenotype selection for linolenic acid using genetic markers. Crop Sci. 46:747–750.

Bilyeu, K., L. Palavalli, D. Sleper, and P. Beuselinck. 2005. Mutations in soybean microsomal omega-3 fatty acid desaturase genes reduce linolenic acid concentration in soybean seeds. Crop Sci. 45:1830–1836.

Bilyeu, K., J.D. Gillman, and A. LeRoy. 2011. Novel *FAD3* mutant allele combinations produce soybeans containing 1% linolenic acid in the seed oil. Crop Sci. 51:259–264.

Boerma, H.R., R.S. Hussey, D.V. Phillips, E.D. Wood, and S.L. Finnerty. 1992. Registration of 'Cook' soybean. Crop Sci. 32:497.

Boerma, H.R., R.S. Hussey, D.V. Phillips, E.D. Wood, G.B. Rowan, and S.L. Finnerty. 1997. Registration of ‘Benning’ soybean. Crop Sci. 37:1982.

Boerma, H.R., R.S. Hussey, D.V. Phillips, E.D. Wood, G.B. Rowan, and S. L. Finnerty 2000. Registration of 'Boggs' soybean. Crop Sci. 40:294–295.

Brace, R.C., W.R. Fehr, and S.R. Schnebly. 2011. Agronomic and seed traits of soybean lines with high oleate concentration. Crop Sci. 51:534–541.

Brink, K., C.F. Chui, R.F. Cressman, P. Garcia, N. Henderson, B. Hong, C.A. Maxwell, K. Meyer, J. Mickelson, K.L. Stecca, C.W. Tyree, N. Weber, W. Zeng, and C.X. Zhong. 2014. Molecular characterization, compositional analysis, and germination evaluation of a high-oleic soybean generated by the suppression of expression. Crop Sci. 54:2160–2174.

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.

Camacho-Roger, A.M. 2006. Molecular markers and genes associated with low palmitic and low linolenic acid content in N97-3681-11 and N97-3708-13 soybean lines. NCSU Institutional Repository http://www.lib.ncsu.edu/resolver/1840.16/2909.

Cardina, J. and J.E. Hook. 1989. Factors influencing germination and emergence of Florida beggarweed (*Desmodium totruo-sum*). Weed Tech. 3:402–407.

Carter, T.E., J.W. Burton, D.T. Bowman, Z. Cui, X. Zhou, M.R. Villagarcia, A.S. Niewoehner, and M.O. Fountain. 2003. Registration of ‘N7001’ Soybean Registration by CSSA. Crop Sci. 43:1126–1127.

Chappell, A.S. and K.D. Bilyeu. 2006. A *GmFAD3A* mutation in the low linolenic acid soybean mutant C1640. Plant Breeding 125:535–536.

Chatterton, N. and A. Kadish. 1969. A temperature gradient germinator. Agron. J. 61:643–644.

Cherry, J.H., L. Bishop, P.M. Hasegawa, and H.R. Lefflert. 1985. Differences in the fatty acid composition of soybean seed produced in northern and southern areas of the U.S.A. Phytochemistry 24:237–241.

Cherry, J.H., L. Bishop, N. Leopold, C. Pikaard, and P.M. Hasegawa. 1984. Patterns of fatty acid deposition during development of soybean seed. Phytochemistry 23:2183–2186.

Dierking, E. 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. 2007. Breeding for modified fatty acid composition in soybean. Crop Sci. 47:S72–S87.

Fehr, W.R. and E.G. Hammond. 2000. Reduced linolenic acid production in soybean. U.S. Patent No. 6133509. Date issued: 17 Oct.

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 Biotechnol. J. 7:411–421.

Grey, T.L., J.P. Beasley, T.M. Webster, and C.Y. Chen. 2011. Peanut seed vigor evaluation using a thermal gradient. J. Agron. 2011:7.

Hammond, E.G. and W.R. Fehr. 1983. Registration of A5 germplasm line of soybean (reg. No. GP44). Crop Sci. 23:192–192.

Howell, R.W. and F.I. Collins. 1957. Factors affecting linolenic and linoleic acid content of soybean oil. Agron. J. 49:593–597.

Kinney, A.J. 1996. Development of genetically engineered soybean oils for food applications. J. Food Lipids 3:273–292.

Keim, P., T.C. Olson, and R.C. Shoemaker. 1988. A rapid protocol for isolating soybean DNA. Soybean Genet. Newsl. 15:150–152.

Knowlton, S. 1999. Soybean oil having high oxidative stability. U.S. Patent No. 5981781. Date issued: 9 November

Kumar, V., A. Rani, S. Solanki, and S.M. Hussain. 2006. Influence of growing environment on the biochemical composition and physical characteristics of soybean seed. J. Food Comp. Anal. 19:188–195.

La, T.C., S.M. Pathan, T. Vuong, J.D. 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.

Lee, J.D., K.D. Bilyeu, and J.G. Shannon. 2007. Genetics and breeding for modified fatty acid profile in soybean seed oil. J. Crop Sci. Biotech 10:201–210.

Lee, J.D., M. Woolard, D.A. Sleper, J.R. Smith, V.R. Pantalone, C.N. Nyinyi, A. Cardinal, and J.G. Shannon. 2009. Environmental effects on oleic acid in soybean seed oil of plant introductions with elevated oleic concentration. Crop Sci. 49:1762–1768.

Maestri, D.M., D.O. Labuckas, J.M. Meriles, A.L. Lamarque, J.A. Zygadlo, and C.A. Guzman. 1998. Seed composition of soybean cultivars evaluated in different environmental regions. J. Sci. Food Agric. 77:494–498.

Mohapatra, S.C. and W.H. Johnson. 1986. A thermogradient incubator for seed germination studies. Agron. J. 78:351–356.

Oliva, M.L., J.G. Shannon, D.A. Sleper, M.R. Ellersieck, A.J. Cardinal, R.L. Paris, and J.D. Lee. 2006. Stability of fatty acid profile in soybean genotypes with modified seed oil composition. Crop Sci. 46:2069–2075.

Pham, A.T., K. Bilyeu, P. Chen, R. Boerma, and Z. Li. 2014. Characterization of the *fan1* locus in soybean line A5 and development of molecular assays for high-throughput genotyping of *FAD3* genes. Mol. Breeding 33:895–907.

Pham, A.T., J.D. 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 Biology 10:195.

Pham, A.T., J.D. Lee, J.G. Shannon, and K.D. 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.

Reddy, B.V.S., A.A. Kumar, and S. Ramesh. 2007. Sweet sorghum: a water saving bio-energy crop. International Conference on Linkages between Energy and Water Management for Agriculture in Developing Countries. 29–30 Jan. IWMI, ICRISAT Campus, Hyderabad, India

Reinprecht, Y., S.Y. Luk-Labey, J. Larsen, V.W. Poysa, K. Yu, I. Rajcan, G.R. Ablett, and K.P. Pauls. 2009. Molecular basis of the low linolenic acid trait in soybean EMS mutant line RG10. Plant Breeding 128:253–258.

Reinprecht, Y., V.W. Poysa, I. Rajcan, G.R. Ablett and K.P. Pauls. 2006. Agronomic performance of soybean with seed lipoxygenase nulls and low linolenic acid content. Can. J. Plant Sci. 86:379–387.

Rennie, B., D. and J.W. Tanner. 1991. New allele at the *Fan* locus in the soybean line A5. Crop Sci. 31:297–301.

Ross, A.J., W.R. Fehr, G.A. Welke, and S.R. Cianzio. 2000. Agronomic and seed traits of 1%-linolenate soybean genotypes. Crop Sci. 40:383–386.

Schlueter, J.A., J.Y. Lin, S.D. Schlueter, I.F. Vasylenko-Sanders, S. Deshpande, J. Yi, M. O'Bleness, B.A. Roe, R.T. Nelson, B.E. Scheffler, S.A. Jackson, and R.C. Shoemaker. 2007. Gene duplication and paleopolyploidy in soybean and the implications for whole genome sequencing. BMC Genomics 8:330.

Shi, Z., N. Bachleda, A. 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.

Sudaric, A., D. Simic, and M. Vrataric. 2005. Characterization of genotype by environment interactions in soybean breeding programmes of southeast Europe. Plant Breeding 125:191–194.

Thomas, A. 2000. Fats and fatty oils. Ullmann's encyclopedia of industrial chemistry. John Wiley and Sons, Hoboken, NJ.

Wilcox, J.R. and J.F. Cavins. 1986. Registration of C1640 soybean germplasm. Crop Sci. 26:209–210.

Wilcox, J.R., A.D. Nickell, and J.F. Cavins. 1993. Relationships between the *fan* allele and agronomic traits in soybean. Crop Sci. 33:87–89.

Wolf, R.B., J.F. Cavins, R. Kleiman, and L.T. Black. 1982. Effect of temperature on soybean seed constituents: oil, protein, moisture, fatty acids, amino acids and sugars. J. Am. Oil Chem. Soc. 59:230–232.

CHAPTER 3

IDENTIFYING *FATB1A* DELETION THAT CAUSES REDUCED PALMITIC ACID CONTENT IN SOYBEAN N87-2122-4 TO DEVELOP A FUNCTIONAL MARKER FOR MARKER-ASSISTED SELECTION[1]

---

[1] Nicole Bachleda, Anh Pham, and Zenglu Li. Submitted to *Crop Science*, 8/14/2015.

Abstract

Palmitic acid is a major saturated fatty acid in soybean oil, and consumption of saturated fat is linked to a risk of coronary diseases. Development of soybean (*Glycine max*) cultivars with reduced palmitic acid content is an important goal of soybean breeding. The *FATB1a* gene was previously found to be responsible for reduced palmitic acid in the soybean line N87-2122-4. The objective of this research was to characterize the *FATB1a* gene in N87-2122-4 and develop a breeder-friendly, functional marker to facilitate marker-assisted selection and improve breeding efficiency for reduced palmitic soybeans. With the availability of soybean genetic maps, reference genome, and gene annotations, an approximate 254 kb deleted genomic region, including the *FATB1a* gene was identified. Based on the gene deletion information, we developed a TaqMan marker and tested it with a segregating F_2 population that consisted of 140 individual plants derived from 'Cook' x N87-2122-4. Analyses of the fatty acids in this population were conducted using gas chromatograph. The marker performed well and accounted for 57% of the phenotypic variation. The marker was also validated using a panel of 121 diverse soybean lines with known fatty acid profiles. The result indicated that the marker can be used effectively in marker-assisted breeding for reduced palmitic acid in soybean.

Keywords: Reduced palmitic acid, TaqMan assay, Functional marker, Marker-assisted selection

Introduction

Soybean [*Glycine max* (L.) Merr.] is a major global oil crop and accounted for 55% of U.S. vegetable oil consumption in 2014 (http://soystats.com, 2015). This widely consumed oil is made up of five major fatty acids: palmitic (110 g kg^{-1}), stearic (40 g kg^{-1}), oleic (250 g kg^{-1}), linoleic (520 g kg^{-1}), and linolenic (80 g kg^{-1}) acids (Fehr, 2007). It has been long considered that consuming saturated fat, including palmitic and stearic acids, increased cholesterol and therefore the risk of coronary artery disease (Lawrence, 2013; Steinberg, 2005). A growing body of evidence is challenging whether saturated fat is directly associated with coronary events (Puaschitz et al., 2015). However, the American Heart Association still recommends limiting saturated fat consumption and instead replacing it with mono- and polyunsaturated fats. Since soybean oil accounts for more than half of the vegetable oil consumption in the USA (http://soybase.org) and contains 15% saturated fat, it is an excellent candidate to trim saturated fats from the American diet. As palmitic acid is a major saturated fatty acid in soybean oil, reducing the amount of this fatty acid in soybean seed oil would improve the oil quality for human consumption. Hence, there have been many attempts to develop reduced palmitic acid soybeans.

Modified palmitic acid soybean has been achieved through chemical mutagenesis, hybridization, and recurrent selection. Using ethyl methanesulfonate (EMS) on ‘Century,’ Wilcox et al. (1990) developed a reduced palmitic line C1726 (85 g kg^{-1}) and an increased palmitic line C1727 (172 g kg^{-1}). Fehr et al. (1991) selected a palmitic mutant line A1937NMU-173 (68 g kg^{-1}), designated as A22 from cultivar ‘A1937’ using N-nitroso-N-methyl-urea. C1726 possesses a recessive *fap1* allele on chromosome (Chr) 9 (Bubeck et al., 1989), while A22 carries a recessive *fap3* allele on Chr 5 (Fehr et al., 1991). Hybridization between A22 and C1726 resulted in an F_2 progeny with the genotype of *fap1fap1fap3fap3* that produced 44 g kg^{-1} palmitic acid (Fehr et al., 1991). The *fap2* allele has been found responsible for increased palmitic acid in the line C1727 through a premature stop codon in the *GmKASIIA* gene, a β-ketoacyl-acyl-carrier protein synthase responsible for synthesizing the palmitoyl-acyl carrier protein (16:0-ACP) (Anai et al., 2012; Beuselinck et al., 2006; Erickson et al., 1988).

Additional reduced palmitic lines have also been developed through a recurrent selection method, such as lines N87-2122-4, N79-2077, N94-2575, and C1943 (Burton et al., 1994; Burton et al., 1998). Soybean line N87-2122-4 has palmitic acid of 53 g kg^{-1} and is an F_6–derived line from a cross between N78-2245 and N79-2077 (Burton et al., 1994). Parental line N78-2245 has normal palmitic acid levels (100 g kg^{-1}), but increased oleic acid (383 g kg^{-1}) and decreased linoleic (426 g kg^{-1}) and linolenic acids (55 g kg^{-1}). In contrast, parental line N79-2077 has decreased palmitic acid (60 g kg^{-1}), which was derived from parents PI 90406, PI 92567, and N69-2774 in the fifth cycle of a high-oleic recurrent selection experiment (Burton et al., 1994). Soybean line RG3 with the reduced palmitic acid of ~45 g kg^{-1}, derived from the cross C1726 x Ellp2, is another source of soybean with reduced palmitic acid (Primomo et al., 2002). Soybean line RG3 possesses mutations at both *fap1* and *fap3-ug* loci and a nonsense mutation in the *fap3-ug/fatb1a* gene and a splice site mutation in *fap1/KASIII* gene was responsible for reduced palmitic acid. Both mutations are amicable to single nucleotide polymorphism (SNP) based markers. Gillman et al. (2014) was able to develop SimpleProbe molecular marker assays for these two genes.

Since its release in 1994, soybean line N87-2122-4 has been a source for the reduced palmitic trait in soybean for improving nutritional quality of soybean oil. The allele conferring reduced palmitic acid in N87-2122-4 that was inherited from its parent, N79-2077 was designated as *fap*$_{nc}$. Previous research by Wilson et al. (2001) linked the *fap*$_{nc}$ allele carried in N79-2077-12 to a genomic deletion of 16:0-ACP thioesterase enzyme (designated *FATB*) using Northern blot assays to correlate the amount of transcripts to the number of *fap*$_{nc}$ alleles. Using F_2 and $F_{2:3}$ populations derived from 'Benning' x N87-2122-4 and SSR markers, one major QTL mapped on Linkage Group (LG) A1 (Chr 5) and one minor QTL on LG M (Chr 7) were reported to be associated with reduced palmitic acid in soybean N87-2122-4 (Li et al., 2002). These QTL have additive effects and explained 51% of the total phenotypic variation for palmitic acid content in the F_2 population and 43% of the total phenotypic variation in the $F_{2:3}$ families. By isolating the full length cDNAs from three unique *FATB* genes, *GmFATB1a, GmFATB1b*, and *GmFATB2a*, Cardinal et al. (2007) indicated that the isoform designated as *GmFATB1a* represented the specific gene deleted in the lines carrying the *fap*$_{nc}$ allele, which is responsible for reduced palmitic acid; however, the location and size of the deletion in the genome was unknown. They also developed agarose-gel based allele specific primers at the *GmFATB1a* locus to genotype plants from the populations segregating for the *fap*$_{nc}$ allele. These agarose-gel based markers were scored as present or absent and the homozygous mutant lines were inferred based on the absence of PCR fragments. Developing a diagnostic, codominant, high-throughput marker for the reduced palmitic trait from N87-2122-4 will greatly improve the speed and efficiency in incorporating the reduced palmitic trait into elite soybean lines. Hence, the objectives of this study were to: i) characterize the genomic deletion of the *FATB1a* gene in the soybean line N87-2122-4 using a PCR-based strategy and ii) develop TaqMan assay based markers for detection of the deletion of the *FATB1a* gene in line N87-2122-4 to be used for high-throughput marker-assisted selection of the reduced palmitic acid trait in soybean breeding programs.

Materials and Methods

Population Development and Validation Panel

A population derived from ‘Cook’ x N87-2122-4 was selected for this study. The cross was made in 2002 in Athens, GA and advanced to the F_2 generation. N87-2122-4 is an F_6–derived line from a cross between N78-2245 (normal palmitic, high oleic acid) and N79-2077 (reduced palmitic acid) and it has palmitic acid of 53 g kg^{-1}(Burton, 1994). Both N78-2245 and N79-2077 were developed from the same fatty acid recurrent selection population at the USDA-ARS research station in Raleigh, NC (Burton et al., 1983; Burton et al., 1994). ‘Cook’ is a high yielding cultivar with multiple disease and nematode resistance, and normal palmitic acid derived from an F_5 plant of a cross between ‘Braxton’ and ‘Young’ (Boerma et al., 1992). The F_2 population of ‘Cook’ x N87-2122-4 was used to test the genetic association of the marker and palmitic acid content, and usefulness in discerning heterozygous lines from the homozygous mutants or wild types. Due to poor germination, only 158 out of 386 planted seed from the ‘Cook’ x N87-2122-4 population germinated and grew for leaf tissue collection and DNA extraction as well as subsequent marker genotyping.

To determine the robustness of the marker in different genetic backgrounds, 121 diverse soybean lines were selected (Table 3.2). These soybean lines were selected based on diverse pedigrees and availability of fatty acid profile. These lines consisted of 33 out of 35 soybean ancestors defined by Gizlice et al. (1994), elite soybean cultivars and lines, and other plant introductions (PIs).

Identification of the Genomic Deletion of *FATB1a* Gene in N87-2122-4

To search for the deleted genomic region in N87-2122-4, the 4 kb *FATB1a* gene plus 300 kb upstream and 300 kb downstream sequence, totaling 604 kb, was obtained from Phytozome website (Glycine max Wm82.a2.v1; http://phytozome.jgi.doe.gov/). Four rounds of primer pairs with a total of 58 primer pairs were designed at upstream and downstream locations of the *FATB1a* gene based on soybean reference genome (Wm82.a2.v1) and gene annotation sequence data from Phytozome and Soybase (http://soybase.org/) websites. Primers were designed using Primer3Plus software (Untergasser et al., 2012). Secondary structures of the primers designed in Primer3Plus were examined using Primer Express 3.0.1 software (Life Technologies, Carlsbad, CA, USA) in order to select best primer pairs with ideal secondary structures and matching melting temperatures for PCR. The specificity of primers was tested by blasting the sequences in SoyBase to ensure the primers would not produce false positives or multiple products. Primers were synthesized at Sigma-Aldrich (Saint Louis, MO) and tested with N87-2122-4 (reduced palmitic acid) and control lines (normal palmitic acid), ‘Benning,’ ‘Bossier,’ ‘Cook,’ and ‘Davis’ (Fig. 3.1).

DNA extraction from N87-2122-4, ‘Cook’, and the F_2 population of ‘Cook’ x N87-2122-4 began by collecting young trifoliolate leaves from seedlings. The tissue samples were then dried in an incubator at 55°C for approximately 24 h. DNA was subsequently extracted using the hexadecyltrimethylam-monium bromide (CTAB) procedure from Keim et al. (1988). After DNA extraction, 149 of the 158 samples had good quality DNA for subsequent marker testing.

The PCR reactions for deletion identification were performed in 20 µL reactions containing 50–100 ng DNA template, 0.5 µM final concentration of primers, 3.5 mM $MgCl_2$, 200 µM dNTPs, and 0.2 µL Flexi Promega Taq polymerase Flexi Buffer (Promega Corporation, Madison, WI, USA). PCR conditions used were as follows: 5 min at 95°C, 35 cycles of 15 s at 95°C, 20 s at 60°C, and 30 s at 72°C. After PCR, products were examined on a 1% agarose gel by electrophoresis by combining 5 µL PCR product with 2 µL 5x loading dye and running at 150 V for 30 min. PCR fragments were visualized under UV light.

For sequencing, 120 PCR fragments were produced from four replicated samples of N87-2122-4 and two controls, 'Benning' and 'Bossier,' using forward and reverse primers from 145, 148, and 147 kb upstream and 104, 105, 106, 107, and 108 kb downstream. Additionally, the control products were obtained from the primer pair 145F-US and 104R-DS, however, these PCR products were very large and had poor read quality, so additional primers were used to generate reduced size PCR fragments for sequencing. N87-2122-4 was also included in the tests of these additional primers, however, the deletion caused a lack of PCR products. All sequencing was performed at the Georgia Genomics Facility at the University of Georgia, Athens campus. Sequences were assembled using Geneious version 7.1.2 software (http://www.geneious.com) (Kearse et al., 2012).

Development of TaqMan Marker and Genotyping

With the identified deletion information, a TaqMan marker, designated as GSM369, was designed using PrimerExpress 3.0.1 software based on the location of the deletion region of the *FATB1a* gene in N87-2122-4. GSM369 was composed of one "mutant" and one "wild type" set of primers and probes. The "wild type" probe is labeled with VIC dye that aligns to the DNA sequence within the deletion region, and therefore cannot bind to the region containing the deletion in a mutant line. Each probe has its own set of forward and reverse primers.

Each F_2 seed of the 'Cook' x N87-2122-4 population was chipped with a razor blade for fatty analysis and the remaining part of the seed was grown for leaf tissue collection and subsequent marker genotyping. Fatty acid profiling was conducted using ¼ seed chips from the F_2 population, while the remaining viable ¾ seeds containing embryonic axis were grown in a greenhouse located on the University of Georgia campus in Athens, GA. The ¾ seed were planted in sand flats in December 2013 and tissue collection was conducted approximately two weeks after planting. DNA was extracted using a CTAB extraction method (Keim et al., 1988).

The PCR reactions for TaqMan maker genotyping were performed in 384-well plates at 4 µL per well containing approximately 40 ng DNA, 1.8 µL of 2 × TaqMan Universal MasterMix II (Life Technologies, Carlsbad, CA, USA), and 0.2 µL of 5 × assay mix (0.223 µM of each amplification primer, 0.10 µM FAM mutant probe, and 0.05 µM VIC wild type probe at final concentration) (Pham et al., 2014).

Genotyping the F_2 population derived from 'Cook' x N87-2122-4 with the marker GSM369 was performed using Bio-Rad DNA Engine Tetrad2 (Hercules, CA) under PCR conditions of 10 min at 95°C, and 50 cycles of 20 s at 92°C and 1 min at 58°C. Two parents N87-2122-4 and 'Cook', each with two replicates were included as controls.

Fatty Acid Profiling and Data Analysis

The fatty acid profiling was done with the ¼ F_2 seed chips of the F_2 population derived from 'Cook' x N87-2122-4 at the DNA facility located at the Iowa State University, Ames, IA, following protocols reported by Beuselinck et al. (2006). For the majority of validation lines, fatty acid data were obtained through the USDA Germplasm Resources Information Network (GRIN) (http://www.ars-grin.gov). However, for the lines that did not have fatty acid data available, gas chromatography (GC) analyses were performed at either the DNA facility of Iowa State University, Ames, IA or the USDA-ARS Molecular Genetics Laboratory, Columbia, MO.

Fatty acid content is reported as a proportion of total fatty acid present in extracted oil by a GC method.   Statistical analyses were performed using a general linear model (GLM) in RStudio (www.rstudio.com), an open source software interface for R 3.0.2 (http://www. r-project.org).

Results

Identification of the Deletion of *FATB1a* Genomic Region in the Soybean Line N87-2122-4

Based on the initial sequence of the *FATB1a* gene plus 300 kb upstream and 300 kb downstream obtained from the Phytozome website (Glycine max Wm82.a2.v1; http://phytozome.net), 28 primer pairs were first designed to be every 5 kb within the region of 70 kb upstream and 70 kb downstream of the genes. No PCR products were found for N87-2122-4 within the selected regions, while PCR products were observed for ‘Cook’ and ‘Benning’ (the wild-type phenotype, control samples) indicating the deletion might be larger than this scope.

Subsequently, 10 primer pairs were designed to amplify PCR products every 50 kb from 100 to 300 kb upstream and downstream of the *FATB1a* gene regions. In both upstream and downstream regions, there were PCR products for N87-2122-4 from 150 to 300 kb, but not at 100 kb upstream, while there were PCR products for controls for all of the primers in these regions. The regions from 100–150 kb upstream and downstream of the *FATB1a* were further investigated.

Eight primer pairs were designed for upstream or downstream regions to amplify PCR products at every 10 kb in the region from 100–150 kb upstream and downstream of the

FATB1a. For the upstream region, primers at all positions of 110, 120, 130, and140 kb yielded PCR products for controls, but not for N87-2122-4. For the downstream region, all four primers gave PCR products for both N87-2122-4 and controls.

Twelve primer pairs were further designed for the region 140–150 kb (5 primer pairs at 143, 144, 145 147, and 148 kb) upstream and 100–110 kb (7 primer pairs at 101, 103, 104, 105, 106, 107, and 108 kb) downstream of the *FATB1a* gene (Fig. 3.1). For the upstream region, primers at 145 kb upstream and beyond gave PCR products for all of the samples, while primer pairs from 143–144 kb showed PCR products in the control genotypes, but not in the mutant. For the downstream region, primers from 104–108 kb downstream produced PCR products for all samples (N87-2122-4 and controls), but primers at 101–103 kb showed PCR products for control DNAs, but not for N87-2122-4. Missing fragments at this genomic region in N87-2122-4 indicated the deletion and the deletion region was narrowed to 145 kb upstream and 104 kb downstream of the gene, for a deletion of approximate 254 kb in length.

The forward primers of the primer pair at 145 kb upstream and the reverse primer of the primer pair at 104 kb downstream were then used for PCRs. The primer pair of 145F-US and 104R-DS successfully produced approximate 1.4 kb product for N87-2122-4, but no PCR product for the controls. Based on the result, it can be determined that the genomic deletion in line N87-2122-4 begins 145 kb upstream and ends 104 kb downstream of the gene, and is from 1,277,909 to 1,023,799 bp with an interval of 254,110 bp on Chr 5. According to the annotation information from Soybase (Soybase.org), this deleted genomic region contains 19 genes (Table 3.3). Acyl-ATP thioesterase gene (Glyma.05g012300) reported by Wilson et al. (2001) is included in this region.

Genetic Association Analysis of TaqMan Marker and Fatty Acid Content

Based on sequence data, we designed a TaqMan assay around the deletion region of *FATB1a* in N87-2122-4, designated as GSM369. To detect the wild type genotypes, the primers and probe were designed by targeting the “wild type” sequence at the deleted genomic region. Likewise, the “mutant” FAM probe binds to the region that bridges the deletion, and therefore cannot bind in the lines that do not contain the deletion. The “mutant” probe also has its separate set of forward and reverse primers to detect the mutant allele (Fig. 3.2).

Testing of the TaqMan marker GSM369 was initially performed on the parents of the population, ‘Cook’ and N87-2122-4. The ‘Cook’ allele (normal palmitic acid) clearly separated from mutant allele of N87-2122-4 (reduced palmitic acid) in the clusters along the VIC-axis (X-axis) and FAM-axis (Y-axis,), respectively. The primers and probes successfully targeted the expected regions at *FATB1a* locus and optimum PCR assay conditions had been reached.

Genotyping the F_2 population derived from ‘Cook’ x N87-2122-4 with GSM369 showed a distinct separation of three clusters, two homozygous and one heterozygous (Fig. 3.3). The allelic segregation of GSM369 in the F_2 population of ‘Cook’ x N87-2122-4 was tested with a Chi-squared test, which fitted the expected segregation ratio 1:2:1 of this allele in an F_2 generation (P-value=0.80). Statistical association analysis of marker genotypes and palmitic acid determined that the three categories (homozygous mutant, heterozygous, and homozygous wild type) are statistically different ($P < 0.001$) with an R^2 value of 0.57 (Table 3.1 and 3.4.)

To validate if this TaqMan marker can be used in diverse genetic backgrounds for genotyping and selection, a panel of 121 diverse soybean lines with known palmitic acid were genotyped using the marker GSM369. Of these lines, ‘Benning low lin/low palm’ is a near-isoline of cultivar ‘Benning’ (Boerma et al., 1997), which carries the reduced palmitic and linolenic traits introgressed from C1640 and N87-2122-4, respectively. The marker genotyping results completely matched with the phenotypes of those lines, thus demonstrating the robustness of this TaqMan marker to be used at different genetic backgrounds for marker-assisted selection (Fig. 3.4 and Table 3.2).

Discussion

In this study, we have determined that deletion at the *FATB1a* locus in N87-2122-4 on Chr 5 which caused reduced palmitic acid levels ranged from 1,277,909 to 1,023,799 bp with an interval of 254,110 bp using a stepwise PCR based genomic approach. A TaqMan marker, GSM369, was successfully developed based on the deletion of the *FATB1a*. The robustness of this marker has been confirmed with the F_2 population derived from ‘Cook’ x N87-2122-4 and a diverse panel of soybean lines. GSM369 genotypes accounted for 57% of the phenotypic variation of palmitic acid in the F_2 population. This marker has been used to reliably select the mutant allele at *FATB1a* locus from N87-2122-4 in soybean breeding for reduced palmitic acid. Fatty acids are linked through the biosynthesis pathway. In this study, the palmitic acid is significantly negatively correlated with oleic and linoleic acids, respectively ($P<0.05$), since the palmitic acid is the first fatty acid produced during the fatty acid synthesis. This is consistent with the result reported by Primomo et al. (2002). Additionally, stearic and oleic acids, oleic and linoleic acids, and linoleic and linolenic acids are all significantly ($p<0.05$) negatively correlated (Table 3.4). Therefore, decrease of palmitic acid in soybean oil could help increase oleic acid content, which has provided better health benefit (Hunter et al., 2010).

Most marker detection assays are typically used for SNPs (Giancola et al., 2006; Shen et al., 2009). In this study, we were able to develop a marker to detect the deletion of *FATB1a* allele found in N87-2122-4 and differentiate homozygote and heterozygote genotypes. Similar approach was also used by Pham et al. (2014) to detect the deletion allele from A5 for reduced linolenic acid content and Fedick et al. (2012). In Pham et al. (2014) study, they reported 6.4 kb deletion of *FAD3A* gene in soybean line A5. In contrast, the deletion (approximate 254 kb) in N87-2122-4 from this study was much larger than expected and 19 annotated genes (Table 3.3) were located in the deletion region. It is not clear at this point if the yield, seed composition, and agronomic traits are affected by this big deletion when a line possesses this *FATB1a* mutant allele. We are evaluating an NIL containing *FATB1a* allele introgressed from N87-2122-4 to determine the impacts. By use of TaqMan assays to design the marker targeting the deletion region and with the mutant probe sitting directly over the bridge of the deletion (Fig. 3.2), we have successfully developed such marker for genotyping. To determine if the primer and probe combination that we developed to target the wild type allele can work across different genetic backgrounds, we tested the primer and probe combination across a diverse germplasm panel. The diverse panel of lines included 31 out of 35 soybean ancestors defined by Gizlice et al. (1994), which contributed 95% of genes found in the modern soybean cultivars based on the pedigree analysis. With the genotyping results from 121 diverse soybean lines, we confirmed the function and robustness of this marker (Table 3.2), suggesting that the marker GSM369 unlikely produces false homozygous mutant results if the individual is heterozygous at the *FATB1a* locus.

Cardinal et al. (2007) reported an allele specific marker to identify the fap_{nc} allele. However, the marker was agarose gel-based and scored as presence or absence of the “wild type” allele. The mutant genotypes were scored based on the absence of the PCR fragment and the marker was not able to differentiate the “wild type” homozygous and heterozygous individuals. Therefore, it cannot be used in backcrossing populations, which only have heterozygous and homozygous wild type groups. With the technological advancements of fluorescence-based DNA makers, this new TaqMan marker GSM369 has the ability to discern between all three groups (Fig. 3.3 and 3.4) and provides reliable results in high throughput genotyping.

GSM369 accounted for 57% of the variation in the mean palmitic acid between homozygous mutants, heterozygotes, and homozygous wild types from a population derived from ‘Cook’ x N87-2122-4. Previously, Cardinal et al. (2007) developed an allele specific marker for the fap_{nc} that accounted for up to 70% of variation for palmitic acid in a segregating population and they also used a second population for testing, in which the marker accounted for 61.8% of phenotypic variation. Our result is consistent with the result previously reported. The difference might be due to the genetic background or environmental effects on the palmitic fatty acid level. In addition, the *FATB1a* gene is from the major QTL reported by Li et al. (2002) in a population of Benning x N87-2122-4 and another minor QTL on Chr 7 was found to have an additive effect with the major QTL on Chr 5. The minor QTL was not followed in this study, so we did not determine potential additive effects.

One point of hindrance when testing GSM369 was getting the initial PCR assay and conditions optimized. The typical assay mix using equal amounts of both probes produced distorted, inconclusive graphs with no clear segregation and slanted towards the wildtype. This problem was effectively overcome by doubling the mutant probe. Additionally, 40 PCR cycles is typically enough to get good amplification and clear allele segregation, however, we found 50 cycles made excellent improvements in marker allele clustering.

In summary, with the development of the robust TaqMan marker GSM369, high-throughput marker-assisted selection can be deployed in breeding programs for selection of reduced palmitic soybeans using N87-2122-4 as a trait donor. Efficient marker-assisted selection can easily be implemented on any populations, regardless of genetic backgrounds for a rapid selection before the palmitic acid content is even known. This allows for early selection as well as for backcross breeding.

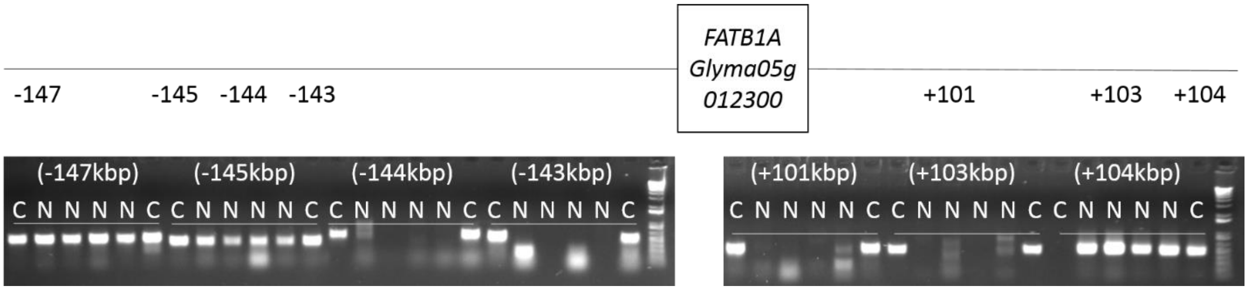

Figure 3.1. Agarose-gel images of the PCR products of N87-2122-4 (N) and controls (C) in final round to scout a deletion of *FATB1a* genomic region in N87-2122-4. Distances are in kb. Lane 1: control; Lane 2–5: N84-2122; Lane 6: control

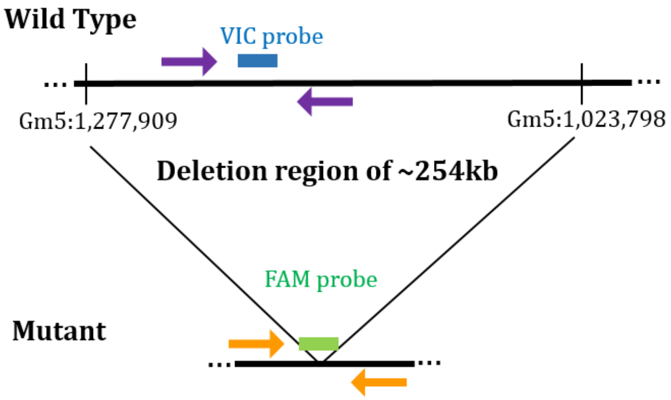

Figure 3.2. Visual representation of a deletion region in N87-2122-4 to illustrate the TaqMan marker design. Deletion including the *FATB1a* gene begins at Gm5:1,277,909 and ends at Gm5:1,023,798. Wild Type VIC probe sits within the deletion region and cannot bind to the mutant sequence; likewise, mutant FAM probe binds over the beginning and end of the deletion and cannot completely bind to the wild type sequence.

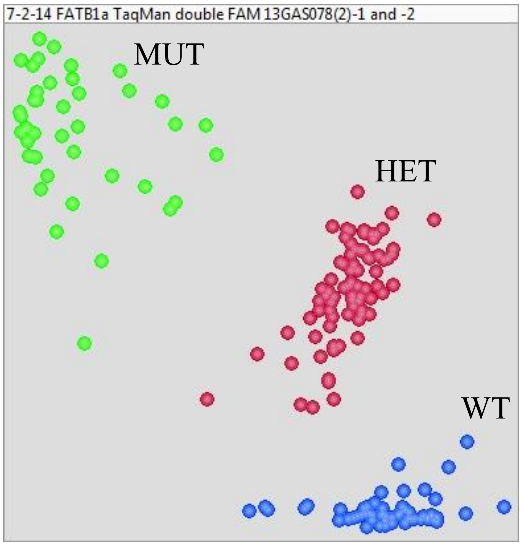

Figure 3.3. TaqMan maker GSM369 genotyping graph on the F_2 population derived from ‘Cook’ x N87-2122-4. MUT= homozygous mutant alleles from N87-2122-4; HET=heterozygous alleles from both N87-2122-4 and Cook; and WT=homozygous wild type alleles from ‘Cook’.

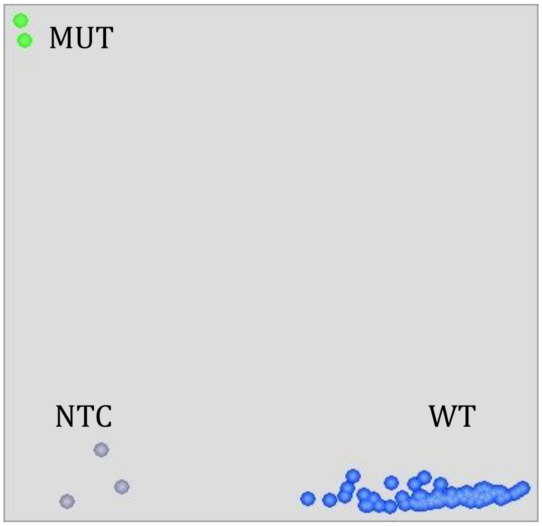

Figure 3.4. A representative TaqMan marker, GSM369 genotyping graph on a panel of 121 soybean lines. MUT = homozygous mutant alleles from N87-2122-4; WT = homozygous wild type alleles; and NTC = No template control.

Table 3.1. Significance test of the *FATB1a* genotypes defined by the marker GSM369 and fatty acids in the F_2 population of 'Cook' x N87-2122-4.

***FATB1a* genotype†**	***N*‡**	**Palmitic acid**	**Stearic acid**	**Oleic acid**	**Linoleic acid**	**Linolenic acid**
		g kg^{-1}				
MUT	35	66	35	305	545	48
HET	80	92	37	282	536	54
WT	39	110	41	272	526	50
Significance		***	***	**	NS§	NS
R^2 (%)		57.3	9.8	4.0	0.3	1.5

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

† *FATB1a* genotypes are determined by the marker GSM369. MUT = homozygous genotype for the genomic deletion at *FATB1a* locus; HET = heterozygous genotype *FATB1a* locus; and WT = wild type genotype at *FATB1a* locus.

‡ *N* = number of plants with a specific genotype.

§ NS stands for non-significant.

Table 3.2 Fatty acid profile and marker genotypes at *FATB1a* locus of 121 diverse soybean lines

Line Name	Palmitic acid	Stearic acid	Oleic acid	Linoleic acid	Linolenic acid	Marker-GSM369 genotype†	Origin	Maturity Group
	g kg^{-1}							
N87-2122-4	57	38	377	515	37	MUT	USA	VI
PI 157440	118	35	211	570	66	WT	South Korea	V
PI 171451	129	39	185	566	81	WT	Japan	VII
PI 200538	145	44	227	520	65	WT	Japan	VIII
PI 227224	144	39	201	550	66	WT	Japan	VII
PI 227687	119	38	201	557	86	WT	Japan	VIII
PI 229358	119	39	224	535	83	WT	Japan	VII
PI 230976	116	31	184	584	85	WT	Japan	VI
PI 230977	114	55	264	503	64	WT	Japan	VII
PI 398469	122	30	211	559	78	WT	South Korea	VI
PI 416937	131	31	206	546	87	WT	Japan	VI
PI 471938	122	34	236	548	60	WT	Nepal	V
PI 594401B	98	40	278	513	72	WT	China	III
PI 594403	95	34	229	563	79	WT	China	IV
PI 594427C	118	35	265	502	80	WT	China	VI
PI 594470C	121	34	182	577	85	WT	China	VIII
PI 594596	124	34	213	556	73	WT	China	VI
PI 594661	120	29	186	582	83	WT	China	V
PI 594662A	121	32	151	592	105	WT	China	V
PI 594753A	113	46	158	588	95	WT	China	VII
PI 203398	114	37	178	574	97	WT	Brazil	VIII
PI 417360	111	24	503	325	37	WT	Japan	V
PI 506582	117	28	445	363	48	WT	Japan	V
PI 506965	112	26	228	555	79	WT	Japan	VI
PI 567024	121	47	179	569	85	WT	Indonesia	VIII
PI 567058D	126	31	162	574	107	WT	Indonesia	IX
PI 567129	127	42	184	553	93	WT	Indonesia	IX
PI 605781A	125	42	236	530	67	WT	Vietnam	VI
PI 605854B	125	33	238	548	57	WT	Vietnam	V
PI 605865B	130	29	237	546	58	WT	Vietnam	V
PI 605891A	126	28	251	538	56	WT	Vietnam	V
PI 197182	115	33	177	585	90	WT	Malaysia	VIII
PI 200487	122	36	162	597	84	WT	Japan	VIII

PI 200488	116	42	195	570	77	WT	Japan	VIII
PI 221717	113	27	193	592	76	WT	South Africa	VI
PI 416826A	114	35	134	632	86	WT	Japan	VIII
PI 416873B	125	38	164	599	73	WT	Japan	VIII
PI 417116	143	44	212	536	65	WT	Japan	VII
PI 417120	115	38	194	577	76	WT	Japan	VIII
PI 417126	110	36	202	570	82	WT	Japan	VIII
PI 417134	110	44	226	551	69	WT	Japan	VIII
PI 417503	125	33	247	521	74	WT	Brazil	VI
PI 471931	127	30	210	570	62	WT	Nepal	V
PI 506695	118	25	197	589	71	WT	Japan	VI
PI 567031B	126	53	255	492	74	WT	Indonesia	VIII
PI 567056A	129	52	221	506	91	WT	Indonesia	VIII
PI 567059	120	51	294	470	65	WT	Indonesia	V
PI 567104B	132	39	222	511	96	WT	Indonesia	IX
PI 567123A	129	53	214	521	83	WT	Indonesia	VIII
PI 567190	129	33	270	486	82	WT	Vietnam	VI
PI 594172A	129	32	198	569	72	WT	Japan	VII
PI 605773	132	27	212	557	73	WT	Vietnam	V
PI 605791A	128	34	283	493	62	WT	Vietnam	VI
PI 605829	122	34	274	510	60	WT	Vietnam	V
PI 605838	121	34	338	452	55	WT	Vietnam	V
PI 605885B	126	31	251	531	61	WT	Vietnam	V
PI 605891B	119	27	297	504	52	WT	Vietnam	VI
PI 594796	146	38	156	552	108	WT	China	VIII
PI 507008	115	37	210	573	66	WT	Japan	VII
PI 417119	116	41	212	557	75	WT	Japan	VIII
PI 200445	116	31	179	595	79	WT	Japan	VIII
PI 578457A	117	46	255	519	64	WT	Vietnam	VIII
PI 507009	115	28	206	582	69	WT	Japan	VI
PI 594149	123	32	185	582	77	WT	Japan	VIII
PI 567034	127	43	247	511	72	WT	Indonesia	VIII
PI 417128	136	45	233	525	62	WT	Japan	VII
PI 416886	108	45	235	544	68	WT	Japan	VIII
PI 567090	128	38	170	562	102	WT	Indonesia	IX
PI 567039	126	33	245	518	78	WT	Indonesia	VII
PI 567020A	119	50	185	569	78	WT	Indonesia	VIII
PI 518295	109	38	209	556	88	WT	Taiwan	VII

PI 619083	120	41	304	450	85	WT	South Korea	V
PI 416937	141	31	206	546	87	WT	Japan	VI
TN04-5321	127	40	200	552	82	WT	USA	V
Woodruff	129	33	196	547	95	WT	USA	VII
Benning	121	42	193	553	90	WT	USA	VII
Benning high-pro	135	32	179	554	100	WT	USA	VII
Benning low lin/low palm	48	40	196	667	48	MUT	USA	VII
Cook	121	52	205	546	76	WT	USA	VII
Boggs	123	31	199	564	82	WT	USA	VII
NC-Raleigh	154	39	163	572	72	WT	USA	VI
Hutcheson	127	41	182	556	94	WT	USA	VI
SQ970252_S17-2-1	117	32	244	506	101	WT	Canada	0
SQ970252_S17-2-3	116	37	251	487	110	WT	Canada	0
S12-11641	77	31	817	32	43	WT	USA	V
S12-11711	86	30	797	36	51	WT	USA	V
RMS X S11-184	127	38	167	509	159	WT	USA	VI
LM-335	146	91	185	514	64	WT	USA	VI
RMS X S11-183	133	43	149	520	155	WT	USA	VI
PI 548298	112	36	232	534	86	WT	China	III
PI 548438	110	49	206	567	67	WT	North Korea	VI
PI 548302	123	32	230	515	101	WT	Japan	II
PI 548311	112	36	204	561	88	WT	Canada	0
PI 548318	118	31	287	485	78	WT	China	III
PI 548325	109	29	252	529	80	WT	Russia	0
PI 548456	111	29	250	528	83	WT	North Korea	VI
PI 548348	116	36	237	523	88	WT	China	III
PI 548461	114	33	247	533	74	WT	LA, USA	VIII
PI 548657	105	54	223	543	74	WT	NC, USA	VII
PI 548352	102	26	407	394	70	WT	North Korea	III
PI 548360	117	40	241	521	82	WT	North Korea	II
PI 548362	116	39	190	552	103	WT	USA	III
PI 548379	124	29	286	467	95	WT	China	0
PI 153217	119	29	237	520	96	WT	Belgium	0
PI 548391	129	36	231	520	83	WT	China	II

PI 548477	115	33	280	498	74	WT	USA	VI
PI 548402	128	32	201	517	122	WT	China	V
PI 548603	110	44	315	464	66	WT	USA	IV
PI 548484	110	41	187	580	82	WT	North Korea	VI
PI 548406	116	36	222	544	82	WT	China	II
PI 548485	110	40	190	568	92	WT	China	VII
PI 548488	131	37	218	511	103	WT	China	V
FC33243	109	35	238	535	83	WT	Unknown	IV
PI 080837	122	39	255	504	81	WT	Japan	IV
PI 180501	115	33	217	554	81	WT	Germany	0
PI 438471	117	29	282	494	78	WT	Sweden	0
PI 438477	119	27	250	524	80	WT	Sweden	0
FC31745	114	29	326	454	77	WT	Unknown	VI
PI 240664	124	38	243	508	87	WT	Philippines	X
PI 88788	122	36	189	561	92	WT	China	III
PI 548445	104	42	249	517	87	WT	China	VII
PI 548356	107	31	353	429	80	WT	North Korea	II

† MUT = homozygous genotype for the genomic deletion at *FATB1a* locus; and WT = wild type genotype at *FATB1a* locus.

Table 3.3. The 19 genes located within the deleted genomic region from 1,277,909 to 1,023,799 bp on chromosome 5 at *FATB1a* locus in a reduced palmitic line N87-2122-4†.

Gene Name	**Start position‡**	**Stop position**	**Description**
Glyma.05g012600	1,179,295	1,185,153	Glycosyl hydrolases family 17, plant protein of unknown function (DUF936)
Glyma.05g011600	1,043,861	1,046,592	Leucine-rich repeat receptor-like protein kinase
Glyma.05g011700	1,056,287	1,063,382	Sodium-bile acid cotransporter
Glyma.05g011800	1,070,013	1,079,448	WD40 repeat protein
Glyma.05g011900	1,081,225	1,092,755	ATP-binding cassette transporter
Glyma.05g012000	1,102,788	1,103,854	Integral membrane protein DUF92
Glyma.05g012100	1,106,315	1,107,216	(ancestor identifier)
Glyma.05g012200	1,109,904	1,113,717	Protein phosphatase 2A regulatory B subunit (B56 family)
Glyma.05g012300	**1,127,438**	**1,131,632**	**Acyl-ATP thioesterase**
Glyma.05g012400	1,151,734	1,171,083	DNAJ homolog subfamily C member
Glyma.05g012500	1,173,106	1,176,514	(ancestor identifier)
Glyma.05g012700	1,188,556	1,193,182	Met-10+ like-protein
Glyma.05g012800	1,198,706	1,199,770	Histone deacetylase 15
Glyma.05g012900	1,200,699	1,206,341	Histone deacetylase domain
Glyma.05g013000	1,208,979	1,210,626	Myb-like DNA-binding domain
Glyma.05g013100	1,213,775	1,217,715	Docking domain of Afi1 for Arf3 in vesicle trafficking
Glyma.05g013200	1,218,838	1,227,836	Transmembrane Fragile-X-F protein
Glyma.05g013300	1,241,512	1,244,692	TCP family transcription factor
Glyma.05g013400	1,263,240	1,282,578	Transcription initiation factor TFIID

† The data was based on gene model gff3 files from Soybase (http://Soybase.org).
‡ Physical positions were based on the genome assembly version of Glyma.Wm82.a2.v1

Table 3.4. Pearson's correlation coefficient between seed fatty acids in an F_2 population derived from Cook x N87-2122-4.

	Stearic acid	Oleic acid	Linoleic acid	Linolenic acid
Palmitic acid	0.14	-0.21*	-0.20*	0
Stearic acid		-0.23**	0.08	-0.09
Oleic acid			-0.88***	-0.06
Linoleic acid				-0.19*

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

*** Significant at the 0.001 probability level.

References

Anai, T., T. Hoshino, N. Imai, and Y. Takagi. 2012. Molecular characterization of two high-palmitic-acid mutant loci induced by X-ray irradiation in soybean. Breed Sci. 61:631–638.

Beuselinck, P.R., D.A. Sleper, and K.D. Bilyeu. 2006. An assessment of phenotype selection for linolenic acid using genetic markers. Crop Sci. 46:747–750.

Boerma, H.R., R.S. Hussey, D.V. Phillips, E.D. Wood, and S.L. Finnerty. 1992. Registration of 'Cook' soybean. Crop Sci. 32:497.

Boerma, H.R., R.S. Hussey, D.V. Phillips, E.D. Wood, G.B. Rowan, and S.L. Finnerty. 1997. Registration of ‘Benning’ soybean. Crop Sci. 37:6.

Bubeck, D.M., W.R. Fehr, and E.G. Hammond. 1989. Inheritance of palmitic and stearic acid mutants of soybean. Crop Sci. 29:652–656.

Burton, J.W., R.F. Wilson, and C.A. Brim. 1983. Recurrent selection in soybeans. IV. Selection for increased oleic acid percentage in seed oil. Crop Sci. 23:744–747.

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.

Burton, J.W., J.R. Wilcox, R.F. Wilson, W.P. Novitzky, and G.J. Rebetzke. 1998. Registration of low palmitic acid soybean germplasm lines N94-2575 and C1943. Crop Sci. 38:1407.

Cardinal, A.J., J.W. Burton, A.M. Camacho-Roger, J.H. Yang, R.F. Wilson, and R.E. Dewey. 2007. Molecular analysis of soybean lines with low palmitic acid content in the seed oil. Crop Sci. 47:304–311.

Erickson, E.A., J.R. Wilcox, and J.F. Cavins. 1988. Inheritance of altered palmitic acid percentages in two soybean mutants. J. Hered. 79:465–468.

Fedick, A., J. Su, and N.R. Treff. 2012. Development of TaqMan allelic discrimination based genotyping of large DNA deletions. Genomics 99:127–131.

Fehr, W.R. 2007. Breeding for modified fatty acid composition in soybean. Crop Sci. 47:S72–S87.

Fehr, W.R., G.A. Welke, S.R. Cianzio, D.N. Duvick, and E.G. Hammond. 1991. Inheritance of reduced palmitic acid content in seed oil of soybean. Crop Sci. 31:88–89.

Gillman, J., A. Tetlow, K. Hagely, J. Boersma, A. Cardinal, I. Rajcan, and K. Bilyeu. 2014. Identification of the molecular genetic basis of the low palmitic acid seed oil trait in soybean mutant line RG3 and association analysis of molecular markers with elevated seed stearic acid and reduced seed palmitic acid. Mol. Breeding 34:447–455.

Giancola, S., H.I. McKhann, A. Berard, C. Camilleri, S. Durand, P. Libeau, F. Roux, X. Reboud, I.G. Gut, and D. Brunel. 2006. Utilization of the three high-throughput SNP genotyping methods, the GOOD assay, Amplifluor, and TaqMan, in diploid and polyploid plants. Theor. Appl. Genet. 112:1115–1124.

Gizlice, Z., T. E. Carter, and J. W. Burton. 1994. Genetic base for North-American public soybean cultivars released between 1947 and 1988. Crop Sci. 34:1143–1151.

Hunter, J.E., J. Zhang, and P.M. Kris-Etherton. 2010. Cardiovascular disease risk of dietary stearic acid compared with *trans*, other saturated, and unsaturated fatty acids: a systematic review. Am. J. Clin. Nutr. 91:46–63.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Mentjies, and A. Drummond. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.

Keim, P., T.C. Olson, and R.C. Shoemaker. 1988. A rapid protocol for isolating soybean DNA. Soybean Genet. Newsl. 15:150–152.

Lawrence, G.D. 2013. Dietary fats and health: dietary recommendations in the context of scientific evidence. Adv. Nutr. 4:294–302.

Li, Z., R.F. Wilson, W.E. Rayford, and H.R. Boerma. 2002. Molecular mapping genes conditioning reduced palmitic acid content in N87-2122-4 soybean. Crop Sci. 42:373–378.

Puaschitz, N.G., E Strand, T.M. Norekval, J. Kierkes, L. Dahl, G. Frodahl, T. Svingen, J. Assmus, H. Schartum-Hansen, J. Oyen, E.K.R. Pedersen, C.A. Drevon, G.S. Tell, and O. Nygard. 2015. Dietary intake of saturated fat is not associated with risk of coronary events or mortality in patients with established coronary artery disease. J. Nutr. 145(2):299–305.

Pham, A.T., K. Bilyeu, P. Chen, R. Boerma, and Z. Li. 2014. Characterization of the *fan1* locus in soybean line A5 and development of molecular assays for high-throughput genotyping of *FAD3* genes. Mol. Breeding 33:895–907.

Primomo, V.S., D.E. Falk, G.R. Ablett, J.W. Tanner, and I. Rajcan. 2002. Inheritance and interaction of low palmitic and low linolenic soybean. Crop Sci. 42:31–36.

Schnebly, S.R., W. Fehr, G. Welke, E.G. Hammond, and D.N. Duvick. 1994. Inheritance of reduced and elevated palmitate in mutant lines of soybean. Crop Sci. 34:829–833.

Shen, G.Q., K.G. Abdullah, and Q.K. Wang. 2009. The TaqMan method for SNP genotyping. Methods Mol. Biol. 578:293–306.

Steinberg, D. 2005. Thematic review series: the pathogenesis of atherosclerosis. An interpretive history of the cholesterol controversy: Part II: The early evidence linking hypercholesterolemia to coronary disease in humans. J. Lipid Res. 46:179–190.

Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B.C. Faircloth, M. Remm, and S.G. Rozen. 2012. Primer3-new capabilities and interfaces. Nucleic Acids Res. 40(15):e115.

Wilcox, J.R. and J.F. Cavins. 1990. Registration of C1726 and C1727 soybean germplasm with altered levels of palmitic acid. Crop Sci. 30:240.

Wilson, R.F., T.C. Marquardt, W.P. Novitzky, J.W. Burton, J.R. Wilcox, A.J. Kinney, and R.E. Dewey. 2001. Metabolic mechanisms associated with alleles governing the 16:0 concentration of soybean oil. J. Am. Oil Chem. Soc. 78:335–340.

CHAPTER 4

SUMMARY

Typical soybean oil is composed of 110 g kg^{-1} palmitic, 40 g kg^{-1} stearic, 250 g kg^{-1} oleic, 520 g kg^{-1} linoleic, and 80 g kg^{-1} linolenic acids. Hydrogenation of soybean oil is done to increase shelf life and improve cooking stability. In 2006, the U.S. Food and Drug Administration began requiring *trans* fat labeling and in 2013 affirmed that *trans* fats are not generally regarded as safe. Palmitic acid is a major saturated fatty acid in soybean oil, and consumption of saturated fat is linked to a risk of coronary diseases. With advancement of DNA marker technologies, molecular breeding has enabled plant breeders to effectively select desired fatty acid profile to develop soybean cultivars with better oil quality and improved health benefits for consumers. The objectives of this study were to: i) evaluate effects of high oleic and low linolenic acid alleles on yield, fatty acids, protein, and oil contents in Georgia environments, ii) determine if these traits affect seed germination in soybean, iii) characterize the *FATB1a* gene in N87-2122-4, and iv) develop a breeder-friendly, diagnostic marker to facilitate marker-assisted selection and improve breeding efficiency for reduced palmitic soybean.

High oleic and low linolenic acid contents in soybean seed are the key compositional traits that improve oxidative stability and greatly increase oil functionality and shelf life. Using a marker-assisted selection method, we have developed near-isolines (NILs) of G00-3213 and G00-3880 for high oleic and low linolenic traits, respectively through backcrossing. We have yield-tested these NILs that have various combinations of *FAD2-1A* and *FAD2-1B*, or different

FAD3A alleles and compared them to the normal fatty acid NILs derived from the same backcrossing populations. The results indicated that NILs with both homozygous mutant *FAD2-1A* and *FAD2-1B* alleles produce an average of 806 g kg^{-1} oleic acid content and NILs with homozygous mutant *FAD3A* alleles could reduce the linolenic acid content to 37 g kg^{-1}. The results also demonstrated that possessing these mutant alleles did not cause a yield reduction. Furthermore, seed germination tests across 12 temperatures showed that modified seed fatty acid composition for oleic and linolenic acids in general did not have a major impact on seed germination. However, there is a possible reduction in seed germination vigor when high oleic seed are planted in cold soil. The observed significant differences ($p<0.05$) of the germination rates between low linolenic and normal linolenic NILs were dependent on the genetic background. The mutant *FAD2-1A* and *FAD2-1B* alleles did not hinder either seed or plant development.

The *FATB1a* gene was previously found to be responsible for reduced palmitic acid in the soybean line N87-2122-4. With the availability of soybean genetic maps, reference genome, and gene annotation, an approximate 254 kb deletion region including the *FATB1a* gene was identified. Based on the gene deletion information, we successfully developed a TaqMan marker and tested it with a segregating F_2 population that consisted of 140 individual plants derived from ‘Cook’ x N87-2122-4. The marker performed well and accounted for 57% of the phenotypic variation. The marker was also validated using a panel of 121 diverse soybean lines with known fatty acid profiles. The result indicated that the marker can be used effectively in marker-assisted breeding for reduced palmitic acid in soybean early generation selection or in backcrossing of the reduced palmitic acid trait in soybean.