ABSTRACT

BACHLAVA, ELENI. Genetic Control of High Oleic Acid Seed Content in Soybean. (Under the direction of Andrea J. Cardinal and Ralph E. Dewey.)

The increase of oleic acid content in soybean [*Glycine max* (L.) Merr.] oil improves its nutritional value and oxidative stability; therefore, the incorporation of the high oleate trait in soybean germplasm is required for the continued commercial success of soybean oil. Little is known, however, about the response to selection for higher oleate content, which is dependent on its heritability in breeding populations, and the correlated responses of other fatty acids and agronomic traits to selection for oleate content, which depends on their genetic correlations with oleate. Furthermore, the lack of knowledge on the genetic factors underlying oleate variation in soybean oilseeds hampers the use of marker assisted selection for high oleate content in soybean breeding programs. Three soybean populations segregating for major and minor oleate genes, which were grown in replicated multi-environment trials, were studied in order to shed light on the research questions above.

The results of this study suggested that oleate heritability was sufficiently high that early generation selection can be effective when practiced on unreplicated lines grown at a single environment. Moreover, this study indicated that selection for higher oleate content will result in lower linoleate, linolenate and palmitate content in soybean oil. Also, the significant negative correlation between oleate content and yield implied that the development of high oleate germplasm may be hindered by

lower yields. The isoforms of the *FAD2-1* and *FAD2-2* genes, which encode the microsomal ω-6 desaturase enzymes that catalyze the desaturation of oleic acid to linoleic acid during fatty acid biosynthesis, were mapped in order to investigate their cosegregation with the oleate quantitative trait loci (QTLs) identified. The *FAD2-1A* and *FAD2-1B* isoforms mapped on linkage groups O and I, respectively, while the closely linked *FAD2-2A* and *FAD2-2B* isoforms mapped on linkage group L of soybean genome. An oleate QTL with moderate effects was reported in the proximity of *FAD2-1B* isoform on linkage group I. Also, an oleate QTL with moderate effects was detected on linkage group F, which interacted epistatically with the QTL on linkage group I. In conclusion, no major oleate QTLs were identified and the *FAD2-1A*, *FAD2-1B* and *FAD2-2B* isoforms did not encode for the major oleate genes that contributed to the elevated oleic acid content of the experimental lines used in this study.

Genetic Control of High Oleic Acid Seed Content in Soybean

by

Eleni Bachlava

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Crop Science

Raleigh, North Carolina

2008

APPROVED BY:

Dr. Andrea J. Cardinal Co-chair of Advisory Committee	Dr. Ralph E. Dewey Co-chair of Advisory Committee
Dr. José M. Alonso	Dr. Joseph W. Burton

Dr. Jason A. Osborne

BIOGRAPHY

Eleni Bachlava was born in Athens, Greece during the fall of 1979, the eldest child of Dimitrios and Maria Bachlava. Being always fascinated by the natural sciences and with no interest whatsoever in social sciences, probably due to several years of required attendance at ancient greek language, literature and history courses, she was relieved to graduate from Evangeliki Experimental High School and start her studies in the Department of Agricultural Biotechnology at the Agricultural University of Athens. After five years of dedicated study and research in the Laboratory of Genetics under the advising of Dr. Andreas Katsiotis, she had no doubt that the best way to contribute to society and satisfy her aspirations was to pursue a career in research. Rather unexpectedly, she found an opportunity to work as an intern in the Department of Agronomy at Purdue University under Dr. Lauren McIntyre, and without further thought she decided to take it. After an amazing summer at West Lafayette, IN, in 2003, Eleni returned to the United States in the summer of 2004 in order to pursue her doctoral studies in the Department of Crop Science at North Carolina State University under the advising of Dr. Andrea Cardinal. Since then, she has come to realize that elucidating the genetic factors that control oleic acid content in soybean oilseeds is more challenging than she initially thought. After almost four years of exciting research at North Carolina State University, Eleni

is looking forward to enriching her research experience as a postdoctoral researcher at the University of Georgia.

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. Andrea Cardinal not only for giving me the opportunity to join her research group and conduct exciting research in plant breeding and genetics, but also for her continuous advising, support, and encouragement during my doctoral studies. The quality of this study was greatly influenced by her help in avoiding adversities and in solving problems which arose during the course of the research project, and also by her research ethics and her perseverance for a deeper understanding of this project.

I would also like to thank Dr. Ralph Dewey for his collaboration and contribution to this project, part of which was conducted in his laboratory. I am grateful to Dr. Joseph Burton for his helpful insights and for supplying the germplasm used in this study, to Dr. Cavell Brownie for her numerous suggestions on research questions pertained to her expertise, as well as to Dr. José Alonso and Dr. Jason Osborne for serving on my advisory committee. Also, I would like to thank all the professors from different educational levels who have encouraged me or inspired me, one way or another, to appreciate academia and to pursue an advanced degree.

I greatly acknowledge the help of the current and past members of Dr. Cardinal's group: Dr. Jérôme Auclair, Ana Maria Camacho-Roger, Sanbao Wang and Rebecca Whetten. I also greatly appreciate the assistance and suggestions of Carol

Griffin, and all members of Dr. Dewey's research group, as well as William Novitzky, during the course of this research.

At last, many thanks to all friends that I was fortunate to meet in NC and had so much fun with. I am so happy that we shared a part of our "journeys" and I really hope to stay in touch. I am especially thankful to all you that our friendship has already lasted so long and your support has not changed, although so far away. I am also grateful to my family, especially my parents, Maria and Dimitrios, and my brother, Antonis, for their love and understanding. I deeply wish that in the near future I will be able to share more time with you. Finally, I would like to thank Ricardo for being there to share my excitement and disappointment, encourage me keep my motivation at all times, and make me smile even those days that "everything went wrong".

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (continued)

Page

TABLE OF CONTENTS (continued)

Page

LIST OF TABLES

Page

II. Mapping Genes Encoding Microsomal ω-6 Desaturase Enzymes and Their Cosegregation with QTL Affecting Oleate Content in Soybean

LIST OF TABLES (continued)

Page

III. Heritability of Oleic Acid Seed Content in Soybean Oil and its Genetic Correlation with Fatty Acid and Agronomic Traits

LIST OF TABLES (continued)

Page

LIST OF TABLES (continued)

Page

LIST OF FIGURES

Page

I. INTRODUCTION

Soybean: an Oilseed Crop

Soybean [*Glycine max* (L.) Merr] is a crop of major importance since it provides much of the worlds' protein for livestock feed and oil for human consumption. According to the United States Department of Agriculture statistics (2006) soybean produces 29.4% of the world's edible vegetable oil and, therefore, outnumbers all other edible oilseeds' production. The United States is the leading soybean producer representing 40.1% of the world production. The Unites States is also the major exporter of the commodity in world markets, but follows Brazil and Argentina in soybean oil exports, and Argentina in soybean meal exports (Wilcox, 2004). Apart from the soy-based applications in the food and feed industry, alternative industrial products include adhesives, coatings, printing inks, lubricants and plastics. Furthermore, the use of soybean for biodiesel production undertakes not only the conservation of fossil fuels, but also the enhancement of human and environmental safety (http://www.soynewuses.org/, verified 11/03/2007).

Seed Metabolic Pathways: from CO_2 to Seed Oil

Seed is comprised of storage components necessary for germination and growth of young seedlings before the establishment of photosynthetic mechanism. Storage reserves constitute more than 90% of the seed dry weight and consist of

carbohydrates such as starch, oils in the form of triacylglycerols (TAGs) and proteins. The relative proportions of seed carbohydrate, oil and protein varies among species. Among these reserves, oil is the most efficient form of energy storage containing at least twice the amount of energy, either per-volume or per-weight, in comparison with the energy stored in protein or starch (Huang, 1992).

The metabolic pathway for the conversion of carbohydrates to seed oil initiates with the accumulation of photosynthate and terminates with the synthesis of oil, which is stored in the form of TAGs synthesized by the sequential acylation of glycerol backbones. Therefore, the major pathways involved in the conversion of sucrose into fatty acids (Fig. 1) include the glycolytic pathway in the cytosol, the transfer of intermediates into the plastid, starch biosynthesis and degradation in the plastid, the glycolytic pathway in the plastid, the oxidative pentose phosphate cycle, the plastid pyruvate dehydrogenase complex, as well as fatty acid biosynthesis (White et al., 2000). The flux through different reactions among these pathways indicates that oil metabolism cannot be restricted to fatty acid and TAG biosynthesis since the abundance of substrates for fatty acid biosynthesis is dependent on the preceding pathways. Although this review is focused on the biosynthesis of oilseed TAGs, it should be noted that the fatty acid biosynthesis pathway contributes to the production of either storage lipids, in the form of TAG, which have fatty acids attached to all three positions of the glycerol backbone, or membrane lipids, which have fatty acids attached to the first and second position, and a polar head group to the third position

of the glycerol backbone. Membrane lipids, commonly referred as glycerolipids, are divided into glycolipids and phospholipids, depending on the type of the polar head groups attached (Cardinal, 2008).

Fatty acids in oilseeds, as precursors for the synthesis of TAGs, are mainly provided by the eukaryotic pathway in the endoplasmic reticulum (Browse and Somerville, 1991). However, the contribution from the prokaryotic pathway in the plastids cannot be ruled out (Fig. 2). It has been suggested that disruption of one of the pathways is compensated by the fatty acid biosynthesis through the other pathway and that intermediates are transferred between the compartments; although, enzymes reside exclusively in either the endoplasmic reticulum or the plastids (Somerville and Browse, 1996). It should be noted that the initial steps of the eukaryotic fatty acid pathway for the biosynthesis of the primary 16 and 18 carbon precursors take place in the plastids of the developing seeds. The eukaryotic fatty acid biosynthetic pathway initiates with the formation of the malonyl - coenzyme A (CoA), the main carbon donor for fatty acid synthesis, from acetyl-CoA, catalyzed by acetyl-CoA carboxylase (ACCase). ACCase is implicated in the flux control of the pathway through a feedback inhibition mechanism (Ohlrogge and Jaworski, 1997). The first condensation reaction of malonyl - acyl carrier protein (ACP) and acetyl-CoA is catalyzed by the ketoacyl-ACP synthetase (KAS) III. The next several condensations, catalyzed by KAS I, lead to the production of the 16 carbon palmitoyl-ACP. Next, the elongation of the palmitoyl-ACP to the 18 carbon stearoyl-ACP takes place, which

requires KAS II activity. Stearoyl-ACP desaturase is the first enzyme that catalyzes the formation of the monounsaturated fatty acids from the saturated fatty acids produced by the condensation reactions (Ohlrogge and Browse, 1995). The primary products of fatty acid biosynthesis, that is the palmitoyl-ACP and oleoyl-ACP, are synthesized in the plastids of the developing seeds. Finally, polyunsaturated fatty acids are produced through the subsequent desaturation reactions catalyzed by the membrane-bound desaturase enzymes of either the plastids or the endoplasmic reticulum. During the eukaryotic fatty acid biosynthetic pathway, the ω-6 and ω-3 desaturase enzymes which reside in the endoplasmic reticulum, act on the fatty acids esterified to the first and second position of phosphatidyl choline (PC) (Cardinal, 2008).

Oil is stored in the form of TAGs in the developing oilseeds. TAGs are synthesized through the Kennedy pathway with the esterification of all three positions of the glycerol backbones with fatty acids. Initially, fatty acids are transferred from ACP to glycerol-3-phosphate (G3P). G3P acyl transferase catalyzes the formation of lysophosphatidic acid (LPA). LPA is converted to phosphatidic acid (PA) by LPA acyltransferase. Diacylglycerol (DAG) is released during PA dephosphorylation, catalyzed by PA-phosphatase. Finally, DAG-acyltransferase catalyzes the synthesis of TAGs with the transfer of a fatty acid to the third position of DAG. The five predominant fatty acids in soybean oilseed TAGs are palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2) and linolenic (18:3); however, more than 300 different

fatty acids have been detected in plants (Ohlrogge and Browse, 1995). As previously mentioned, in many oilseeds fatty acids are not immediately available for TAG biosynthesis but are initially incorporated into PC, which serves as a substrate for 18:1 and 18:2 desaturation (Browse and Somerville, 1991).

The differences in seed oil content among species have raised the question of how seed oil quantity is determined. Proposed mechanisms implicate either the abundance of fatty acid precursors for TAG synthesis or the demand for fatty acids in response to the relative amounts of metabolic intermediates (Ohlrogge and Jaworski, 1997). Another important consideration is the differential oil content and fatty acid composition among the different tissues of an oilseed; that is the seed coat, the endosperm, the cotyledons, the hypocotyle and the radicle (Li et al., 2006). Furthermore, various factors, such as light and vernalization may have an impact on seed oil content and fatty acid composition (Li et al., 2006; O'Neill et al., 2003).

Ultimately, the TAGs in the mature seed are stored in discrete sub-cellular organelles, the oil bodies. Oil bodies contain 94-98% TAGs, 0.6-2% phospholipids (PL) and 0.6-3% proteins and their average size, which varies among species and within seed, is affected by environmental and nutritional factors. Oil bodies are surrounded by a single layer membrane, which consists of PL and specialized proteins, known as oleosins, and maintain a hydrophilic surface. Oleosins contribute to the stability, synthesis and metabolism of oil bodies. It has been suggested that oleosins are only required for the long-term storage of oil bodies. Caleosins are also

found in oil bodies and; presumably, they are involved in the calcium mediated fusion of oil bodies. Oil bodies are finally degraded during germination, leading to TAG hydrolysis to glycerol and fatty acids, which cover the need for carbohydrates during seedling growth (Frandsen et al., 2001; Huang, 1992).

Genetic Control of the Unsaturated Fatty Acid Biosynthesis

Research conducted on *Arabidopsis thaliana* (L.) Heynh. shed light into the genes controlling the enzymatic activities of ω-6 and ω-3 desaturases with the assignment of the *fad2* and *fad3* loci, encoding the desaturation of oleic to linoleic and linoleic to linolenic acid, respectively (Browse et al., 1993; Miquel and Browse, 1992). The *fad2* and *fad3* genes were isolated and mapped on chromosome III and chromosome II of the *Arabidopsis* genome (Arondel et al., 1992; Yadav et. al, 1993; Okuley et al., 1994). The location of *fad2* and *fad3* genes coincided with oleate and linoleate quantitative trait loci (QTLs) in *Arabidopsis* (Hobbs et al., 2004). However, the approximately eightfold larger genome size of soybean in comparison to *Arabidopsis* as well as the ancient polyploid nature of the soybean genome, which is highly duplicated and consists of 40-60% repetitive sequences, implies a greater complexity of the genetic factors underlying enzymatic activities of the fatty acid biosynthesis pathway in soybean (Arumuganathan and Earle, 1991; Jackson et al., 2006).

Three different alleles *Ol*, *ol* and *ol*^{a} residing at a single locus were initially implicated in the control of the oleate trait in soybean (Rahman et al., 1998; Takagi and Rahman, 1996). However, two different genes, designated as *FAD2-1* and *FAD2-2*, were assigned to the microsomal ω-6 desaturase activity. The *Ol* allele that conferred elevated oleate content was later assigned to a deletion in *FAD2-1* gene (Kinoshita et al., 1998). It was suggested that the seed-specific *FAD2-1* is responsible for the desaturation of storage lipids from oleate to linoleate in the developing seed; while, *FAD2-2* is involved in the desaturation of membrane lipids, both in the seed and in vegetative tissues (Heppard et al., 1996). Latest reports confirm the existence of two distinct isoforms of the *FAD2-1* gene, *FAD2-1A* and *FAD2-1B*, which differ in stability at high temperatures (Tang et al., 2005). *FAD2-1A* and *FAD2-1B*, map on linkage groups O and I of soybean genome, respectively. Furthermore, three isoforms were reported for *FAD2-2*, designated as *FAD2-2A*, *FAD2-2B* and *FAD2-2C*. *FAD2-2A* and *FAD2-2B* were localized on linkage group L (Bachlava et al., 2008; Schlueter et al., 2007).

At least three genetic loci were associated with linolenate trait, designated as *fan*, *fan2* and *fan3* (Fehr et al., 1992; Fehr and Hammond, 1998; Wilcox and Cavins, 1987). Three different alleles have been assigned to the *fan* locus (Fehr et al., 1992; Rahman et al.,1994a; Stojsin, 1998; Wilcox and Cavins, 1987) which is linked to the *Idh2* (isocitrate dehydrogenase) locus (Rennie et al., 1988). Moreover, *fanx* (Rahman and Takagi, 1997) and *fanx*^{a} alleles (Rahman et al., 1998) are not allelic with the *fan*

locus; however, whether *fanx* or *fanx*^{a} loci differ from *fan2* and *fan3* was not clarified. Indeed, four independent microsomal ω-3 fatty acid desaturases, *FAD3-1a*, *FAD3-1b*, *FAD3-2a* and *FAD3-2b* have been identified (Anai et al., 2005; Bilyeu et al., 2003). The *fanx*^{a} locus corresponded to the *FAD3-1a* isoform, and the *fan* locus corresponded to the *FAD3-1b* isoform (Anai et al., 2005), which is equivalent to the *FAD3A* isoform, according to the nomenclature of Bilyeu et al. (2003). The *FAD3A* isoform was mapped on linkage group B2 and cosegregated with a major linolenate QTL (Brummer et al., 1995). A *FAD3A* specific marker has been developed by Camacho-Roger (2006).

Breeding Objectives and Efforts Related to the Unsaturated Fatty Acid Composition

Oilseed industry currently aims at lower levels of *trans* isomers of the unsaturated fatty acids in order for the soybean oil to meet the guidelines of the United States Food and Drug Administration, to retain consumer acceptance and enhance its competitive position in domestic and global markets (Wilson et al., 2002). *Trans* isomers are formed during hydrogenation, which confers long-term oxidative stability to soybean oil. Alternatively, the need for oil hydrogenation can be diminished if the level of the polyunsaturated fatty acids, linoleate and linolenate, which are susceptible to autoxidation, is reduced. Also, increase in oleate content can

contribute to the oxidative stability of soybean oil and the improvement of its nutritional value (Wilson, 2004).

Thereby, breeders have successfully incorporated the high oleate and low linolenate traits into soybean germplasm with traditional plant breeding methods. The low linolenate line N78-2245 was developed through recurrent selection (Wilson et al., 1981) and the low linolenate line N85-2176 was derived from the cross of N98-2245 with PI123440 (Burton et al., 1989). The low linolenate cultivar 'Soyola' was developed from the cross of N87-2117-3 with 'Brim' (Burton et al., 2004). Also, the elevated oleate line N98-4445A, derived from the cross of N94-2473 x (N93-2007-4 x N92-3907), was recently released (Burton et al., 2006). The later, as well as its sister line N97-3363-3, were utilized in this study for the development of segregating populations for the oleate trait. Germplasm with elevated oleate and reduced linolenate content, was also developed by chemical or x-ray mutagenesis and includes, among others, the lines C1640 (Wilcox et al., 1984), A5 (Hawkins et al. 1983), M23 (Rahman et al., 1994b) and M11 (Takagi and Rahman, 1995). Recently, it was revealed that the high oleate content in M23 is due to the deletion of the *FAD2-1A* isoform (Sandhu et al., 2007). Genetic manipulation of *FAD2-1* gene has also resulted in transgenic soybean lines with high oleate and low linolenate content using different molecular approaches (Kinney, 1995; Kinney and Knowlton, 1998; Buhr et al., 2002). However, soybean industry is still in need for soybean germplasm with

elevated oleate and reduced linolenate content that is adapted to different growing regions.

Quantitative Trait Loci Controlling the Unsaturated Fatty Acid Content

A consensus linkage map of soybean genome has been developed by the integration of densely saturated genetic maps (Cregan et al., 1999; Song et al., 2004). With the consensus map as a guideline, molecular markers have been utilized for the mapping of QTLs that explain the observed variation in oleate, linoleate and linolenate contents. Currently, SoyBase (http://soybeanbreederstoolbox.org/; verified August, 2007) contains 19 QTLs conditioning the unsaturated fatty acid composition, including six, six and seven QTLs for oleate, linoleate and linolenate traits, respectively, which were mapped in diverse soybean populations grown at different environments. The majority of these QTLs were mapped for $F_{2:3}$ lines derived from the cross of a normal oleate normal linolenate *Glycine max* line with a normal oleate high linolenate *G. soja* line using RFLPs, isoenzymes and morphological markers (Diers and Shoemaker, 1992). In this study, QTLs for oleate coincided with QTLs for linoleate with opposite effects. Oleate QTLs were detected on linkage groups A1, E and B2, linoleate QTL were mapped on linkage groups A1 and E, and linoleate QTLs were found on linkage groups E, K and L. Moreover, the *fan* locus, corresponding to the *FAD3A* isoform, was mapped on linkage group B2 and cosegregated with a major

linolenate QTL (Brummer et al., 1995) when tested in an F_2 population derived from the cross of the low linolenate *G. max* mutant C1640 with the high linolenate *G. soja* accession PI479750. Recently, six oleate QTLs were mapped on linkage groups A1, D2, G and L in a soybean population consisting of $F_{2:3}$ lines from the cross of the normal oleate G99-6725 with the high oleate N00-3350, a single plant selection of N98-4445A (Monteros et al., 2004). In addition, several minor QTLs have been detected for oleate (linkage groups D1b, L and E), linoleate (linkage groups F, L and E) and linolenate (linkage groups F, L, E and G) traits. Among them, the QTLs mapped on linkage groups L and E coincided for all the unsaturated fatty acid traits (Hyten et al., 2004; Panthee et al., 2006).

Environmental Influence on the Unsaturated Fatty Acid Composition

The stability of the unsaturated fatty acid composition across a range of environments, which differs among genotypes with altered fatty acid profiles, is critical for the release of improved soybean germplasm. Some of the elevated oleate breeding lines, such as N98-4445A and N97-3363 that were utilized in this study, are unstable across environments and their oleate variation is largely attributed to temperature changes during oil deposition (Oliva et al., 2006; Wilson et al., 2002). However, previous studies suggested that mutants with increased oleate and reduced linolenate content exhibited higher stability than lines with normal fatty acid levels

(Primomo et al., 2002). Carver et al. (1986) also indicated that lines with decreased linolenate content produced through recurrent selection were less sensitive to environment; therefore, they suggested the direct selection for that trait with limited testing across environments in order to accelerate genetic progress.

Environmentally induced variation in the unsaturated fatty acid composition of soybean seed is primarily due to temperature effects. Linolenate content was found to be inversely correlated with temperature changes (Howell and Collins, 1957). It has been suggested that linolenate content is affected by temperature during the pod-filling stage, which coincides with oil deposition (Burton et al., 1983; Wilcox et al., 1993). Apart from temperature effects, oleate accumulation patterns were found to differ among soybean germplasm (Ishikawa et al., 2001). Several studies attempted to interpret the changes in the unsaturated fatty acid composition due to temperature effects. Wilcox and Cavins (1992) reported that linolenic acid was affected more than the other fatty acids by the planting date, which may result in different temperatures during oil deposition. Burton (1991) suggested that the effects of temperature on linolenate content may be caused by changes in oleate desaturation and the subsequent changes in linoleate desaturation. Indeed, Cheesbrough (1989) reported the dramatic reduction in the oleyl and linoleyl desaturase activities at 35ºC. Also, Tang et al. (2005) revealed that the *FAD2-1A* enzyme degrades at high growth temperatures (30ºC), and post-translational modifications may affect its enzymatic activity. Byfield and Upchurch (2007) conducted expression studies on different

soybean germplasm and growth conditions, and suggested that transcript accumulation significantly decreased for both *FAD2-1A* and *FAD2-1B* isoforms at day/night growth temperatures of 30/26ºC. However, Heppard et al. (1996) showed that the transcript levels of *FAD2-1*, *FAD2-2* and *FAD6* remain constant when soybean plants are grown at day/night temperatures between 18/12 and 32/28ºC. Schlueter et al. (2007) detected significant increase in expression only for the *FAD2-2C* isoform when developing pods were grown at 18/12ºC day/night temperatures.

Several environmental factors have been also reported to have an effect on the fatty acid profiles of soybean oilseeds. Differential accumulation of oleic and linoleic acid has been observed at different plant nodes. Seeds developed on the plant apex had lower linoleate and higher oleate content than seeds developed on basal nodes (Bennett et al., 2003). Furthermore, maternal effects were considered responsible for the environmentally induced variation in the unsaturated fatty acid traits (Singh and Hadley, 1968). Using cultivar 'Essex' and the mid-oleate line N78-224, which have similar flowering and maturity profiles, as rootstock and scion by grafting both genotypes in a reciprocal fashion indicated that the rootstock genotype affected the oleate content of seed derived from the scions. This observation implied that maternal effects can be due to translocating factors that cross the graft union, which could originate from the leaf tissue of the maternal plant (Carver et al., 1987).

Objectives of the Study

This study attempts to dissect oleate inheritance and shed light onto the genetic factors implicated in oleate biosynthesis of soybean oilseeds. The objectives of this study were to: (i) estimate oleate heritability in soybean populations segregating for the oleate trait; (ii) estimate the genetic correlations between oleate trait and other fatty acid and agronomic traits; (iii) identify QTLs for oleic acid seed content in populations segregating for the trait; (iv) localize the isoforms of the *FAD2-1* and *FAD2-2* genes in the soybean genome; and (v) investigate whether the oleate desaturase isoforms cosegregate with oleate QTLs. The research questions addressed will elucidate the progress of selection for the high oleate trait, the correlated responses of selection for high oleate content, as well as the genetic factors underlying oleate variation in soybean.

REFERENCES

Anai, T., T. Yamada, T. Kinoshita, S.M. Rahman, and Y. Takagi. 2005. Identification of corresponding genes for three low-α-linolenic acid mutants and elucidation of their contribution to fatty acid biosynthesis in soybean seed. Plant Sci. 168:1615-1623.

Arondel, V., B. Lemieux, I. Hwang, S. Gibson, H.M. Goodman, and C.R. Somerville. 1992. Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. Science. 258: 1353-1355.

Arumuganathan, K., and E. Earle. 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. Report. 9:208-218.

Bachlava, E., R.E. Dewey, J. Auclair, S. Wang, J.W. Burton, and A.J. Cardinal. 2008. Molecular mapping of genes encoding microsomal omega-6 desaturase enzymes and their cosegregation with QTL affecting oleic acid content in soybean. Crop Sci. (in press).

Bennett, J.O., A.H. Krishnan, W.J. Wiebold, and H.B. Krishnan. 2003. Positional effect on protein and oil content and composition of soybeans. J. Agric. Food Chem. 51:6882-6886.

Bilyeu, K.D., L. Palavalli, D.A. Sleper, and P.R. Beuselinck. 2003. Three microsomal omega-3 fatty-acid desaturase genes contribute to soybean linolenic acid levels. Crop Sci. 43:1833-1838.

Browse, J., and C. Somerville. 1991. Glycerolipid synthesis: biochemistry and regulation. Ann. Rev. Plant Physiol. Plant Mol. Biol. 42:467-506.

Browse, J., M. McMonn, D.W. James Jr, and M. Miquel. 1993. Mutants of Arabidopsis deficient in the synthesis of α-linolenate. J. Biol. Chem. 268:16345-16351.

Brummer, E.C., A.D. Nickell, J.R. Wilcox, and R.C. Shoemaker. 1995. Mapping the *Fan* locus controlling linolenic acid content in soybean oil. J. Hered. 86:245-247.

Buhr, T., S. Sato, F. Ebrahim, A.Q. Xing, Y. Zhou, M. Mathiesen, B. Schweiger, A. Kinney, P. Staswick, and T. Clemente. 2002. Ribozyme termination of RNA transcripts down-regulate seed fatty acid genes in transgenic soybean. Plant J. 30:155-163.

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, C.A. Brim, and R.W. Rinne. 1989. Registration of soybean germplasm lines with modified fatty acid composition of seed oil. Crop Sci. 29:1583-1583.

Burton, J.W. 1991. Recent developments in breeding soybeans for improved oil quality. Fat Sci. Technol. 93:121-128.

Burton, J.W., R.F. Wilson, W. Novitzky, and T.E. Carter Jr. 2004. Registration of 'Soyola' soybean. Crop Sci. 44:687-688.

Burton, J.W., R.F. Wilson, G.J. Rebetzke, and V.R. Pantalone. 2006. Registration of N98-4445A mid-oleic soybean germplasm line. Crop Sci. 46:1010-1012.

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

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. M.S. diss. North Carolina State Univ., Raleigh, NC.

Cardinal, A.J. 2008. Molecular genetics and breeding for fatty acid manipulation in soybean. Plant Breed. Rev. 30 (in press).

Carver, B.F., J.W. Burton, T.E. Carter, Jr., R.F. Wilson. 1986. Responses to environmental variation of soybean lines selected for altered unsaturated fatty acid composition. Crop Sci. 26:1176-1181.

Carver, B.F., J.W. Burton, and R.F. Wilson. 1987. Graft-transmissible influence on fatty acid composition of soybean seed. Crop Sci. 27:53-56.

Cheesbrough, T.M. 1989. Changes in the enzymes for fatty acid synthesis and desaturation during acclimation of developing soybean seeds to altered growth temperature. Plant Physiol. 90:760-764.

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

Diers, B.W., and R.C. Shoemaker. 1992. Restriction fragment length polymorphism analysis of soybean fatty acid content. J. Am. Oil Chem. Soc. 69:1242-1244.

Fehr, W.R., G.A. Welke, E.G. Hammond, D.N. Duvick, and S.R. Cianzio. 1992. Inheritance of reduced linolenic acid content in soybean genotypes A16 and A17. Crop Sci. 32:903-906.

Fehr, W.R., E.G. Hammond. 1998. Reduced linolenic acid production in soybeans. U.S. Patent 5850030 1998.

Frandsen, G.I., J. Mundy, and J.T.C. Tzen. 2001. Oil bodies and their associated proteins, oleosin and caleosin. Physiol. Plant. 112:301-307.

Hawkins, S.E., W.R. Fehr, and E.G. Hammond. 1983. Resource allocation in breeding for fatty acid composition of soybean oil. Crop Sci. 23:900-904.

Heppard, E.P., A.J. Kinney, K.L. Stecca, G. Miao. 1996. Developmental and growth temperature regulation of two different microsomal omega-6 desaturase genes in soybeans. Plant Physiol. 110:311-319.

Hobbs, D.H., J.E. Flintham, and M.J. Hills. 2004. Genetic control of storage oil synthesis in seeds of Arabidopsis. Plant Physiol. 136:3341-3349.

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

Huang, A.H.C. 1992. Oil bodies and oleosins in seeds. Ann. Rev. Plant Physiol. Plant Mol. Biol. 43:177-200.

Hyten, D.L., V.R. Pantalone, A.M. Saxton, M.E. Schmidt, and C.E. Sams. 2004. Molecular mapping and identification of soybean fatty acid modifier quantitative trait loci. J. Am. Oil Chem. Soc. 81:1115-1118.

Ishikawa, G., H. Hasegawa, Y. Takagi, and T. Tanisaka. 2001. The accumulation pattern in developing seeds and its relation to fatty acid variation in soybean. Plant Breed. 120:417- 423.

Jackson, S.A., D. Rokhsar, G. Stacey, R.C. Shoemaker, J. Schmutz, and J. Grimwood. 2006. Toward a reference sequence of the soybean genome: a multi-agency effort. Crop Sci. 46:S-55-61.

Kinney, A.J. 1995. Improving soybean seed quality, p. 101-113 Induced Mutations and Molecular Techniques for Crop Improvement: Proceedings. International Atomic Energy Agency (IAEA), Vienna, Austria, 19-23 June.

Kinney, A.J., and S. Knowlton. 1998. Designer oils: the high oleic acid soybean, p. 193-213, *In* S. H. S. Roller, ed. Genetic Modification in the Food Industry. Chapman and Hall, London.

Kinoshita, T., S.M. Rahman, T. Anai, and Y. Takagi. 1998. Genetic analysis of restriction fragment length polymorphism on fatty acid synthesis in soybean mutants and their progenies: II. High oleic acid mutants with two microsomal omega-6 fatty acid desaturase cDNAs as probes. Bull. Fac. Agric., Saga Univ. 83:37-42.

Li, Y.H., F. Beisson, M. Pollard, and J. Ohlrogge. 2006. Oil content of *Arabidopsis* seeds: the influence of seed anatomy, light and plant-to-plant variation. Phytochemistry 67:904-915.

Miquel, M., and J. Browse. 1992. *Arabidopsis* mutants deficient in polyunsaturated fatty acid synthesis: biochemical and genetic characterization of a plant oleoyl-phosphatidylcholine desaturase. J. Biol. Chem. 267:1502-1509.

Monteros, M.J., J.W. Burton, and H.R. Boerma. 2004. Molecular mapping and confirmation of fatty acid QTL in N00-3350 soybean. ASA Annual Meeting, Seattle, WA. (http://www.cropsoil.uga.edu/soylab/presentations.html; verified 12/13/2007).

Ohlrogge, J., and J. Browse. 1995. Lipid biosynthesis. Plant Cell. 7:957-970.

Ohlrogge, J.B., and J.G. Jaworski. 1997. Regulation of fatty acid synthesis. Ann. Rev. Plant Physiol. Plant Mol. Biol. 48:109-136.

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.

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.

O'Neill, C.M., S. Gill, D. Hobbs, C. Morgan, and I. Bancroft. 2003. Natural variation for seed oil composition in Arabidopsis thaliana. Phytochemistry. 64:1077-1090.

Panthee, D.R., V.R. Pantalone, and A.M. Saxton. 2006. Modifier QTL for fatty acid composition in soybean oil. Euphytica. 152:67-73.

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. 42:37-44.

Rahman, S.M., Y. Takagi, K. Miyamoto, and T. Kawakita. 1994a. Inheritance of low linolenic acid content in soybean mutant line M-5. Breed. Sci. 44:379-382.

Rahman, S.M., Y. Takagi, K. Kubota, K. Miyamoto, and T. Kawakita. 1994b. High oleic acid mutant in soybean induced by x-ray irradiation. Biosci. Biotech. Biochem. 58:1070-1072.

Rahman, S.M., and Y. Takagi. 1997. Inheritance of reduced linolenic acid content in soybean seed oil. Theor. Appl. Genet. 94:299-302.

Rahman, S.M., T. Kinoshita, T. Anai, S. Arima, and Y. Takagi. 1998. Genetic relationships of soybean mutants for different linolenic acid contents. Crop Sci. 38:702-706.

Rennie, B.D., J. Zilka, M.M. Cramer, and W.D. Beversdorf. 1988. Genetic analysis of low linolenic acid levels in the soybean line PI361088b. Crop Sci. 28:655-657.

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

Schlueter, J.A., I.F. Vasylenko-Sanders, S. Deshpande, J. Yi, M. Siegfried, B.A. Roe, S.D. Schlueter, B.E. Scheffler, and R.C. Shoemaker. 2007. The *FAD2* gene family of soybean: Insights into the structural and functional divergence of a paleopolyploid genome. The Plant Genome, a Suppl. to Crop Sci. 47:S14-S26.

Singh, B.B., and H.H. Hadley. 1968. Maternal control of oil synthesis in soybeans. Crop Sci. 8:622.

Somerville, C., and J. Browse. 1996. Dissecting desaturation: Plants prove advantageous. Trends Cell Biol. 6:148-153.

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

Stojsin, D., B.M. Luzzi, G.R. Ablett, and J.W. Tanner. 1998. Inheritance of low linolenic acid level in the soybean line RG10. Crop Sci. 38:1441-1444.

Takagi, Y., and S.M. Rahman. 1995. Variation of different fatty acids in mutants in comparison with natural soybean varieties. Bull. Fac. Agric., Saga Univ. 79:23-27.

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

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

White, J.A., T. Todd, T. Newman, N. Focks, T. Girke, O.M. de Ilarduya, J.G. Jaworski, J.B. Ohlrogge, and C. Benning. 2000. A new set of *Arabidopsis* expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil. Plant Physiol. 124:1582-1594.

Wilcox, J.R., J.F. Cavins, and N.C. Nielsen. 1984. Genetic alteration of soybean oil composition by a chemical mutagen. J. Am. Oil Chem. Soc. 61:97-100.

Wilcox, J.R., and J.F. Cavins. 1987. Gene symbol assigned for linolenic acid mutant in soybean. J. Hered. 78:410-410.

Wilcox, J.R., and J.F. Cavins. 1992. Normal and low linolenic acid soybean strains: response to planting date. Crop Sci. 32:1248-1251.

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.

Wilcox, J.R. 2004. World distribution and trade of soybean, p. 1-13, *In* H. R. Boerma and J. E. Specht, eds. Soybeans: Improvement, Production and Uses, 3rd ed. American Society of Agronomy, Inc., Madison, WI.

Wilson, R.F., J.W. Burton, and C.A. Brim. 1981. Progress in the selection for altered fatty acid composition in soybeans. Crop Sci. 21:788-791.

Wilson, R.F., J.W. Burton, V.R. Pantalone, and R.E. Dewey. 2002. New gene combinations governing saturated and unsaturated fatty acid composition in soybean p. 95-113, *In* T. M. Kuo and H. W. Gardner, eds. Lipid Biotechnology. Marcel Dekker, Inc., IL.

Wilson, R.F. 2004. Seed composition, p. 621-669, *In* H. R. Boerma and J. E. Specht, eds. Soybeans: Improvement, Production, and Uses, 3rd ed. American Society of Agronomy, Inc., Madison, WI.

Yadav, N.S., A. Wierzbicki, M. Aegerter, C.S. Caster, L. Perez-Grau, A.J. Kinney, W.D. Hitz, J.R. Booth Jr, B. Schweiger, K.L. Stecca, S.M. Allen, M. Blackwell, R.S. Reiter, T.J. Carlson, S.H. Russell, K.A. Feldmann, J. Pierce, and J. Browse. 1993. Cloning of higher plant omega-3 fatty acid desaturases. Plant Physiol. 103:467-476.

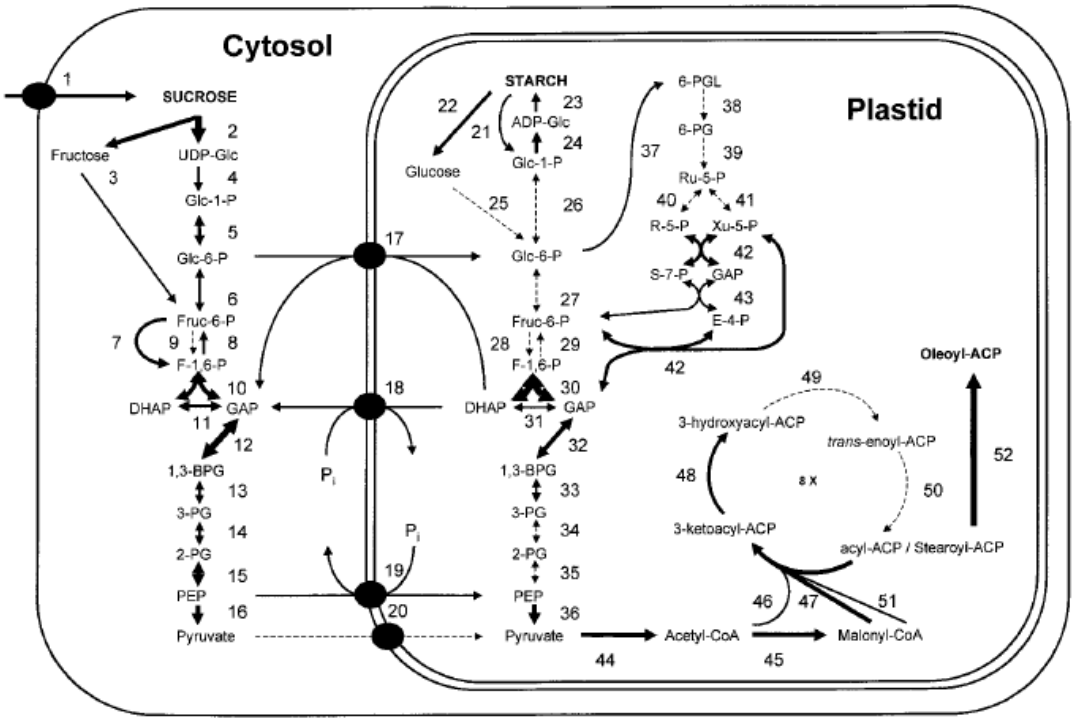

Figure 1. Schematic representation of metabolic pathways in an oil storing cell of a developing *Arabidopsis* embryo. The reactions correspond to the glycolytic pathway for the conversion of the imported sucrose in the cytosol (1-16), the transfer of intermediates into the plastid (17-20), starch biosynthesis and degradation in the plastid (21-26), the glycolytic pathway in the plastid (27-36), the oxidative pentose phosphate cycle (37-42), the plastid pyruvate dehydrogenase complex (44) and fatty acid biosynthesis (45-52) (White et al,. 2000).

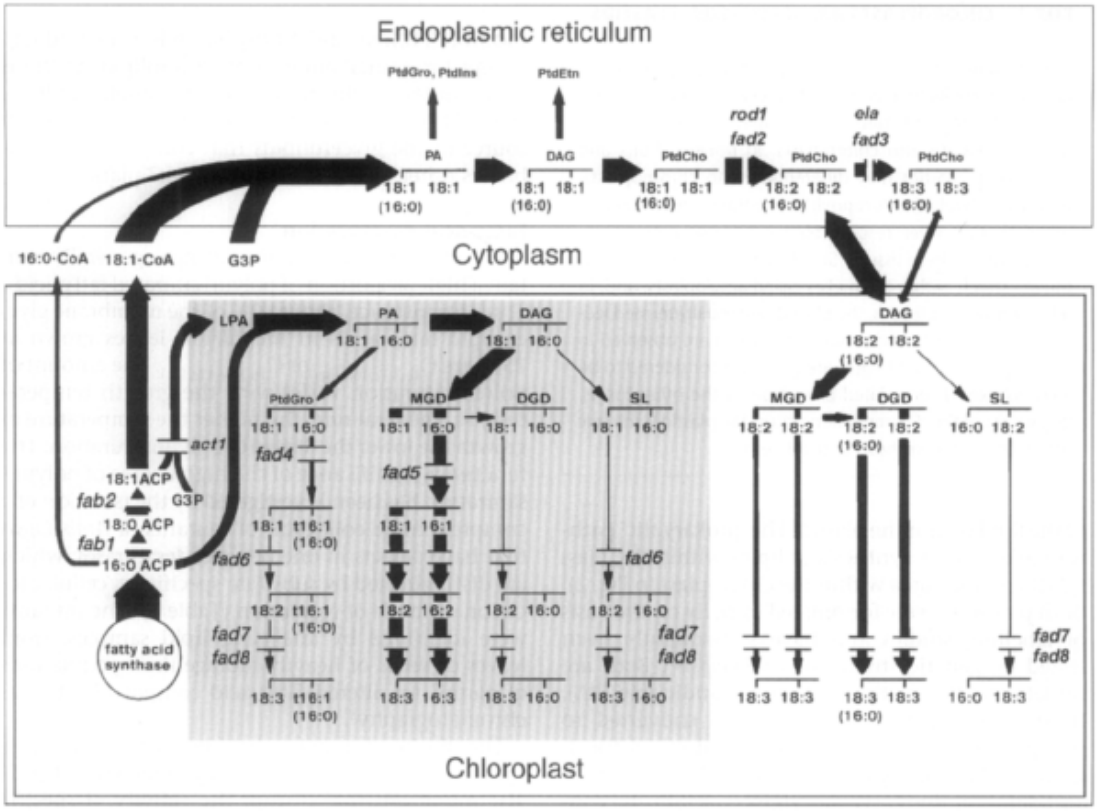

Figure 2. Glycerolipid biosynthesis in the leaves of *Arabidopsis.* The shaded area corresponds to the prokaryotic pathway. The width of the arrows represents the relative flux among reactions of the prokaryotic and eukaryotic pathways (Somerville and Browse, 1996).

II. Mapping Genes Encoding Microsomal ω-6 Desaturase Enzymes and Their Cosegregation with QTL Affecting Oleate Content in Soybean

Eleni Bachlava[1], Ralph E. Dewey[1], Jérôme Auclair[1], Sanbao Wang[1], Joseph W. Burton[1, 2], and Andrea J. Cardinal[1]

[1] Department of Crop Science, North Carolina State University, Campus Box 7620, Raleigh, NC 27695-7620

[2] USDA-ARS, 3127 Ligon St, Soybean and Nitrogen Fixation Research Unit, Raleigh, NC 27607

Formatted for: Crop Science

ABSTRACT

The microsomal ω-6 desaturase enzymes, which catalyze the desaturation of oleic acid to linoleic acid during fatty acid biosynthesis, are encoded by the *FAD2-1* and *FAD2-2* genes in soybean. Breeders aim to incorporate the high oleate trait into soybean germplasm in order to improve the nutritional value and oxidative stability of soybean oil. The objectives of this study were to map the isoforms of the *FAD2-1* and *FAD2-2* genes and investigate the association of these genetic loci with the oleate phenotype in three populations segregating for oleate content. The populations were grown in replicated multi-environment field trials. According to linkage analysis conducted for two of the populations, *FAD2-1A* and *FAD2-1B* mapped on linkage groups O and I, respectively, while the closely linked *FAD2-2A* and *FAD2-2B* isoforms mapped on linkage group L. Oleate quantitative trait loci (QTLs) with minor effects were detected in the proximity of *FAD2-1B* and possibly *FAD2-2B* on linkage groups I and L. QTL affecting maturity were also detected on chromosomal regions adjacent to the *FAD2* genes. The genotyping assays developed for the *FAD2-1A*, *FAD2-1B* and *FAD2-2B* isoforms, as well as their linked SSR markers, can be used in soybean breeding programs for the elevation of oleic acid seed content through marker-assisted selection.

INTRODUCTION

Soybean (*Glycine max* (L.) Merr) is an oilseed crop of major importance providing much of the world's edible oil. According to the United States Department of Agriculture statistics (2006) soybean oilseeds produce 29.4% of the world's edible vegetable oil, while the United States represent 40.1% of the world's soybean production. Yet, soybean oil is lacking in some quality characteristics. Oil quality is determined by the relative composition of saturated and unsaturated fatty acids in seed triacylglycerols. The average fatty acid composition of soybean oil is 110 g kg^{-1} palmitic acid (16:0), 40 g kg^{-1} stearic acid (18:0), 230 g kg^{-1} oleic acid (18:1), 530 g kg^{-1} linoleic acid (18:2) and 90 g kg^{-1} linolenic acid (18:3) (Wilson, 2004). Increased concentration of the monounsaturated oleic acid would improve the nutritional value and oxidative stability of soybean oil. This would also reduce the need for hydrogenation that leads to undesirable *trans* fatty acids (Wilson, 2004). Thus, extensive research has been conducted to incorporate the high oleic acid trait into soybean germplasm with traditional plant breeding methods (Wilson et al., 1981; Burton et al., 1989; Burton et al., 2006) and genetic engineering approaches (Kinney, 1995; Kinney & Knowlton, 1998; Buhr et al., 2002).

Fatty acids are the precursors for the synthesis of triacylglycerols (TAGs), which constitute the storage lipid reserves of the developing oilseeds. Polyunsaturated fatty acids of the seed TAGs are synthesized primarily by ω-3 and ω-6 desaturase

enzymes located in the endoplasmic reticulum; however, a small percentage is also derived from comparable enzymes localized in the plastids. The desaturase enzymes of the endoplasmic reticulum act on the fatty acids esterified to the first and second backbone position of phosphatidylcholine (PC), whereas the plastid localized desaturases predominantly utilize galactolipids as substrates (Ohlrogge and Browse, 1995).

Plant fatty acid desaturase genes were initially elucidated in *Arabidopsis*. The desaturation of oleic to linoleic acid in the endoplasmic reticulum was shown to be controlled by the *fad2* gene, encoding the oleoyl-PC ω-6 desaturase activity (Miquel & Browse, 1992; Okuley et al., 1994). Subsequently, genes designated *FAD2-1* and *FAD2-2* were reported to encode comparable activities in soybean (Heppard et al., 1996). *FAD2-1* was expressed only during seed development and *FAD2-2* was constitutively expressed during seed development and vegetative growth. Therefore, it was suggested that the seed-specific *FAD2-1* is primarily responsible for the desaturation of oleic to linoleic acid for fatty acids destined for seed storage lipids, while *FAD2-2* is involved in the desaturation of membrane lipids in vegetative tissues, playing a lesser role in storage oil biosynthesis. Moreover, the existence of at least two copies of each of the *FAD2-1* and *FAD2-2* genes was proposed (Heppard et al., 1996).

Recently, the existence of two distinct isoforms of the soybean *FAD2-1* gene (designated *FAD2-1A* and *FAD2-1B*), encoding microsomal ω-6 desaturase enzymes

that differ in stability at high temperatures, was confirmed (Tang et al., 2005). *FAD2-1A* and *FAD2-1B* map on linkage groups O and I, respectively, in the soybean genome. Both *FAD2-1A* and *FAD2-1B* produce alternatively spliced transcripts that do not change the amino acid residues of the encoded ω-6 desaturase enzymes (Schlueter et al., 2007). For *FAD2-2* genes, which also encode microsomal ω-6 desaturase enzymes, three isoforms have been reported, designated *FAD2-2A*, *FAD2-2B* and *FAD2-2C*. *FAD2-2A* and *FAD2-2B* are localized on linkage group L (Schlueter et al., 2007).

Environmental factors affect the unsaturated fatty acid content of soybean oilseeds. Temperature exerts an effect on the unsaturated fatty acid composition during the pod-filling stage that coincides with oil deposition. Planting and maturity dates lead to changes in the fatty acid profiles due to different temperatures during oil deposition (Burton et al., 1983; Wilcox & Cavins, 1992). For example, the elevated oleic lines N98-4445A and N97-3363-3, which were used in this study, are unstable across environments and their oleate variation is largely attributed to growth temperature differences (Oliva et al., 2006; Wilson et al., 2002). Although temperature effects during oil deposition were implicated in the altered fatty acid composition of soybean oilseeds, the expression patterns of the isoforms encoding the ω-6 desaturase enzymes, with the exception of *FAD2-2C*, showed no change in the level of transcript accumulation at different temperatures during seed development that could explain the changes in the levels of unsaturated fatty acids in seed

triacylglycerols (Heppard et al., 1996; Schlueter et al., 2007). However, a recent study of the ω-6 desaturase enzymes encoded by the *FAD2-1A* and *FAD2-1B* genes provided evidence of differential enzymatic stability at high temperatures (Tang et al., 2005). Apart from temperature effects, coordination of metabolite flux through the reactions of glycerolipid biosynthesis in the endoplasmic reticulum coupled with the *de novo* fatty acid biosynthesis pathway in the plastids underlie the complexity of the factors (both genetic and environmental) that can ultimately determine oleate composition of the storage oil.

Although the location of the isoforms encoding the ω-6 desaturase enzymes was recently revealed, there is minimal information regarding whether the isoforms of *FAD2-1* and *FAD2-2* genes cause the observed genotypic variation for oleate content in populations in which this trait is segregating. As an initial step toward unraveling the complexity of oleate biosynthesis and accumulation in soybean, the effects of the *FAD2-1* and *FAD2-2* isoforms (and their alleles), as well as their epistatic and environmental interactions, need to be investigated (Cardinal, 2008). The objectives of this study were to: (i) identify single nucleotide polymorphisms (SNPs) in the coding or upstream regions of the isoforms of *FAD2-1* and *FAD2-2* genes for the soybean lines under study; (ii) develop isoform- and allele-specific markers; (iii) map the individual isoforms of the *FAD2-1* and *FAD2-2* genes in the soybean genome; and (iv) investigate whether segregation at *FAD2-1* and *FAD2-2* loci is associated

with changes in oleic acid seed content in three soybean populations segregating for oleate content.

MATERIALS AND METHODS

Development of Soybean Populations

Three soybean populations, designated as FAE, FAF and FAS, were developed by single seed descent (Brim, 1966). FAE consists of 721 F_5- derived lines from the cross of N98-4445A x N97-3525. FAF represents 118 F_5- derived lines from the cross of N97-3363-3 x PI423893. FAS consists of 231 F_3- derived lines from the cross of N98-4445A x PI423893. N98-4445A is a high oleic (563.1 g kg^{-1}), low linoleic acid line developed by the USDA-ARS in Raleigh, NC (Burton et al., 2006). N98-4445A originated as a plant selection from the cross N94-2473 x (N93-2007-4 x N92-3907) and is a sister line of N97-3363-3. PI423893 is a plant introduction with mid-oleic acid (305.8 g kg^{-1}) seed oil content and unknown genetic background (USDA-ARS National Plant Germplasm System, http://www.ars-grin.gov/npgs /searchgrin.html, verified 05/08/2007). N97-3525 is the low palmitate, low linolenate cultivar Satelite, which was derived from the cross of Soyola x {Brim (2) x [N88-431(2) x (N90-2013 x C1726)]} (Cardinal et al., 2007). The FAF and FAS populations are segregating for oleate genes and the reduced linolenate *fan*(PI123440) allele; while, the FAE population is segregating for oleate genes and the fap_{nc} reduced palmitate allele.

Experimental Design

The FAE and FAF populations were planted separately in sets within replications experimental designs with two replications in each location. Fifteen sets of seven rows by seven columns were randomly assigned in each replication, location, and year for the FAE population. Five sets of five rows by five columns were randomly assigned in each replication, location, and year for the FAF population. In each population, before lines were randomly assigned to each set, they were divided into five arbitrary groups that differed in maturity according to the lines maturity phenotype in 2004. Each maturity group was represented at least once in each set of the FAE and FAF populations, and the maturity groups that included the majority of the segregating lines were represented two to three times in each set. Once lines were randomly assigned to a set, they remained in that set in all replications, locations, and years. Within each set, the maturity groups were randomly assigned to a column to facilitate mechanical harvest. Then, the lines within each maturity group were randomly assigned to plots within each column. The parental lines N97-3363-3 and PI423893 in FAF population and the parental line N97-3525 in FAE population were randomly assigned to one set and, depending on their maturity date, to a maturity group within that set and, finally, to a plot within the set, as explained above. The N98-4445A parental line of the FAE population was assigned to a maturity group in all the sets and randomly assigned a plot within each set as

explained above. The FAE population was planted in four-row plots at Clinton, Kinston and Plymouth, NC, in 2005 and 2006. The FAF population was grown at Clinton and Kinston, NC, in 2005 and at Clinton, Kinston, Clayton and Plymouth, NC, in 2006. The FAF population was planted in four-row plots, with the exception of Clayton in 2006, which was planted in one-row plots.

The FAS population was planted in five-seed hills at Clayton, NC, in a sets within replications design with three replications in 2006. Sets were randomly assigned to each replication, and lines within each set were randomly assigned to hills. The parental lines, N98-4445A and PI423893, were randomly assigned within each of the five sets in each replication.

Phenotypic Evaluation

Maturity date was recorded for FAE, FAF and FAS populations at the R8 reproductive stage as days after planting (DAP) (Fehr and Caviness, 1977). For the FAE and FAF populations mature soybean seed was harvested mechanically from the two middle rows of each experimental plot or each single row plot for the FAF population at Clayton in 2006. Approximately 10 g of seed were subsampled from each plot for the evaluation of fatty acid composition. For the FAS population 5 g of seed were subsampled from the seed harvested from each hill. Fatty acid composition was evaluated by gas chromatography, as described by Burkey et al. (2007).

Statistical Analysis

Statistical analysis of the sets within replications design was conducted using Proc MIXED in SAS 9.1 (SAS Institute, 2004). Environments, sets, replications, lines, as well as their interactions (Appendix 1), were considered as random effects. Best linear unbiased predictors (BLUPs) were obtained for all traits and lines of the FAE, FAF and FAS populations (Appendix 2). The degrees of freedom for FAF and FAS populations were calculated using the Satterthwaite approximation, while for FAE population they were derived with the containment method in order to reduce the computing power required. Analysis of the FAE and FAF populations was conducted separately for each environment, as well as combined across all the environments for fatty acid traits and maturity date. The phenotypic data of the FAF population from Clinton, NC, in 2005 were discarded due to numerous missing data points and heterogeneity of variance in comparison with the other environments. The analysis of the FAS population for a single environment was performed as described above.

Simple sequence repeat (SSR) markers and allele-specific markers were tested for association with fatty acid phenotypes, derived as BLUPs, in FAE, FAF, and FAS populations using Single Factor Analysis in Proc GLM of SAS 9.1. The additive genetic effects were obtained by the CONTRAST statement in Proc GLM of SAS 9.1. Dominance genetic effects were also estimated for the FAS population. The

means of the homozygous genotypic classes for each trait in FAE and FAF populations or the homozygous and heterozygous genotypes in FAS population, as well as their standard errors, were obtained with the LSMEANS statement in Proc GLM of SAS 9.1. The difference of the means of the homozygous genotypic classes, as well as its standard error and its significance were determined with the ESTIMATE statement in Proc GLM SAS of 9.1. In order to investigate whether maturity date affected the associations between fatty acid traits and molecular markers, Single Factor Analysis was also conducted with maturity date included as a covariate in the model.

Genotypic Evaluation and Linkage Mapping

Genomic DNA was extracted from leaf tissue of approximately five to ten plants for each of the lines of FAF and FAE populations, collected from Kinston (2005) and Clayton (2004), respectively. The DNA isolation was conducted in a 96-well plate format with the Gentra PureGene DNA purification kit (Gentra Systems, Minneapolis, MN). Genomic DNA was isolated from leaf tissue (Keim et al., 1988) of each line of the FAS population from all the plants in each hill grown at Clayton (2006). One hundred forty-seven polymorphic SSR markers of the consensus linkage map (Cregan et al., 1999; Song et al., 2004) were genotyped across the 20 linkage groups of the soybean genome in the FAF population. Only SSR markers flanking the

genomic regions of interest, as discussed below, were genotyped in the FAE and FAS populations (Table 1). The reactions were performed in a 384-well PTC-200 Peltier Thermal Cycler (MJ Research, Waltham, MA). Thermocycling conditions were 95ºC for 2 min and 38 cycles of 92ºC for 1 min, 49ºC for 1 min and 68ºC for 1 min and 30 s. The amplification products were resolved on 4% SFR agarose gels (Amresco, Solon, OH) with ethidium bromide staining in 1x Tris-Borate-EDTA buffer, or 6.5% polyacrylamide gels in a LICOR 4300 DNA Analysis System (LICOR Biosciences, Lincoln, NE) using M13-tailed unmodified primers and IRD-700 or IRD-800 labeled M13 oligonucleotides.

Linkage analysis was performed with JoinMap 3.0 (Van Ooijen and Voorrips, 2001). The Kosambi's mapping function was used for linkage map construction and linkage was declared with a maximum recombination frequency of 0.4 and a minimum logarithm of odds (LOD) score of 3.0. Linkage analysis was also conducted with MAPMAKER/EXP 3.0 (Lander et al., 1987), as previously described by Cardinal et al. (2001), in order to verify the insufficient linkage declared in JoinMap 3.0 and the differential placement of *FAD2-2B* in FAF and FAS populations, as explained below. Linkage analysis with MAPMAKER/EXP 3.0 was performed using the same LOD and recombination thresholds as linkage analysis with JoinMap 3.0; however, the heterozygote genotypes were discarded in order for the F_5-derived lines of FAF population to be analyzed as recombinant inbred lines, and the F_3-derived lines of FAS population were actually analyzed as F_2 lines. Therefore, linkage

mapping results from both analyses are discussed herein, but only the linkage analysis with JoinMap 3.0 is presented (Figure 2).

Database Search and Polymerase Chain Reaction Primer Design

The coding regions of *FAD2-1A* and *FAD2-1B* were initially amplified with isoform-specific primers designed from the GenBank (Benson et al., 2002) accessions AB188250 and AB188251, respectively. For the isoforms of *FAD2-2*, *in silico* analysis was conducted on the soybean expressed sequence tags (Shoemaker et al., 2002) deposited in GenBank using BLASTN searches. Three distinct isoforms were predicted and isoform-specific primers were designed. The specific primers that amplify the coding region of each isoform are: FAD2-1A_d_F (5'-GTG TGG CCA AAG TGG AAG TT-3') and FAD2-1A_b_R (5'-CAA AGC TCC CTT CAG CCA GT-3'), FAD2-1B_a_F (5'-GTG TGG CCA AAG TTG AAA T-3') and FAD2-1B_b_R (5'-AAG CAC CCC TCA GCC AAT-3'), FAD2-2B_c_F (5'-CCT CAG TTT AGT CTC AGC CAG ATT-3') and FAD2-2B_c_R (5'-TGG GCT TTA TTG CCT TTG TC-3'). The specific primers that amplify the upstream region of *FAD2-2B* isoform are: FAD2-2B_UP_1_F (5'-ACG CAG TGA ATC AAA TGA CAA-3') and FAD2-2B_UP_2_R (5'-CTT CTT AAT CTG GCT GAG ACT AAA CT-3'). Isoform and primer designations for the isoforms of *FAD2-1* genes follow the nomenclature of

Tang et al. (2005); while, for the isoforms of *FAD2-2* genes follow the nomenclature of Schlueter et al. (2007).

Sequencing of *FAD2* Genes, Single Nucleotide Polymorphism Detection and Allele-Specific Marker Development

Genomic DNA was isolated from bulked leaf tissue (Keim et al., 1988) from approximately 10 plants of N98-4445A, N97-3363-3, PI423893 lines and 'Brim'. The amplification of the various *FAD2-1* and *FAD2-2* genes was carried out in a MyCycler thermal cycler (Bio-Rad, Hercules, CA). The polymerase chain reactions (PCRs) were performed using the Expand High Fidelity PCR system (Roche Applied Science, Indianapolis, IN). Each 50-µl reaction contained 0.8 U polymerase, 20 pmol of each of the forward and reverse primers, 187 µM dNTPs and 100 ng of genomic template DNA. The cycling program included a 5-min step at 94 ºC, 30 amplification cycles of 30-s denaturation at 94 ºC, annealing at 55 ºC for 30 s, and extension at 72 ºC for 2 min, followed by an extension step for 7 min at 72 ºC step. The amplification products were cloned into pCR 2.1 cloning vectors (Invitrogen, Carlsbad, CA) and sequenced (Iowa State DNA facility, Ames, IA). Pairwise alignments of the sequences obtained for the isoforms of *FAD2-1* and *FAD2-2* genes were performed for the detection of SNPs among the soybean lines N98-4445A, N97-3363-3, PI423893 and 'Brim'. Single nucleotide polymorphisms leading to restriction site

polymorphisms were identified and cleaved amplified polymorphic sequence (CAPS) markers were developed for *FAD2-1B* and *FAD2-2B* using the *Hpy*CH4III and *Alu*I endonucleases, respectively.

The amplified fragments of *FAD2-1B* and *FAD2-2B* isoforms for CAPS analysis were obtained with the following primers: FAD2-1B_c_F (5'-ATC CAA AGT TGC ATG GTA CAC-3'), FAD2-1B_b_R (5'-AAG CAC CCC TCA GCC AAT-3'); FAD2-2B_UP_1_F (5'-ACG CAG TGA ATC AAA TGA CAA-3'), FAD2-2B_UP_1_R (5'-TGC TAT CGT TAC CCA CAT ACG A-3'). Cycling conditions were the same as described above. Amplification of genomic DNA was conducted with a 12-µl PCR reaction containing 6 pmol of each of the forward and reverse primers, 0.5 U Taq polymerase (New England Biolabs, Ipswich, MA), 10x buffer (10 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$, pH 8.3), 234 µM dNTPs and 75 ng genomic DNA. Next, 10 µl of the amplified product was digested in a-22 µl reaction at 37 ºC for 2 h. For the digestion of *FAD2-1B* fragments the reaction contained 2.5 U *Hpy*CH4III, buffer (20 mM Tris-acetate, 50 mM potassium acetate, 10 mM magnesium acetate, 1 mM dithiothreitol, pH 7.9) and 5 µg BSA. The *FAD2-2B* amplification products were digested in reactions containing of 5 U *Alu*I, buffer (10 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, pH 7.9) and 5 µg BSA. Restriction fragments were resolved in 4% SFR agarose gels, as described above.

In the FAF population, the *FAD2-1A* isoform was genotyped with the allele-specific primer extension (ASPE) assay (University of California-Davis Genome Center, Davis, CA) in a Luminex platform (Luminex Corp., Austin, TX) using the primers FAD2-1A_ASPE1 (5'-TCA ATT ACT TCA CTT TAA TCC TTT gcc aaa gtg gaa gtt caa g-3') for N97-3363-3 and FAD2-1A_ASPE2 (5'-TCA TTC ATA TAC ATA CCA ATT CAT gcc aaa gtg gaa gtt caa c-3') for PI423893. The primers include the SNP-specific sequence (lowercase letters) coupled to the oligonucleotides associated with the fluorescent microspheres (uppercase letters). In the FAS population, *FAD2-1A* was genotyped with the single-base extension (SBE) assay in a Luminex platform (Luminex Corporation, Austin, TX) according to Chen et al. (2000) (data not shown).

RESULTS AND DISCUSSION

Identification of the Isoforms of the *FAD2-2* Gene

A database query indicated that there are three isoforms for the *FAD2-2* gene, represented within GenBank accessions AB188253 and AC166091 (data not shown). The AC166091 accession corresponds to a bacterial artificial chromosome (BAC) containing two independent *FAD2-2* genes. Recently, Schlueter et al. (2007) designated the isoform in AB188253 (also found in a BAC corresponding to accession number AC166742) as *FAD2-2C* and the two isoforms in AC166091 as *FAD2-2A* and *FAD2-2B*. Because *FAD2-2A* and *FAD2-2B* are located on the same BAC, they should map to the same chromosomal region in the soybean genome. Since *FAD2-2A* has an approximately 100-base pair (bp) deletion within the coding region of the enzyme, it most likely encodes a non-functional enzyme. Therefore, *FAD2-2A* was not considered a candidate gene that could explain differences in the ω-6 desaturase activity in our populations and was not studied further. Indeed, transcript analysis of *FAD2-2A* recently provided evidence that this isoform may not be expressed (Schlueter et al., 2007). Also, gene-specific primers designed for *FAD2-2C* amplified only part of the isoform's coding region, which did not reveal any SNPs (data not shown) among the genotypes used in this study. Consequently, further information for *FAD2-2C* isoform could not be obtained.

Single Nucleotide Polymorphism Detection

The isoform-specific primers were used for the amplification of *FAD2-1A*, *FAD2-1B* and *FAD2-2B* from genomic DNA of lines N98-4445A, N97-3363-3 and PI423893. Amplification products were cloned, sequenced and aligned for the parental soybean lines of interest and the cultivar Brim, leading to the detection of SNPs in coding or upstream regions of each isoform. One SNP originating from PI423893 results in the change of the 22nd amino acid residue of the predicted protein of *FAD2-1A* from glycine to arginine due to a substitution of a guanine with a cytosine in the first position of the codon. This is one of the four amino acid residues that differ between the FAD2-1A and FAD2-1B enzymes in the N-terminal region. Interestingly, the location of the SNP in the coding region of *FAD2-1A* coincides with a region responsible for the instability of the enzyme at high temperatures (Tang et al., 2005).

For *FAD2-1B*, a total of five SNPs were detected in the coding region, two of which result in non-synonymous amino acid changes. The substitution of a cytosine with a thymine in the second position of the 86th codon of *FAD2-1B* causes an amino acid change from serine to phenylalanine in PI423893. Also, the 126th amino acid residue of the FAD2-1B enzyme in N98-4445A and N97-3363-3 changes from valine to methionine due to the substitution of a guanine with an adenine at the first position of this codon. Three additional nucleotide polymorphisms that do not change the

amino acid residues of the predicted FAD2-1B enzyme were also detected in PI423893 (Figure 1). These are located in the third positions of codons 219, 223 and 228. All non-synonymous SNPs in *FAD2-1A* and *FAD2-1B* were evaluated for potential implications in protein function using the SIFT program (Ng and Henikoff, 2003). None of the non-synonymous amino acid changes, however, were predicted to be deleterious.

Amplification and sequencing of the coding regions of *FAD2-2B* revealed no polymorphisms among the soybean lines of interest. Therefore, primers were designed to amplify a region immediately upstream of *FAD2-2B*. One SNP was detected 641 bp upstream of the putative start codon of *FAD2-2B*, corresponding to a polymorphism of a thymine versus cytosine.

Development of Allele-Specific Molecular Markers for *FAD2-1B* and *FAD2-2B*

The detected nucleotide substitutions were examined for possible restriction site polymorphisms coinciding with the location of the SNPs in order to develop molecular assays that could readily distinguish among the alleles found in N98-4445A, N97-3363-3 and PI423893. The substitution of a thymine with a cytosine in the 223rd codon of *FAD2-1B* generates an additional restriction site for the *Hpy*CH4III endonuclease in PI423893 (position 172, Figure 1), and thus enabled the development of a CAPS marker. Digestion of a 395-bp *FAD2-1B* specific fragment

from PI423893 generates fragments of 172, 90, 70 and 63 bp, whereas the same product derived from N98-4445A or N97-3363-3 CAPS yields fragments of 235, 90 and 70 bp (data not shown). The additional *Hpy*CH4III restriction sites that do not correspond to nucleotide substitutions within the amplified fragment serve as digestion controls.

The unique SNP in the upstream region of *FAD2-2B* results in the loss of an *Alu*I restriction site in the fragment amplified from PI423893. The primers used for CAPS analysis amplify a 494-bp fragment in the upstream region of *FAD2-2B*. The *Alu*I digestion of this amplification product from N98-4445A and N97-3363-3 generates fragments of size 173, 147 and 174 bp, in contrast to fragments of 320 and 174 bp when genomic DNA from PI423893 is used as template (data not shown).

The SNP identified in the coding region of *FAD2-1A*, did not alter restriction endonuclease cleavage sites. Therefore, an ASPE assay was designed with SNP-specific primers in order to distinguish between the N97-3363-3 and PI423893 *FAD2-1A* alleles in the FAF population. In the FAS population, an SBE assay was designed according to Chen et al. (2000) to distinguish between the N98-4445A and PI423893 *FAD2-1A* alleles using fluorescent microspheres in a Luminex platform. The segregation patterns of the CAPS markers for *FAD2-1B* and *FAD2-2B*, as well as of the ASPE and SBE assays for *FAD2-1A*, were utilized for linkage analysis and association tests with fatty acid content in the FAF and FAS populations.

Mapping the Isoforms of *FAD2-1* and *FAD2-2* Genes

Linkage analysis was conducted for 147 polymorphic SSR markers genotyped in the FAF population, which cover the 20 linkage groups of soybean genome, as well as the CAPS, ASPE and SBE assays with both JoinMap 3.0 (Van Ooijen and Voorrips, 2001) and MAPMAKER/EXP 3.0 (Lander et al., 1987). Since for the linkage analysis performed with MAPMAKER/EXP 3.0 the recombination frequency among markers and the map distances were slightly biased by analyzing the FAF F_5-derived population as a recombinant inbred line population, only the linkage analysis performed with JoinMap 3.0 is presented herein (Figure 2). It should be noted that linkage could not be declared among all markers on the same linkage group probably due to the relatively small size of the FAF population. The FAS population was used to validate the localization of the *FAD2-1A*, *FAD2-1B* and *FAD2-2B* isoforms.

Mapping of *FAD2-1A* at the end of linkage group O near the SSR marker sat_108, and of *FAD2-1B* in the interval of sat_268 and satt354 on linkage group I, for both FAF and FAS populations, coincide with the locations proposed by Schlueter et al. (2007). The congruency in the localization of *FAD2-1A* and *FAD2-1B* verifies the efficiency of the developed assays and the efficacy of linkage analysis as a mapping tool. However, a discrepancy in the linkage analysis of FAF and FAS populations was revealed on linkage group L, both with JoinMap 3.0 and MAPMAKER/EXP 3.0. The *FAD2-2A* and *FAD2-2B* isoforms mapped in the interval

of the SSR markers sat_340 and satt462 for the FAF population; while, for the FAS population they were localized in the interval of sat_340 and satt006. Considering the difference in FAF and FAS populations' sizes, which implies the limited mapping precision of FAF population, and the location proposed by Schlueter et al. (2007), *FAD2-2A* and *FAD2-2B* most likely map within the interval of sat_340 and satt006.

Associations between Molecular Markers and Fatty Acid Content or Maturity Date

Overview of cosegregation analysis

Segregation of the *FAD2-1A*, *FAD2-1B* and *FAD2-2B* allele-specific markers, as well as SSR markers in the vicinity of the isoforms on linkage groups O, I and L, was used to test for association with the content of unsaturated fatty acids in the oil and with maturity date (Table 1). Fatty acids and maturity date BLUPs were obtained for each line from the analysis of each environment separately and from the combined analysis across all environments using mixed models in SAS 9.1. BLUPs for each trait were used in Single Factor Analysis with *FAD2-1A*, *FAD2-1B* and *FAD2-2B* allele-specific and SSR marker classes as the independent variables. Although maturity effects were significant for the majority of fatty acid traits and environments tested in the FAE, FAF and FAS populations, they were not accounted in the computation of BLUPs. Instead, Single Factor Analysis was also performed with

maturity date included as a covariate, in order to determine whether the associations between fatty acid content and molecular markers were affected by maturity. All markers that were significantly associated ($P < 0.05$) with the unsaturated fatty acid content and with maturity date were included in Tables 2, 3 and 4.

Cosegregation analysis in the proximity of *FAD2-1A*, *FAD2-1B* and *FAF2-2B*

On linkage group I, segregation of the *FAD2-1B* CAPS marker was associated with changes in oleate and linoleate content in the FAS population. *FAD2-1B* CAPS marker explained 3.3 and 3.5% of oleate and linoleate genetic variation, when maturity date was included as a covariate (Table 3). The genetic locus coincided with a maturity quantitative trait locus (QTL), which explained 6.0% of the genetic variation for maturity. The high-oleate and low-linoleate alleles were inherited from PI423893 and the difference between the means of the homozygous genotypic classes was -13.6 and 11.6 g kg^{-1}, respectively (Table 3). The allele for late maturity was also inherited from PI423893 as expected from the positive correlation between maturity and oleic acid seed content in FAS population (data not shown). The SSR markers satt354 and sat_268, located 4.91 and 6.07 cM from *FAD2-1B* isoform, were also supportive of the minor oleate QTL in the proximity of *FAD2-1B* in both FAS and FAE populations (Tables 3 and 4). No significant association was detected on linkage group I for the *FAD2-1B* CAPS marker or the flanking SSR markers with fatty acid phenotypes in the FAF population. Thus, the minor oleate QTL in the vicinity of

FAD2-1B was detected in the FAS and FAE populations, but was not validated by the FAF population that has the smallest population size.

Analysis of the FAE population indicated the existence of a minor oleate QTL in the proximity of the *FAD2-2B* isoform on linkage group L. The oleate QTL, which mapped on sat_340 (13.86 cM from *FAD2-2B* according to the linkage map of FAS population) explained 2.2% of oleate genotypic variation and it was detected in all six environments tested for FAE population (data not shown). The high oleate allele at this locus was inherited from the N98-4445A parental line and the means of the homozygous genotypic classes differed by 18.8 g kg^{-1} (Table 4). The minor oleate QTL on linkage group L was not detected in the FAF or FAS populations possibly due to their smaller population size, which impacts the power of QTL detection and the magnitude of the QTL effect (Beavis, 1994). Quantitative trait loci with minor effects for oleic, linoleic and linolenic acid traits have been reported on linkage group L at the genomic region where the *FAD2-2A* and *FAD2-2B* isoforms map (Hyten et al., 2004); however, the population used for that study did not segregate for major oleic acid genes, since it was derived from the cross of normal-oleate parental lines, and maturity effects were not taken into consideration.

On linkage group O, segregation of the *FAD2-1A* ASPE and SBE markers was not associated with changes in unsaturated fatty acid content in the FAF and FAS populations, which suggests the lack of oleate QTLs in the proximity of *FAD2-1A*.

Cosegregation of other genetic loci mapped on linkage groups O, I and L

On linkage group O, analysis suggests a minor oleate QTL near the SSR markers sat_108 and satt153, which explained 2.4 and 0.6% of oleate genotypic variation in FAS and FAE populations, respectively (Tables 3 and 4). The satt153 locus coincided with a QTL for maturity explaining 8.9% of maturity genotypic variation in the FAE population. The late-maturity and low-oleate alleles were inherited from N97-3525, as expected by the negative genetic correlation between maturity and oleic acid seed content in the late-maturing FAE population (data not shown). However, after accounting for maturity effects in the model, N97-3525 contributed the high-oleate allele (Table 4), which favors the hypothesis of two linked QTLs controlling maturity and oleate content near the SSR marker satt153.

Apart from the minor oleate QTL near *FAD2-2B* on linkage group L, minor QTLs for oleate content were also mapped on the SSR markers satt652 and satt462. The SSR marker satt652, which is located 65.47 cM away from *FAD2-2B* and 19.67 cM from satt462 according to the linkage map of the FAS population, explained 5.2% of the genotypic variation for the oleate trait in the FAF population (Table 2). The association between oleate and satt652 was significant at the $P < 0.01$ level in each of the five environments tested for the FAF population (data not shown), as well as in the combined analysis across all environments (Table 2). The oleate QTL near satt652 on linkage group L was further confirmed with the FAE population, where satt652 explained 0.5% of the genetic variation for the oleate trait (Table 4). The

significant association between satt652 and oleate content was consistent across two of the six environments when maturity was used as covariate (data not shown). The high-oleate allele of the detected QTL was inherited from the N97-3363-3 and N98-4445A alleles in the FAF and FAE populations, respectively. In addition, minor QTLs for linoleate and linolenate traits were detected near satt652, when maturity effects were accounted. These results confirm other oleate QTL mapping studies currently in progress (H. R. Boerma, personal communication, 2007). In the FAS population, satt652 was not associated with oleate content but explained 1.6% of linolenate genetic variation (Table 3).

Also, on linkage group L, the SSR marker satt462 explained 4.6% of the genotypic variation for the oleate trait in the FAF population (Table 2). The segregation of satt462 was significantly associated with oleate content at the $P < 0.05$ level in four of the five environments tested (data not shown), as well as in the overall environment analysis (Table 2). The FAE population provided additional evidence for an oleate QTL in the proximity of satt462. When maturity effects were taken into account, satt462 explained 1.2% of the genetic variation for oleic acid content in the seed oil (Table 4). The associations declared for satt462 marker and oleate content in the FAE population were consistent across all six environments tested when maturity was used as covariate (data not shown). The high-oleate alleles of the QTL near satt462 on linkage group L were inherited from the N97-3363-3 and the N98-4445A parental lines in the FAF and FAE populations, respectively. The FAS population did

not confirm the association of oleic acid content with the SSR markers sat_340 and satt462. In the FAF population, satt462 explained 7.9% of the genotypic variation for linolenate content (Table 2), but the QTL conditioning linolenate content could not be validated in the FAE or FAS populations.

CONCLUSIONS

Oleate QTLs with minor effects were detected in the proximity of the *FAD2-1B* and possibly *FAD2-2B* oleate desaturase genes. Thus, the results of this study suggest that no major oleate QTLs cosegregate with the *FAD2-1A*, *FAD2-1B* and *FAD2-2B* isoforms of the genes encoding the ω-6 desaturase enzymes. Therefore, the *FAD2-1A*, *FAD2-1B* and *FAD2-2B* isoforms do not appear to encode for the major oleate genes that contribute to the elevated oleic acid content of the N98-4445A and N97-3363-3 lines. Further research on the *FAD2-2C* isoform as well as other candidate genes implicated in oleate biosynthesis, such as the plastidial desaturase enzymes, is critical in order to substantiate the results presented and shed light on the genetic factors that control oleic acid content in soybean oilseeds.

ACKNOWLEDGEMENTS

The authors would like to thank William Novitzky at the USDA Nitrogen Fixation Laboratory, Raleigh, NC for providing training and equipment for the fatty acid analysis, Perry Cregan and Song Qijian at the USDA Soybean Genomics and Improvement Laboratory, Beltsville, MD, for conducting the SBE assay, as well as Carol Griffin for her technical expertise. This project was funded by the USDA Cooperative State Research, Education and Extension Service National Research

Initiative, Plant Genome Program (award number 2003-35300-13190), and by the United Soybean Board.

REFERENCES

Beavis, W.D. 1994. The power and deceit of QTL experiments: Lessons from comparative QTL studies. *In* D. B. Wilkinson, ed. 49th Annual Corn and Sorghum Research Conference. American Seed Trade Association, Chicago, IL.

Benson, D.A., I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, B.A. Rapp, and D.L. Wheeler. 2002. GenBank. Nucleic Acids Res. 30:17-20.

Brim, C.A. 1996. A modified pedigree method of selection in soybean. Crop Sci. 6:220.

Buhr, T., S. Sato, F. Ebrahim, A.Q. Xing, Y. Zhou, M. Mathiesen, B. Schweiger, A. Kinney, P. Staswick, and T. Clemente. 2002. Ribozyme termination of RNA transcripts down-regulate seed fatty acid genes in transgenic soybean. Plant J. 30:155-163.

Burkey, K.O., F.L. Booker, W.A. Pursley, and A.S. Heagle. 2007. Elevated carbon dioxide and ozone effects on peanut: II. Seed yield and quality. Crop Sci. 47:1488-1497.

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, C.A. Brim, and R.W. Rinne. 1989. Registration of soybean germplasm lines with modified fatty acid composition of seed oil. Crop Sci. 29:1583-1583.

Burton, J.W., R.F. Wilson, G.J. Rebetzke, and V.R. Pantalone. 2006. Registration of N98-4445A mid-oleic soybean germplasm line. Crop Sci. 46:1010-1012.

Cardinal, A.J., M. Lee, N. Sharopova, W.L. Woodman-Clikeman, and M.J. Long. 2001. Genetic mapping and analysis of quantitative trait loci for resistance to stalk tunneling by the european corn borer in maize. Crop Sci. 41:835-845.

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. 47:304-310.

Cardinal, A.J. 2008. Molecular genetics and breeding for fatty acid manipulation in soybean. Plant Breed. Rev. 30 (in press).

Chen, J.W., M.A. Iannone, M.S. Li, J.D. Taylor, P. Rivers, A.J. Nelsen, K.A. Slentz-Kesler, A. Roses, and M.P. Weiner. 2000. A microsphere-based assay for multiplexed single nucleotide polymorphism analysis using single base chain extension. Genome Res. 10:549-557.

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

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

Heppard, E.P., A.J. Kinney, K.L. Stecca, and G. Miao. 1996. Developmental and growth temperature regulation of two different microsomal omega-6 desaturase genes in soybeans. Plant Physiol. 110:311-319.

Hyten, D.L., V.R. Pantalone, A.M. Saxton, M.E. Schmidt, and C.E. Sams. 2004. Molecular mapping and identification of soybean fatty acid modifier quantitative trait loci. J. Am. Oil Chem. Soc. 81:1115-1118.

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

Kinney, A.J. 1995. Improving soybean seed quality, p. 101-113 Induced Mutations and Molecular Techniques for Crop Improvement: Proceedings. International Atomic Energy Agency (IAEA), Vienna, Austria, 19-23 June.

Kinney, A.J., and S. Knowlton. 1998. Designer oils: The high oleic acid soybean, p. 193-213. *In* S. H. S. Roller, ed. Genetic Modification in the Food Industry. Chapman and Hall, London.

Kinney, A.J., E.B. Cahoon, and W.D. Hitz. 2002. Manipulating desaturase activities in transgenic crop plants. Biochem. Soc. Trans. 30:1099.

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

Miquel, M., and J. Browse. 1992. *Arabidopsis* mutants deficient in polyunsaturated fatty acid synthesis: biochemical and genetic characterization of a plant oleoyl-phosphatidylcholine desaturase. J. Biol. Chem. 267:1502-1509.

Ng, P.C., and S. Henikoff. 2003. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 31:3812-3814.

Ohlrogge, J. and Browse, J. 1995. Lipid biosynthesis. Plant Cell 7: 957-970.

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.

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.

SAS Institute. 2004. SAS/STAT software. Release 9.1.3. SAS Inst., Cary, NC.

Schlueter, J.A., I.F. Vasylenko-Sanders, S. Deshpande, J. Yi, M. Siegfried, B.A. Roe, S.D. Schlueter, B.E. Scheffler, and R.C. Shoemaker. 2007. The FAD2 gene family of soybean: Insights into the structural and functional divergence of a paleopolyploid genome. The Plant Genome, a Suppl. to Crop Sci. 47:S14-S26.

Shoemaker, R., P. Keim, L. Vodkin, E. Retzel, S.W. Clifton, R. Waterston, D. Smoller, V. Coryell, A. Khanna, J. Erpelding, X.W. Gai, V. Brendel, C. Raph-Schmidt, E.G. Shoop, C.J. Vielweber, M. Schmatz, D. Pape, Y. Bowers, B. Theising, J. Martin, M. Dante, T. Wylie, and C. Granger. 2002. A compilation of soybean ESTs: Generation and analysis. Genome 45:329-338.

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

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

Van Ooijen, J.W., and R.E. Voorrips. 2001. Joinmap® 3.0: Software for the calculation of genetic linkage maps, Wageningen, The Netherlands.

Wilcox, J.R., and J.F. Cavins. 1992. Normal and low linolenic acid soybean strains: Response to planting date. Crop Sci. 32:1248-1251.

Wilson, R.F., J.W. Burton, and C.A. Brim. 1981. Progress in the selection for altered fatty acid composition in soybeans. Crop Sci. 21:788-791.

Wilson, R.F., J.W. Burton, V.R. Pantalone, and R.E. Dewey. 2002. New gene combinations governing saturated and unsaturated fatty acid composition in soybean p. 95-113, In T. M. Kuo and H. W. Gardner, eds. Lipid Biotechnology. Marcel Dekker, Inc., IL.

Wilson, R.F. 2004. Seed composition, p. 621-669. *In* H. R. Boerma and J. E. Specht, eds. Soybeans: Improvement, Production, and Uses, 3rd ed. American Society of Agronomy, Inc., Madison, WI.

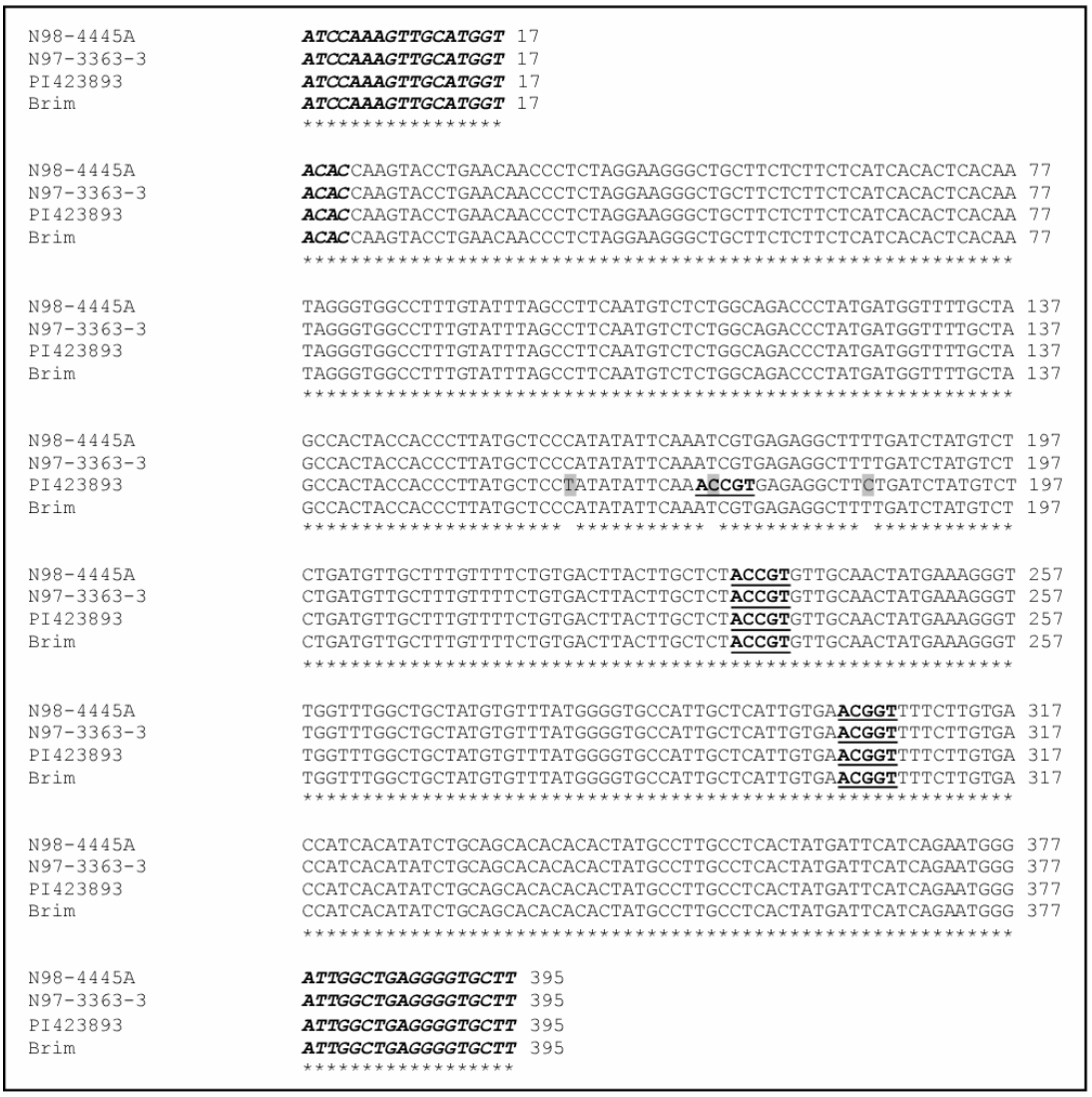

Figure 1. Alignment of the 395-bp fragment amplified for *FAD2-1B* CAPS analysis in N98-4445A, N97-3363-3, PI423893 and 'Brim'. The primers FAD2-1B_c_F and FAD2-1B_b_R designed for amplification are italicized. The restriction sites for *Hpy*CH4III are underlined and the three single nucleotide polymorphisms (SNPs) are highlighted. The substitution from thymine to cytosine in one of the SNPs generates an additional restriction site for *Hpy*CH4III endonuclease in PI423893.

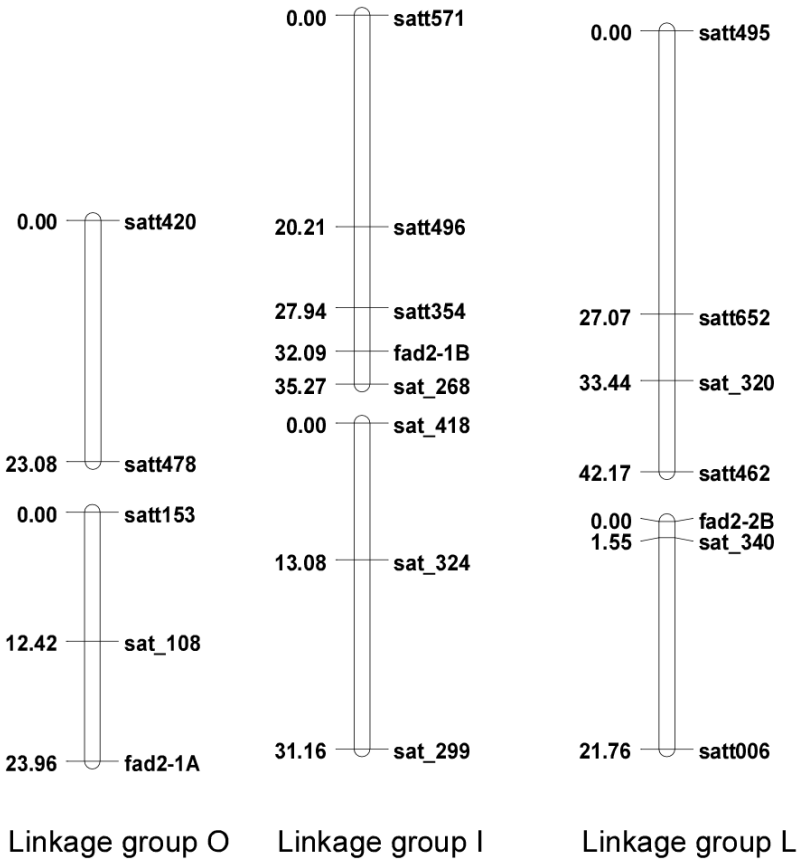

Figure 2. Genetic map of linkage groups I, O and L for the FAF population (N97-3363-3 x PI423893) constructed with JoinMap 3.0. The estimated locations of *FAD2-1A*, *FAD2-2A* and *FAD2-2B* are presented. It should be noted that mapping of *FAD2-2B* on linkage group L in the FAF population is not in agreement with linkage analysis in the FAS population. The *FAD2-2B* isoform most likely maps within the interval of the simple sequence repeat markers sat_340 and satt006 on linkage group L, as indicated by the linkage map of FAS population (N98-4445A x PI423893). The *FAD2-2A* isoform also co-localizes with *FAD2-2B* on linkage group L. Genetic distances were estimated with Kosambi's mapping function. Linkage groups were drawn on an equal scale.

Table 1. Molecular assays designed for the isoforms *FAD2-1A*, *FAD2-1B* and *FAD2-2B* as well as simple sequence repeat markers mapped on linkage groups I, O and L that were tested for associations with the unsaturated fatty acids' content in the FAF (N97-3363-3 x PI423893), FAE (N98-4445A x N97-3525), and FAS (N98-4445A x PI423893) populations.

Linkage group I	**FAF**	**FAE**	**FAS**	**Linkage group L**	**FAF**	**FAE**	**FAS**	**Linkage group O**	**FAF**	**FAE**	**FAS**
FAD2-1B CAPS[†]	x[‡]		x	*FAD2-2B* CAPS	x		x	*FAD2-1A* ASPE/SBE[†]	x		x
satt571	x			satt495	x			satt420	x		
satt496	x			satt652	x	x	x	satt478	x		
satt354	x	x	x	sat_320	x			satt153	x	x	x
sat_268	x	x	x	satt462	x	x	x	sat_108	x		x
sat_418	x			sat_340	x	x	x				
sat_324	x			satt006	x						
sat_299	x										

[†]CAPS, cleaved amplified polymorphic sequence; ASPE, allele specific primer extension; SBE, single-base extension.
[‡]Molecular assays and simple sequence repeat markers tested.

Table 2. The means and standard errors (in parentheses) of homozygous genotypic classes for maturity (R8), oleate (18:1), linoleate (18:2) or linolenate (18:3) content, the difference between means, its standard error (in parentheses), and its significance, as well as the R^2 for the difference of the means for molecular markers significantly associated with maturity, oleate, linoleate and linolenate content in the FAF (N97-3363-3 x PI423893) population.

Trait	Marker (linkage group)	AA[†] (g kg^{-1})	BB[†] (g kg^{-1})	Difference	R^2 (%)
R8[‡]	sat_340 (L)	*115.3 (0.5)*[§]	113.1 (0.6)	2.2 (0.8)**	6.3
	FAD2-2B (L)	*115.6 (0.6)*[§]	112.6 (0.6)	3.0 (0.8)***	10.8
	satt006 (L)	115.3 (0.5)	112.4 (0.7)	2.8 (0.9)**	9.1
	sat_268 (I)	113.6 (0.6)	115.4 (0.7)	-1.8 (0.9)*	3.7
18:1	satt462 (L)	416.8 (7.9)	391.6 (9.6)	25.1 (12.4)*	3.8
	satt652 (L)	420.5 (8.2)	394.8 (8.0)	25.7 (11.4)*	4.1
18:1 / R8[¶]	satt462 (L)	417.2 (7.4)	389.1 (9.1)	28.2 (11.7)*	4.6
	satt652 (L)	421.0 (7.5)	391.7 (7.4)	29.3 (10.6)**	5.2
18:2 / R8[¶]	satt652 (L)	402.7 (6.6)	425.8 (6.6)	-23.1 (9.4)*	4.3
18:3	satt462 (L)	44.0 (1.4)	49.5 (1.6)	-5.4 (2.1)*	5.8
18:3 / R8[¶]	satt462 (L)	43.6 (1.3)	50.1 (1.6)	-6.4 (2.0)**	7.9

*, **, *** Significant at the 0.05, 0.01 and 0.001 probability levels, respectively.
[†]AA, lines homozygous for N97-3363-3 allele; BB, lines homozygous for PI423893 allele.
[‡]R8, maturity (days after planting).
[§]Genotypic classes scored as dominant for N97-3363-3 allele are italicized. The mean and standard deviation for the dominant genotypic class corresponds to the pool of homozygous and heterozygous genotypes.
[¶]Maturity used as a covariate during Single Factor Analysis.

Table 3. The means and standard errors (in parentheses) of homozygous and heterozygous genotypic classes for maturity (R8), oleate (18:1), linoleate (18:2) or linolenate (18:3) content, the difference between means of homozygous genotypic classes, its standard error (in parentheses), and its significance, as well as the R^2 for the difference of the means for molecular markers significantly associated with maturity, oleate, linoleate and linolenate content in FAS (N98-4445A x PI423893) population.

Trait	Marker (linkage group)	AA[†] (g kg^{-1})	BB[†] (g kg^{-1})	HH[†] (g kg^{-1})	Difference	R^2 (%)
R8[‡]	sat_340 (L)	102.3 (0.4)	99.8 (0.4)	100.9 (0.6)	2.5 (0.6)***	7.6
	FAD2-2B (L)	*102.2 (0.4)*[§]	100.0 (0.4)		2.2 (0.5)***	7.3
	satt462 (L)	102.0 (0.4)	100.5 (0.4)	99.8 (0.8)	1.5 (0.6)**	3.0
	satt652 (L)	102.0 (0.4)	100.5 (0.4)	101.0 (0.6)	1.6 (0.6)**	3.0
	satt354 (I)	100.2 (0.4)	102.0 (0.4)	101.4 (0.6)	-1.8 (0.6)**	4.0
	FAD2-1B (I)	100.1 (0.4)	102.0 (0.4)[¶]		-2.0 (0.5)***	6.0
	sat_268 (I)	100.2 (0.4)	102.3 (0.4)	100.8 (0.6)	-2.1 (0.6)***	5.7
18:1	satt354 (I)	326.1 (3.5)	343.6 (3.8)	338.8 (5.7)	-17.5 (5.2)***	4.9
	FAD2-1B (I)	326.9 (3.4)	342.4 (3.4)		-15.5 (4.7)**	4.6
	sat_268 (I)	328.4 (3.5)	344.3 (3.4)	333.1 (5.6)	-15.9 (5.2)**	3.9
	sat_108 (O)	331.2 (3.7)	344.2 (3.6)	322.1 (5.8)	-12.9 (5.2)* [-31.2 (12.7)*][#]	2.6
18:1 / R8[††]	satt354 (I)	327.0 (3.5)	342.6 (3.8)	338.5 (5.7)	-15.6 (5.2)**	3.7
	FAD2-1B (I)	327.8 (3.4)	341.5 (3.3)		-13.6 (4.8)**	3.3
	sat_268 (I)	329.3 (3.5)	343.0 (3.9)	333.4 (5.6)	-13.7 (5.3)*	2.7
	sat_108 (O)	331.2 (3.7)	343.7 (3.6)	323.2 (5.8)	-12.5 (5.2)* [-28.6 (12.7)*]	2.4

Table 3. Continued.

Trait	Marker (linkage group)	AA† (g kg^{-1})	BB† (g kg^{-1})	HH† (g kg^{-1})	Difference	R^2 (%)
18:2	satt354 (I)	466.6 (2.9)	453.1 (3.2)	457.8 (4.8)	13.6 (4.3)**	4.3
	FAD2-1B (I)	466.4 (2.8)	454.2 (2.7)		12.2 (3.9)**	4.1
	sat_268 (I)	466.0 (2.9)	451.6 (3.2)	461.9 (4.6)	14.4 (4.3)***	4.6
18:2 / R8	satt354 (I)	466.2 (3.0)	453.5 (3.2)	458.0 (4.8)	12.7 (4.4)**	3.6
	FAD2-1B (I)	466.1 (2.9)	454.4 (2.8)		11.6 (4.1)**	3.5
	sat_268 (I)	465.7 (2.9)	452.0 (3.3)	461.8 (4.6)	13.7 (4.5)**	3.9
18:3	sat_340 (L)	55.2 (1.2)	59.1 (1.4)	58.9 (1.8)	-3.9 (1.9)*	1.9
	satt462 (L)	56.3 (1.1)	59.7 (1.3)	53.8 (2.6)	-3.4 (1.7)*	1.7
	satt652 (L)	55.4 (1.3)	59.8 (1.2)	55.3 (1.8)	-4.3 (1.8)*	2.4
18:3 / R8	satt652 (L)	55.9 (1.3)	59.5 (1.2)	55.3 (1.8)	-3.6 (1.8)*	1.6

*, **, *** Significant at the 0.05, 0.01 and 0.001 probability levels, respectively.

†AA, lines homozygous for N98-4445A allele; BB, lines homozygous for PI423893 allele; HH, lines heterozygous for N98-4445A and PI423893 alleles.

‡ R8, maturity (days after planting).

§Genotypic classes scored as dominant for N98-4445A allele are italicized. The mean and standard deviation for the dominant genotypic class corresponds to the pool of homozygous and heterozygous genotypes.

¶Genotypic classes scored as dominant for PI423893 allele are underlined. The mean and standard deviation for the dominant genotypic class corresponds to the pool of homozygous and heterozygous genotypes.

#Dominance effects estimated as the difference between means of the heterozygous genotypic class from the average of the two homozygous genotypic classes, its standard error, as well as its significance (in square brackets).

††Maturity used as a covariate during Single Factor Analysis.

Table 4. The means and standard errors (in parentheses) of homozygous genotypic classes for maturity (R8), oleate (18:1), linoleate (18:2) or linolenate (18:3) content, the difference between means, its standard error (in parentheses) and its significance, as well as the R^2 for the difference of the means for molecular markers significantly associated with maturity, oleate, linoleate and linolenate traits in FAE (N98-4445A x N97-3525) population.

Trait	Marker (linkage group)	AA† (g kg^{-1})	BB† (g kg^{-1})	Difference	R^2 (%)
R8‡	satt153 (O)	126.2 (0.6)	133.3 (0.6)	-7.1 (0.9)***	8.9
	satt652 (L)	129.4 (0.6)	131.4 (0.8)	-2.0 (1.0)*	0.6
18:1	sat_340 (L)	438.1 (4.4)	425.8 (4.3)	12.3 (6.1)*	0.9
	satt652 (L)	439.6 (3.0)	423.1 (4.6)	16.5 (5.5)**	1.4
	satt153 (O)	441.8 (3.6)	427.7 (3.7)	14.1 (5.1)**	1.2
	satt354 (I)	429.3 (3.8)	441.6 (3.9)	-12.3 (5.5)*	0.9
	satt462 (L)	439.7 (4.0)	427.7 (3.8)	12.0 (5.5)*	0.8
18:1 / R8§	sat_340 (L)	441.5 (3.5)	422.7 (3.4)	18.8 (4.9)***	2.2
	satt462 (L)	440.4 (3.2)	425.8 (3.0)	14.6 (4.4)**	1.2
	satt652 (L)	438.8 (2.4)	428.4 (3.6)	9.9 (4.4)*	0.5
	satt153 (O)	428.9 (3.0)	439.1 (3.0)	-10.2 (4.3)*	0.6
	satt354 (I)	430.5 (3.0)	440.4 (3.1)	-9.9 (4.3)*	0.6
18:2	satt462 (L)	417.8 (3.7)	428.7 (3.5)	-10.9 (5.1)*	0.8
	sat_340 (L)	417.9 (4.0)	431.0 (3.9)	-13.1 (5.6)*	1.3
	satt652 (L)	418.5 (2.8)	432.9 (4.2)	-14.3 (5.1)**	1.2
	satt153 (O)	416.9 (3.3)	429.1 (3.4)	-12.1 (4.7)**	1.0

Table 4. Continued.

Trait	Marker (linkage group)	AA† (g kg^{-1})	BB† (g kg^{-1})	Difference	R^2 (%)
18:2 / R8	satt462 (L)	417.3 (3.1)	430.4 (2.9)	-13.1 (4.2)**	1.2
	sat_340 (L)	414.9 (3.3)	433.8 (3.2)	-18.8 (4.6)***	2.6
	satt652 (L)	419.8 (2.3)	428.4 (3.5)	-8.6 (4.2)*	0.4
	satt153 (O)	428.1 (2.8)	419.4 (2.9)	8.7 (4.1)*	0.5
18:3	satt153 (O)	32.4 (0.2)	33.2 (0.2)	-0.9 (0.3)**	1.6
	satt354 (I)	33.3 (0.2)	32.2 (0.2)	1.1 (0.3)***	2.4
	sat_268 (I)	33.1 (0.2)	32.5 (0.2)	0.6 (0.3)*	0.9
18:3 / R8	satt354 (I)	33.2 (0.2)	32.3 (0.2)	0.9 (0.2)***	1.9
	satt462 (L)	32.6 (0.2)	33.2 (0.2)	-0.6 (0.3)*	0.6
	sat_268 (I)	33.1 (0.2)	32.5 (0.2)	0.6 (0.3)*	0.8
	satt652 (L)	33.1 (0.1)	32.4 (0.2)	0.7 (0.3)**	1.0

*, **, *** Significant at the 0.05, 0.01 and 0.001 probability levels, respectively.
†AA, lines homozygous for N98-4445A allele; BB, lines homozygous for N97-3525.
‡R8, maturity (days after planting).
§Maturity used as a covariate during Single Factor Analysis.

III. Heritability of Oleic Acid Content in Soybean Seed Oil and its Genetic Correlation with Fatty Acid and Agronomic Traits

Eleni Bachlava[1], Joseph W. Burton[1, 2], Cavell Brownie[3], Sanbao Wang[1], Jérôme Auclair[1] and Andrea J. Cardinal[1]

[1] Department of Crop Science, North Carolina State University, Raleigh, NC 27695-7620

[2] USDA-ARS, Soybean and Nitrogen Fixation Research Unit, Raleigh, NC 27607

[3] Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203

Formatted for: Crop Science

ABSTRACT

Oleate content is important for the nutritional value and oxidative stability of soybean seed oil. Response to selection for higher oleate content depends on its heritability in breeding population. Oleate content also depends on its genetic correlation to other fatty acids and agronomic traits. The objective of this study was to estimate heritability of oleate content and to determine the correlation of oleate with other fatty acids and agronomic traits in three soybean populations segregating for major and minor oleate genes grown in multiple environments. One of the populations consisted of 721 lines, providing excellent precision for estimation of the genetic parameters. Heritability for oleate content was sufficiently high that early generation selection can be effective when practiced on unreplicated lines grown at a single environment. Significant negative correlations were observed between oleate and linoleate, oleate and linolenate as well as oleate and palmitate in all three populations. Significant positive correlations were detected between palmitate and stearate in one population segregating for oleate genes and fap_{nc} and *fap1* alleles which reduce palmitate content. In the same population we also observed a significant negative correlation between yield and oleate content, and positive correlations between yield and linoleate, linolenate and palmitate contents.

INTRODUCTION

Soybean oil quality is determined by the relative composition of fatty acids in seed triacylglycerols. Oleic acid is the monounsaturated fatty acid in soybean seed oil with an average concentration of 230 g kg^{-1} (Wilson, 2004). An increase in oleate concentration is required for the improvement of the nutritional value and oxidative stability of soybean oil. The polyunsaturated fatty acids, linoleate and linolenate, are susceptible to oxidation; therefore, soybean oil is hydrogenated in order to prevent the development of unfavorable odors and flavors (Mounts et al., 1988). Hydrogenation leads to the synthesis of *trans* fatty acids in soybean oil that expose consumers to several health risks (Wilson et al., 2002). Breeding for higher oleate and lower linolenate content not only will diminish the need for oil hydrogenation, but will also facilitate the commercial success of soybean oil under the new guidelines of the U.S. Food and Drug Administration (Wilson et al., 2002).

Soybean lines with elevated oleic acid content have been developed. N78-2245 was derived through recurrent selection and contained 510 g kg^{-1} oleic acid and 42 g kg^{-1} linolenic acid (Wilson et al., 1981). The soybean lines N85-2124 and N85-2176 possessed 320 - 440 g kg^{-1} oleic acid and 33 – 40 g kg^{-1} linolenic acid (Burton et al., 1989). Recently, N98-4445A was released with a concentration of 549 g kg^{-1} oleic and 29 g kg^{-1} linolenic acid (Burton et al., 2006).

Two oleic acid mutant lines, M23 and M11, with oleate concentration of 461 g kg^{-1} and 359 g kg^{-1}, respectively, were developed by x-ray irradiation of the cultivar 'Bay', and allelism tests suggested that oleic acid content is monogenically inherited (Rahman et al., 1994; Rahman et al., 1998). The high oleate trait in M23 was inherited as a partially recessive *ol* allele; while, ol^a allele in M11 was completely dominant to the *ol* allele, and both alleles were located at the same locus (Takagi and Rahman, 1996; Rahman et al., 1998). Genetic variation in oleate content was also affected by modifier genes (Alt et al., 2005). Minor and major oleate quantitative trait loci (QTL) identified in soybean genome, explained approximately 4 to 28% of the observed variation for the trait (Diers and Shoemaker, 1992; H. R. Boerma, personal communication, 2007).

The *Fad2-1* and *Fad2-2* genes, coding for the microsomal ω-6 desaturase enzymes, which catalyze the desaturation of oleic to linoleic acid during fatty acid biosynthesis are also implicated in the genetic variation of oleate content. The *ol* allele for elevated oleic content in M23 was due to a deletion of the *Fad2-1* gene (Kinoshita et al., 1998). Genetic manipulation of *Fad2-1* gene has resulted in transgenic soybean lines with high oleic acid content (840 – 880 g kg^{-1}). The high oleate phenotype was stable across different environments and the yield of the transgenic plants was comparable with elite soybean lines (Kinney and Knowlton, 1998). Recently, two isoforms were reported for the *Fad2-1* gene and three isoforms for the *Fad2-2* gene (Tang et al., 2005; Schlueter et al., 2007). The isoforms vary in

expression for different plant tissues and stability at different temperatures during the period of oil deposition (Heppard et al., 1996; Tang et al., 2005; Schlueter et al., 2007). This demonstrates the complexity of oleate biosynthesis and the need to identify the specific genes responsible for oleate genotypic variation (Cardinal, 2008).

In addition to the quantitative nature of inheritance, breeders have to cope with the lack of stability of some oleate genotypes across environments. Some of the current mid-oleic breeding lines, such as N98-4445A and N97-3363 utilized in this study, are unstable across environments. Their oleate variation is largely attributed to temperature changes during the seed filling stage (Oliva et al., 2006; Wilson et al., 2002). Also, maternal effects may be partly responsible for the environmentally induced variation in oleate content (Carver et al., 1987).

Heritability estimates for the oleate trait were previously estimated from F_2 plants and $F_{2:3}$ lines derived from the cross of the mid-oleic mutant M23 and cultivar 'Archer', which has a normal oleate profile (Alt et al., 2005). Narrow-sense heritability estimates were 0.33 on an F_3 seed basis and 0.44 on an F_2 plant basis. Broad-sense heritability estimates for the 88 $F_{2:3}$ lines ranged from 0.37 to 0.46 among different environments on a plot basis whereas heritability on an entry mean basis was 0.82 (Alt et al., 2005). The realized oleate heritability estimated from four cycles of recurrent mass selection for high oleate content was 0.21 (Burton et al., 1983). However, there are no heritability estimates for oleate content in soybean populations segregating for major and minor oleate genes. Improved estimates could

be obtained by tests performed with larger samples in replicated tests across environments.

Beside the lack of oleate heritability estimates, correlations of oleate with other fatty acid and agronomic traits have not been studied in populations segregating for oleate content. Previous studies reported a negative correlation of oleate with linoleate and linolenate traits explained by the flux through the fatty acid biosynthesis pathway (Ohlrogge and Browse, 1995). This negative correlation was confirmed by indirect responses to recurrent mass selection for increased oleate content (Burton et al., 1983). Also, in the $F_{2:3}$ population of M23 x 'Archer', family means for oleate content were negatively correlated with palmitate (-0.47), stearate (-0.38), linoleate (-0.97) and linolenate (-0.45) traits (Alt et al., 2005). Estimates of the genotypic correlations of oleate content with other fatty acid traits have not been reported (Cardinal, 2008).

The objectives of this study were: (i) to estimate heritability of oleate content for three soybean breeding populations segregating for seed oleate level grown in replicated, multi-environment trials, in order to determine the response to selection for high oleate trait and the extent of phenotypic evaluations required for efficient selection; and (ii) to estimate the phenotypic and genotypic correlations between oleate and other fatty acid as well as agronomic traits in order to predict indirect responses to selection for high oleate content.

MATERIALS AND METHODS

Population Development

Three soybean populations, designated as FAE, FAF and FAS, were developed by single seed descent (Brim, 1966). The FAE population consisted of 721 F_5-derived lines from the cross of N98-4445A x 'Satelite', a low palmitate, low linolenate cultivar (Cardinal et al., 2007). The FAF population consisted of 118 F_5-derived lines from the cross of N97-3363-3 x PI423893. N97-3363-3 is a sister line of N98-4445A (Burton et al., 2006). FAS population consisted of 231 F_3- derived lines from the cross N98-4445A x PI423893. N98-4445A is a high oleate (563.1 g kg^{-1}), low linoleate line developed by the USDA-ARS in Raleigh, NC (Burton et al., 2006). PI423893 is a plant introduction with mid-oleate (305.8 g kg^{-1}) seed oil content (USDA-ARS National Plant Germplasm System, http://www.ars-grin.gov/npgs/searchgrin.html, verified 05/08/2007). The FAF and FAS populations were segregating for oleate genes and the reduced linolenate *fan*(PI123440) allele. The FAE population was segregating for oleate genes as well as for the fap_{nc} and *fap1* reduced palmitate alleles, but was fixed for the reduced linolenate *fan*(PI123440) allele.

Experimental Design

The FAE and FAF populations were planted separately in sets within replications experimental designs (Hallauer and Miranda, 1988) with two replications in each location. Fifteen seven-row by seven-column sets were randomly assigned in each replication, location, and year for the FAE population. Five five-row by five-column sets were randomly assigned in each replication, location, and year for the FAF population. Before lines in each population were randomly assigned to each set, they were divided into four groups in FAF population and five groups in FAE population according to maturity. Each maturity group was represented at least once in each set of the FAE and FAF populations, and the maturity groups that included the majority of the segregating lines were represented two to three times in each set. Once lines were assigned to a set, they remained in this set in all replications, locations, and years. Within each set, the maturity groups were randomly assigned to a column to facilitate mechanical harvest. Then, the lines were randomly assigned to plots within each column. The parental lines N97-3363-3 and PI423893 in FAF population and the parental line 'Satelite' in FAE population were randomly assigned to one set and, depending on their maturity date, to a maturity group within that set (rewrite) and finally to a plot within the set as explained previously. The N98-4445A parental line in FAE population was assigned to a maturity group in all the sets and randomly assigned to a plot within each set as previously explained. The FAE

population was planted in four 4.88 m rows with 0.97 m spacing between rows. Only the two middle-rows, end-trimmed to 3.96 m, were harvested at Clinton, Kinston and Plymouth, NC, in 2005 ($F_{5:7}$) and 2006 ($F_{5:8}$). The FAF population was grown at Clinton and Kinston, NC, in 2005 ($F_{5:7}$) and at Clinton, Kinston, Clayton and Plymouth, NC, in 2006 ($F_{5:8}$). The FAF population was planted in four-row plots, as previously described for FAE population, with the exception of Clayton in 2006 that was planted in single 3.048 m row plots.

For both FAE and FAF populations, phenotypic data were also collected from individual F_2 plants and $F_{5:6}$ lines. F_2 seeds descended from a common F_1 plant were planted in separate rows at Clayton in 2000 and $F_{2:3}$ seed was harvested from individual F_2 plants. The $F_{5:6}$ lines of the FAE and FAF populations were planted in two-row plots in a modified augmented design (Lin and Poushinsky, 1983) at Clayton in 2004. N98-4445A was used as the systematic check in the center of each main plot, which consisted of 21 subplots. Cultivars ‘Manokin’, ‘Williams’ and ‘Holladay’ were used as secondary checks in five randomly selected main plots within each of the five blocks.

The FAS population was planted in five-seed hills at Clayton in a sets within replications design with three replications in 2006. The order of sets was randomized in each replication, and lines within each set were randomly assigned to hills. The parental lines, N98-4445A and PI423893, were randomly assigned within each of the five sets in each replication.

Phenotypic Evaluation

Flowering and maturity dates were recorded for FAE and FAF populations at the R2 and R8 reproductive stages, respectively, as days after planting (Fehr and Caviness, 1977). Plant height was also measured for the FAF and FAE populations as the distance from the soil to the plant apex in cm. Yield and moisture content were evaluated only for the FAE population from the two middle rows of each plot after end-trimming and mechanical harvest. Yield was adjusted to 13% moisture content. Mature soybean seed was harvested mechanically from the two middle rows of each four-row plot or each one-row plot for the FAF population at Clayton in 2006. Approximately 10 g of seed were subsampled from each plot for the evaluation of oil content and fatty acid composition in the FAE and FAF populations. Only maturity at the R8 reproductive stage (Fehr and Caviness, 1977), recorded as days after planting, was evaluated for the FAS population. Oil content and fatty acid composition were estimated from approximately 5 g of seed, which were subsampled from the seed harvested from each hill of the FAS population. Seed oil content was determined by nuclear magnetic resonance (NMR) using a Maran pulsed NMR (Resonance Instruments, Witney, Oxfordshire, UK) and fatty acid composition was evaluated by gas liquid chromatography using a Model 6890 GC (Agilent Technologies, Inc., Wilmington, DE), as described by Burkey et al. (2007).

Statistical Analysis

Statistical analyses of the sets within replications designs for FAE, FAF and FAS populations were conducted with Proc MIXED in SAS 9.1 (SAS Institute, 2004). Environments, sets, replications, lines, as well as their interactions (Appendix 1), were considered random effects. Best linear unbiased predictors (BLUPs) were obtained for all traits and lines of the FAE, FAF and FAS populations (Appendix 2). The degrees of freedom for FAF and FAS populations were calculated using the Satterthwaite approximation. Degrees of freedom for FAE population were derived with the containment method in order to reduce the computing power required. The phenotypic data of the FAF population from Clinton, NC, in 2005 were discarded due to excess missing data and greater error variance compared to the other environments. The analysis of the FAS population for a single environment was performed as described above. Statistical analysis of the modified augmented design was conducted with Proc GLM in SAS 9.1 according to Scott and Milliken (1993).

Heritability, expressed as the regression of $F_{5:6}$ offspring on $F_{2:3}$ progenitors, was estimated with Proc REG in SAS 9.1 (Appendix 3). Heritability estimates of the F_3- and F_5- derived lines on a plot mean and an entry means basis, and their standard errors (Appendix 4), were obtained according to Holland et al. (2003). Genotypic and phenotypic correlations (Appendix 5) were estimated according to Holland (2006).

RESULTS AND DISCUSSION

Comparison of Genetic Variance and Heritability Estimates

Estimates of oleate genetic variation were consistently high for all three populations (Tables 1, 2 and 3) because FAF, FAS and FAE were segregating for several oleate genes. The genetic by environment and error variance components for oleate trait were at least fourfold smaller than the genetic variance components; however, their magnitude varied among populations.

Heritability estimates for oleate content on a plot mean basis were higher in the FAF than the FAS population (Tables 1 and 2), but lower in the FAE than the FAS population (Tables 1 and 3). According to Cockerham (1983), the numerator of the heritability estimator for the F_5- derived populations included an approximately 25% larger proportion of additive variance than the estimator for the F_3- derived population. However, the FAE population had a larger error variance component due to its larger size in comparison to the FAF and FAS populations. In the FAS population heritability estimates were confounded by the genotype x environment interaction since the population was grown in a single environment.

The high heritability estimates for oleate trait on a plot mean basis in the F_5-derived populations suggest that selection for oleate content can be effective if the number of environments studied is compromised. However, sampling more

replications and environments would improve oleate response to selection as demonstrated by the higher entry mean basis heritabilities (Tables 2 and 3). Also, the high oleate heritability estimates on a plot basis for the F_3- derived population, FAS, suggest that effective selection for oleate trait can be practiced in early generations. According to Bernardo (2003), the expected correlation between the phenotypic mean of F_3- derived lines and genotypic mean of the descendant homozygous lines approaches 0.80 when oleate heritability equals 0.76 (Table 1). Therefore, early generation selection for oleate trait should be effective when practiced on unreplicated F_3- derived lines.

The parent-offspring heritability estimates in both FAF and FAE populations were significantly smaller than the heritability estimates on a plot mean and on an entry mean basis (Tables 2 and 3) since the later estimators included an 87% larger proportion of additive variance (Cockerham, 1983). It should be noted that parent and progeny from each population were evaluated in different environments; therefore, the parent-offspring heritability estimator was not confounded by the genotype x environment covariance (Nyquist, 1991).

Heritability estimates for the FAF population were also obtained based on variance components estimated with maturity date (R8) or maturity date within environment included as covariates in the mixed models. Although maturity effects explained a significant part of oleate variation in all environments, oleate heritability estimates for the FAF population differed by less than one percentage point (data not

shown) when maturity effects were ignored. Thus, accounting for maturity effects was not considered necessary for the heritability estimates reported herein.

Heritability estimates for oleate content in all three populations were higher that those reported by Alt et al. (2005). Differences could be attributed to the segregation of oleate genes that vary in number, nature or genetic effects among populations, leading to differences in the genetic variance components. Furthermore, differences can be due to the genetic covariance components among relatives utilized for heritability estimation. Alt et al. (2005) reported heritability for F_2-derived lines, while estimators reported herein refer to F_3- and F_5-derived lines. According to Cockerham (1983), these heritability estimators include a 50% and an 87% larger proportion of additive variance, in the FAS population, and in the FAF and FAE populations, respectively, than in the study by Alt et al. (2005). Differences in the error variance can also explain part of this discrepancy. Oleate heritability in all three populations was also higher than the realized heritability estimate reported by Burton et al. (1983). The heritability estimators reported herein include a 100% and an approximately 150% larger proportion of additive variance in the FAS population and in the FAF and FAE populations, respectively, compared with the study by Burton et al. (1983). Differences in the error variance are also expected since in that study oleate was measured on single male-sterile plants.

The greatest differences in the estimates of variance components for fatty acid traits among the three populations were observed for palmitate and linolenate. The

FAE population (Table 3) had significantly higher genetic variance for palmitate and lower genetic variance for linolenate than the FAF and FAS populations (Tables 1 and 2). This is due to segregation of the FAF and FAS populations for the reduced linolenate *fan*(PI123440) allele. The FAE population segregated for the fap_{nc} and *fap1* reduced palmitate alleles and was fixed for the *fan*(PI123440) allele. This also explains the relatively higher linolenate heritability estimates and lower palmitate heritability estimates observed for the FAF and FAS populations. The heritability estimates for palmitate are in agreement with those reported by Cardinal and Burton (2007) for three different populations segregating for the *fap1*, fap_{nc} and *fan* alleles.

The variance components estimates for agronomic traits evaluated in the three populations were similar with the exception of maturity date, which differed among the three populations (Tables 1, 2 and 3) due to their different genetic backgrounds. The three populations were segregating for at least some different maturity genes since PI423893, the parental line used for the development of FAF and FAS populations, is a MGIII line, and 'Satelite', the parental line of FAE population, is a MGVI line. It should also be noted that in the FAF and FAE populations (Tables 2 and 3) height, maturity and flowering dates have high heritability estimates both on a plot and an entry mean basis. However, the heritability estimate for yield in the FAE population (Table 3) was relatively low on a plot basis, indicating that multi-environment trials are essential when selecting for yield. Low heritability estimates for yield have been previously reported (Burton, 1987).

Comparison of Genotypic and Phenotypic Correlations among Fatty Acid Traits

In all three populations, significant negative genotypic and phenotypic correlations were observed between oleate and linoleate, as well as oleate and linolenate traits (Tables 4, 5 and 6). These results are in agreement with the previous studies (Alt et al., 2005) and with the current knowledge of fatty acid biosynthesis (Ohlrogge and Browse, 1995). These results are also strengthened by the outcome of recurrent mass selection for high oleate content, which led to the increase of mean oleate content from 25 to 33% and the correlated decrease of linolenate content from 8 to 6% and linoleate content from 53 to 47% (Burton at al. 1983). Although the magnitude of the correlations reported from different studies differs due to the diverse genetic background of the populations and the different methods used for their estimation, the strongest correlation was consistently observed between oleate and linoleate.

Oleate was negatively correlated with palmitate in all three populations, but oleate was significantly correlated with stearate only in the FAE population (Table 6). Pearson's correlation estimates of oleate with palmitate and stearate, reported by Alt et al. (2005), were both negative. Significant negative correlations between oleate and palmitate were also reported by Rebetzke et al. (1998) and Rebetzke et al. (2001). Rebetzke et al. (1996) showed that oleate was negatively correlated with palmitate, but not significantly correlated with stearate, in a study evaluating the stability of

palmitate and stearate content in 30 soybean lines with early and late maturity profiles derived from the base, 3rd and 6th cycles of recurrent selection for high oleate content. However, recurrent mass selection for high oleate content (Burton at al., 1983) resulted in no change in palmitate and stearate contents. Cardinal and Burton (2007) detected negative correlations between oleate and palmitate, as well as between oleate and stearate, only in one of three F_4- derived populations segregating for the *fap*$_{nc}$, *fap1* and *fan* alleles. Also, Primomo et al. (2002) reported significant negative correlations between oleate and palmitate only in one of six F_2 soybean populations segregating for different palmitate and linolenate alleles.

In the FAE population, significant positive genotypic and phenotypic correlations were detected between palmitate and stearate (Table 6). This observation is in agreement with Cardinal and Burton (2007), where three F_4-derived populations, segregating for the *fap*$_{nc}$, *fap1* and *fan* alleles, revealed significant positive correlations between palmitate and stearate contents. Cardinal and Burton (2007) suggested that the *FATB1a* gene, which encodes for the palmitate thioesterase in soybean and is deleted in the *fap*$_{nc}$ allele, is putatively active toward oleoyl acyl carrier protein (ACP) substrates and, therefore, provides a possible explanation for the positive correlation between palmitate and stearate.

Comparison of Genotypic and Phenotypic Correlations between Fatty Acid and Agronomic Traits

Significant positive genotypic and phenotypic correlations were observed between oleate and maturity date for both FAF and FAS populations (Tables 4 and 5), while significant negative genotypic and phenotypic correlations were found between oleate and maturity date, and between oleate and flowering date for the FAE population (Table 6). As expected, the correlations of linoleate and linolenate with flowering and maturity dates were all significant in the opposite direction in each population. These correlations can be interpreted by the maturity profiles of the three populations. Lines of the FAF and FAS populations matured earlier on average (Tables 1 and 2) than lines of the FAE population (Table 3), since the parental line used for their development belongs to an earlier maturity group. Early maturing lines of the FAF and FAS populations had reduced oleate content due to the high temperatures during oil deposition, which exceeded the 30ºC; while, late maturing lines of the FAE population had reduced oleate content due to low temperatures during oil deposition (data not shown).

Oil content appeared to have significant negative genotypic and phenotypic correlations with oleate and maturity date for the FAF population as well as with flowering date for the FAS population (Tables 4 and 5). In the FAE population, oil content was negatively correlated with both flowering and maturity dates (Table 6).

However, for the FAE population no correlation was observed between oil and oleate, but oil was negatively correlated with linolenate. The negative correlations of flowering and maturity dates with oil content can be explained since early flowering and maturity leads to higher temperatures during the period of oil deposition, which results in elevated oil content (Martin et al., 1986; Wolf et al., 1982).

Significant negative phenotypic and genotypic correlations between yield and oleate (Table 6) were shown in the FAE population. As expected, yield was positively correlated, both phenotypically and genotypically, with linoleate and linolenate. Since yield was significantly positively correlated with flowering and maturity dates, there is no reason to assume that the negative correlation between yield and oleate is due to maturity effects. Positive correlations between yield and maturity were also reported by Cherrak et al. (2003) for a population segregating for palmitate and linolenate genes. Furthermore, significant positive phenotypic and genotypic correlations were observed between yield and palmitate, which is in agreement with the study of three F_4-derived populations, segregating for the *fap*$_{nc}$, *fap1* and *fan* alleles, reported by Cardinal and Burton (2007). Cardinal and Burton (2007) suggested that the significant positive correlation between yield and palmitate is due to the genetic effect of the major palmitate alleles *fap*$_{nc}$ and *fap1* on yield. It should be noted that the FAE population is also segregating for the major palmitate alleles *fap*$_{nc}$ and *fap1*.

CONCLUSIONS

The oleate heritability estimates reported herein suggest that effective selection can be practiced on unreplicated F_3-derived lines grown at a single environment. Therefore, early generation selection can be conducted with limited multi-environment testing. Results of this study indicate that selection for higher oleate content will result in lower linoleate, linolenate and palmitate content in soybean oil. The significant negative correlation between oleate content and yield also suggests that the development of high oleate germplasm, which is currently on demand due to its nutritional value and oxidative stability, may be hindered by lower yields.

ACKNOWLEDGEMENTS

The authors would like to thank William Novitzky at the USDA-ARS Soybean and Nitrogen Fixation Research Unit, Raleigh, NC, for providing training and equipment for the fatty acid analysis, and James B. Holland for the review of this manuscript. This project was funded by the USDA Cooperative State Research, Education and Extension Service National Research Initiative, Plant Genome Program (award number 2003-35300-13190), and by the United Soybean Board.

REFERENCES

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

Bernardo, R. 2003. On the effectiveness of early generation seletion in self-pollinated crops. Crop Sci. 43:1558-1560.

Brim, C.A. 1996. A modified pedigree method of selection in soybean. Crop Sci. 6:220.

Burkey, K.O., F.L. Booker, W.A. Pursley, and A.S. Heagle. 2007. Elevated carbon dioxide and ozone effects on peanut: II. Seed yield and quality. Crop Sci. 47:1488-1497.

Burton, J.W. 1987. Quantitative genetics: Results relevant to soybean breeding, p. 211-247, *In* J. R. Wilcox, ed. Soybeans: Improvement, Production, and Uses, 2nd ed. American Society of Agronomy, Inc., Madison, WI.

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, C.A. Brim, and R.W. Rinne. 1989. Registration of soybean germplasm lines with modified fatty acid composition of seed oil. Crop Sci. 29:1583-1583.

Burton, J.W., R.F. Wilson, G.J. Rebetzke, and V.R. Pantalone. 2006. Registration of N98-4445A mid-oleic soybean germplasm line. Crop Sci. 46:1010-1012.

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. 47:304-310.

Cardinal, A.J. 2008. Molecular genetics and breeding for fatty acid manipulation in soybean. Plant Breed. Rev. 30 (in press).

Cardinal, A.J., and J.W. Burton. 2007. Correlations between palmitate content and agronomic traits in soybean populations segregating for the *fap1*, fap_{nc} and *fan* alleles. Crop Sci. 47:1804-1812.

Carver, B.F., J.W. Burton, and R.F. Wilson. 1987. Graft-transmissible influence on fatty acid composition of soybean seed. Crop Sci. 27:53-56.

Cherrak, C.M., V.R. Pantalone, E.J. Meyer, D.L. Ellis, S.L. Melton, D.R. West, and J.R. Mount. 2003. Low-palmitic, low-linolenic soybean development. J. Am. Chem. Soc. 80:539-543.

Cockerham, C.C. 1983. Covariances of relatives from self-fertilization. Crop Sci. 23:1177-1180.

Diers, B.W., and R.C. Shoemaker. 1992. Restriction fragment length polymorphism analysis of soybean fatty acid content. J. Am. Chem. Soc. 69:1242-1244.

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

Hallauer, A.R., and J.B. Miranda. 1988. Quantitative genetics in maize breeding. 2nd ed. Iowa State University Press, Ames, IA.

Hawkins, S.E., W.R. Fehr, and E.G. Hammond. 1983. Resource allocation in breeding for fatty acid composition of soybean oil. Crop Sci. 23:900-904.

Heppard, E.P., A.J. Kinney, K.L. Stecca, and G. Miao. 1996. Developmental and growth temperature regulation of two different microsomal omega-6 desaturase genes in soybeans. Plant Physiol. 110:311-319.

Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martinez. 2003. Estimating and interpreting heritability for plant breeding: An update. Plant Breed. Rev. 22:9-112.

Holland, J.B. 2006. Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation with SAS Proc MIXED. Crop Sci. 46:642-654.

Kinney, A.J., and S. Knowlton. 1998. Designer oils: the high oleic acid soybean, p. p. 193-213, *In* S. H. S. Roller, ed. Genetic Modification in the Food Industry. Chapman and Hall, London.

Kinoshita, T., S.M. Rahman, T. Anai, and Y. Takagi. 1998. Genetic analysis of RFLP on fatty acid synthesis in soybean mutants and their progenies: II. High oleic acid mutants with two microsomal omega-6 fatty acid desaturase cDNAs as probes. Bull. Fac. Agric., Saga Univ. 83:37-42.

Lin, C.S., and G. Poushinsky. 1983. A modified augmented design for an early stage of plant selection involving a large number of test lines without replication. Biometrics 39:553-561.

Martin, B.A., R.F. Wilson, and R.W. Rinne. 1986. Temperature effects upon the expression of a high oleic acid trait in soybean. J. Am. Oil Chem. Soc. 63:346-352.

Mounts, T.L., K. Warner, G.R. List, R. Kleiman, W.R. Fehr, E.G. Hammond, and 88. Effect of altered fatty acid composition on soybean oil stability. J. Am. Chem. Soc. 65:624-628.

Nyquist, W.E. 1991. Estimation of heritability and prediction of selection response in plant populations. Crit. Rev. Plant Sci. 10:235-322.

Ohlrogge, J., and J. Browse. 1995. Lipid biosynthesis. Plant Cell 7:957-970.

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.

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.

Rahman, S.M., Y. Takagi, K. Miyamoto, and T. Kawakita. 1994. Inheritance of low linolenic acid content in soybean mutant line M-5. Breed. Sci. 44:379-382.

Rahman, S.M., T. Kinoshita, T. Anai, S. Arima, and Y. Takagi. 1998. Genetic relationships of soybean mutants for different linolenic acid contents. Crop Sci. 38:702-706.

Rebetzke, G.J., V.R. Pantalone, J.W. Burton, B.F. Carver, and R.F. Wilson. 1996. Phenotypic variation for saturated fatty acid content in soybean. Euphytica. 91:289-295.

Rebetzke, G.J., J.W. Burton, T.E. Carter, and R.F. Wilson. 1998. Changes in agronomic and seed characteristics with selection for reduced palmitic acid content in soybean. Crop Sci. 38:297-302.

Rebetzke, G.J., V.R. Pantalone, J.W. Burton, T.E. Carter, and R.F. Wilson. 2001. Genetic background and environment influence palmitate content of soybean seed oil. Crop Sci. 41:1731-1736.

SAS Institute. 2004. SAS/STAT software. Release 9.1.3. SAS Inst., Cary, NC.

Schlueter, J.A., I.F. Vasylenko-Sanders, S. Deshpande, J. Yi, M. Siegfried, B.A. Roe, S.D. Schlueter, B.E. Scheffler, and R.C. Shoemaker. 2007. The *FAD2* gene family of soybean: insights into the structural and functional divergence of a paleopolyploid genome. The Plant Genome, a Suppl. to Crop Sci. 47:S14-S26.

Scott, R.A., and G.A. Milliken. 1993. A SAS program for analyzing augmented randomized complete block designs. Crop Sci. 33:865-867.

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

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

Wilson, R.F., J.W. Burton, and C.A. Brim. 1981. Progress in the selection for altered fatty acid composition in soybeans. Crop Sci. 21:788-791.

Wilson, R.F., J.W. Burton, V.R. Pantalone, and R.E. Dewey. 2002. New gene combinations governing saturated and unsaturated fatty acid composition in soybean p. 95-113, *In* T. M. Kuo and H. W. Gardner, eds. Lipid Biotechnology. Marcel Dekker, Inc., IL.

Wilson, R.F. 2004. Seed composition, p. 621-669, *In* H. R. Boerma and J. E. Specht, eds. Soybeans: Improvement, Production, and Uses, 3rd ed. American Society of Agronomy, Inc., Madison, WI.

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.

Table 1. Estimates of overall mean, range of best linear unbiased predictors (BLUPs) and standard errors (in parentheses) of the difference between the BLUPs for two random lines, estimates and standard errors (in parentheses) of genetic and error variance components, as well as heritability (h^2) estimates on a plot mean and an entry mean basis, and their standard errors (in parentheses), for palmitate (16:0), stearate (18:0), oleate (18:1), linoleate (18:2), linolenate (18:3), oil, and maturity date (R8) in the FAS population (N98-4445A x PI423893).

	16:0 (g kg^{-1})	**18:0 (g kg^{-1})**	**18:1 (g kg^{-1})**	**18:2 (g kg^{-1})**	**18:3 (g kg^{-1})**	**Oil (g kg^{-1})**	**R8 (DAP)†**
Mean of BLUPs	108.6	39.1	335.1	459.9	57.3	183.3	101.2
Range of BLUPs	89.0-124.3	32.4-49.2	253.3-457.3	355.3-520.1	35.5-86.2	165.1-203.8	92.6-110.9
SE of BLUP differences‡	4.13	2.21	18.65	15.27	3.82	5.11	3.14
Var(g)§ (SE)	44.8 (0.52)	10.4 (0.13)	1496.7 (15.93)	1047.5 (11.03)	160.2 (1.57)	74.8 (0.87)	2119.3 (26.29)
Var(e)§ (SE)	23.6 (0.18)	7.25 (0.05)	471.4 (3.54)	302.7 (2.27)	13.6 (0.10)	44.9 (0.32)	1673.0 (11.95)
h^2 plot mean basis (SE)	0.65 (0.033)	0.59 (0.038)	0.76 (0.025)	0.78 (0.023)	0.92 (0.009)	0.63 (0.034)	0.56 (0.038)
h^2 entry mean basis (SE)	0.82 (0.022)	0.77 (0.027)	0.88 (0.014)	0.89 (0.013)	0.97 (0.004)	0.81 (0.022)	0.76 (0.028)

†Days after planting

‡The standard error of BLUPs difference was calculated as the weighted average standard error of the difference between the BLUPs of two random lines from the same set and two random lines from different sets

§Genetic variance [Var(g)] and error variance [Var(e)] components

Table 2. Estimates of overall mean, range of best linear unbiased predictors (BLUPs) and standard errors (in parentheses) of the difference between the BLUPs for two random lines, estimates and standard errors (in parentheses) of genetic, genetic by environment, and error variance components, as well as estimates of heritability (h^2) and their standard errors (in parentheses) on a parent-offspring regression, plot mean and entry mean basis for palmitate (16:0), stearate (18:0), oleate (18:1), linoleate (18:2), linolenate (18:3), flowering date (R2), maturity date (R8), oil, and height in the FAF population (N97-3363-3 x PI423893).

	16:0 (g kg^{-1})	**18:0 (g kg^{-1})**	**18:1 (g kg^{-1})**	**18:2 (g kg^{-1})**	**18:3 (g kg^{-1})**	**R2 (DAP)†**	**R8 (DAP)†**	**Oil (g kg^{-1})**	**Height (cm)**
Mean of BLUPs	103.8	29.4	410.6	410.7	45.6	47.0	114.3	191.3	67.4
Range of BLUPs	84.9-122.7	24.9-33.8	301.2-551.4	297.7-519.0	26.0-74.4	39.8-51.0	103.6-122.7	170.2-207.0	29.8-97.5
SE of BLUP differences‡	2.78	1.04	16.78	14.77	2.66	2.34	2.00	2.85	4.36
Var(g)§ (SE)	75.1 (1.07)	3.0 (0.01)	3558.2 (49.37)	2708.3 (37.58)	109.5 (1.52)	6.8 (0.99)	21.3 (3.16)	60.0 (0.86)	323.4 (44.65)
Var(gxe)§ (SE)	2.8 (0.13)	1.2 (0.02)	333.6 (4.64)	275.1 (3.57)	4.9 (0.07)	1.2 (0.15)	5.5 (0.69)	10.6 (0.14)	26.1 (3.17)
Var(e)§ (SE)	24.1 (0.15)	2.3 (0.05)	560.7 (3.56)	402.8 (2.56)	7.4 (0.05)	1.7 (0.11)	7.2 (0.46)	16.1 (0.10)	32.4 (2.07)
h^2 parent-offspring (SE)	0.22 (0.12)	0.17 (0.09)	0.25 (0.14)	0.26 (0.14)	0.43 (0.13)	-	-	-	-
h^2 plot mean basis (SE)	0.74 (0.029)	0.46 (0.042)	0.80 (0.024)	0.80 (0.024)	0.90 (0.013)	0.70 (0.032)	0.63 (0.037)	0.69 (0.033)	0.85 (0.019)

Table 2. Continued.

	16:0 (g kg^{-1})	**18:0 (g kg^{-1})**	**18:1 (g kg^{-1})**	**18:2 (g kg^{-1})**	**18:3 (g kg^{-1})**	**R2 (DAP)†**	**R8 (DAP)†**	**Oil (g kg^{-1})**	**Height (cm)**
h^2 entry mean basis (SE)	0.96 (0.006)	0.86 (0.022)	0.96 (0.005)	0.96 (0.005)	0.98 (0.002)	0.94 (0.009)	0.92 (0.013)	0.94 (0.009)	0.97 (0.004)

†Days after planting

‡The standard error of BLUPs difference was calculated as the weighted average standard error of the difference between the BLUPs of two random lines from the same set and two random lines from different sets

§Genetic variance [Var(g)], genetic by environment [Var(gxe)] and error variance [Var(e)] components

Table 3. Estimates of overall mean, range of best linear unbiased predictors (BLUPs) and standard errors (in parentheses) of the difference between the BLUPs for two random lines, estimates and standard errors (in parentheses) of genetic, genetic by environment, and error variance components, as well as estimates of heritability (h^2) and their standard errors (in parentheses) on a parent-offspring regression, plot mean and entry mean basis for palmitate (16:0), stearate (18:0), oleate (18:1), linoleate (18:2), linolenate (18:3), flowering date (R2), maturity date (R8), oil, yield, and height in the FAE population (N98-4445A x 'Satelite').

	16:0 (g kg^{-1})	**18:0 (g kg^{-1})**	**18:1 (g kg^{-1})**	**18:2 (g kg^{-1})**	**18:3 (g kg^{-1})**	**Oil (g kg^{-1})**	**R2 (DAP)†**	**R8 (DAP)†**	**Yield (kg ha^{-1})**	**Height (cm)**
Mean of BLUPs	72.9	36.9	436.8	420.7	32.7	191.6	52.8	129.7	1452.2	93.5
Range of BLUPs	33.6-107.1	26.8-48.5	266.1-631.9	235.0-572.3	24.3-66.4	168.2-218.2	42.8-65.6	106.5-155.5	560.8-2125.7	34.0-159.7
SE of BLUP differences‡	2.41	1.50	21.44	19.01	1.24	3.43	1.00	2.24	176.03	6.2
Var(g)§ (SE)	267.0 (1.42)	17.6 (0.10)	4280.9 (24.05)	3673.8 (20.38)	11.8 (0.07)	65.5 (0.38)	25.5 (1.37)	137.4 (7.38)	76997 (5067)	482 (26.3)
Var(gxe)§ (SE)	7.6 (0.04)	4.3 (0.02)	808.4 (3.54)	622.2 (2.81)	1.9 (0.01)	22.1 (0.09)	0.9 (0.07)	6.4 (0.36)	60440 (2821)	45 (3.9)
Var(e)§ (SE)	15.5 (0.04)	4.6 (0.01)	1078.2 (2.43)	888.5 (2.00)	4.8 (0.01)	28.0 (0.06)	3.4 (0.08)	11.0 (0.29)	94574 (2104)	80 (3.3)
h^2 parent-offspring (SE)	0.38 (0.022)	0.44 (0.038)	0.22 (0.018)	0.22 (0.018)	0.27 (0.030)	-	-	-	-	-
h^2 plot mean basis (SE)	0.92 (0.004)	0.66 (0.013)	0.69 (0.013)	0.71 (0.012)	0.64 (0.014)	0.57 (0.015)	0.86 (0.007)	0.89 (0.006)	0.33 (0.016)	0.79 (0.010)

Table 3. Continued.

	16:0 (g kg^{-1})	**18:0 (g kg^{-1})**	**18:1 (g kg^{-1})**	**18:2 (g kg^{-1})**	**18:3 (g kg^{-1})**	**Oil (g kg^{-1})**	**R2 (DAP)†**	**R8 (DAP)†**	**Yield (kg ha^{-1})**	**Height (cm)**
h^2 entry mean basis (SE)	0.99 (0.001)	0.94 (0.003)	0.95 (0.003)	0.95 (0.003)	0.94 (0.003)	0.92 (0.005)	0.98 (0.001)	0.98 (0.001)	0.81 (0.011)	0.96 (0.002)

†Days after planting
‡The standard error of BLUPs difference was calculated as the weighted average standard error of the difference between the BLUPs of two random lines from the same set and two random lines from different sets.
§Genetic variance [Var(g)], genetic by environment [Var(gxe)] and error variance [Var(e)] components

Table 4. Genotypic (above diagonal) and phenotypic (below diagonal) correlation coefficients, and their standard errors, among palmitate (16:0), stearate (18:0), oleate (18:1), linoleate (18:2), linolenate (18:3), maturity date (R8), oil, and yield in the FAS population (N98-4445A x PI423893).

	16:0 (g kg^{-1})	**18:0 (g kg^{-1})**	**18:1 (g kg^{-1})**	**18:2 (g kg^{-1})**	**18:3 (g kg^{-1})**	**R8 (DAP)†**	**Oil (g kg^{-1})**
16:0 (g kg^{-1})		-0.010 (0.085)	-0.800‡ (0.031)	0.659‡ (0.047)	0.248‡ (0.071)	-0.414‡ (0.072)	0.320‡ (0.074)
18:0 (g kg^{-1})	0.023 (0.053)		-0.029 (0.081)	-0.044 (0.081)	-0.049 (0.077)	0.538‡ (0.067)	-0.300‡ (0.077)
18:1 (g kg^{-1})	-0.727‡ (0.025)	-0.096 (0.055)		-0.917‡ (0.012)	-0.292‡ (0.066)	0.199‡ (0.079)	-0.495‡ (0.062)
18:2 (g kg^{-1})	0.574‡ (0.037)	0.010 (0.055)	-0.926‡ (0.009)		-0.084 (0.072)	-0.128 (0.080)	0.541‡ (0.058)
18:3 (g kg^{-1})	0.219‡ (0.056)	-0.034 (0.057)	-0.308‡ (0.055)	-0.030 (0.061)		-0.200‡ (0.076)	0.031 (0.076)
R8 (DAP)†	-0.328‡ (0.046)	0.409‡ (0.042)	0.114‡ (0.053)	-0.064 (0.054)	-0.125‡ (0.055)		-0.451‡ (0.068)
Oil (g kg^{-1})	0.290‡ (0.049)	-0.290‡ (0.048)	-0.389‡ (0.046)	0.427‡ (0.044)	0.022 (0.058)	-0.489‡ (0.040)	

†Days after planting

‡Correlation coefficient significantly different from zero, when | correlation | > 1.96 * SE

Table 5. Genotypic (above diagonal) and phenotypic (below diagonal) correlation coefficients, and their standard errors, among palmitate (16:0), stearate (18:0), oleate (18:1), linoleate (18:2), linolenate (18:3), flowering date (R2), maturity date (R8), oil, and height in the FAF population (N97-3363-3 x PI423893).

	16:0 (g kg^{-1})	**18:0 (g kg^{-1})**	**18:1 (g kg^{-1})**	**18:2 (g kg^{-1})**	**18:3 (g kg^{-1})**	**R2 (DAP)†**	**R8 (DAP)†**	**Oil (g kg^{-1})**	**Height (cm)**
16:0 (g kg^{-1})		0.110 (0.105)	-0.565‡ (0.068)	0.401‡ (0.084)	0.380‡ (0.085)	0.002 (0.102)	-0.222‡ (0.099)	-0.054 (0.101)	0.011 (0.100)
18:0 (g kg^{-1})	0.110 (0.062)		0.111 (0.105)	-0.166 (0.103)	-0.063 (0.104)	0.250‡ (0.101)	0.390‡ (0.096)	0.111 (0.106)	0.491‡ (0.083)
18:1 (g kg^{-1})	-0.539‡ (0.054)	0.061 (0.065)		-0.968‡ (0.006)	-0.422‡ (0.081)	0.185 (0.098)	0.405‡ (0.087)	-0.436‡ (0.032)	-0.003 (0.100)
18:2 (g kg^{-1})	0.376‡ (0.066)	-0.124 (0.064)	-0.970‡ (0.005)		0.222‡ (0.094)	-0.181 (0.098)	-0.373‡ (0.090)	0.489‡ (0.078)	0.007 (0.100)
18:3 (g kg^{-1})	0.335‡ (0.071)	-0.042 (0.069)	-0.440‡ (0.068)	0.259‡ (0.079)		-0.194‡ (0.096)	-0.335‡ (0.092)	0.083 (0.099)	-0.114 (0.098)
R2 (DAP)†	-0.014 (0.073)	0.119 (0.061)	0.147‡ (0.074)	-0.142 (0.074)	-0.155‡ (0.077)		0.759‡ (0.047)	-0.243‡ (0.096)	0.478‡ (0.079)
R8 (DAP)†	-0.193‡ (0.069)	0.207‡ (0.058)	0.334‡ (0.065)	-0.306‡ (0.066)	-0.270‡ (0.071)	0.572‡ (0.043)		-0.201‡ (0.100)	0.447‡ (0.084)
Oil (g kg^{-1})	-0.036 (0.073)	0.002 (0.063)	-0.359‡ (0.019)	0.406‡ (0.062)	0.064 (0.079)	-0.147‡ (0.070)	-0.197‡ (0.067)		0.160 (0.098)
Height (cm)	0.009 (0.079)	0.296‡ (0.059)	-0.007 (0.082)	0.009 (0.083)	-0.085 (0.086)	0.370‡ (0.065)	0.317‡ (0.067)	0.136 (0.076)	

†Days after planting
‡Correlation coefficient significantly different from zero, when | correlation | > 1.96 * SE

Table 6. Genotypic (above diagonal) and phenotypic (below diagonal) correlation coefficients, and their standard errors, among palmitate (16:0), stearate (18:0), oleate (18:1), linoleate (18:2), linolenate (18:3), flowering date (R2), maturity date (R8), oil, yield, and height in the FAE population (N98-4445A x 'Satelite').

	16:0 (g kg^{-1})	**18:0 (g kg^{-1})**	**18:1 (g kg^{-1})**	**18:2 (g kg^{-1})**	**18:3 (g kg^{-1})**	**Oil (g kg^{-1})**	**R2 (DAP)†**	**R8 (DAP)†**	**Yield (kg ha^{-1})**	**Height (cm)**
16:0 (g kg^{-1})		0.301‡ (0.036)	-0.191‡ (0.037)	-0.085 (0.039)	0.053 (0.039)	-0.016 (0.040)	0.061 (0.038)	0.081‡ (0.038)	0.457‡ (0.034)	0.274‡ (0.036)
18:0 (g kg^{-1})	0.265‡ (0.028)		-0.454‡ (0.032)	0.322‡ (0.036)	0.305‡ (0.036)	-0.090‡ (0.040)	0.509‡ (0.029)	0.580‡ (0.027)	0.196‡ (0.042)	0.517‡ (0.030)
18:1 (g kg^{-1})	-0.236‡ (0.030)	-0.399‡ (0.023)		-0.959‡ (0.003)	-0.660‡ (0.023)	-0.013 (0.040)	-0.589‡ (0.026)	-0.636‡ (0.024)	-0.507‡ (0.034)	-0.325‡ (0.036)
18:2 (g kg^{-1})	-0.001 (0.031)	0.286‡ (0.025)	-0.968‡ (0.002)		0.621‡ (0.025)	0.044 (0.040)	0.552‡ (0.027)	0.594‡ (0.026)	0.393‡ (0.037)	0.225‡ (0.037)
18:3 (g kg^{-1})	0.083‡ (0.030)	0.250‡ (0.024)	-0.627‡ (0.016)	0.586‡ (0.017)		-0.358‡ (0.036)	0.538‡ (0.028)	0.536‡ (0.028)	0.321‡ (0.039)	0.291‡ (0.036)
Oil (g kg^{-1})	0.004 (0.029)	-0.128‡ (0.025)	-0.032 (0.026)	0.061‡ (0.026)	-0.304‡ (0.022)		-0.392‡ (0.034)	-0.397‡ (0.034)	0.125‡ (0.043)	-0.302‡ (0.037)
R2 (DAP)†	0.060 (0.034)	0.387‡ (0.024)	-0.469‡ (0.022)	0.445‡ (0.023)	0.403‡ (0.023)	-0.271‡ (0.025)		0.920‡ (0.006)	0.208‡ (0.041)	0.659‡ (0.022)
R8 (DAP)†	0.069‡ (0.034)	0.426‡ (0.023)	-0.490‡ (0.022)	0.464‡ (0.023)	0.386‡ (0.024)	-0.291‡ (0.025)	0.824‡ (0.009)		0.190‡ (0.041)	0.708‡ (0.020)
Yield (kg ha^{-1})	0.244‡ (0.021)	0.005 (0.021)	-0.218‡ (0.020)	0.173‡ (0.021)	0.155‡ (0.020)	0.110‡ (0.020)	0.092‡ (0.022)	0.107‡ (0.023)		0.156‡ (0.041)
Height (cm)	0.238‡ (0.031)	0.378‡ (0.024)	-0.239‡ (0.027)	0.165‡ (0.029)	0.210‡ (0.027)	-0.190‡ (0.026)	0.552‡ (0.021)	0.598‡ (0.019)	0.136‡ (0.022)	

Table 6. Continued.

†Days after planting
‡Correlation coefficient significantly different from zero, when | correlation | > 1.96 * SE

IV. Mapping and Comparison of Quantitative Trait Loci for Oleic Acid Content in Two Segregating Soybean Populations

Eleni Bachlava[1], Ralph E. Dewey[1], Joseph W. Burton[1, 2] and Andrea J. Cardinal[1]

[1] Department of Crop Science, North Carolina State University, Raleigh, NC 27695-7620

[2] USDA-ARS, Soybean and Nitrogen Fixation Research Unit, Raleigh, NC 27607

Formatted for: Crop Science

ABSTRACT

Soybean produces 29.4% of the world's edible vegetable oil. An important determinant of the nutritional value and the oxidative stability of soybean oil is the oleic acid content. Elevation of the oleate content levels leads to the improvement of soybean oil quality. However, the lack of knowledge on the genetic factors underlying oleate variation in soybean seeds hampers the use of marker assisted selection in soybean breeding programs. The objective of this study was to identify oleate quantitative trait loci (QTLs) in two soybean populations segregating for oleic acid content, and to determine whether the detected QTLs cosegregate with the isoforms of the *FAD2-1* and *FAD2-2* genes, which encode the oleate desaturase activity in the endoplasmic reticulum. This study revealed an oleate QTL with moderate effects on linkage group F in the proximity of the simple sequence repeat marker sat_309, which was confirmed in both populations across all environments tested. Furthermore, this study reported an oleate QTL with moderate effects in the proximity of *FAD2-1B* isoform on linkage group I, which interacted epistatically with the oleate QTL on linkage group F. Oleate QTLs with moderate effects were also detected on linkage groups A2 and N only in one of the populations under study. Minor QTLs on linkage groups E, L, A1 and D2 confirmed previous mapping studies for oleate content in soybean.

INTRODUCTION

Soybean [*Glycine max* (L.) Merr.] is the leading oilseed crop in the United States in terms of gross vegetable oil production and economic importance (Wilcox, 2004). In order to retain its competitive position in domestic and global markets, the oilseed industry has to abide by the guidelines of the United States Food and Drug Administration, which raise the need for lowering the levels of *trans* isomers found in food products (Wilson et al., 2002). *Trans* isomers are formed during hydrogenation, a process which enhances the long-term oxidative stability of soybean oil. The need for oil hydrogenation can be diminished by reducing the levels of the polyunsaturated linoleic and linolenic acids, which are susceptible to oxidation through the concomitant elevation of the levels of the oxidatively stable monounsaturated oleic acid (Wilson, 2004). Breeders have successfully incorporated the high oleate and low linolenate traits into soybean germplasm through conventional breeding (Wilson et al., 1981; Burton et al., 1989; Burton et al., 2006) and genetic engineering (Kinney, 1995; Kinney and Knowlton, 1998; Buhr et al., 2002). However, the majority of the genetic factors that cause the observed variation in oleate and linoleate contents remain unknown.

Previous studies reported quantitative trait loci (QTLs) controlling the unsaturated fatty acid content in soybean. Currently, 19 QTLs, mapped in diverse soybean populations grown at different environments, are designated in SoyBase

(http://soybeanbreederstoolbox.org/; verified 12/08/2007) that condition the unsaturated fatty acid content. These include six oleate QTLs on linkage groups A1, B2 and E, six linoleate QTLs on linkage groups A1, B1, E, K and L, and seven linolenate QTLs on linkage groups E, K and L. The majority of these QTLs were detected for an $F_{2:3}$ population derived from the cross of a normal oleate normal linolenate *Glycine max* line with a normal oleate high linolenate *G. soja* line using RFLPs, isoenzymes and morphological markers (Diers and Shoemaker, 1992). Oleate QTLs coincided with linoleate QTLs with opposing effects. In a recent stusy, six oleate QTLs were mapped on linkage groups A1, D2, G and L in a soybean population of $F_{2:3}$ lines derived from the cross of the normal oleate G99-6725 with the high oleate N00-3350, a single plant selection of N98-4445A (Monteros et al., 2004). Finally, several minor QTLs have been detected for oleate (linkage groups D1b, L and E), linoleate (linkage groups F, L and E) and linolenate (linkage groups F, L, E and G) traits. The QTLs on linkage groups L and E conditioned both mono- and polyunsaturated fatty acid traits and, as expected from the flux of the fatty acid biosynthesis pathway, they had opposite effects (Hyten et al., 2004; Panthee et al., 2006).

Besides the QTLs that condition oleate, linoleate and linolenate, the *GmFAD2* and *GmFAD3* genes, encoding the microsomal ω-6 and ω-3 desaturase enzymes, are implicated in the genetic control of the unsaturated fatty acid content. The two isoforms of *FAD2-1* gene, *FAD2-1A* and *FAD2-1B*, which are primarily responsible

for the oleate desaturation of storage lipids in developing seeds, were mapped on linkage groups O and I of soybean genome, respectively. Functional isoforms of the *FAD2-2* gene are involved in oleate desaturation of membrane lipids both in the developing seeds and vegetative tissues. *FAD2-2B*, as well as the non-functional isoform *FAD2-2A*, were localized on linkage group L, while *FAD2-2C* has not been mapped (Heppard et al., 1996; Schlueter et al., 2007; Bachlava et al., 2008). Of the known isoforms of the *FAD3* gene, whose product is responsible for linoleate desaturation, *FAD3A* was mapped on linkage group B2 and cosegregated with the *fan* locus, a major linolenate gene (Brummer et al., 1995; Bilyeu et al., 2003; Spencer et al., 2004). The map positions of *FAD3B* and *FAD3C*, as well as a fourth distinct *FAD3* isoform (Anai et al., 2005) have yet to be reported.

The objectives of this study were to: (i) map QTLs for oleic acid seed content in two soybean populations segregating for the trait; (ii) determine their genetic effects and the amount of oleate variation they explain; (iii) report the epistatic interactions among the QTLs, as well as the QTL by environment interactions; and (iv) investigate whether the detected QTLs cosegregate with the isoforms of the *FAD2-1* and *FAD2-2* genes that were previously mapped on the soybean genome (Schlueter et al., 2007; Bachlava et al., 2008).

MATERIALS AND METHODS

Population Development and Experimental Design

Two soybean populations, denoted as FAF and FAS, were developed by single seed descent (Brim, 1966). The FAF population consisted of 118 F_5- derived lines from the cross of N97-3363-3 x PI423893 and the FAS population consisted of 231 F_3- derived lines from the cross of N98-4445A x PI423893. N98-4445A is a high oleate (563.1 g kg^{-1}) low linoleate line, which was developed by the USDA-ARS in Raleigh, NC (Burton et al., 2006). N98-4445A originated as a plant selection from the cross N94-2473 x (N93-2007-4 x N92-3907) and is a sister line of N97-3363-3. PI423893 is a mid oleate (305.8 g kg^{-1}) plant introduction with unknown genetic background (USDA-ARS National Plant Germplasm System, http://www.ars-grin.gov/npgs/searchgrin.html; verified 12/08/2007). Apart from oleate genes, the FAF and FAS populations are also segregating for the reduced linolenate *fan*(PI123440) allele.

The FAF population was planted in a sets within replications experimental design (Hallauer and Miranda, 1988) with two replications in each location, as described by Bachlava et al. (2008). The population was grown at Clinton and Kinston, NC, in 2005 and at Clinton, Kinston, Clayton and Plymouth, NC, in 2006. Each experimental line was planted in a four-row plot, with the exception of Clayton

in 2006, where each line was planted in a one-row plot. The FAS population was planted in five-seed hills in a sets within replications design with three replications at Clayton in 2006 and 2007, as described by Bachlava et al. (2008).

Phenotypic Evaluation and Statistical Analysis

Maturity date was recorded at the R8 reproductive stage (Fehr and Caviness, 1977) as days after planting for all environments where the FAF population was grown, and for the FAS population in 2006. Late planting of the FAS population in 2007 resulted in insufficient variation in the maturity dates of the experimental lines. Mature soybean seed was harvested mechanically from the two middle rows of each experimental plot of the FAF population or from each one-row plot at Clayton in 2006. Approximately 10 g of seed were subsampled from each plot for the evaluation of fatty acid composition. For the FAS population 5 g were subsampled from the seed harvested from each hill. Fatty acid composition was evaluated by gas chromatography, as described by Burkey et al. (2007).

Statistical analyses of the phenotypic data for the FAF and FAS populations, which were both planted in sets within replications designs, were conducted using Proc MIXED in SAS 9.1 (SAS Institute, 2004). Environments, sets, replications, lines, as well as their interactions (Appendix 1), were considered random effects. Best linear unbiased predictors (BLUPs) were obtained for maturity and fatty acid traits,

for each experimental line of the FAF and FAS populations (Appendix 2). The phenotypic data of the FAF population from Clinton in 2005 were discarded due to excess missing data and greater error variance compared to the other environments. Analysis of both the FAF and FAS populations was conducted separately for each environment, as well as combined across all the environments. In order to test whether accounting for maturity effects influences the QTL detected herein, BLUPs that included maturity as a covariate in the mixed model were also derived.

Genotypic Evaluation and Linkage Mapping

For each of the FAF experimental lines, genomic DNA was extracted from leaf tissue collected at Kinston in 2005 from approximately ten plants. Genomic DNA was isolated using the Gentra PureGene DNA purification kit (Gentra Systems, Minneapolis, MN). For each of the FAS experimental lines, genomic DNA was isolated from leaf tissue collected at Clayton in 2006 from all plants of each hill using a modified CTAB protocol described by Keim et al. (1988).

A total of 164 polymorphic simple sequence repeat (SSR) markers covering the 20 linkage groups of soybean genome, according to the consensus linkage map (Cregan et al,. 1999; Song et al., 2004), were genotyped for the FAF population. The FAS population was genotyped with 88 SSR markers flanking the genomic regions where QTLs were detected for the FAF population or previous studies, as well as in

the proximity of the isoforms of *FAD2-1* and *FAD2-2* genes. The polymerase chain reactions were performed in a 384-well PTC-200 Peltier Thermal Cycler (MJ Research, Waltham, MA). Thermocycling conditions were 95ºC for 2 min and 38 cycles of 92ºC for 1 min, 49ºC for 1 min and 68ºC for 1 min 30 s. The amplification products were resolved on 4% SFR agarose gels (Amresco, Solon, OH) with ethidium bromide staining in 1x Tris-Borate-EDTA buffer, or 6.5% polyacrylamide gels in a LICOR 4300 DNA Analysis System (LICOR Biosciences, Lincoln, NE) using M13-tailed unmodified primers and IRD-700 or IRD-800 labeled M13 oligonucleotides.

Furthermore, isoform- and allele- specific markers were developed for *FAD2-1A*, *FAD2-1B* and *FAD2-2B*. *FAD2-1A* was genotyped with an allele specific primer extension (ASPE) assay in the FAF population and with a single base extension (SBE) assay in the FAS population. *FAD2-1B* and *FAD2-2B* were genotyped in both populations using cleaved amplified polymorphic sequence (CAPS) markers. The marker development and mapping of *FAD2-1A*, *FAD2-1B* and *FAD2-2B* to the soybean genome are described by Bachlava et al. (2008). Also, the FAF population was genotyped with the *FAD3A* specific marker developed by Camacho-Roger (2006).

Linkage analysis of the FAF population was conducted with MAPMAKER/EXP 3.0 (Lander et al., 1987) according to Cardinal et al. (2001), using a minimum likelihood of odds (LOD) score of 3.0 and a maximum Kosambi distance of 40 cM. The heterozygote genotypes were discarded in order for the F_5-

derived lines to be analyzed as recombinant inbred lines. Linkage analysis of the FAS population was conducted with JoinMap 3.0 (Van Ooijen and Voorrips, 2001), since analyzing the F_3-derived lines as F_2 lines with MAPMAKER/EXP 3.0 resulted in the distortion of map distances. For the JoinMap 3.0 analysis, the linkage map was constructed using the Kosambi's mapping function and linkage was declared with maximum recombination frequency of 0.4 and minimum LOD score of 3.0.

Quantitative Trait Locus Mapping

QTL analysis was conducted with WinQTL Cartographer version 2.5 (Wang et al., 2005) for maturity and fatty acid traits across environments, as well as in each environment separately, using the BLUPs derived for each experimental line of the FAF and FAS population. Initially, the marker genotypes were tested for association with the phenotypic traits using single factor analysis (SFA), a simple linear regression approach. Next, in the FAS population, the composite interval mapping (CIM) procedure was implemented using likelihood ratio (LR) thresholds derived from 1000 permutations at the 0.05 significance level. However, in the FAF population, CIM analysis was conducted with an LR threshold of 11.5 since the high permutation thresholds did not allow the detection of QTL. The "forward and backward regression" method was used for cofactor selection and the genome scans were conducted with window size of 10 cM and walk speed of 2 cM. The CIM

models for each trait and environment were used for the multiple interval mapping (MIM) procedure using the "scan through CIM analysis" option. The models were refined using the "optimize positions", "search for new QTLs" and "test existing QTLs" options, for main effects, in both populations, and epistatic interactions, in the FAS population. MIM models with the minimum Bayesian Information Criterion (BIC) were chosen and "search for new QTLs" was conducted once, since additional searches resulted in the detection of QTL with very small effects. Final models were confirmed using the "MIM forward selection" option. The QTL effects and the proportion of the variation explained from each QTL of the final model was outputted using the "summary" option. For the FAF population, only additive effects were fitted in the model due to the advanced inbreeding generation.

QTL mapping for the FAF population was also conducted with PLABQTL version 1.2 (Utz and Melchinger, 2006), which utilizes a multiple regression procedure in contrast to the maximum likelihood approach of WinQTL Cartographer. CIM analysis was performed using the "cov SEL" statement for cofactor selection and LOD threshold derived from 1000 permutations at the 0.05 significance level. Akaike's Information criterion (AIC) was used for the selection of the best CIM model. Due to limitations in the number of characters for the datasets analyzed by PLABQTL software, supplementary QTL analysis for the FAS population was conducted with MapQTL 5 (Van Ooijen, 2004) using the default settings. Markers selected during the interval mapping (IM) were used as cofactors for the multiple-

QTL model (MQM) mapping. QTL by environment interactions were tested using Proc GLM in SAS 9.1 for the markers that mapped in the proximity of the detected QTLs in the FAF and FAS populations. ESTIMATE and CONTRAST statements were also obtained in order to investigate the contribution of each environment to the interaction.

RESULTS AND DISCUSSION

Linkage Mapping

Two mapping populations, denoted as FAF and FAS, were developed by single seed descent (Brim, 1966) and consisted of 118 F_5- derived lines from the cross of N97-3363-3 x PI423893 and of 231 F_3- derived lines from the cross of N98-4445A x PI423893, respectively. For the FAF population, a linkage map was constructed with 151 SSR markers covering the 20 linkage groups of soybean genome. Segregation distortion was observed for 6 markers at the 0.0025 significance level; while, 13 of the 164 genotyped SSR markers could not be mapped. The length of the linkage map was 1646 cM, which suggests an adequate coverage of the genome considering that the consensus linkage map is 2536 cM in length (Song et al., 2004). The order of the majority of SSR markers coincided with the consensus genetic map of soybean (Cregan et al., 1999; Song et al., 2004). Due to insufficient linkage, as a result of the lack of polymorphic SSR markers or inadequate recombination due to the modest population size, the map contained 13 linkage gaps. The average distance between SSR markers approximately 14.6 cM. The FAS population was selectively genotyped around the genomic regions where QTLs associated with oleic acid content mapped in the FAF population or previous studies, and in the proximity of the isoforms of *FAD2-1* and *FAD2-2* genes. Eight of the 88 genotyped SSR markers remained

unmapped. The length of the genomic region covered was 964 cM, and the SSR markers mapped on average 17.8 cM apart.

Quantitative Trait Locus Mapping Overview

Six oleate QTLs were identified for each of the FAF and FAS populations, explaining 54.9 and 57.4% of the genotypic variation for oleic acid content, respectively (Tables 1 and 2). The QTL with the largest effect was mapped on linkage group F and explained 16.5 and 18.3% of the respective oleate variation in the FAF and FAS populations (Tables 1 and 2). However, it should be noted that due to the modest population sizes, the genetic effects are likely to be overestimated (Beavis, 1994). The high oleate allele of the QTL on linkage group F was inherited from PI423893 in both FAF and FAS populations, whereas the high oleate alleles were inherited from N97-3363-3 for the remaining QTLs detected in the FAF population. An equal number of QTLs inherited the high oleate allele from N98-4445A and PI423893 in the FAS population. With the exception of the oleate QTL on linkage group F, the QTLs identified in the FAF and FAS populations explained less than 12% of oleate variation, which emphasizes the quantitative nature of oleate trait that is controlled by several minor QTLs.

The comparison of the oleate QTLs identified across environments suggested a lack of congruency between the mapping results of the FAF and FAS populations

(Tables 1 and 2). However, comparison of the oleate QTLs identified in each environment tested for both populations (Tables 3 and 4) revealed that the QTLs mapped in the 'Clayton 2006' environment for the FAF population (linkage groups D2 and M) were also identified for the FAS population, which was grown only at the 'Clayton 2006' and 'Clayton 2007' environments. Moreover, QTLs that were identified across environments for the FAF population, but not detected at 'Clayton 2006' (linkage groups N and E), were also absent from the FAS population (Tables 3 and 4). In general, the oleate QTLs mapped for each of the FAF and FAS populations were consistent across environments (Tables 3 and 4) and their effects were similar in direction and magnitude (data not shown).

Apart from the strong environmental influence on the oleate trait, the discrepancies observed between the QTLs identified for the FAF and FAS populations may be due to the limitations of QTL mapping as an approach to identify the genetic factors controlling quantitative traits. The drawbacks of QTL mapping include the lack of congruency among mapped QTLs due to the effects of different environments, genetic backgrounds, number and type of marker loci, level of inbreeding and data analysis techniques (Beavis, 1994). Also, the power of QTL detection is dependent on the size of the mapping populations and the heritability of the traits (Bernardo, 2002). In this study, the two mapping populations were derived from crosses of sister maternal lines with the same paternal plant introduction; therefore they are expected to segregate for common oleate genes. The heritability of

the oleate trait on a plot mean basis was previously reported to be 0.80 and 0.76, and on an entry mean basis was 0.96 and 0.88 for the FAF and FAS populations, respectively (Bachlava et al., 2007). It should be noted, however, that the heritability estimates of the FAS population were calculated only for the 'Clayton 2006' environment. Therefore, the identification of different oleate QTLs for the FAF and FAS populations may be related to differences in the level of inbreeding and the size of the populations, in addition to the effects of different environments.

Five linoleate QTLs were detected for each of the FAF and FAS populations, explaining 60.1 and 46.4% of the genotypic variation for linoleate content, respectively (Tables 1 and 2). All linoleate QTLs corresponded to QTLs identified for oleate content for each of the two mapping populations and had opposite effects, as expected from the strong negative genetic correlations between the two traits that was higher than -0.90 in both populations (Bachlava et al., 2007). Although current QTL mapping analysis alone cannot differentiate whether the cause of the correlated effects for the two traits is due to linkage or pleiotropy, our understanding of the flux through the fatty acid biosynthesis pathway favors the latter.

In contrast, only one linolenate QTL was simultaneously mapped for oleate and linoleate traits in the FAF population, and linolenate shared no common QTLs with oleate or linoleate traits in the FAS population, as expected from the low genetic correlations of linolenate with oleate and linoleate content (Bachlava et al., 2007). A major linolenate QTL was identified on linkage group B2 that cosegregated with the

FAD3A gene, which has been previously mapped to the *fan* locus (Brummer et al., 1995; Bilyeu et al., 2003; Spencer et al., 2004). As shown in Tables 1 and 2, the linolenate QTL on linkage group B2 explained 35.9% of the observed variation in the FAF population and 23.2% in the FAS population. Significant QTL by environment interactions were detected for the linolenate QTL on linkage group B2 (specifically the *FAD3A* marker) in the FAF population. The additive effects of *FAD3A,* estimated using Proc GLM in SAS 9.1, varied from 14.869 g kg^{-1} in the 'Kinston 2005' environment to 22.413 g kg^{-1} in the 'Clayton 2006' environment.

Comparisons of the number and the effects of QTLs identified using MIM in WinQTL Cartographer with those identified using the SFA and CIM analysis illustrate the increased power of QTL detection of MIM procedure (Tables 1 and 2). Moreover, the QTLs detected using the CIM procedure of PLABQTL for the FAF population and the MQM procedure of MapQTL for FAS coincided with the majority of the QTLs mapped using the CIM analysis of WinQTL Cartographer. For the FAF population, the CIM procedure of PLABQTL identified the oleate QTLs on linkage groups F, E, N and A1; while, for the FAS population, the MQM procedure of MapQTL identified the oleate QTLs on linkage groups F, G and I (data not shown).

Comparison of Quantitative Trait Loci across Populations and Mapping Studies

One QTL reported herein was confirmed in both populations used in this study, and four QTLs confirmed QTL identified in previous mapping studies for the unsaturated fatty acid content in soybean. Comparisons among QTLs identified from different studies were conducted according to Cardinal et al. (2001), with the assumption that QTLs that mapped less than 20 cM apart were common. The QTL for oleate and linoleate traits on linkage group F was detected for both mapping populations of this study and was stable across all environments tested (Tables 3 and 4). It should be noted that the locations where the oleate QTL mapped for the FAF and FAS populations are adjacent but not identical marker intervals. This inconsistency is probably due to the different SSR markers genotyped for the two populations (satt206 was not polymorphic in the FAS population) and the difference between the two linkage maps as a consequence of the different population sizes and generation advancement of the two populations.

A maturity QTL was also mapped in the proximity of the oleate QTL on linkage group F in the FAF population (Table 1), which explained approximately 15.7% of the genotypic variation for maturity date and was identified in three of the five environments tested (data not shown). However, QTL mapping for oleate and linoleate traits using BLUPs accounting for maturity effects revealed that the QTL on

linkage group F is not an artifact of maturity differences among the experimental lines that could lead to different temperatures during the period of oil deposition and, therefore, different fatty acid profiles. Also, no QTL for maturity was identified on linkage group F in the FAS population, where the QTLs for oleate and linolenate traits were consistently detected in both environments tested (Table 4). Significant QTL by environment interactions were detected for the QTL on linkage group F (SSR marker sat_133) only in the FAS population for both oleate and linoleate contents. The ranges of the additive effects of the sat_133 marker for oleate and linoleate, which were estimated with Proc GLM in SAS 9.1, were -32.089 to -57.185 g kg^{-1} and 28.938 to 49.962 g kg^{-1} in the 'Clayton 2006' and 'Clayton 2007' environments, respectively.

Furthermore, the minor oleate QTL detected on linkage group E for the FAF population (Table 1) confirmed previously reported results of Panthee et al. (2006) that identified QTLs for oleate, linoleate and linolenate in the proximity of SSR marker satt263. This oleate QTL explained 4.5% of oleate genotypic variation across environments in the FAF population (Table 1) and varied from 3.2 to 5.0% in separate environments (data not shown). The high oleate allele was inherited from N97-3363-3 and resulted in an increase of 15.243 g kg^{-1} in oleic acid content (Table 1). The minor oleate QTL on linkage group E was not detected for the FAS population, but the absence of the QTL from the 'Clayton 2006' environment of the

FAF population (Table 4) implies that it may be due to the environmental effect of this location on the QTL.

Another QTL that validates previous studies is that on linkage group L, which was detected for oleate and linoleate traits in the 'Clayton 2007' environment of FAS population and mapped in the interval of the SSR markers satt006 and sat_113 (Table 4). The QTL on linkage group L coincided with an oleate QTL reported by Monteros et al. (2004) and QTLs for oleate, linoleate and linolenate traits detected by Hyten et al. (2004), which were mapped within the same marker interval. The N98-4445A allele in this study, as well as the N00-3350 allele in the study of Monteros et al. (2004) increased the oleate content. The minor additive and dominance genetic effects of this locus explained less that 4% of oleate and linoleate variation in the FAS population (data not shown). A QTL for maturity was also mapped in the same marker interval in the 'Clayton 2006' environment of the FAS population and across environments for the FAF population. The maturity QTL on linkage group L was previously identified by Orf et al. (1999) near the gene for growth habit *Dt1*. However, the lack of significant maturity differences across the experimental lines, due to late planting of the FAS population in 2007, implies that the QTL on linkage group L actually explains the observed variation for oleate content.

On linkage group A1, oleate and linoleate QTLs were detected in the proximity of the RFLP markers A082_1, A104_1 and A170_1 by Diers and Shoemaker (1992). Monteros et al. (2004) confirmed the existence of a minor QTL around the SSR

marker satt200. In our study, two distinct oleate QTLs were identified for the FAF population on linkage group A1. The oleate QTL located in the interval of satt225 and sat_217, which was detected for the 'Clayton 2006', 'Clinton 2006' and 'Plymouth 2006' environments, coincided with the previously reported QTL on linkage group A1 (Table 3). This QTL explained 2.3, 3.9 and 3.5% of oleate variation in the three environments, respectively (data not shown). The N97-3363-3 allele increased oleate content, which is in agreement with the study by Monteros et al. (2004) where the high oleate allele was inherited from N00-3350, a single plant selection of N98-4445A.

The results of this study also confirm the presence of a minor oleate QTL on linkage group D2, which was identified by Monteros et al. (2004). The QTL on linkage group D2 was mapped in the interval of the SSR markers sat_300 and satt301 for the 'Clayton 2006' environments in the FAF and FAS populations (Tables 3 and 4), and explained 4.3 and 5.0% of the observed variation for oleate content in the two populations, respectively. The high oleate allele for the QTL on linkage group D2 was inherited from the N97-3363-3 and N98-4445A lines in the FAF and FAS populations, respectively, and the N00-3350 line in the study by Monteros et al. (2004).

Cosegregation with the *FAD2-1* and *FAD2-2* Isoforms

The QTL identified on linkage group I for the FAS population, which explained 9.8% of oleate variation and 7.1% of linoleate variation across environments (Table 2), is located less than 3 cM from the *FAD2-1B* gene, previously mapped in the interval between satt354 and sat_268 (Bachlava et al., 2008). These results suggested that allelic variants of *FAD2-1B* cosegregated with the oleate and linoleate QTLs detected for the FAS population. The location and the effect of the QTLs were consistent in the 'Clayton 2006' and 'Clayton 2007' environments (Table 4). The high oleate allele was inherited from PI423893 and resulted in the increase of oleate content by 16.387 g kg^{-1} across environments (Table 2). The results of this study are in agreement with Bachlava et al. (2008), where SFA revealed a minor QTL near *FAD2-1B*. The oleate and linoleate QTL also coincided with a maturity QTL of smaller effect; however, the QTLs for both traits were detected when the analysis accounted for maturity effects in the 'Clayton 2006' environment. In the 'Clayton 2007' environment, the lack of variation in maturity among the experimental lines was further evidence that the QTL on linkage group I explained the observed variation in oleate and linoleate contents. 'Additive by additive' epistatic interactions with minor effects were identified between oleate and linoleate QTLs on linkage groups I and F and were consistent across environments (Table 4). Selection of the PI423893 allele of the oleate QTL on linkage group I and the same allele of the QTL

on linkage group F can lead to a further increase in oleate content by 7.630 g kg^{-1} across environments (Table 2). Additionally, another oleate and linoleate QTL was identified in the FAF population on linkage group I and mapped in the interval of sat_324 and sat_299 for all environments tested (Table 3), approximately 47 cM from the QTL in the proximity of *FAD2-1B*.

An oleate and linoleate QTL was detected on linkage group O for the FAS population and mapped at least 18 cM from the *FAD2-1A* gene (Bachlava et al., 2008). The QTL was identified in both environments tested for the FAS population and contributed minor additive and dominance effects that each explained less that 3% of the observed variation. However, the lack of significant association between oleate content and the *FAD2-1A* marker implies that the QTL on linkage group O does not cosegregate with the *FAD2-1A* isoform. Moreover, the locations of the aforementioned QTLs on linkage group L do not coincide with the *FAD2-2A* and *FAD2-2B* genes, which were previously mapped in the proximity of sat_340 (Bachlava et al., 2008).

CONCLUSIONS

This study revealed a QTL for oleic acid content with moderate effects located on linkage group F in the proximity of the SSR marker sat_309, which was confirmed by both the FAF and FAS populations across all environments tested. Since the high oleate allele was inherited from PI423893 in both populations, the QTL on linkage group F can contribute to the further increase of oleate content of the high oleate lines N97-3363-3 and N98-4445A. Moreover, oleate QTLs with moderate effects were reported on linkage groups A2 and N. These QTLs were not validated by the FAS population possibly due to environmental effects that were evident in the case of the QTL on linkage group N, due to its absence from the 'Clayton 2006' environment. Minor oleate QTLs were also detected on the linkage groups E, L, A1 and D2, confirming previous mapping studies for oleic acid content in soybean (Diers and Shoemaker, 1992; Hyten et al., 2004; Monteros et al., 2004; Panthee et al., 2006).

Furthermore, this study reported an oleate QTL with moderate effects in the proximity of *FAD2-1B* isoform on linkage group I, previously suggested by Bachlava et al. (2008). Although the oleate QTL, which was detected only in the FAS population, coincided with a maturity QTL on linkage group I, further analysis suggested that the oleate QTL is not an artifact of the variation in maturity dates among the experimental lines in the 'Clayton 2006' environment. Interestingly, 'additive by additive' epistatic interactions were observed between the oleate QTL on

linkage group F, which exhibited the largest additive effects in both mapping populations used in this study, and the oleate QTL near *FAD2-1B* isoform on linkage group I. Further investigation is required in order to verify the cosegregation of *FAD2-1B* with the oleate QTL on linkage group I and determine its stability across environments.

In conclusion, an overall view of the QTL mapping studies, to date, shows that oleic acid content is a complex quantitative trait controlled by several minor QTLs that are not always stable across environments and mapping populations. Therefore, marker assisted selection for oleate content is not likely to be effective, unless a few loci with relatively large effects, that are consistent across breeding populations, are identified (Holland, 2004).

ACKNOWLEDGEMENTS

The authors would like to thank William Novitzky at the USDA-ARS Soybean and Nitrogen Fixation Research Unit, Raleigh, NC for providing training and equipment for the fatty acid analysis. This project was funded by the USDA Cooperative State Research, Education and Extension Service National Research Initiative, Plant Genome Program (award number 2003-35300-13190), and by the United Soybean Board.

REFERENCES

Anai, T., T. Yamada, T. Kinoshita, S.M. Rahman, and Y. Takagi. 2005. Identification of corresponding genes for three low-linolenic acid mutants and elucidation of their contribution to fatty acid biosynthesis in soybean seed. Plant Sci. 168:1615-1623.

Bachlava, E., J. Auclair, J.W. Burton, and A.J. Cardinal. 2007. Heritability of oleic and linolenic acid seed content and their genetic correlation with quality and agronomic traits in soybean. ASA Annual Meeting, New Orleans, LA. (http://a-c-s.confex. com/a-c-s/2007am/techprogram/S4012.HTM; verified 02/07/2008).

Bachlava, E., R.E. Dewey, J. Auclair, S. Wang, J.W. Burton, and A.J. Cardinal. 2008. Molecular mapping of genes encoding microsomal omega-6 desaturase enzymes and their cosegregation with QTL affecting oleic acid content in soybean. Crop Sci. (in press).

Beavis, W.D. 1994. The power and deceit of QTL experiments: lessons from comparative QTL studies, *In* D. B. Wilkinson, ed. 49th Annual Corn and Sorghum Research Conference. American Seed Trade Association. Chicago, IL.

Bernardo, R. 2002. Breeding for quantitative traits in plants Stemma Press, Woodbury, MN.

Bilyeu, K.D., L. Palavalli, D.A. Sleper, and P.R. Beuselinck. 2003. Three microsomal omega-3 fatty-acid desaturase genes contribute to soybean linolenic acid levels. Crop Sci. 43:1833-1838.

Brim, C.A. 1996. A modified pedigree method of selection in soybean. Crop Sci. 6:220.

Brummer, E.C., A.D. Nickell, J.R. Wilcox, and R.C. Shoemaker. 1995. Mapping the *Fan* locus controlling linolenic acid content in soybean oil. J. Hered. 86:245-247.

Buhr, T., S. Sato, F. Ebrahim, A.Q. Xing, Y. Zhou, M. Mathiesen, B. Schweiger, A. Kinney, P. Staswick, and T. Clemente. 2002. Ribozyme termination of RNA transcripts down-regulate seed fatty acid genes in transgenic soybean. Plant J. 30:155-163.

Burkey, K.O., F.L. Booker, W.A. Pursley, and A.S. Heagle. 2007. Elevated carbon dioxide and ozone effects on peanut: II. Seed yield and quality. Crop Sci. 47:1488-1497.

Burton, J.W., R.F. Wilson, C.A. Brim, and R.W. Rinne. 1989. Registration of soybean germplasm lines with modified fatty acid composition of seed oil. Crop Sci. 29:1583-1583.

Burton, J.W., R.F. Wilson, G.J. Rebetzke, and V.R. Pantalone. 2006. Registration of N98-4445A mid-oleic soybean germplasm line. Crop Sci. 46:1010-1012.

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. M.S. diss. North Carolina State Univ., Raleigh, NC.

Cardinal, A.J., M. Lee, N. Sharopova, W.L. Woodman-Clikeman, and M.J. Long. 2001. Genetic mapping and analysis of quantitative trait loci for resistance to stalk tunneling by the european corn borer in maize. Crop Sci. 41:835-845.

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

Diers, B.W., and R.C. Shoemaker. 1992. Restriction fragment length polymorphism analysis of soybean fatty acid content. J. Am. Oil Chem. Soc. 69:1242-1244.

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

Gilsinger, J.J. 2008. Reciprocal effects and selection for altered fatty acid composition of soybean oil. Ph.D. diss. North Carolina State Univ., Raleigh, NC.

Hallauer, A.R., and J.B. Miranda. 1988. Quantitative genetics in maize breeding. 2nd ed. Iowa State University Press, Ames, IA.

Heppard, E.P., A.J. Kinney, K.L. Stecca, and G. Miao. 1996. Developmental and growth temperature regulation of two different microsomal omega-6 desaturase genes in soybeans. Plant Physiol. 110:311-319.

Holland, J.B. 2004. Implementation of molecular markers for quantitative traits in breeding programs-challenges and opportunities "New directions for a diverse planet". Proceedings of the 4th International Crop Science Congress, Brisbane, Australia.

Hyten, D.L., V.R. Pantalone, A.M. Saxton, M.E. Schmidt, and C.E. Sams. 2004. Molecular mapping and identification of soybean fatty acid modifier quantitative trait loci. J. Am. Oil Chem. Soc. 81:1115-1118.

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

Kinney, A.J. 1995. Improving soybean seed quality, p. 101-113 Induced Mutations and Molecular Techniques for Crop Improvement: Proceedings. International Atomic Energy Agency (IAEA), Vienna, Austria, 19-23 June.

Kinney, A.J., and S. Knowlton. 1998. Designer oils: the high oleic acid soybean, p. 193-213, *In* S. H. S. Roller, ed. Genetic Modification in the Food Industry. Chapman and Hall, London.

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

Monteros, M.J., J.W. Burton, and H.R. Boerma. 2004. Molecular mapping and confirmation of fatty acid QTL in N00-3350 soybean. ASA Annual Meeting, Seattle, WA. (http://www.cropsoil.uga.edu/soylab/presentations.html; verified 12/13/2007).

Orf, J.H., K. Chase, T. Jarvik, L.M. Mansur, P.B. Cregan, F.R. Adler, and K.G. Lark. 1999. Genetics of soybean agronomic traits: I. Comparison of three related recombinant inbred populations. Crop Sci. 39:1642-1651.

Panthee, D.R., V.R. Pantalone, and A.M. Saxton. 2006. Modifier QTL for fatty acid composition in soybean oil. Euphytica. 152:67-73.

SAS Institute Inc. 2004. SAS/STAT software. Release 9.1.3. SAS Institute Inc., Cary, NC.

Schlueter, J.A., I.F. Vasylenko-Sanders, S. Deshpande, J. Yi, M. Siegfried, B.A. Roe, S.D. Schlueter, B.E. Scheffler, and R.C. Shoemaker. 2007. The *FAD2* gene family of soybean: Insights into the structural and functional divergence of a paleopolyploid genome. The Plant Genome, a Suppl. to Crop Sci. 47:S14-S26.

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

Spencer, M.M., D. Landau-Ellis, E.J. Meyer, and V.R. Pantalone. 2004. Molecular markers associated with linolenic acid content in soybean. J. Am. Oil Chem. Soc. 81:559-562.

Utz, H.F., and A.E. Melchinger. 2006. PLABQTL version 1.2: A computer program to map QTL. Institute of Plant Breeding, Seed Science, and Population Genetics, Univeristy of Hohenheim, Stuttgart, Germany. (https://www.uni-hohenheim.de/ plantbreeding/ software/; verified 12/13/2007).

Van Ooijen, J.W., and R.E. Voorrips. 2001. Joinmap® 3.0 : Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, The Netherlands.

Van Ooijen, J.W. 2004. MapQTL® 5: Software for the mapping of quantitative trait loci in experimental populations. Kyazma, BV, Wageningen, The Netherlands.

Wang, S., C.J. Basten, and Z.B. Zeng. 2007. Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State Univ., Raleigh, NC (http://statgen. ncsu.edu/qtlcart/ WQTLCart.htm; verified 12/13/2007).

Wilcox, J.R. 2004. World distribution and trade of soybean, p. 1-13, *In* H. R. Boerma and J. E. Specht, eds. Soybeans: Improvement, Production and Uses, 3rd ed. American Society of Agronomy, Inc., Madison, WI.

Wilson, R.F., J.W. Burton, and C.A. Brim. 1981. Progress in the selection for altered fatty acid composition in soybeans. Crop Sci. 21:788-791.

Wilson, R.F., J.W. Burton, V.R. Pantalone, and R.E. Dewey. 2002. New gene combinations governing saturated and unsaturated fatty acid composition in soybean p. 95-113, *In* T. M. Kuo and H. W. Gardner, eds. Lipid Biotechnology. Marcel Dekker, Inc., IL.

Wilson, R.F. 2004. Seed composition, p. 621-669, *In* H. R. Boerma and J. E. Specht, eds. Soybeans: Improvement, Production, and Uses, 3rd ed. American Society of Agronomy, Inc., Madison, WI.

Table 1. Quantitative trait loci (QTLs) for oleate, linoleate, linolenate and maturity date detected with multiple interval mapping (MIM) in the FAF population (N96-3363-3 x PI423893) based on best linear unbiased predictors (BLUPs) derived across environments. The linkage group (LG) and the marker interval where each QTL was mapped, as well as its size, the distance of the QTL peak from the most nearby marker (underlined), the additive effect of the QTL, and the proportion of variation explained by each QTL (R^2) is reported.

Trait	LG	Marker interval	Interval size (cM)	Distance from nearby marker (cM)	Additive Effect† (g kg^{-1})	R^2 (%)
Oleate	‡N	satt022 – sat_304	14.00	4.01	21.041**	11.2
	‡A2	sat_294 – satt333	13.20	5.19	21.827**	11.6
	‡F	satt206 – sat_309	23.20	7.19	-25.512***	16.5
	I	sat_324 – sat_299	14.60	5.10	16.726**	7.4
	A1	satt050 – satt619	18.90	0.01	12.646	3.7
	E	sat_273 – satt651	16.10	0.01	15.243	4.5
					Total $R^{2\,\S}$ (%)	54.9
Linoleate	‡N	satt022 – sat_304	14.00	2.01	-15.525***	9.4
	A2	satt333 – satt233	31.40	10.01	-24.897**	20.6
	‡F	satt206 – sat_309	23.20	6.00	22.061***	16.7
	I	sat_324 – sat_299	14.60	6.60	-16.502**	10.1
	A1	satt225 – sat_217	4.10	0.09	-10.026	3.3
					Total $R^{2\,\S}$ (%)	60.1
Linolenate	‡L	satt495 – sat_301	9.40	5.39	2.602**	6.1
	‡A2	satt409 – sat_294	5.20	1.20	-2.336***	5.6
	D1b	satt634 – sat_423	45.60	11.59	-3.338**	10.6
	B2	satt687 – *FAD3A*	17.30	3.29	-6.334***	35.9
	A1	satt684 – sat_368	9.80	1.79	-2.231*	5.1
					Total $R^{2\,\S}$ (%)	63.3

Table 1. Continued.

Trait	LG	Marker interval	Interval size (cM)	Distance from nearby marker (cM)	Additive Effect[†] (g kg^{-1})	R^2 (%)
Maturity	[‡] B2	*FAD3A* – satt066	17.00	2.99	1.518***	13.5
	[‡] F	sat_309 – sat_133	3.30	1.29	-1.767***	15.7
	L	satt495 – sat_301	9.40	4.00	-1.080*	5.3
	[‡] L	*FAD2-2B* – satt561	14.70	4.01	1.362***	10.7
	I	satt571 – satt496	18.90	6.01	-1.134	4.9
	B1	sat_270 – satt509	9.90	0.01	0.959*	4.8
					Total R^2 [§] (%)	54.9

*, **, *** Significance at the 0.05, 0.01 and 0.001 level of the marker closest to the peak of the QTL (underlined) estimated with single factor analysis (SFA)

[†]Additive effect of the QTL was estimated as the difference of the homozygous N97-3363-3 and PI423893 with MIM analysis. Positive additive effects indicate that the N97-3363-3 allele increases the value of the trait

[‡]QTL detected with composite interval mapping (CIM) analysis using likelihood ratio threshold LR = 11.5

[§]Total variation explained when all QTLs are simultaneously fitted in the MIM model

Table 2. Quantitative trait loci (QTLs) for oleate, linoleate, linolenate and maturity date detected with multiple interval mapping (MIM) in the FAS population (N98-4445A x PI423893) based on best linear unbiased predictors (BLUPs) derived across environments. The linkage group (LG) and the marker interval where each QTL was mapped, as well as its size, the distance of the QTL peak from the most nearby marker (underlined), the additive and/or dominance effect of the QTL, and the proportion of variation explained by each QTL (R^2) is reported.

Trait	LG	Marker interval	Interval size (cM)	Distance from nearby marker (cM)	Effect (A/D)[†] (g kg^{-1})	R^2 (%)
Oleate	M	satt220 – sat_226	27.76	11.76	38.802 (D)*	11.7
	[‡] G	satt472 – satt288	19.80	1.96	11.152 (A)**	5.0
	D2	sat_300 – satt301	7.90	0.08	10.084 (A)*	2.7
	O	sat_108 – satt153	11.99	5.98	- 8.401 (A)*/ -17.726 (D)	2.8 / 2.7
	[‡] I	satt354 – *FAD2-1B*	4.91	2.12	-16.387 (A)***	9.8
	[‡] F	sat_133 – sat_309	8.94	0.10	-22.912 (A)***	18.3
	I x F [#]				7.630 (A x A)	1.7
					Total $R^{2\,\S}$ (%)	57.4
Linoleate	G	satt472 – satt288	19.80	0.01	-9.232 (A)**	4.0
	D2	satt002 – satt669	19.84	6.01	-10.818 (A)**	5.9
	O	*FAD2-1A* – sat_108	24.02	6.01	7.582 (A)*/ 15.096 (D)	3.1 / 3.0
	[‡] I	satt354 – *FAD2-1B*	4.91	0.90	10.745 (A)***	7.1
	[‡] F	sat_133 – sat_309	8.94	0.10	20.199 (A)***	20.9
	I x F [#]				-6.263 (A x A)	1.6
					Total $R^{2\,\S}$ (%)	46.4

Table 2. Continued.

Trait	LG	Marker interval	Interval size (cM)	Distance from nearby marker (cM)	Effect (A/D)† (g kg^{-1})	R^2 (%)
Linolenate	‡ B2	sat_355 – satt066	24.33	0.32	-6.074 (A)***	23.2
	J	satt244 – sat_395	28.10	10.09	12.144 (D)	16.3
					Total R^{2} § (%)	39.7
Maturity¶	M	sat_226 - sat_256	16.11	2.01	1.097 (A)***	6.6
	I	*FAD2-1B* – sat_268	5.96	0.01	-1.038 (A)***	6.5
	‡ O	satt478 – satt420	25.84	1.83	-1.027 (A)***/ -1.656 (D)	7.1 / 2.3
	‡ L	satt006 – sat_113	11.83	5.32	2.285 (A)***	26.1
	B2	sat_355 – satt066	24.33	8.32	0.666 (A)***/ 2.269 (D)	2.9 / 4.7
	M x L #				-1.073 (A x A)	1.9
					Total R^{2} § (%)	59.2

*, **, *** Significance at the 0.05, 0.01 and 0.001 level of the marker closest to the peak of the QTL (underlined) estimated with single factor analysis (SFA)

†Additive effect of the QTL was estimated as the difference of the homozygous N98-444A and PI423893 genotypes; while, dominance effect was estimated as the difference of the heterozygous genotypes from the mean of the homozygous N98-444A and PI423893 genotypes with MIM analysis. Positive additive effects indicate that the N98-4445A allele increases the value of the trait. Additive and dominance effects of the QTL were both included in the table when each effect explained more than 1.5% of the variation

‡QTL detected with composite interval mapping (CIM) analysis using the permutation likelihood ratio threshold

§Total variation explained when all QTLs were simultaneously fitted in the MIM model

Table 2. Continued.

¶QTL analysis for maturity date was conducted with BLUPs derived from a single environment (Clayton, 2006)
#Epistatic interactions between the designated QTLs

Table 3. Comparison of quantitative trait loci (QTLs) for oleate detected with multiple interval mapping (MIM) based on best linear unbiased predictors (BLUPs) derived across environments with oleate QTLs detected based on BLUPs derived separately for each environment where FAF population (N97-3363-3 x PI423893) was tested. The marker interval, where each QTL was mapped, and the marker closer to the peak of the QTL (underlined) are reported.

LG[†]	Across[‡]	Kinston 2005	Kinston 2006	Clayton 2006	Clinton 2006	Plymouth 2006
N	satt022 - sat_304	satt022 - sat_304	satt022 - sat_304		satt022 - sat_304	
A2	sat_294 - satt333	sat_294 - satt333	sat_294 - satt333	satt333 - satt233	sat_294 - satt333	sat_294 - satt333
F	satt206 - sat_309	satt206 - sat_309	satt206 - sat_309	satt206 - sat_309	satt206 - sat_309	satt206 - sat_309
I	sat_324 - sat_299	sat_418 - sat_324	sat_418 - sat_324	sat_324 - sat_299		sat_324 - sat_299
A1[§]	satt050 - satt619	satt050 - satt619	satt300 - satt050	satt225 - sat_217	satt225 - sat_217 satt050 - satt619	satt225 - sat_217
E	sat_273 - satt651	sat_273 - satt651	sat_273 - satt651		sat_273 - satt651	
				D2[¶] (sat_300 - satt301) **M**[¶] (sat_226 - satt220)	**J**[¶] (sat_224 - sat_395)	**M**[¶] (sat_226 - satt220)

[†] Linkage group
[‡] QTL detected using BLUPs derived across environments

Table 3. Continued.

[§] Two distinct QTL positions were reported on linkage group A1 in different environments, but both QTLs were simultaneously fitted only in one of the environments tested (Clinton 2006)
[¶] QTL that were not detected using BLUPs derived across environments

Table 4. Comparison of quantitative trait loci (QTLs) for oleate detected with multiple interval mapping (MIM) based on best linear unbiased predictors (BLUPs) derived across environments with oleate QTLs detected based on BLUPs derived separately for each environment where FAS population (N98-4445A x PI423893) was tested. The marker interval, where each QTL was mapped, and the marker closer to the peak of the QTL (underlined) are reported.

LG[†]	Across[‡]	Clayton 2006	Clayton 2007
M	satt220 - sat_226	satt220 - sat_226	
G	satt472 - satt288	satt472 - satt288	satt472 - satt288
D2[§]	sat_300 - satt301	sat_300 - satt301	satt135 - satt002
O	sat_108 - satt153	sat_108 - satt153	*FAD2-1A* - sat_108
I	satt354 - *FAD2-1B*	satt354 - *FAD2-1B*	satt354 - *FAD2-1B*
F	sat_133 - sat_309	sat_133 - sat_309	sat_133 - sat_309
L[¶]			satt006 - sat_113
	I x F[#]	**I x F**[#]	**I x F**[#] **G x L**[#]

[†] Linkage group
[‡] QTL detected using BLUPs derived across environments
[§] Two distinct QTL positions were reported on linkage group D2 in 2006 and 2007
[¶] QTL that was not detected using BLUPs derived across environments
[#] Additive x additive epistatic effects

V. CONCLUSIONS

In this study, three soybean populations, that are segregating for major and minor oleate genes, were grown in replicated multi-environment trials and used for the estimation of oleate heritability and its genetic correlations with other fatty acid and agronomic traits. Heritability of oleate content has not been previously estimated in soybean populations that are segregating for oleate content and have adequate experimental lines replicated and tested across environments. One of the populations used in this study consisted of 721 random lines, providing better precision for the estimation of genetic parameters than previous studies in soybean. Moreover, estimates of the genotypic correlations of oleate content with other traits have not been reported (Cardinal, 2008). This study reported that oleate heritability was sufficiently high for unreplicated lines grown at a single environment and proposed that effective early generation selection can be conducted with limited replication and testing across environments. Genotypic and phenotypic correlations suggested that selection for higher oleate content will result in lower linoleate, linolenate and palmitate content in soybean oil, which is in agreement with previous reports (Burton at al. 1983; Rebetzke et al., 1998; Alt et al., 2005; Cardinal and Burton, 2007), as well as with the current knowledge on fatty acid biosynthesis (Ohlrogge and Browse, 1995). This study also revealed the significant negative correlation between yield and oleate content, and positive correlations between yield and linoleate, linolenate and palmitate contents, suggesting that the development of high oleate soybean germplasm may be hampered by an indirect effect on yield.

This study focused also on the mapping of the isoforms of *FAD2-1* and *FAD2-2* genes, which encode the microsomal ω-6 desaturase enzymes that catalyze the desaturation of oleate to linoleate during the fatty acid biosynthesis in the endoplasmic reticulum. The *GmFAD2* genes were considered candidates to explain the observed oleate variation; therefore, the association of these genetic loci with oleate phenotypes was investigated in the three segregating populations. Furthermore, oleate QTLs were identified, for two of the three populations used in this study, in order to shed light into the genetic factors underlying oleate variation in soybean. Genotyping assays were developed for the *FAD2-1A*, *FAD2-1B* and *FAD2-2B* isoforms. Linkage analysis revealed that *FAD2-1A* and *FAD2-1B* map on linkage groups O and I, respectively, and the tightly linked *FAD2-2A* and *FAD2-2B* map on linkage group L, confirming previous findings of Schlueter et al. (2007). An oleate QTL with moderate effects was detected on linkage group F in the proximity of the SSR marker sat_309 for both populations across all environments tested. In addition, *FAD2-1B* isoform cosegregated with an oleate QTL on linkage group I in one of the two segregating populations, while oleate phenotypes were associated with a nearby SSR marker in the third population studied herein. It should be noted that the oleate QTL with moderate effects in the proximity of *FAD2-1B* isoform interacted epistatically with the oleate QTL on linkage group F. Minor oleate QTLs were also detected on the linkage groups E, L, A1 and D2, confirming previous mapping studies

for oleate content in soybean (Diers and Shoemaker, 1992; Hyten et al., 2004; Monteros et al., 2004; Panthee et al., 2006).

The results of this study suggested that the *FAD2-1A*, *FAD2-1B* and *FAD2-2B* isoforms do not encode for major oleate genes that contribute to the elevated oleic acid content of the N98-4445A and N97-3363-3 lines. Additional research on the *FAD2-2C* isoform, as well as other candidate genes implicated in oleate biosynthesis, is necessary in order to substantiate these results. Further investigation is also required in order to verify the cosegregation of *FAD2-1B* with the oleate QTL on linkage group I and determine its stability across environments. However, the results of this study, as well as relevant previous studies, suggest that oleate content is a complex quantitative trait controlled by several minor QTLs that are not always stable across environments and mapping populations, which renders marker assisted selection for oleate content ineffective, to date.

REFERENCES

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

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.

Cardinal, A.J., and J.W. Burton. 2007. Correlations between palmitate content and agronomic traits in soybean populations segregating for the *fap1*, fap_{nc} and *fan* alleles. Crop Sci. 47:1804-1812.

Cardinal, A.J. 2008. Molecular genetics and breeding for fatty acid manipulation in soybean. Plant Breed. Rev. 30: (in press).

Diers, B.W., and R.C. Shoemaker. 1992. Restriction fragment length polymorphism analysis of soybean fatty acid content. J. Am. Oil Chem. Soc. 69:1242-1244.

Hyten, D.L., V.R. Pantalone, A.M. Saxton, M.E. Schmidt, and C.E. Sams. 2004. Molecular mapping and identification of soybean fatty acid modifier quantitative trait loci. J. Am. Oil Chem. Soc. 81:1115-1118.

Monteros, M.J., J.W. Burton, and H.R. Boerma. 2004. Molecular mapping and confirmation of fatty acid QTL in N00-3350 soybean ASA Annual Meeting, Seattle, WA. (http://www.cropsoil.uga.edu/soylab/presentations.html; verified 12/13/2007).

Ohlrogge, J., and J. Browse. 1995. Lipid biosynthesis. Plant Cell 7:957-970.

Panthee, D.R., V.R. Pantalone, and A.M. Saxton. 2006. Modifier QTL for fatty acid composition in soybean oil. Euphytica. 152:67-73.

Rebetzke, G.J., J.W. Burton, T.E. Carter, and R.F. Wilson. 1998. Changes in agronomic and seed characteristics with selection for reduced palmitic acid content in soybean. Crop Sci. 38:297-302.

Schlueter, J.A., I.F. Vasylenko-Sanders, S. Deshpande, J. Yi, M. Siegfried, B.A. Roe, S.D. Schlueter, B.E. Scheffler, and R.C. Shoemaker. 2007. The *FAD2* gene family of soybean: Insights into the structural and functional divergence of a paleopolyploid genome. The Plant Genome, a Suppl. to Crop Sci. 47:S14-S26.

VI. APPENDICES

Appendix 1. Expected mean squares and degrees of freedom (DF) for all sources of variance fitted in the mixed model for the analysis of the sets within replications design across environments. The table corresponds to the FAF population tested in five environments with two replications per environment, five sets per replication and 25 genotypes per set. Expected mean squares and degrees of freedom are calculated for balanced data.

Source of Variance	DF	Expected Mean Squares
Env[†]	4	$2\,\sigma^2_{EG(S)}$[‡] $+ 25\,\sigma^2_{SR(E)} + 100\,\sigma^2_{SE} + 125\,\sigma^2_{R(E)} + 2500\,\sigma^2_{E} + \sigma^2_{e}$
Rep (Env)	5	$25\,\sigma^2_{SR(E)} + 125\,\sigma^2_{R(E)} + \sigma^2_{e}$
Set	4	$2\,\sigma^2_{EG(S)} + 10\,\sigma^2_{G(S)} + 25\,\sigma^2_{SR(E)} + 100\,\sigma^2_{SE} + 2500\,\sigma^2_{S} + \sigma^2_{e}$
Set*Env	16	$2\,\sigma^2_{EG(S)} + 25\,\sigma^2_{SR(E)} + 100\,\sigma^2_{SE} + \sigma^2_{e}$
Set*Rep (Env)	20	$25\,\sigma^2_{SR(E)} + \sigma^2_{e}$
Genotype (Set)	120	$2\,\sigma^2_{EG(S)} + 10\,\sigma^2_{G(S)} + \sigma^2_{e}$
Env*Genotype (Set)	480	$2\,\sigma^2_{EG(S)} + \sigma^2_{e}$
Error	600	σ^2_{e}

[†] Environment

[‡]The subscipts of the variance components correspond to environments (E), replications (R), sets (S), genotypes (G) and error (e)

Appendix 2. SAS code for the estimation of best linear unbiased predictors (BLUPs) for experimental lines grown in replicated mutli-environment trials in a sets within replications design. The code estimates BLUPs for oleate content in the FAF population.

```
PROC IMPORT OUT= work.a
    DATAFILE= "C:\Documents and Settings\eleni bachlava\FAF_dataset.xls"
    DBMS=EXCEL REPLACE;
    GETNAMES=YES;
RUN;

%macro all_BLUPs(trait);

proc mixed data=a;
class env sets rep geno_id;
model &trait = /ddfm=satterth solution;
random env rep(env) sets sets*env sets*rep(env) geno_id(sets) env*geno_id(sets)/solution;
ods listing exclude solutionF solutionR;
ods output solutionF = &trait.F solutionR = &trait.R;
run;

data m; set &trait.F; if Effect = "Intercept"; grd_mn = Estimate; dummy = 1;
proc print;
run;

data e; set &trait.R; if Effect = "Geno_ID(Sets)"; rnd_effect = Estimate; dummy = 1;
proc print;
run;

data &trait.blup;
merge e m; by dummy;
BLUP = grd_mn + rnd_effect;
&trait=BLUP;
keep geno_id grd_mn rnd_effect &trait;
proc sort data= &trait.blup; by geno_id;
proc print data= &trait.blup;
run;

proc export data=&trait.blup outfile= 'C:\Documents and Settings\BLUPs_oleic.xls'
DBMS=EXCEL REPLACE;
SHEET = &trait;
run;

%mend all_BLUPs;
%all_BLUPs (oleic)
run;
```

Appendix 3. SAS code for the estimation of parent-offspring heritabilities expressed as the regression of $F_{5:6}$ offspring on $F_{2:3}$ progenitors. The code refers to the heritability of oleate content in the FAF population.

```
PROC IMPORT OUT= a
    DATAFILE= "C:\Documents and Settings\FAF_dataset.xls"
    DBMS=EXCEL REPLACE;
    GETNAMES=YES;
RUN;
%macro heritability(FA);
TITLE1 "Regression for &FA";
%let F2 = F2_;
%let F56 = F56_;

proc reg data=a;
model &F56&FA=&F2&FA;
plot &F56&FA * &F2&FA / cframe=ligr;
run;

%mend heritability;
%heritability(oleic)
run;
```

Appendix 4. SAS code for the estimation of heritabilities on a plot and an entry means basis using variance components according to Holland et al. (2003). The code refers to the heritability of oleate content in the FAF population.

```
PROC IMPORT OUT= a
    DATAFILE= "C:\Documents and Settings\FAF_dataset.xls"
    DBMS=EXCEL REPLACE;
    GETNAMES=YES;
RUN;

%macro heritability(trait);
proc mixed data = a asycov;
class env rep sets geno_id;
model &trait = ;
random env rep(env) sets sets*env sets*rep(env) geno_id(sets) env*geno_id(sets);
ods listing exclude AsyCov CovParm;
ods output asycov = covmat covparms = estmat;
run;
proc iml;
start seh(V, C, LG, LP, H, SE);
Vp = LP`*V;
Vg = LG`*V;
H = VG/Vp;
d = (1/Vp)*(LG - (LP*H));
VH = d`*C*d;
SE = sqrt(VH);
finish seh;

use estmat; read all into v; use covmat; read all into c;
C = C(|1:nrow(C), 2:ncol(C)|);
print C;
LG = {0, 0, 0, 0, 0, 1, 0, 0};
LP = {0, 0, 0, 0, 0, 1, 1, 1};
call seh(V, C, LG, LP, H, SE);
print "Heritability on a Plot Basis &trait", H, SE;
e = 4.829;
p = 9.247;
LP = 0//0//0//0//0//1//(1/e)//(1/p);
call seh(V, C, LG, LP, H, SE);
print "Heritability on an Entry Mean Basis &trait", H, SE;
quit;
run;

%mend heritability;
%heritability(oleic)
run;
```

Appendix 5. SAS code for the estimation of genotypic and phenotypic correlations and their standard errors according to Holland (2006). The code refers to the genotypic and phenotypic correlation of oleate with linoleate content in the FAF population.

```
PROC IMPORT OUT= a
     DATAFILE= "C:\Documents and Settings\FAF_dataset.xls"
     DBMS=EXCEL REPLACE;
    GETNAMES=YES;
RUN;

data d; length trait $ 9; set a;
trait = "oleic"; y = oleic; output;
trait = "linoleic"; y = linoleic; output;
run;

%macro corr(trait1, trait2);
data f; set d; if trait = "&trait1" or trait = "&trait2";
proc mixed data = f asycov;
class trait env sets rep geno_id;
model y = env(trait) rep(env*trait) sets(trait) sets*env(trait) sets*rep(env*trait);
random trait/subject= geno_id(sets) type=un;
random trait/subject= env*geno_id(sets) type=un;
repeated trait/sub = rep*geno_id(sets*env) type=un;
ods output covparms = estmat asycov = covmat;
run;

proc iml;
use estmat; read all into e;
use covmat; read all into cov;
C = cov(|1:nrow(cov), 2:ncol(cov)|);
print e cov C;
run;

CovG = e(|2,1|);
VG1 = e(|1,1|);
VG2 = e(|3,1|);
CovP = CovG + e(|5,1|) + e(|8,1|);
VP1 = VG1 + e(|4,1|) + e(|7,1|);
VP2 = VG2 + e(|6,1|) + e(|9,1|);

start correl(C, CovG, VG1, VG2, CovP, VP1, VP2, RG, RP, SERG, SERP);
RG = CovG/sqrt(VG1*VG2);

dg = (-1/(2*VG1))//(1/CovG)//(-1/(2*VG2))//0//0//0//0//0//0;
varrg = (RG**2)*dg`*C*dg; serg = sqrt(varrg);
RP = CovP/sqrt(VP1*VP2);
*Make the derivate vector for rp;
d1p = -1/(2*VP1);
```

Appendix 5. Continued.

```
d2p = 1/CovP;
d3p = -1/(2*VP2);
dp= d1p//d2p//d3p//d1p//d2p//d3p//d1p//d2p//d3p;
varrp = (RP**2)*dp`*C*dp; serp = sqrt(varrp);
finish correl;
call correl(C, CovG, VG1, VG2, CovP, VP1, VP2, RG, RP, SERG, SERP);
print "Genotypic Correlation Between &trait1 and &trait2";
print RG serg;
print "Phenotypic Correlation Between &trait1 and &trait2";
print RP serp;
quit; run;

%mend corr;
%corr(oleic,linoleic);
run;
```