

MAPPING AND CONFIRMATION OF SOYBEAN QUANTITATIVE TRAIT LOCI
FOR OLEIC ACID CONTENT AND REACTION TO ASIAN SOYBEAN RUST

by

MARIA JOSE MONTEROS

Under the Direction of H. ROGER BOERMA

ABSTRACT

Soybean, *Glycine max* (L.) Merrill, is a major commodity traded in world markets and is currently the world's primary oilseed crop. Increasing oleic acid content of soybean oil would reduce the need for hydrogenation, which creates unhealthy *trans* fatty acids. Six oleic acid quantitative trait loci (QTL) from the mid-oleic soybean line N00-3350 have been mapped and confirmed on linkage groups A1, D2, G, and L using SSR markers. Additional sequence-based markers have been mapped to these regions in the soybean genome. Asian soybean rust (ASR) caused by *Phakopsora pachyrhizi*, was reported for the first time in the USA in 2004 and has the potential to cause considerable losses in soybean yield. A novel source of resistance to ASR from the cultivar Hyuuga has been mapped to LG-C2 using SSR and SNP markers, and its location has been confirmed in an independent population. The identification of molecular markers closely linked to the identified oleic acid QTL and the ASR resistance gene from Hyuuga will facilitate the use of marker-assisted selection (MAS) in soybean breeding programs to increase the oleic acid content in soybean seed and

develop ASR resistant soybean cultivars with desirable agronomic performance adapted to the various production regions of the USA.

INDEX WORDS: Asian soybean rust, fatty acid, *Glycine max*, marker-assisted selection, oleic acid, *Phakopsora pachyrhizi*, soybean, SSR, SNP, SSCP, QTL mapping.

MAPPING AND CONFIRMATION OF SOYBEAN QUANTITATIVE TRAIT LOCI
FOR OLEIC ACID CONTENT AND REACTION TO ASIAN SOYBEAN RUST

by

MARIA JOSE MONTEROS

B.S., Universidad del Valle, Guatemala, 2000

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in
Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

© 2006

Maria Jose Monteros

All Rights Reserved

MAPPING AND CONFIRMATION OF SOYBEAN QUANTITATIVE TRAIT LOCI
FOR OLEIC ACID CONTENT AND REACTION TO ASIAN SOYBEAN RUST

by

MARIA JOSE MONTEROS

Major Professor: H. Roger Boerma

Committee: Joseph Bouton
Jeffrey Dean
Steven Knapp
Wayne Parrott

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2006

DEDICATION

To my parents

ACKNOWLEDGEMENTS

I would like to thank people who have helped me throughout the development of my research, including Jennie Alvernaz, David Hulburt, Erin Legget, Dale Wood, Gina Rowan, Earl Baxter, Joseph Head, Herschel Chambers and all of my colleagues. I value the interactions and exchange of ideas with David Walker, Ali Missaoui, Jennifer Yates, and Bo-Keun Ha. I am thankful for the contribution from Adam Ball in collecting data and his willingness to work and learn. I am grateful to collaborators at other institutions including Perry Cregan, Randy Nelson, Mariangela Hungria, and Daniel Phillips. I would like to acknowledge funding from Dr. Glenn Burton's family, the Tinker Foundation, United Soybean Board, and the Georgia Agricultural Experiment Stations.

I appreciate the contribution to my graduate experience from the members of my committee Dr. Joseph Bouton, Dr. Jeffrey Dean, Dr. Steven Knapp and Dr. Wayne Parrott. I would especially like to thank Dr. Roger Boerma for giving me the opportunity to work in his lab, and the freedom to investigate many areas in soybean research. I value all of his insights and encouragement to explore other areas of my professional growth. I have thoroughly enjoyed working with someone with such enthusiasm and passion for their work. I can't thank you enough for everything. I have learned so much from you.

I would like to thank Frank for his support and understanding throughout this process. I am incredibly grateful to my friends and family, specially my parents for their continued support and encouragement to reach my goals throughout my life. To all of you, thank you!

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES.....	ix
LIST OF FIGURES.....	xii
CHAPTER	
1 INTRODUCTION.....	1
2 LITERATURE REVIEW	14
Oils and fats	14
Fatty acids	14
Vegetable oils.....	16
Soybean oil	18
Role of fats in human health.....	20
Hydrogenation and <i>trans</i> fats	21
Lipid synthesis in plants	22
Breeding for fatty acid content	24
Transgenics with altered fatty acid content	28
Environmental effects in variation in oleic acid content.....	29
Molecular markers.....	31
Quantitative trait loci (QTL)	32
Marker-assisted selection (MAS)	35
Asian soybean rust (ASR).....	37

References.....	42
3 MOLECULAR MAPPING AND CONFIRMATION OF QTL ASSOCIATED WITH OLEIC ACID CONTENT IN N00-3350 SOYBEAN	60
Abstract.....	61
Introduction	62
Materials and Methods.....	65
Results and Discussion.....	69
References.....	77
4 DISCOVERY AND MAPPING OF SEQUENCE-BASED MARKERS ASSOCIATED WITH OLEIC ACID CONTENT IN N00-3350 SOYBEAN	93
Abstract.....	94
Introduction	95
Materials and Methods.....	98
Results and Discussion.....	104
References.....	110
5 MAPPING AND CONFIRMATION OF THE 'HYUUGA' RED-BROWN LESION RESISTANCE GENE FOR ASIAN SOYBEAN RUST	124
Abstract.....	125
Introduction	126
Materials and Methods.....	128
Results and Discussion.....	134
References.....	139
6 SUMMARY	147

APPENDICES	150
1 PLANT INTRODUCTIONS (PI'S) WITH MID-OLEIC ACID CONTENT	150
2 FINE MAPPING A RESISTANCE GENE TO ASIAN SOYBEAN RUST FROM THE CULTIVAR HYUUGA.....	152
Introduction	152
Objectives	153
Materials and Methods.....	153
Results and Discussion.....	156
References.....	160
3 OLIGONUCLEOTIDES FOR OLEIC ACID QTL	169
4 LIST OF ABBREVIATIONS.....	172

LIST OF TABLES

	Page
Table 2.1: Fatty acids most commonly found in plants.....	57
Table 2.2: Fatty acid composition of different types of oil.....	58
Table 3.1: Mean fatty acid content of the parents and mean range of progeny from the G99-G725 × N00-3350 population grown in an Athens, GA greenhouse and in the field at Isabela, PR	83
Table 3.2: SSR markers associated with the oleic acid content in 259 lines from G99- G725 × N00-3350	84
Table 3.3: Mean oleic acid content for SSR markers associated with putative oleic acid QTL in 259 lines from G99-G725 × N00-3350	85
Table 3.4: SF-ANOVA for markers associated with soybean fatty acid content in G99- G725 × N00-3350	86
Table 3.5: Mean fatty acid content of parents and the mean range of 231 progeny lines from the G99-G3438 × N00-3350 population used for confirmation	87
Table 3.6: Marker regression analysis for oleic acid content using 231 lines from G99- G3438 × N00-3350	88
Table 3.7: Marker associations with oleic acid content from the sub-samples from G99- G3439 × N00-3350 in Athens	89
Table 4.1: <i>G. max</i> sequences from genes in the fatty acid biosynthetic pathway	116
Table 4.2: <i>G. max</i> UniGene sets from candidates in the fatty acid biosynthetic pathway	117

Table 4.3: Oligonucleotides developed from sequence-tagged-sites in linkage groups from the soybean genome associated with oleic acid QTL	118
Table 4.4: SF-ANOVA marker associations with oleic acid content from G99-G725 × N00-3350.....	119
Table 4.5: SSCP and SNP markers from fatty acid gene-based sequences and the linkage group in which they were mapped	120
Table 5.1: Field and greenhouse evaluations for type of lesion, severity, and average number of lesions of RILs from Dillon × Hyuuga and Benning × Hyuuga	143
Table 5.2: QTL mapping of ASR severity from the field and lesion number in the greenhouse using RILs from Dillon × Hyuuga	144
Table A.1.1: Fatty acid content for PI's and checks grown in the greenhouse	150
Table A.1.2: Plant Introduction SSR marker amplicon sizes on LG-A1, D2, G, and L.	151
Table A.2.1: SNP markers tested in the region between Satt307 and Satt460 on LG-C2	162
Table A.2.2: SNP genotypes of mapping parents and sources of Asian soybean rust resistance	163
Table A.2.3: Asian soybean rust reaction of Hyuuga and previously reported sources of resistance	164
Table A.3.1: Oligonucleotide sequences for SNP markers associated with fatty acid gene sequences	169
Table A.3.2: Oligonucleotide sequences for SSCP markers from genes in the fatty acid biosynthetic pathway	170

Table A.3.3: Sequences from genes in the fatty acid biosynthetic pathway from other

species 171

LIST OF FIGURES

	Page
Figure 2.1: Fatty acid synthesis and glycerolipid synthetic pathways in soybean	59
Figure 3.1: Pedigree of N00-3350	90
Figure 3.2: QTL likelihood plots from interval mapping for oleic acid QTL using 259 lines from the G99-G725 $\times$ N00-3350 population	91
Figure 3.3: Composite interval mapping for oleic acid QTL using 259 lines from G99- G725 $\times$ N00-3350 and the combined oleic acid content	92
Figure 4.1: Outline of the fatty acid biosynthetic pathway	121
Figure 4.2: Genetic linkage map of the G99-G725 $\times$ N00-3350 $F_{2:3}$ population in linkage groups with oleic acid QTLs	122
Figure 4.3: Genetic linkage map for LG-D1b, LG-K, and LG-O of the G99-G725 $\times$ N00- 3350 $F_{2:3}$ population	123
Figure 5.1: Genetic linkage map of a region of soybean LG-C2 containing a locus associated with the type of lesions (tan, red-brown, or mixed) caused by Asian soybean rust.	145
Figure 5.2: QTL likelihood plots for ASR severity and lesion number from the Dillon $\times$ Hyuuga RILs	146
Figure A.2.1: Molecular mapping of the SNP marker BARC-010457-00640	165
Figure A.2.2: Graphical genotypes of Dillon $\times$ Hyuuga RILs	166
Figure A.2.3: Graphical genotypes of Benning $\times$ Hyuuga RILs	167
Figure A.2.4: Pedigree of the Brazilian line FT-2.....	168

CHAPTER 1

INTRODUCTION

Somewhere there is something incredible waiting to be known. –Carl Sagan

The world's population continues to grow at an increasing rate, driving agricultural practices towards a more efficient use of the available resources. Progress in our ability to provide adequate amounts of food, fiber, feed, and fuel from domesticated crop plants has been possible due to the successes of agricultural scientists and farmers. Earlier progress in plant improvement has resulted from selection based almost entirely on the phenotype. The role of science and technology in improving our food supply increases as new technology becomes available and is more widely used (Stuber et al., 1999). Plant breeders can apply new technologies as well as use the information that becomes available to improve a crop's tolerance to abiotic stresses, keep up with pathogen evolution, maintain and increase yields, and alter nutrient components based on changes in consumer preferences.

Historically, agriculture has been influenced by the movement of seeds and plants from one area of the world to another. Countries that produce certain crops are often not within the same geographical area from which these crops originated (Fehr, 1993). Vavilov (1951) originally proposed eight centers for the origin of crop species. Harlan (1975) later proposed three centers (Near East, China, and Mesoamerica) and three non-centers (Africa, Southeast Asia and Pacific Islands, and South America). Information on taxonomy, genetics, ecology, geography, history, geology, and paleobotany is needed to determine the origin and dispersal of cultivated plants. The domestication of plants may have occurred in their center of origin or in other parts of the world (Fehr, 1993).

ORIGIN

Cultivated soybean, *Glycine max* (L.) Merr., is believed to have originated in China (Hymowitz and Newell, 1981). Soybean is self-pollinated and is propagated commercially by seed (Fehr, 1989). The domestication process is

believed to have taken place during the Shang dynasty (1500-110 B.C.) or maybe earlier. Evidence suggests that the domesticated soybean emerged sometime after that in the eastern half of northern China during the Zhou dynasty. By the first century AC it is believed that soybean reached central and south China as well as the Korean peninsula (Hymowitz, 1970). Chinese legend says that Emperor Shen Nong, the Father of Agriculture and Medicine, reported the first use of soybean in a herbal concoction. Between the first century and the 15th century, sea and land trade routes became established, and tribes from China began migrating. The migration and acceptance of soybean seed as a stable food, promoted the introduction of soybean to Japan, Korea, Indonesia, the Philippines, Vietnam, Thailand, Malaysia, Myanmar, Nepal, and northern India, where land races eventually developed, making these regions a secondary gene center (Hymowitz, 1990; Hymowitz and Newell, 1980). Since soybean's domestication, individual farm families have continuously grown and selected the crop for specific traits, giving rise to specific land races that have been developed in East Asia (Hymowitz, 2004).

Samuel Bowen brought soybean from China to North America in 1765 and asked Henry Yonge, the surveyor General of the Colony of Georgia, to plant soybean on Bowen's farm near Savannah, GA (Hymowitz, 2004). Another early introduction of soybean to North America was by Benjamin Franklin. During 1770, he sent seeds to a botanist named John Bartram, who planted them in his garden, near Philadelphia, PA (Hymowitz and Harlan, 1983). In 1851, soybean reached Illinois and spread through the "Corn Belt" (Hymowitz, 1987).

The genus *Glycine* is divided into two subgenera: *Glycine* and *Soja* (Moench) F. J. Herm. (Hymowitz and Newell, 1981). *Glycine max* (L.) Merr., is a true domesticate in that it would not exist without human intervention. Cultivated soybean is an annual domesticated crop (Hymowitz, 2004). Soybean is morphologically variable, as can be seen from the variation among land races from East Asia. These land races are a valuable source of genetic diversity maintained in germplasm collections. Evolutionary studies and genome analysis suggest that soybean [*G. max* subgenus *soja*] is an ancient tetraploid, which later

became diploidized (Hadley and Hymowitz, 1973). Segmental duplication has been detected in several regions of soybean chromosomes, and is believed to have contributed to the duplicated nature of the soybean genome. The subgenus *soja* is believed to have experienced an additional round of genome duplication, and has been referred to as an “ancient polyploid” (Soltis et al., 1993). Restriction fragment length polymorphism (RFLP) marker data show that large areas of the soybean genome have undergone genome duplication in addition to the previous suggested tetraploidization event (Shoemaker et al., 1996).

SOYBEAN PRODUCTION

Soybean is a major commodity traded in world markets and is currently the world's primary oilseed crop (Sonka et al., 2004). Soybean is grown commercially in more than 35 countries, but most of the production occurs in the USA, Brazil, Argentina, and China (Fehr, 1989; Wilcox, 2004). Soybean is a major economic crop in North America, Europe, and in South America. In the last 50 years, the USA has been the world's leading producer of soybean, with over 75 million megagrams (Mg) of soybean produced on average during 2000/2002. As of 2002, the USA was still the largest producer and exporter of whole soybean worldwide (FAS, 2002). Brazil is the largest producer of soybean in South America, with 31 million Mg produced in 1998/1999. In 2003, Brazil contributed 26.8% of the world's soybean production (Wilcox, 2004). Increased use of soybean for livestock feed, meal, and vegetable oil has stimulated an increase in soybean production (Hatje, 1989).

SOYBEAN USES

In the USA, soybean was grown primarily as a forage crop until 1941, when the number of hectares of grain harvested first exceeded the area harvested for forage. Since then, the area grown as forage has declined, and today, the crop is grown almost exclusively for its seed. Currently, soybean is grown mainly for its protein and oil content. Soybean seed contains about 40%

protein and 20% oil (Fehr, 1987), and the levels of these components are negatively correlated (Diers et al., 1992; Lee et al., 1996; Chung et al., 2003). Soybean protein is used primarily as a livestock feed, but is also important for many food products and industrial applications. The oil is used for human consumption as margarine, shortenings, and other fat and oil products, as well as nonfood applications (Fehr, 1987; Glaudemans et al., 1998). The 176 million Mg of soybean produced in 2001 was 35% of the world total oilseed production (Wilcox, 2004). The fatty acid composition of soybean is related to the flavor, stability, and nutritional value of the oil (Mensik et al., 1994). The predominant fatty acids in soybean are palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid (Töpfer et al., 1995). Current soybean cultivars contain 160 to 280 g kg⁻¹ oleic acid (USDA, ARS, National Genetic Resources Program, 2004). Research priorities to target fatty acid profiles with the greatest market for expansion have been described. These include soybean oil with high oleic and low linolenic acid content used for cooking and baking, oil with much higher oleic acid concentration for use in lubricant manufacturing, hydraulic oil base stocks, and soy diesel, and soybean oil with an increased amount of long chain polyunsaturated fatty acids as dietary supplements (Kinney, 2004; Wilson, 2004).

Increasing oleic acid content of soybean oil would result in a decrease of the total saturated fatty acid content and reduce the need for hydrogenation, which is used to improve the oxidative stability of the oil (Hayakawa et al., 2000). Industrial hydrogenation can be relatively expensive, and produces *trans* fatty acids, which have a reduced nutritional value (Wilcox et al., 1984). Soybean breeders are working to increase the amounts of desirable fatty acids, such as oleic acid, and reduce the levels of saturated fats and *trans* isomers of unsaturated fatty acids (Pantalone et al., 2004).

GERMPLASM RESOURCES AND VARIATION

Modern soybean cultivars were developed from a narrow genetic base (Carter et al., 2004). Pedigree analysis determined that 80% of the genes found in public soybean cultivars released between 1947 and 1988 were derived from

13 ancestral lines (Gizlice et al., 1996). Analysis of soybean cultivars using RFLPs generally detects only two alleles at most loci (Keim et al., 1989). In contrast, a group of 20 inbred lines of maize (*Zea mays* L.) was found to average 4.5 RFLP alleles for each locus (Melchinger et al., 1990). Breeding has reduced the genetic diversity among elite breeding lines and cultivars relative to that among the founding ancestors (Gizlice et al., 1993).

Pedigree analysis has shown that northern germplasm (cultivars from Canada and the northern USA), originated from a different genetic base than cultivars from the southern USA (southern germplasm) (Gizlice et al., 1993). The separation of northern and southern elite germplasm has been shown by RFLP analysis of a selected number of elite lines (Keim et al., 1992).

SOYBEAN DISEASES

The health of soybean plants is important for profitable production. The management of soybean diseases is facilitated by planting cultivars that possess resistance to the prevalent pathogens in a particular region. Some soybean diseases are widely distributed, while others are geographically limited (Hartman et al., 1999; Wyllie and Scott, 1988). Fungi, nematodes, viruses, bacteria, and phytoplasmae are known to cause diseases in soybean (Grau et al., 2004; Tolin and Lacy, 2004; Niblack et al., 2004). Worldwide, the most important disease-causing nematodes in soybean are the soybean cyst nematode, *Heterodera glycines* Ichinohe, and some of the root-knot nematodes, *Meloidogyne* spp. (Koenning et al., 1999). Examples of viral diseases that affect soybean include members of the *Tospovirus*, *Potyvirus*, *Cucumovirus*, *Comoviridae*, and *Bromoviridae* families (Tolin and Lacy, 2004). Two of the most prominent bacterial diseases in soybean include bacterial pustule caused by *Xanthomonas campestris* pv *glycines*, and bacterial blight caused by *Pseudomonas syringae* pv. *glycinea* (Kennedy and Sinclair, 1993).

Significant soybean disease problems are caused by more than 40 fungal pathogens worldwide (Hartman et al., 1999). Fungal species of *Pythium*, *Fusarium*, *Macrophomina*, *Rhizoctonia*, *Diaporthe*, *Sclerotinia*, *Phakospora*, and

others are pathogenic on soybean (Grau et al., 2004). Fungal pathogens vary greatly in their associated soybean yield loss and frequency of occurrence (Wrather et al., 1995). Soybean rust is a devastating disease of soybean, which can cause considerable losses in yield (Ogle et al., 1979). *Phakopsora pachyrhizi*, the causal agent of 'Asian soybean rust', is more aggressive than *P. meibomia* (Sinclair and Hartmann, 1999). *P. pachyrhizi* is considered a potential threat to the food supply and its presence in the USA was predicted to have serious repercussions throughout the economy (Kuchler et al., 1984). Prior to 2004, ASR was not present in the continental USA. Therefore, assessment of the effects of soybean rust on U.S. soybean cultivars was performed only in Bio Safety Level 3 containment facilities or in countries where ASR was already established. In November 2004, the disease was first reported in the USA in plots at a Louisiana State Univ. research station near Baton Rouge (Schneider et al., 2005). Since its initial detection, it has been found on soybean in Florida, Georgia, Alabama, North Carolina, South Carolina, Tennessee, Arkansas, Texas, Mississippi, Kentucky, Louisiana, Missouri, Virginia, Indiana, and Illinois (USDA, 2006). Fungicides are available for control of some fungal pathogens in soybean, but economic considerations and environmental concerns limit their application (Grau et al., 2004). Therefore, breeding strategies for resistant varieties remains a viable alternative to the use of fungicides.

SOYBEAN OIL

Soybean oil used for human consumption is subject to U.S. Food and Drug Administration (FDA) guidelines for health claims made on labels. FDA labeling regulations in accordance with the Nutritional Labeling and Education Act of 1990 requires that a "low-saturated" vegetable oil have less than 7% total saturated fatty acids (US Food and Drug Administration, 1999). Currently, soybean oil contains about 15% saturated fat (Wilson et al., 2002). The process of hydrogenation is used to improve the oxidative stability of soybean oil, but in the process creates *trans* fatty acids, which have undesirable health effects (Mazur et al., 1999). These findings have prompted the FDA to require that, in

addition to the saturated fat content, information on the amount of *trans* fatty acids must be included on the Nutrition Facts panel of a product's label (FDA, 2004). The effective date of this labeling mandate was January 1, 2006 (CFSAN, 2003). A viable alternative to hydrogenation would be to produce soybean oil with a more favorable fatty acid composition (i.e., 500 to 550 g kg⁻¹ oleic acid and less than 30 g kg⁻¹ linolenic acid) (Wilson et al., 2002). Researches at the USDA-ARS in Raleigh, NC developed the line N98-4445A, which contains 500 to 600 g kg⁻¹ oleic acid content (Burton et al., 2006). N00-3350 is a single plant selection from the mid-oleic acid line N98-4445A.

PROJECT OBJECTIVES

The overall objectives of this research were: i) to map and confirm quantitative trait loci (QTLs) associated with oleic acid content in N00-3350 soybean, ii) develop and map sequence-based molecular markers from fatty acid pathway genes and sequence-tagged sites in regions of the soybean genome previously associated with oleic acid QTLs, and iii) map and verify QTLs associated with Asian soybean rust resistance caused by *Phakopsora pachyrhizi*.

REFERENCES

- 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.* 46:1010-1011.
- Carter, T.E., R.L. Nelson, C.H. Sneller, and Z. Cui. 2004. Genetic diversity in soybean. p. 303-416. *In* H.R. Boerma, J.E. Specht (ed.) *Soybeans: Improvement, production and uses*. 3rd ed. ASA, CSSA, and SSSA. Madison, WI, USA.
- CFSAN, 2003. Center for Food Safety and Applied Nutrition. Office of Nutritional Products, Labeling, and Dietary Supplements. Available at <http://www.crohns.org/governments/cfsan.htm> (Verified 28 Oct. 2006).
- Chung, J., H.L. Babka, G.L. Graef, P.E. Staswick, D.J. Lee, P.B. Cregan, R.C. Shoemaker, and J.E. Specht. 2003. The seed protein, oil, and yield QTL on soybean linkage group I. *Crop Sci.* 43:1053-1067.
- Diers, B.W., P. Keim, W.R. Fehr, and R.C. Shoemaker. 1992. RFLP analysis of soybean seed protein and oil content. *Theor. Appl. Genet.* 83:608-612.
- FAS, 2002. FOP 12-02 Oilseeds: Markets and Trade. Foreign Agricultural Service, U. S. Department of Agriculture, Washington, D.C. December 2002.
- FDA. 2004. Food labeling and nutrition. U.S. Food and Drug Administration. Available at <http://www.cfsan.fda.gov/~dms/lab-cat.html#transfat> (Verified 28 Oct, 2006).
- Fehr, W.R. 1987. Soybean. p. 533-576. *In* W.R. Fehr (ed.) *Principles of Cultivar Development*. Vol. 2: Crop Species. Macmillan Publishing Company, USA.
- Fehr, W.R. 1989. Soybean. p. 283-300. *In* G. Robbelen, R.K. Downey, and A. Ashri (ed.) *Oil crops of the world*. McGraw-Hill Publishing Co. New York, NY, USA.
- Fehr, R.W. 1993. *Principles of cultivar development*. Vol. 1: Theory and technique. Macmillan Publishing Co., New York, NY, USA.

- Gizlice, Z., T.E. Carter, Jr., and J.W. Burton. 1993. Genetic diversity in North American soybean: I. Multivariate analysis of founding stock and relation to coefficient of parentage. *Crop Sci.* 33:614-620.
- Gizlice, Z., T.E. Carter, Jr., Gerig, T.M., and J.W. Burton. 1996. Genetic diversity patterns in North American public soybean cultivars based on coefficient of parentage. *Crop Sci.* 36:753-765.
- Glaudemans, H.D., M.M. Timmermans, and H. Rijkse. 1998. The world of edible oils. Rabobank International Marketing, Utrecht, the Netherlands.
- Grau, C.R., A.E. Dorrance, J.S. Bond, and J.S. Russin. 2004. Fungal diseases. p. 679-763. *In* H.R. Boerma, J.E. Specht (ed.) Soybeans: Improvement, production and uses. 3rd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- Hadley, H.H., and T. Hymowitz. 1973. Speciation and cytogenetics. p. 97-116. *In* B. E. Caldwell (ed.) Soybeans: Improvement, production and uses. American Society of Agronomy. Madison, WI, USA.
- Harlan, J.R. 1975. Crops and Man. American Society of Agronomy. Madison, WI, USA.
- Hartman, G.L., J.B. Sinclair, and J.C. Rupe (ed.) 1999. Compendium of soybean diseases. 4th edition. APS Press, St. Paul, MN, USA.
- Hatje, G. 1989. World importance of oil crops and their products. p. 1-21. *In* G. Robbelen, R K. Downey, and A. Ashri (ed.) Oil crops of the world. McGraw-Hill Publishing Co. New York, NY, USA.
- Hayakawa, K., Y.Y. Linko, and P. Linko. 2000. The role of *trans* fatty acids in human nutrition. *European J. Lipid Sci. Tech.* 102:419-425.
- Hymowitz, T. 1970. On the domestication of soybean. *Econ. Bot.* 24:408-421.
- Hymowitz, T. 1987. Introduction of the soybean to Illinois. *Econ. Bot.* 41:28-32.
- Hymowitz, T., and J.R. Harlan. 1983. The introduction of soybean to North America by Samuel Bowen in 1765. *Econ. Bot.* 37:371-379.
- Hymowitz, T. and C.A. Newell. 1980. Taxonomy, speciation, domestication, dissemination, germplasm resources, and variation in the genus *Glycine*.

- p. 251-264. *In* R.J. Summerfield and A.H. Bunting (ed.) *Advances in legume science*. Royal Botanic Garden, England.
- Hymowitz, T. and C.A. Newell. 1981. Taxonomy of the genus *Glycine*, domestication and uses of soybean. *Econ. Bot.* 35:272-288.
- Hymowitz, T. 1990. Soybeans: the success story. p. 159-163. *In* J. Janick and J.E. Simon (ed.) *Advances in new crops*. Timber Press, Portland, OR, USA.
- Hymowitz, T. 2004. Speciation and Cytogenetics. p. 97-136. *In* H.R. Boerma and J.E. Specht (ed.) *Soybeans: Improvement, production, and uses*. 3rd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- Keim, P., R.C. Shoemaker, and R.G. Palmer. 1989. Restriction fragment length polymorphism diversity in soybean. *Theor. Appl. Genet* 77:786-792.
- Keim, P., W. Beavis, J. Schupp, and R. Freestone. 1992. Evaluation of soybean RFLP marker diversity in adapted germplasm. *Theor. Appl. Genet.* 85:205-212.
- Kennedy, B.W., and J.B. Sinclair. 1993. Bacterial diseases. p. 4-8. *In* J.B. Sinclair, and P.A. Backman (ed.) *Compendium of soybean diseases*. 3rd edition. APS Press. St. Paul, MN, USA.
- Kinney, A.J. 2004. Modifying soybeans for health and nutrition. p. 44. *In* Abstracts, Soy 2004, Biennial Conference of the Cellular and Molecular Biology of the Soybean. 8-11 Aug. 2004. Univ. of Missouri, CO, USA.
- Koenning, S.R., C. Overstreet, J.W. Noling, P.A. Donald, J.O. Becker, and B.A. Fortnum. 1999. Survey of crop losses in response to phytoparasitic nematodes in the United States for 1994. *J. Nematol.* 31:587-618.
- Kuchler, F., M. Duffy, R.D. Shrum, and W.M. Dowler. 1984. Potential economic consequences of the entry of an exotic fungal pest: The case of soybean rust. *Phytopathology.* 74:916-920.
- Lee, S. H., M.A. Bailey, M.A.R. Mian, T.E. Carter, E.R. Shipe, D.A. Ashley, W.A. Parrott, R.S. Hussey, and H.R. Boerma. 1996. RFLP loci associated with soybean seed protein and oil content across populations and locations. *Theor. Appl. Genet.* 93:649-657.

- Mazur, B., E. Krebbers, and S. Tingey. 1999. Gene discovery and product development for grain quality traits. *Science* 285:372-375.
- Melchinger, A.E., M. Lee, K.R. Lamkey, and W.L. Woodman. 1990. Genetic diversity for restriction fragment length polymorphisms: Relation to estimated genetic effects in maize inbreds. *Crop Sci.* 30:1033-1040.
- Mensink, R.P., E.H. Temme, and G. Hornstra. 1994. Dietary saturated and *trans* fatty acids and lipoprotein metabolism. *Ann. of Med.* 26:461-464.
- Niblack, T.L., G.L. Tylka, and R.D. Riggs. 2004. Nematode pathogens of soybean. p. 821-851. *In* H.R. Boerma, J.E. Specht (ed.) *Soybeans: Improvement, production and uses*. 3rd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- Ogle, H.J., D.E. Byth, and R.J. McLean. 1979. Effect of rust (*Phakopsora pachyrhizi*) on soybean yield and quality in South-eastern Queensland. *Aust. J. Agric. Res.* 30:883-893.
- Pantalone, V.R., D.R. Walker, R.E. Dewey, and I. Rajcan. 2004. DNA marker-assisted selection for improvement of soybean oil concentration and quality. p. 283-311. *In* R. F. Wilson, H.T. Stalker, E.C. Brummer (ed.) *Legume Crop Genomics*. AOCS Press. Champaign, IL, USA.
- Schneider, R.W., C.A. Hollier, H.K. Whitam, M.E. Palm, J.M. McKemy, J.R. Hernández, L. Levy, and R. Devries-Paterson. 2005. First report of soybean rust caused by *Phakopsora pachyrhizi* in the continental United States. *Plant Dis.* 89:774.
- Shoemaker, R.C., K. Polzin, J. Labate, J. Specht, E.C. Brummer, T. Olson, N. Young, V. Concibido, J. Wilcox, J.P. Tamulonis, G. Kochert, and H.R. Boerma. 1996. Genome duplication in soybean (*Glycine* subgenus *soja*). *Genetics*. 144:329-338.
- Sinclair J.B., and G.L. Hartman. 1999. Soybean Rust. p. 25-26. *In* G.L. Hartman, J.B. Sinclair, and J.C. Rupe (ed.) *Compendium of soybean diseases*. 4th edition. APS Press. St. Paul, MN, USA.
- Soltis, P., J. Doyle, and D. Soltis. 1993. Molecular data and polyploid evolution in plants. p. 177-201. *In* P. Soltis, D. Soltis, and J. Doyle (ed.) *Molecular*

- systematics of plants. Routledge, Chapman and Hall. New York, NY, USA.
- Sonka, S.T., K.L. Bender, and D.K. Fisher. 2004. Economics and marketing. p. 919-948. *In* H.R. Boerma, and J.E. Specht (ed.) Soybeans: Improvement, production and uses. 3rd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- Stuber, C.W., M. Polacco, and M.L. Senior. 1999. Synergy of empirical breeding, marker-assisted selection, and genomics to increase crop yield potential. *Crop Sci.* 39:1571-1583.
- Tolin, S.A., and G.H. Lacy. 2004. Viral, bacterial, and phytoplasmal diseases of soybean. p. 765-819. *In* H.R. Boerma, and J. E. Specht (ed.) Soybeans: Improvement, production and uses. 3rd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- Töpfer, R., N. Martini, and J. Schell. 1995. Modification of plant lipid synthesis. *Science.* 268:681-686.
- US Food and Drug Administration. 1999 (FDA). Center for Food Safety and Applied Nutrition. A food labeling guide- Appendix A: Definitions of nutrient content claims. Available at <http://vm.cfsan.fda.gov/~dms/flg-6a.html> (Verified 28 Aug, 2004).
- USDA, ARS, National Genetic Resources Program. *Germplasm Resources Information Network - (GRIN)*. [Online Database] National Germplasm Resources Laboratory, Beltsville, MD. Available at <http://www.ars-grin.gov/cgi-bin/npgs/html/obvalue.pl?51068> (Verified 28 Oct 2006).
- USDA. 2006. Pest information platform for extension and education. National soybean rust commentary. Available at <http://www.sbrusa.net/> (Verified October 20, 2006).
- Vavilov, N.I. 1951. The origin, variation, immunity, and breeding of cultivated plants. Translated from Russian by K. Starr Chester. *Chronica Botanica*, 13:1-366.
- 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. 2004. World distribution and trade of soybean. p. 1-13. *In* H.R. Boerma and J.E. Specht (ed.) Soybeans: Improvement, production, and uses. 3rd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- 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 (ed.) Lipid biotechnology. Marcel Dekker, Inc. IL, USA.
- Wilson, R.F. 2004. Seed composition. p. 621-677. *In* H.R. Boerma, and J.E. Specht (ed.) Soybeans: Improvement, production, and uses. 3rd edition. ASA, CSSA, and SSSA. Madison, Wisconsin, USA.
- Wrather, J.A., A.Y. Chambers, J.A. Fox, W.F. Moore, and G.L. Sciumbato. 1995. Soybean disease loss estimates for the southern United States, 1974 to 1994. *Plant Dis.* 79:1076-1079.
- Wyllie, T.D., and D.H. Scott (ed.) 1988. Soybean diseases of the North Central Region. APS Press, St. Paul, MN, USA.

CHAPTER 2

LITERATURE REVIEW

OILS AND FATS

World production of oil crops has increased considerably due to an increase in demand for oil and fat. To meet the demands for oils from the world's population, further increases in oilseed yield and quality as well as optimal utilization are necessary. Additionally, the rise in production has been promoted by an increase in demand for dairy and meat products, which increases the requirements for high protein animal feeds that can be provided by oilseed crops (Hatje, 1989). The majority of the world's edible fat production comes from vegetable oils (70%), and the rest comes from animal fats (30%), and marine oils (2%) (Vles and Gottenbos, 1989).

Humans obtain their energy from three major types of nutrients: proteins, fats, and carbohydrates. Fat has the highest available energy (9.4 Kcal g^{-1}), while protein and carbohydrate have 4.6 Kcal g^{-1} (Vles and Gottenbos, 1989). In most countries, 90% of the total non-protein energy in the diet comes from fats and carbohydrates. Lipids in the diet also give meals more flavor, act as carriers for vitamins A, D, E and K, and facilitate absorption of these vitamins. The total dietary fat of an average diet in the U.S. usually consists of more than 50% saturated fatty acids (no double bonds), 40% monounsaturated fatty acids (single double bond), and less than 10% polyunsaturated fatty acids (≥ 2 double bonds). Specific values may vary considerably depending on the individual and availability of the foods (Vles and Gottenbos, 1989).

FATTY ACIDS

Plant lipids serve as an energy reserve, form part of cell and organelle membranes, water-proofing and surface protection, protein modification, internal and external signaling molecules, and act as precursors for other components essential for plant metabolism and defense (Somerville et al., 2000). Plants are a significant source of fatty acids used for food, soaps, lubricants, cosmetics, and paints (Ohlrogge, 1994). By definition, a fatty acid is an aliphatic monocarboxylic acid that

can be liberated by hydrolysis from glycerolipids (Wilson, 1987). A fatty acid consists of a hydrocarbon chain with a methyl group at one end and a carboxyl group at the other. Fatty acids are identified by three characteristics: chain length (number of carbons), number of carbon-carbon double bonds, and the location of the first double bond (Lea and Leegood, 1999). Fatty acids are often described by the use of numbers to indicate the number of carbon atoms and the number of double bonds. For example, 18:2 has 18 carbon atoms and two double bonds. The delta symbol may be added, and it indicates that the numbering is with respect to the carboxyl group. Double bond positions may also be designated with respect to the methyl group, and this is done with the omega symbol. For example, omega 6 indicates that the first double bond is on carbon six counting from the methyl end (Gunstone, 1996).

Saturated fatty acids contain the maximum number of hydrogen atoms, while unsaturated fatty acids contain carbon-carbon double bonds (Table 2.1). Monounsaturated fatty acids have a single carbon-carbon double bond, whereas polyunsaturated fatty acids have two or more carbon-carbon double bonds (Lea and Leegood, 1999). The location of the double bond in the carbon chain affects the physiological action of a fatty acid. In oleic acid, the double bond is between the 9th and the 10th carbon atom.

Fatty acids differ in the number of carbon and hydrogen atoms they contain, which causes differences in the nutritional value of each and in their influence on the characteristics of food products. Surveys of the fatty acid composition of seed oils from different plant species have identified more than 200 naturally occurring fatty acids classified into 18 structural classes (Somerville et al., 2000). These classes are defined by the number and position of the double or triple bonds, and the various functional groups attached to them. The most common fatty acids in most plants, including soybean, belong to a small group of C16 and C18 fatty acids that may contain zero to three double bonds (Somerville et al., 2000).

The traditional names of fatty acids derive from the source plant from which a fatty acid was originally isolated. Palmitic acid is named after the oil palm, *Elaeis guineensis* Jacq., and oleic acid after the olive, *Olea europaea* L., in which oleic acid

is the major oil component (Appelqvist, 1989). Oleic acid is more stable than other unsaturated fatty acids, like linoleic and linolenic acids (Mazur et al., 1999).

VEGETABLE OILS

Some components in vegetable oils are essential for the proper functioning of cells and tissues and the metabolic regulation of vital processes in plants and animals (Vles and Gottenbos, 1989). Vegetable oils are used for different purposes depending on their composition. They are used for cooking, salad dressings, confectionary fats, ice cream, and may also be converted into spreadable form, such as margarine (Sommerfeld, 1983).

Two-thirds of the total worldwide edible oil production is obtained from oilseeds (Hatje, 1989). As of 2000, Argentina and Brazil were the two largest exporters of soybean oil, followed by the USA. China, India, and Iran are the largest importers of soybean oil. European countries, with alternative sources of edible oil, import much less soybean oil than Asian countries (Wilcox, 2004). Soybean, sunflower (*Helianthus annuus* L.), canola and rapeseed (*Brassica napus* L.), and oil palm, account for 73% of all vegetable oils produced (Hatje, 1989). Soybean oil alone accounts for approximately 27% of the world's total edible oil production (FAS, 2002; Carter and Wilson, 1998). Oil palm contributes 26% of the world's vegetable oil, and coconut (*Cocos nucifera* L.) contributes 10%. Peanut (*Arachis hypogaea* L.), and cottonseed (*Gossypium hirsutum* L.), each contribute about 7%, while sunflower and olive each contribute 4% or less world oilseed production (Wilcox, 2004). Global soybean oil consumption has increased at a steady rate of about 1 MMT (million metric tons) per year since 1994. Until recently, growth in consumption of low-saturate oils, such as canola and sunflower, kept pace with soybean. However, world consumption of oils with higher oleic acid content grew at a faster rate of 1.8 MMT per year. It is clear that the world's market for oilseed production is continuously changing. It is becoming more competitive, which establishes a motive to tailor soybean seed oil composition in ways that help expand the utility of this commodity (Wilson, 2004).

The fatty acid composition of different types of vegetable oils is variable, and depends on the cultivar (i.e. genetic factors) and environment in the growing season

of the oilseed (Vles and Gottenbos, 1989). The value to industry of a vegetable oil is dependent on its fatty acid content and its ability to be modified and combined with other compounds. The interest of soybean breeders is on seed components because the seed is a convenient organ to store and transport the oil (Pryde and Rothfus, 1989). Soybean seed is also the easiest part of the plant to harvest mechanically, and has the highest concentration of oil of any plant organ.

Most vegetable oils contain large quantities of unsaturated fatty acids with 18 carbon atoms (Table 2.2). Olive oil has about 750 g kg⁻¹ oleic acid and 100 g kg⁻¹ linoleic acid, peanut oil has 500 g kg⁻¹ oleic acid and 300 g kg⁻¹ linoleic acid, corn (*Zea mays* L.) oil has about 550 g kg⁻¹ linoleic and 300 g kg⁻¹ oleic acid, and wild-type sunflower typically produces 120 to 240 g kg⁻¹ oleic acid and 700 to 820 g kg⁻¹ linoleic acid (Vles and Gottenbos, 1989). High oleic sunflower lines originating from an induced mutation (*Ol₁*), produce up to 800 to 940 g kg⁻¹ oleic acid. The seed specific oleate desaturase (FAD2-1), has reduced transcript levels in high oleic acid lines, and co-segregates with *Ol₁* (Schuppert, 2004). A peanut mutant variety containing as much as 800 g kg⁻¹ oleic acid has been reported, and a reduction in FAD2B transcript levels has been associated with the high oleic acid trait (Jung et al., 2000).

The major problem with vegetable oils that contain high concentrations of linoleic acid, and linolenic acid in particular, is flavor stability (Wilson, 1987). Both of these fatty acids may be oxidized by various chemical or enzymatic mechanisms. Stability of the oil refers to the amount of time before the oil becomes rancid due to oxidation (Mercer et al., 1990). The shelf life of the oil is affected by the relative concentrations of specific fatty acids. Saturated fatty acids are generally less susceptible to oxidation than polyunsaturated fatty acids. Oleic acid, a monounsaturated fatty acid, is less susceptible to oxidation during storage and frying than the polyunsaturated fatty acids, and therefore oil with higher oleic acid content maintains a better quality for a longer period of time (Miller et al., 1987; Mercer et al., 1990; O'Bryne et al., 1997). The secondary reaction products from oxidized polyunsaturated fatty acids generally have characteristic flavors that detract from oil flavor quality, especially in liquid cooking oil. To overcome this problem, the oil may be heat-treated to denature proteins, such as lipoxygenases. Additionally, a variety

of antioxidants is added to the oil to inhibit peroxidation of polyunsaturated fatty acids by free radicals, molecular oxygen, and lipoxygenase. Vegetable oils may be partially hydrogenated to lower the levels of linoleic and linolenic acid in the refined product (Wilson, 1987).

To obtain semi-solid spreadable oils, liquid oils are mixed with higher melting temperature fats obtained through a hardening process. During this process part of the unsaturated fatty acids are converted into saturated fatty acids. Some double bonds in the unsaturated fatty acids change their position or stereochemical configuration from the *cis* to the *trans* form (Sommerfeld, 1983). More details on *trans* fatty acids are presented later in this Chapter.

In regard to nutritional dietary quality for *Homo sapiens* and other mammalian species, linoleic acid is an important metabolic precursor for longer chain fatty acids that play a critical role in maintaining good health. Mammals are able to metabolize linoleic acid, but lack the ability to synthesize it. Thus, linoleic acid is recognized as an essential dietary fatty acid. In humans, it is recommended that linoleic acid should contribute about 1% of the daily dietary caloric intake to satisfy the average minimum requirement for essential fatty acids (Wilson, 1987).

SOYBEAN OIL

Soybean seeds contain significant amounts of protein ($\sim 420 \text{ g kg}^{-1}$) and oil ($\sim 180 \text{ g kg}^{-1}$). Simultaneous increases in protein and oil concentration can proceed only to a limited extent, as most experimental evidence shows protein and oil content to be negatively correlated (Burton, 1987; Chung et al., 2003). However, it is possible to improve the levels of certain seed-composition traits, creating value-added or specialty soybeans for use in many food and nonfood applications (Mensik et al., 1994; Brummer, et al., 1997).

Elite soybean cultivars produce seed that average 90 to 110 g kg^{-1} palmitic, 40 to 60 g kg^{-1} stearic, 180 to 260 g kg^{-1} oleic, 500 to 540 g kg^{-1} linoleic, and 70 to 80 g kg^{-1} linolenic acids (Wilcox et al., 1984; Schnebly and Fehr, 1993; Hui, 1996; Table 2.2). Saturated fatty acid content in seed of U.S. soybean cultivars ranges from 100 to 120 g kg^{-1} for palmitic acid (Hawkins et al., 1983; Burton et al., 1994; Cherry et al., 1985) and from 22 to 72 g kg^{-1} for stearic acid (Hymowitz et al., 1972). Soybean oil

contains 550 g kg⁻¹ linoleic acid and 100 g kg⁻¹ linolenic acid (Vles and Gottenbos, 1989). Advances have been made in the development of soybean lines that are inherently low in linoleic and linolenic acid, and in lipoxygenase activity. Discussions on soybean oil quality generally focus on polyunsaturated fatty acid content. Soybean, like most of the major vegetable oil crops, contains a high level of linoleic acid. Unlike the other edible vegetable oils, soybean also has a high level of linolenic acid (Wilson, 1987).

Factors affecting oil quality are the fatty acid composition, iodine value, ratio of oleic to linoleic acid (o/l ratio), and stability of the oil (Bruner et al., 2001). Increasing the oleic acid content of soybean oil would decrease the total saturated fatty acid content and increase the oil quality for human consumption (Hayakawa et al., 2000). Although linolenic acid is an unsaturated fatty acid, autoxidation leads to undesirable odors and flavors (Crapiste et al., 1999; Mounts et al., 1988). The concentration of linolenic acid content is therefore inversely related to soybean oil flavor quality (Mounts et al., 1978). Oleic acid and linoleic acid levels are negatively correlated (Howell et al., 1972; Burton et al., 1983). Linoleic and linolenic acid are obtained from the desaturation of oleic acid (Howell et al., 1972; Wilson et al., 1981). Biochemical evidence indicates that these two polyunsaturated fatty acids are produced by the consecutive desaturation of oleic acid (Wilson et al., 1980) (Fig. 2.1).

Soybean oil used for human consumption is subject to U.S. Food and Drug Administration (FDA) guidelines for health claims made on labels. FDA labeling regulations in accordance with the Nutritional Labeling and Education Act of 1990 requires that a “low-saturated” vegetable oil have less than 7% total saturated fatty acids (FDA, 1999). Currently, soybean oil contains about 15% saturated fat. Because of this, soybean oil has been losing market share to canola and sunflower oils. A viable solution would be to produce soybean oil with a more favorable fatty acid composition (i.e., 500 to 550 g kg⁻¹ oleic acid and less than 30 g kg⁻¹ linolenic acid) (Wilson et al., 2002). This modification should help make soybean oil more attractive as an ingredient to food manufacturers (Wilson, 2004).

The composition of vegetable oil extracted from soybean seed depends on the maturity of the seed as well as its moisture level. Oils in seed also contain small

amounts of non-fatty acid lipids such as sterols (Appelqvist, 1989). Vegetable oils also contain pigments consisting of carotenoids (considered genuine lipids), and chlorophylls. Other components that affect the quality of the oil are the tocopherols and tocotrienols, which are present in small amounts (Appelqvist, 1989).

Soybean oil is one of the most important sources of oil for industrial purposes due to its low cost and abundance. Soybean oil is used for paints, soaps, as a stabilizer in vinyl plastics, in resins, anticorrosion agents, rubber extenders, in cosmetics, flavors, fragrances, fuels and fuel additives, lubricants, plastics, and in fruit-based soft drinks (Fehr, 1989; Pryde and Rothfus, 1989). Eighty percent of the fat in most margarine and 65% of the fat in most shortening comes from soybean oil (Fehr, 1989; DuPont et al., 1991).

Broad-sense heritability is defined as the proportion of the phenotypic variance for a given trait that is strictly due to additive and non-additive genetic variation. It indicates the relative ease with which different traits can be selected under a given testing regime (i.e. numbers of replications and environments), and can when combined with genetic variance be used for the prediction of selection progress (Hanson, 1963; Fehr, 1987). The heritability estimates for oil content in eight studies range from 51% to 89% (Burton, 1987). Heritabilities for seed oil on an entry-mean basis ranged from 84% to 98%, as measured in an F_5 -derived population of 76 recombinant inbred lines (RILs) from the cross of Asgrow A3733 $\times$ PI437088A (Chung et al., 2003).

ROLE OF FATS IN HUMAN HEALTH

Fat and cholesterol in human blood are obtained from the diet, and are also produced by the body (Vles and Gottenbos, 1989). These compounds exist as non-water-soluble cholesterol esters, free cholesterol, and triacylglycerols, and require phospholipids and proteins to be transported. The particles, consisting of fat and protein are lipoproteins, and are divided into four groups based on their density: i) chylomicrons, ii) very low-density lipoproteins (VLDL), iii) low-density lipoproteins (LDLs), and iv) high-density lipoproteins (HDLs). LDL and HDL are the main carriers of cholesterol in the blood (Vles and Gottenbos, 1989). Tissues in the body need some cholesterol, and most of them, except the liver, take up LDL from the

blood to acquire it. LDLs transport the cholesterol into the arterial walls, and with higher amounts of LDL in the blood, more “bad cholesterol” enters the wall of the artery. In contrast, HDL, the “good cholesterol”, transports the excess cholesterol from the cells (arterial wall) to the liver, where cholesterol can be altered and excreted through the bile (Vles and Gottenbos, 1989). Diets high in saturated fatty acids are believed to contribute to an increase in the serum cholesterol levels in the blood (Mensink et al., 1994). High levels of blood cholesterol have been correlated to an increased risk of coronary heart disease (CHD) (Willet, 1994).

Arteriosclerosis is a degenerative disease of the arteries, characterized by fats, mainly cholesterol esters, deposited in the arterial walls. This reduces the diameter of the lumen, affecting blood flow. Arteriosclerosis is the underlying cause of most heart attacks and strokes (Vles and Gottenbos, 1989). Biochemical, epidemiological, and clinical research has shown that higher levels of LDL are the main risk factor for cardiovascular disease caused by arteriosclerosis. The National Research Council recommends a reduction in the saturated fatty acid intake to less than 10% of calories (Willet, 1994). The World Health Organization (WHO) recommends that the fatty acid composition of total dietary fat should be reduced from 50% saturated fatty acids to 30% or less of the dietary fat intake (WHO, 1982). The recommendations on altering the composition of the fat intake are geared towards reducing the cholesterol levels in the plasma, mainly LDL cholesterol. In doing so, the ratio of HDL to LDL is increased, which is believed to lower the risk of CHD (Vles and Gottenbos, 1989).

High oleic acid oils have health-related effects in that they may reduce the incidence of CHD (Thelen and Ohlrogge, 2002). Studies have shown an association with high oleic acid and lower serum cholesterol levels, particularly with LDL (O’Byrne et al., 1997). There is an increasing interest from consumers and the food industry to obtain vegetable oil with high oleic acid and low polyunsaturated fatty acid content (Rahman et al., 2001).

HYDROGENATION AND TRANS FATS

Catalytic hydrogenation of unsaturated lipids, which involves the addition of hydrogen atoms to unsaturated sites on fatty acids, eliminates the double bonds

between carbons. This process is used to improve the oxidative stability of the oil, but in the process creates *trans* fatty acids, which have undesirable health effects (Mazur et al., 1999). *Trans* fatty acids are found in nature either as intermediates in biochemical pathways, or in some seed oils and in fats from ruminant animals, which produce them due to bacterial hydrogenation of dietary unsaturated fatty acids in the rumen (Sommerfeld, 1983).

Hydrogenated soybean oils typically contain lower levels of polyunsaturated fatty acids (PUFA), and the resulting oil has increased oxidative stability. This increases the shelf life of fats and foods that contain them, compared to soybean oils that are only refined, deodorized, and bleached. Yet this apparent solution may give rise to another problem. Initially, the bonds of unsaturated fatty acids in crude vegetable oils are found predominately in a *cis* configuration that introduces a natural bend in the molecule. During “partial hydrogenation” some double bounds may be rearranged so that the hydrogen atoms end up on different sides of the chain, in a configuration called “*trans*” (Wilson, 2004).

The intake of *trans* fats reduces serum levels of HDL and increases the levels of LDL (Willet, 1994). Studies have shown positive associations between levels of *trans* fatty acids, elevated blood levels of LDL, and risk for CHD (Willet, 1994; Hu et al., 1997; Lichtenstein et al., 1999). An increased intake of saturated fat and *trans* unsaturated fat in the diet is associated with an increased risk of CHD (Hu et al., 1997). These findings have prompted the FDA to require that the saturated fat content, and the amount of *trans* fatty acids be listed on the Nutrition Facts panel of a product’s label (FDA, 2004). The effective date of this labeling mandate was January 1, 2006 (CFSAN, 2003). The needs for soybean oil with a lower saturated fat content, and with low levels of *trans* fats are important breeding objectives in the USA (Töpfer et al., 1995; Wilson et al., 2002).

LIPID SYNTHESIS IN PLANTS

Although the fatty acid composition of various soybean plant organs and cultured cells from those organs may be different, the biochemical mechanism for fatty acid synthesis is highly conserved in plant tissues (Harwood, 1988). The fatty acid biosynthetic pathway has been characterized and many of the plant genes

underlying lipid synthesis have been cloned and sequenced (Somerville et al., 2000; Töpfer and Martini, 1994). Fatty acid synthetase (FAS) is a multi-enzyme complex that catalyses the addition of two-carbon fragments from malonyl-CoA to an initial molecule of acetyl-CoA. The first step in fatty acid biosynthesis in the plastid is the formation of malonyl-CoA from acetyl-CoA in a reaction catalyzed by acetyl-CoA carboxylase, ACCase (Harwood, 1988) (Fig. 4.1). In soybean, plastidic ACCase plays an important role in the accumulation of lipids in developing seeds (Kozaki, 1999). Plastidic ACCase contains three functional components: biotin carboxyl carrier protein (BCCP), biotin carboxylase (BC), and a carboxyltransferase (CT) complex (Sugimoto et al., 1989).

To initiate the reaction, the CoA-derivatives are converted to ACP (acyl-carrier protein) derivatives, which are bound to the synthetase complex. Activities from the fatty acid synthase consecutively add two carbon units derived from malonyl-CoA to an acyl chain that is bound to ACP. The final product of the reaction is 16:0-ACP. Palmitoyl-ACP is elongated to C18:0 stearyl-ACP by a specific KAS (ketoacyl-ACP synthase). KAS III is responsible for the initial condensation reaction, and the KAS isoforms I and II, catalyze the subsequent condensation of two carbon subunits. There is now evidence that at least three fatty acid synthetic mechanisms exist in higher plants; 16:0-ACP synthetase which produces 16:0-ACP, 16:0-ACP elongase which produces 18:0-ACP, and 18:0-ACP desaturase which produces 18:1-ACP (Wilson, 1987).

Desaturation is catalyzed by Δ^9 -stearyl-ACP desaturase or SACPD (Δ^9 -DES), which converts stearyl-ACP to oleoyl-ACP. The acyl residues are exported to the cytoplasm and converted to acyl-CoA esters by an acyl-CoA synthetase located in the outer envelope of the plastids. In the endoplasmic reticulum (ER), triacylglycerols are formed by the stepwise acylation of glycerol-3-phosphate (Töpfer et al., 1995). The SACPD gene encodes a soluble enzyme that inserts a double bond at C₉ and converts stearic acid to oleic acid. Two SACPD genes, designated SACPDA and SACPDB have been identified in soybean (Byfield et al., 2006). Stearic acid content in soybean is genetically determined by the *Fas* locus. The gene expression or enzyme activity of SACPD may affect the levels of both stearic and oleic acid in soybean seeds (Byfield et al., 2006).

Fatty acid desaturases are a class of enzymes that catalyze the addition of double bonds into the hydrocarbon chain. Oleoyl-phosphatidyl choline desaturase (Fad2) is needed to convert oleic acid to linoleic acid (Ohlrogge and Browse, 1995). Oleic acid (18:1) is exported from plastids, where it is synthesized, and desaturated to 18:2 and later to 18:3 in the ER (Wilson, 1987). Heppard et al. (1996) reported the isolation of two different cDNA sequences, *Fad2-1* and *Fad2-2*, which encode omega-6 desaturases in soybean. Those sequences were compared with sequences available in the EST database, and were found to be highly homologous to the *Arabidopsis* *Fad2* genes (*Fad2-1* and *Fad2-2*) (Okuley et al., 1994). Coding sequences of these genes were inserted into a vector and used for transformation of an *Arabidopsis fad2-1* mutant. The *Fad2-1* and *Fad2-2* clones were able to complement the *fad2-1* mutant, confirming that both of the genes encode functional microsomal omega-6 desaturases (Heppard et al, 1996).

The Fad3 desaturases introduce the third double bond into linoleic acid to produce linolenic acid (Bilyeu et al., 2003). Soybean possesses three Fad3 genes, *GmFad3A*, *GmFad3B*, and *GmFad3C*. Mutations in two of the three soybean Fad3 genes were accompanied by an increase in linoleic acid content and a reduction in linolenic acid levels (Bilyeu et al., 2005). The fatty acid desaturases are targets for improving the oleic content in plants (López et al., 2000).

BREEDING FOR FATTY ACID CONTENT

Success in drastically altering the fatty acid composition in plants has been achieved with no obvious detrimental effects to the plant (Hammond and Fehr, 1984). It should be possible to significantly modify the proportion of some fatty acids in soybean using traditional plant breeding techniques. However, development of a unique fatty acid composition requires knowledge of the inheritance of the fatty acid composition in soybean. Variability for saturated and unsaturated fatty acid composition in soybean seed has been created using chemical mutagenesis with ethyl methane sulfonate (EMS) or sodium azide, pyramiding of different mutant genes, and genetic engineering (Ohlrogge et al., 1991).

Results obtained from previous studies indicate that some modifications in fatty acid content might be relatively easy to achieve in a breeding program (Khan et

al., 1974; Tai and Young, 1975; Worthington and Hammons, 1971). However, some fatty acids, such as oleic acid and the levels of polyunsaturated fatty acids, are quantitatively inherited in soybean (White et al., 1961; Burton et al., 1983; Carver et al., 1987).

Palmitic Acid

The mean content of palmitic acid in U.S. soybean cultivars is 120 g kg⁻¹ (Fehr et al., 1991), but levels range from 93 g kg⁻¹ to 174 g kg⁻¹ in the *G. max* accessions (USDA, ARS, National Genetic Resources Program, 2004). The soybean lines, N79-2077-12 and N87-2122-4, produce 53 to 60 g kg⁻¹ palmitic acid (Burton et al., 1994). A reduction of palmitic acid content would reduce the amount of saturated fatty acids. Lines N79-2077-12 and N87-2122-4 have reduced palmitate content, and were developed through recurrent selection programs (Burton et al., 1994). C1726 is a breeding line with reduced palmitic acid levels developed by mutagenesis (Erickson et al., 1988a).

Palmitic acid content has been altered by mutations at the *Fap* loci (Erickson et al., 1988b). Other enzymes for which mutations would alter the palmitic acid content in soybean include 3-ketoacyl-ACP synthetase III (KAS-III), 3-keto-acyl-ACP synthetase II (KAS-II), 18:0-ACP desaturase (Δ^9 -DES), 16:0 ACP thioesterase (16:0-ACP TE), 18:1-ACP thioesterase (18:1-ACP TE), glycerol-3-phosphate acyltransferase (G3-PAT), lysophosphatidic acid acyltransferase (LPAAT), and diacylglycerol acyltransferase (DGAT). Modifier genes (i.e., genes with minor effects) associated with palmitic acid biosynthesis may also affect production. For example, the reduced palmitic acid line, N87-2122-4, is believed to have a major gene and a modifier gene (Rebetzke et al., 1998). Minor genes that can change the palmitic acid content of soybean oil by 2 to 23 g kg⁻¹ have been reported (Graef et al., 1988; Horejsi et al., 1994).

Stearic Acid

Stearic acid (18:0) levels in soybean average 30 g kg⁻¹ of the crude oil (USDA, 2004). Stearic acid content has been altered by mutations at the *Fas* locus induced by X-ray or chemical mutagenesis (Rahman et al., 1997). Some *fas* alleles

are believed to be associated with a poor yielding ability, and have been an obstacle in developing soybean cultivars with higher stearic acid (Wilson et al., 2002). A strong negative correlation between stearic and oleic acid has been found (Pantalone et al., 2002). High stearic acid content may be due to a reduced $\Delta 9$ -DES or 18:1-ACP TE activity. The loci for high palmitic and stearic acids are independently inherited. The combined effects of high palmitic and high stearic acid content were associated with a reduction in oleic and linoleic acid content. High stearic acid lines containing 300 g kg⁻¹ have been developed, but the seeds were irregular and failed to grow into plants after germination (Rahman et al., 2003).

Oleic Acid

Seeds of current soybean cultivars possess 180 to 240 g kg⁻¹ oleic acid (Wilcox et al., 1984; Diers et al., 1992; USDA, ARS, National Genetic Resources Program, 2004). It has been reported that higher oleic acid content does not negatively affect yield (Carver et al., 1986). Increases in oleic acid content in soybean using recurrent selection have been correlated with an increase in seed size and early maturity, while seed yield, seed oil, and seed protein content seemed to remain constant. Scientists with the USDA-ARS at Raleigh, NC, developed the germplasm line N78-2245, which produces 510 g kg⁻¹ oleic acid (Wilson et al., 1981; Wilson et al., 2002). The germplasm line, N98-4445A, contains 500 to 600 g kg⁻¹ oleic acid content (Burton, 2006). N00-3350 is a single plant selection from N98-4445A. The following lines are part of its pedigree: C1726, which is a low-palmitic line derived by mutagenesis from the cultivar Century; N79-2077-12, which was selected for increased oleic acid and reduced palmitic acid content; and N87-2122-4, which is a low-palmitic germplasm derived from a cross between N78-2245 and N79-2077 (Burton et al., 1994; Burton et al., 2006). N97-3363-4 is a breeding line that contains 600 g kg⁻¹ oleic acid. It is presumed that this line contains mutations in two different alleles at *Fad* loci, possibly in different *Fad2* genes (Wilson et al., 2002). The oleic acid mutant, M23, contains approximately 500 g kg⁻¹ oleic acid (Rahman et al., 1994).

Studies with *Arabidopsis* have shown that an increase in oleic acid content was accompanied with a reduction of cold tolerance during seed germination,

possibly due to the lack of unsaturated fatty acids in the membranes of these plants (Miquel and Browse, 1994). It is possible to obtain high oleic soybean oil by suppressing delta-12 desaturase, which adds a second double bond to oleic acid to form linoleic acid. The result is a high oleic acid content, with a reduction in linoleic and linolenic acid contents (Diers and Shoemaker, 1992).

Linoleic and linolenic acid

Breeding programs have been initiated to develop soybean cultivars with lower levels of linolenic acid (i.e., 35 g kg⁻¹). Initial efforts have been only partially successful because low linolenic acid content was not maintained in advanced generations, and a significant environmental effect on the linolenic acid content was observed (Hymowitz and Singh, 1987).

At least three independent genetic loci, designated *Fan*, are associated with linolenic acid levels in soybean seed (Wilcox and Cavins, 1987; Fehr et al., 1992; Rahman and Takagi, 1997; Ross et al., 2000; Byrum et al., 1997). Studies indicate that mutations at the *Fan* loci alter the conversion of linoleic to linolenic acid. A mutation in N78-2245 affected ω -6 desaturase activity, and a mutation in PI123440 affected ω -3 desaturase activity. The breeding line N85-2176 was developed from a cross between N78-2245 $\times$ PI123440, and its seed contains 35 g kg⁻¹ linolenic acid (Wilson, 1987). Crosses of N87-2122-4 with PI342434 or PI424031 confirmed that ω -6 desaturase affected the oleic acid concentration, and epistatic gene interactions were identified (Wilson et al., 2002). Mutation breeding was used to develop the soybean line, RG10, which has a linolenic acid content < 25 g kg⁻¹ (Stojšin et al., 1998).

Accumulation of linolenic acid happens only in the seeds of plants that have photosynthetically active chloroplasts during their seed development, such as soybean, flax (*Linum usitatissimum* L.), and rapeseed (Thies, 1970). Linolenic acid has been lowered in soybean, but in no case has it been completely eliminated (Fehr et al., 1992; Rahman et al., 1998), indicating that there may be a biological limitation to obtaining lines completely lacking in linolenic acid. Thies (1970) speculated that since linolenic acid is the main fatty acid component in the thylakoid

membranes, plants completely lacking of linolenic acid could not be developed in species with active seed chloroplasts.

TRANSGENICS WITH ALTERED FATTY ACID CONTENT

In soybean, the conversion of oleic acid to linoleic acid is catalyzed in vegetative tissues by the *Fad2-2* gene product, while *Fad2-2* and *Fad2-1* gene products carry out this same reaction during seed development (Heppard et al., 1996). At the DNA level, *Fad2-2* and *Fad2-1* are 73% identical. Therefore, seed-specific targeting of the *Fad2-1* gene for post-transcriptional gene silencing (PTGS) down-regulates both *Fad2-2* and *Fad2-1* (Vance and Vaucheret, 2001). Transgenic approaches using oleate desaturase (*Fad2*) DNA suppression increased oleic acid content from 200 g kg⁻¹ to about 800 g kg⁻¹ without compromising fatty acid profiles in the vegetative tissues (Yadav, 1995; Kinney, 1996; Kinney and Knowlton, 1997). Transgenic soybean with down-regulated omega-6 desaturase activity and palmitoyl-thioesterase activity (*FatB* gene) exhibited elevated oleic acid and reduced palmitic acid in seed storage lipids (Kinney, 1996).

Yield reductions in transgenic crops with an altered fatty acid content have decreased the initial optimism for producing designer oilseeds as the complexity of the metabolic pathways involved in seed oil biosynthesis has been recognized (Murphy, 1999). The presence of multiple genes can lead to instability of gene expression of transgenes due to RNAi effects (Murphy, 1999). Studies have shown that enzymes further downstream in the metabolic pathways play key roles in regulating the channeling of fatty acids to storage and determining their overall content in the seed (Kinney, 1998). Additionally, the accumulation of high quantities of a given fatty acid in transgenic lines can lead to instability of cell membranes, and protective mechanisms for the breakdown of the novel fatty acid may be activated. An increased appreciation of the importance of fatty acids for storage, as structural components, and as signaling molecules that regulate plant development has resulted (McConn et al., 1997). Although many of the genes encoding enzymes for storage lipid biosynthesis have been cloned, unexpected results have been obtained when these genes are expressed in transgenic plants. Further information is

needed regarding the components of this and other metabolic pathways and their interactions in the plant.

ENVIRONMENTAL EFFECTS ON VARIATION IN FATTY ACID CONTENT

The environment in which the plant is grown affects the degree of polyunsaturated fatty acid synthesis in soybean seed. Temperature is thought to be the major environmental factor that influences the amounts and relative proportions of fatty acids. Developing soybean seeds exposed to high daily temperatures possess higher oleic acid and lower linoleic and linolenic acid contents, and plants exposed to cool temperatures have lower oleic acid and higher levels of linoleic and linolenic acid (Rennie and Tanner, 1989; Burton et al., 1983). Studies evaluating the levels of palmitic and stearic acid at different temperatures indicate that these remained unchanged (Burton et al., 1983; Wolf et al., 1989; Rennie and Tanner, 1989). The alteration of fatty acid composition in response to temperature shifts is relatively fast. Changes in the fatty acid composition of higher plants allow them to acclimate to high or low temperatures. Presumably, increased polyunsaturation of the membrane glycerolipids increases membrane fluidity and may enhance cell function at low temperatures (Wilson, 1987; Kirsch et al., 1997; Hamada et al., 1996; Yamamoto et al. 1992). It has been suggested that as the temperature increases, oxygen becomes less soluble in the cytoplasm and dehydrogenation would not occur as often as it would under cooler temperatures (Wolf et al., 1989).

The enzymes involved in the pathway for fatty acid synthesis are directly affected by temperature. The activities of stearyl, oleoyl, and linoleoyl desaturases are all dramatically altered by changes in growth temperature. Oleoyl and linoleoyl desaturase activities are almost abolished at temperatures higher than 35°C. Linoleoyl and oleoyl desaturases are more active at lower temperatures. The activity of stearyl-ACP desaturase decreased six-fold between 20 and 35°C, while palmitoyl-ACP elongation is mostly unchanged with changes in temperature. It has been suggested that any or all of the desaturases that were evaluated in that study have the potential to be regulatory sites in the pathway (Cheesbrough, 1989).

Arabidopsis mutants for *fad2* are deficient in oleate desaturase activity, and have a decreased concentration of polyunsaturated fatty acids in vegetative tissue

and seed. Exposure of seeds from the mutants to low temperatures negatively affected the germination rates, produced abnormal seedlings, and resulted in reduced accumulation of storage lipids. Mutations that alter seed fatty acid composition produce other changes that may result in plants with less than optimal agronomic performance in the field (Miquel and Browse, 1994).

In plants, the microsomal pathway catalyzed by ω -6 desaturase is the main route for polyunsaturated lipid synthesis (Heppard et al., 1996). The levels of transcripts for *Fad2-1* and *Fad2-2* in transgenic soybean do not increase at low temperature, even though the levels of linoleic and linolenic acid increase as temperature decreases. These results suggest that the increased levels of polyunsaturated fatty acids in plants exposed to low temperatures are not due to enhanced expression of these desaturase genes. In developing soybean seeds, the timing of *Fad2-1* gene expression coincides with that of fatty acid biosynthesis and oil accumulation (Heppard et al., 1996). The melting point of unsaturated fatty acids is lower than their saturated counterparts and therefore may provide greater membrane fluidity for plants to maintain membrane function, even under low temperature growing conditions (Neidleman, 1987). The increased levels of polyunsaturated fatty acids of soybean seeds found in low temperatures are likely caused by translational and post-translational regulation (Cheesbrough, 1989).

The line AN145-66, developed from a recurrent selection population, is hypothesized to have several minor genes that condition the oleic acid content, but these levels are not stable across environments (Primomo et al., 2002). Observed differences in oleic acid content of soybean genotypes across different environments were due to changes in magnitude rather than to changes in rank, indicating that selection for oleic acid percentage is not hampered by environmental interactions, and that progress from selection can be made by testing in relatively few environments (Burton et al., 1983). Significant changes in palmitic and stearic acid levels of genotypes across environments have been reported. It has been surmised that the consistency of QTLs across environments is a result of the control of a trait by a few loci with large effects (Lee et al., 1996a). Factors such as planting date, soil type, cultural practices, precipitation, disease, insects or weeds may also affect fatty acid content in soybean seeds (Dornbos and Mullen, 1992; Primomo et al.,

2002). Drought stress during seed filling only had a small effect on the fatty acid composition of soybean oil (Dornbos and Mullen, 1992).

MOLECULAR MARKERS

Strategically, DNA markers are used to identify genetic factors for traits of interest and to introgress them effectively into elite cultivars more easily and more quickly than can be accomplished by traditional breeding approaches. What constitutes useful alleles depend on the specific objectives of the breeder's program. The selection of a molecular marker system for plant breeding depends on the objectives, the structure of the population, the availability of the marker system, and the cost. Different marker systems have different costs, utility across populations and species, and have different ability to detect polymorphisms. Each marker system has advantages and disadvantages (Staub et al, 1996). Molecular markers that have been commonly used in soybean include restriction fragment length polymorphism (RFLP), simple sequence repeats (SSR), and single nucleotide polymorphism (SNP).

RFLPs are detected using restriction enzymes that cut genomic DNA molecules at specific nucleotide sequences or restriction sites, producing DNA fragments of variable sizes when the restriction sites are polymorphic. This type of marker is a co-dominant marker system in which differences in size are detected (Staub et al., 1996). The initial molecular soybean linkage maps were based on RFLP markers. Complex banding patterns detected with some RFLP probes, and a lack of polymorphism of RFLP loci in elite soybean breeding lines and cultivars drove development of simple sequence repeat (SSR), or microsatellite markers for mapping purposes (Akkaya et al., 1992). Additionally, the use of RFLP markers was expensive, time consuming and seldom yielded results fast enough to be useful as a tool for selection.

Microsatellites are PCR-based markers in which the diversity results from variation in the length of repetitive elements. The majority of soybean SSRs were developed as single-locus markers, and many of them have multiple alleles, making them ideal for creating genetic maps and for defining linkage group homology across mapping populations (Song et al., 2004). Previous studies in maize, tomato

(*Lycopersicum* spp.), and rice (*Oryza sativa* L.) have used SSR markers to investigate genetic phenomena such as epistasis, pleiotropy, and heterosis (Edwards et al., 1987; Eshed and Zamir, 1995; Li et al., 1997).

SNPs are single base differences between homologous DNA fragments and include small insertions and deletions (indels) (Zhu et al., 2003). Fourteen genotypes estimated to have contributed 80.5% of the allelic diversity present in North American soybean cultivars were used to determine the SNP frequency in coding and non-coding soybean DNA sequence (Gizlice et al., 1994; Zhu et al., 2003). In general, nucleotide diversity is higher in random non-coding genomic sequences obtained from BAC clones and SSR flanking regions than in genomic DNA associated with genes. In soybean, the nucleotide diversity has been found to be 2.2 times greater in non-coding DNA closely associated with coding sequence, than in coding sequence (Zhu et al., 2003). Overall, molecular markers can be used to identify which genes have mutations that affect changes in fatty acid composition, and where in the genome they occur (Wilson et al., 2002).

QUANTITATIVE TRAIT LOCI (QTL)

Many economically important plant traits, such as yield, are quantitative, which indicates that they are controlled by multiple genes (Stuber et al., 1999). Improvement of an agronomic trait through breeding is difficult when the trait is genetically complex. Identifying individual genetic components or quantitative trait loci (QTLs) that contribute to the trait may simplify the task (Dudley, 1993). There are examples of individual QTLs being resolved into multiple genetic factors by recombination (Graham et al., 1997; Yamamoto et al., 1998). In plant breeding programs, it may not be critical to determine whether the QTL represents a single factor or a cluster of tightly linked genes (Stuber et al., 1999). However, if a specific QTL is important enough to be cloned, then the chromosomal location of the QTL must be narrowed down to a region of manageable size and this process is called “fine-mapping” (Paterson, 1998). Mapping QTLs underlying a quantitative trait is highly dependent on the magnitude of their effect on the phenotype. To identify a QTL with a small effect, a large population size is needed (Lander and Botstein, 1989).

Genetic maps have been useful for soybean genome analysis. Molecular markers and maps have allowed the identification of soybean genes conditioning QTLs (Mansur et al., 1996; Njiti et al., 1998). High-resolution fine mapping was used as an initial step in identifying a yeast artificial chromosome (YAC) containing a fruit weight QTL in tomato (Alpert and Tanksley, 1996). In breeding programs, manipulating large chromosomal segments may be simpler and more effective in the short term than extracting and manipulating individual genes (Stuber et al., 1999). QTLs can be used in breeding programs without fine mapping, but fine mapping and cloning permit a better understanding of whether one gene or a cluster of genes is involved, and the sequence of a cloned gene could allow discovery of similar genes in soybean and other species.

In most studies aimed at identifying QTLs, stringent probability levels are applied so that there is a low risk of making Type I errors, or identifying false positives. Depending on the trait and the population size, most of the QTLs that are detected have major effects. These QTLs would have high heritabilities, are easily manipulated with traditional breeding practices, and may already be fixed in elite breeding lines (Stuber et al., 1999). Marker-based technologies have already proven powerful for identifying and mapping QTLs. Much progress has been made in identifying major QTLs, and if additional progress in improving complex traits is to be made, marker technology should be used to identify and locate QTLs with moderate or minor effects.

Evaluating larger population sizes and using more markers in the area of interest can increase the resolution of QTL mapping due to an increase in the probability of finding the recombination events that will contribute to more accurate mapping of QTLs (Lander and Botstein, 1989). The number and spacing of mapped genetic markers surrounding the QTL determines the size of the interval to which a QTL can be mapped (Paterson et al., 1990). QTL selection bias is a concept that in experiments in which a QTL is detected, the estimated effect of the QTL will be, on average, larger than its true effect. The expected QTL selection bias is greatest in QTLs with small or moderate effects, and the smaller the QTL effect, the larger the bias (Broman, 2002).

Advances in DNA marker technology, including the development of SSR markers and the development of an integrated soybean genetic linkage map, have facilitated the genetic mapping of quantitative traits in soybean (Cregan et al., 1999). The most recent integrated genetic map for *G. max* contains over 1000 mapped SSR markers (Song et al., 2004). Twelve to 29 additional markers per linkage group have been added to the map of Cregan et al. (1999). Some of these new SSR markers were developed from map-referenced bacterial artificial chromosome (BAC) clones in a strategy to target markers to positions in the genome where markers were scarce.

Inconsistency of some QTLs across environments has been revealed (Paterson et al., 1991). These differences may be due to the heritability of the traits being evaluated, the number of QTLs and their effects, and the crop species. The stability of QTL alleles when transferred to different genetic backgrounds and in different environments is still largely unknown. Efficient selection strategies for improved soybean cultivars require an understanding of the amount of genotypic variation and the magnitude of genotype $\times$ environment ($G \times E$) interactions (Falconer, 1989).

Diers et al. (1992) identified RFLP markers associated with oil and predicted that there are probably many more loci that affect oil content, but which had effects too small to be identified in their study. Studies comparing the position of QTLs for seed oil in different populations indicate that there is population specificity of important QTLs for oil content (Lee et al., 1996b). Interval mapping of seed oil using data from different locations identified seed oil QTLs on different linkage groups than the original study, demonstrating the inconsistency of oil QTLs across locations (Lee et al., 1996b).

In validation studies for seed oil QTLs, an independent population was evaluated in two or three different environments, and only two of three QTLs for oil content were confirmed. These were given the designation “*cq*” for loci that have been confirmed (Fasoula et al., 2004). Results from three independent maize experiments repeated in the same genetic background revealed the inconsistency of the identified QTLs (Beavis, 1994; Beavis et al., 1994). To further emphasize this point, over 900 QTLs for the various quantitative traits in soybean are listed in The

Soybean Breeder's Toolbox (<http://soybeanbreederstoolbox.org/>), but limited information is available on the confirmation of the reported QTLs. The results from verification studies have been inconsistent in confirming all previously identified QTLs for oil (Brummer et al., 1997). Conflicting information for certain traits confirming the reported QTLs in soybean reveals the importance of validation experiments prior to developing breeding strategies based solely on unconfirmed QTLs (Fasoula et al., 2004).

MARKER-ASSISTED SELECTION (MAS)

Molecular markers can be employed for the introgression of QTLs with desirable alleles into improved cultivars using marker-assisted selection (MAS) (Tanksley et al., 1989). MAS uses closely linked, easily identified genetic markers that are closely linked to a QTL being introgressed. By selecting for the appropriate allele at this marker, phenotypic evaluation of progeny can be reduced during the early generations or during backcrossing generations required to introduce the QTL allele of interest.

An effective breeding strategy requires that the genomic regions affecting the QTL of interest be determined accurately, and the value and proximity of genes conditioning other important traits linked to the QTL of interest be considered. Often, the evaluation process to determine a phenotype can be expensive, labor-intensive, and time consuming. Identifying surrogate molecular markers linked to QTLs conditioning the trait can simplify the breeding process and increase our understanding of the genetic basis for these quantitative traits (Orf et al., 1999).

Lande and Thompson (1990) showed through theoretical and analytical investigations that the maximum rate of improvement for first-generation selection can be obtained by integrating both phenotypic and marker data. Statistical limitations on the efficiency of MAS include the precision of the associations between marker loci and the QTLs. On a single trait, using a combination of molecular and phenotypic information, the potential selection efficiency depends on the heritability of the trait, the proportion of the additive genetic variance associated with the QTL, and the selection scheme. The relative efficiency of MAS is greatest for characters with low heritability if a large fraction of the additive genetic variance is

associated with the QTL. Some limitations of MAS in plant improvement programs include the level of linkage disequilibrium in the populations, which affect the number of marker loci and QTL resolution needed, and the sample sizes needed to detect QTLs for traits with low heritabilities (Lande and Thompson, 1990).

Knapp (1988) developed the probability theory for selecting one or more superior genotypes using MAS. A breeder using only phenotypic selection must test up to 16.7 times more progeny than a breeder using MAS to be assured of selecting one or more superior genotypes, depending on the level of the selection goal and the selection intensity. Therefore, MAS can considerably reduce the resources needed to obtain the selection goal of a trait with a low to moderate heritability when the selection goal and selection intensity are high (Knapp, 1998).

Marker-assisted selection allows simultaneous gains for a number of traits, even for traits that are difficult or expensive to characterize (Edwards and Johnson, 1994). Results from population selection studies have shown that quantitative traits can be manipulated using only genotypic marker data. Significant gain in progeny performance was obtained in elite sweet corn breeding populations that had been generated by selection based only on marker genotype information (Stuber and Edwards, 1986).

The relative value of an identified QTL is expected to vary in different genetic backgrounds due to epistasis or recombination that would affect the linkage disequilibrium between the marker and the QTL alleles (Reyna and Sneller, 2001). These novel alleles have potential for incorporation into elite breeding lines, but the possibility of incorporating undesirable agronomic traits through 'linkage drag' must be considered. It is possible that the introduction of a gene for disease resistance into an adapted cultivar may alter physiological processes that can adversely affect yield, due to linkage drag. Linkage drag has been implicated as a limitation to the use of non-domesticated germplasm for the introgression of novel alleles due to introgression of undesirable alleles closely linked to favorable genes. The extent of linkage drag is dependent on many factors including the population size, the number of meiotic generations before applying selection, and the genomic location of the locus of interest (Stam and Zeven, 1981).

ASIAN SOYBEAN RUST (ASR)

Plant rusts, caused by Basidiomycetes of the order Uredinales, are some of the most destructive plant diseases. They are known for their damage on grain crops, like wheat, oat, and barley, and on ornamentals, such as carnation (*Dianthus* spp.), and chrysanthemum (*Chrysanthemum* spp.), but they also affect vegetables and field crops like cotton and soybean (Agrios, 1997). Tropical and subtropical regions in the Eastern and Western Hemispheres have to deal with a devastating disease of soybean, soybean rust (Grau et al., 2004). *Phakopsora pachyrhizi* and *P. meibomia* have been found to be the causal agents of rust in soybean. *P. meibomia* is the causal agent of the 'American' rust disease and has a host range of 66 species, including soybean (Sinclair and Hartman, 1999). This species is native to South America and is present on wild and cultivated legumes from Puerto Rico to southern Brazil (Vakili, 1979). *P. pachyrhizi* is the causal agent of the 'Asian' rust, native to the traditional growing areas in the Orient. The Asian soybean rust (ASR) is considered among the top 25 of the 100 most dangerous exotic pests in the world (Ogle et al., 1979). *P. pachyrhizi* can infect and spread from many wild and cultivated hosts, including many garden legumes (Vakili and Bromfield, 1976). The pathogen can infect soybean any time after germination (Bromfield, 1984). It has a wide host range of over 75 plant species, including soybean, cowpea, kudzu, and other legumes (Rytter et al., 1984; Sinclair and Hartman, 1999). At least nine races of *P. pachyrhizi* have been described and soybean cultivars that are available commercially are susceptible to some, if not all races of the fungus (Burdon and Speer, 1984; Sinclair and Hartman, 1999).

Most ASR lesions occur on the leaves where they are restricted by the leaf veins, but they may be found on the petioles and stems. Water-soaked lesions are the first symptoms, which increase in size and become chlorotic as the disease progresses. The color of the lesions may be grayish brown, tan to dark brown, or reddish brown depending on the virulence of the pathogen, the host genotype, and the age of the lesion. Uredia, which is the fruiting structure of the rust fungi in which urediospores are produced, are found primarily on the underside of the leaves, and they increase in number as the disease progresses (Sinclair and Hartman, 1999).

P. pachyrhizi is known as an obligate parasite, and does not survive in dried or decayed tissues or in the soil. Urediospores survive (in resting or dormant stage) less than 2 d under ambient conditions (Ilag, 1977). Symptoms generally appear from middle to late in the season after a prolonged wet, and cool period that allows for infection and sporulation. Epidemics of rust occur when soybean leaves are infected early in the season. Soybean rust is not seed-borne in soybean (Sinclair and Hartman, 1999). Temperatures ranging from 18 to 26.5°C, 6 to 7 h of continual wetness, and a 12-h dew period are ideal conditions for germination of urediniospores, and subsequent host penetration and lesion development. The host epidermis is penetrated directly by urediniospores and produce hyphae that grow and colonize the mesophyll cells. Germination of the urediniospores and host penetration requires moderate temperatures and relatively high moisture levels (Melching et al., 1989). Uredia have been found to produce urediniospores 9 to 10 d after penetration. The success rate for rust lesion development is affected by weather conditions, and spore viability is affected by ultraviolet light. Spores exposed to sunlight are less viable than spores exposed to cloudy conditions. Rust spores are able to maintain viability without moisture for less than 8 d (Melching et al., 1989).

P. pachyrhizi is prevalent in regions such as Brazil, and parts of Africa (Bromfield, 1984). The first reports of *P. pachyrhizi* in the American continent came from the Rio Paraná region in Paraguay and southern Brazil (Miles et al., 2003). Over the past 3 yr, the rust disease has spread throughout South America wherever soybean has been planted (USDA-ARS, 2004). Asian soybean rust was confirmed north of the equator in an area near Cali, Colombia by USDA-ARS and APHIS in 2004. Prior to 2004, *P. pachyrhizi* was not known to occur naturally in the USA (Hartwig, 1986; Schneider et al., 2005). Asian rust is a highly mobile disease, and rain and air currents can quickly spread spores to other plants, and over long distances (Sinclair and Hartman, 1999), posing a serious threat to soybean production in the USA. Tropical cyclones have the potential to transport *P. pachyrhizi* from the northern South American soybean-growing region directly into the southern USA (Isard et al., 2004). Asian soybean rust has already been reported in at least 15 states in the continental United States including Florida,

Georgia, Alabama, North Carolina, South Carolina, Tennessee, Arkansas, Texas, Mississippi, Kentucky, Louisiana, Missouri, Virginia, Indiana, and Illinois (USDA, 2006).

Dramatic yield losses caused by soybean rust in commercial soybean fields have been described and they range from 13 to 80% (Ogle et al., 1979; Yang et al., 1990; Sinclair and Hartman, 1996). Yield reduction results from the production of fewer pods, less seed per pod, and reduced seed weight (Melching et al., 1989; Sinclair and Hartman, 1999). In addition to the direct yield loss caused by the disease, disruption losses, production and distribution inefficiencies, and ripple effects to the feed and food industries can occur beyond the farm gate, with serious repercussions throughout the economy (Kuchler et al., 1984). A USDA Economic Research Service (ERS) report projected that up to \$640 million to \$1.3 billion in net economic losses were expected during the first year of the pathogen's establishment in the USA (Livingston et al., 2004).

Strategies to control plant diseases include interruption of disease cycles through crop rotation, fungicide application, and crop and cultivar selection (Krupinsky et al., 2002). Certain fungicides can reduce rust damage but may not be cost effective since multiple applications are needed to control the rust disease (Sinclair and Hartman, 1999). Additionally, there is concern from fungicide residues left on food crops and exposure to the consumers. Fungicide applications are more effective for short-term management, need to be applied at the right stage of plant development, and must penetrate the canopy, which can't be effectively done in aerial applications resulting in limited control of the disease (Reid Frederick, USDA-ARS, personal communication). Therefore, breeding strategies for resistant varieties remains a viable alternative to the use of fungicides. To achieve this goal, sources of resistance first need to be identified and if the genes conditioning this resistance can be located in the soybean genome, they can be effectively incorporated in elite soybean cultivars with good agronomic performance.

Until 1986, Hartwig reported that all soybean cultivars grown in the USA had been rated as susceptible to soybean rust. Over 95% of the cultivars for which rust resistance has been assessed are highly susceptible (Burdon and Marshall, 1981). A lack of resistance to the virulent races of *P. pachyrhizi* demonstrates the

vulnerability of the soybean crop to this pathogen. Plant resistance that is race-specific to rust has been identified (Bromfield and Hartwig, 1980; Hartwig and Bromfield, 1983), but no cultivars have been developed that have an acceptable level of resistance to all races of *P. pachyrhizi*. Tolerance to *P. pachyrhizi* in soybean has been identified and used to reduce yield losses (Hartman, 1991).

PI200492 ('Komata') has a single dominant allele (*Rpp*) that confers resistance to an Australian rust isolate (McLean and Byth, 1980). PI462312 ('Ankur') has a single dominant allele for resistance (Singh and Thapliyal, 1977). PI230970 was crossed with a susceptible cultivar and segregation ratios suggest a single dominant allele for resistance (Bromfield and Hartwig, 1980). Intercrossing the three sources of resistance and inoculating them with two rust isolates revealed that the dominant alleles for resistance in PI200492, PI230970, and PI462312 were at different loci (Hartwig and Bromfield, 1983). They suggested that the *Rpp* symbol for PI200492 be changed to *Rpp1*, and assigned *Rpp2* and *Rpp3* to the alleles in PI230970 and PI462312. Later studies showed that the cultivar Bing Nan (PI459025) from China had a single dominant resistance allele (*Rpp4*) at a locus different from the other three resistance alleles (Hartwig, 1986). Other lines suspected of having genes for resistance include PI239871A, PI239987B (*Glycine soja*), PI230971, PI459024, 'Taita Kohsiung no. 5', and 'Tainung no. 4' (Sinclair and Hartman, 1999).

Additional sources of potential resistance to *P. pachyrhizi* from the perennial species *Glycine* have been evaluated in Australia. These accessions include members of the species *G. canescens* F. J. Herm., *G. clandestina* Willd., *G. falcata* Benth., *G. latrobeana* (Meissn.) Benth., *G. tabacina* (Labill.), and *G. tomentella* (Hayata). Although the percent of susceptible accessions ranged from 40 to 73% for some *Glycine* spp., *G. tomentella* and *G. tabacina* accessions included 33% and 32% of highly resistant accessions, respectively. These *Glycine* species represent a source of rust resistance genes that could potentially be used to reduce the vulnerability of the soybean crop to this disease. However, compared to using *G. max* accessions, their incorporation would require significant work (Burdon and Marshall, 1981).

Marker-assisted selection strategies can be used to reduce the time needed to select superior lines with resistance to ASR. The use of molecular markers can

also reduce the amount of donor genome incorporated, and allows for progress to be made in breeding ASR resistant cultivars in regions where the disease is not present without the need to screen them with the *P. pachyrhizi* until the final stages of development.

REFERENCES

- Agrios, G.N. 1997. Plant pathology. 4th edition. Academic Press, London, UK.
- Akkaya, M.S., A.A. Bhagwat, and P.B. Cregan. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132:1131-1139.
- Alpert, K.B., and S.D. Tanksley. 1996. High-resolution mapping and isolation of a yeast artificial chromosome contig containing *fw2.2*: A major fruit weight quantitative trait locus in tomato. *Proc. Natl. Acad. Sci. USA.* 93:15503-15507.
- Appelqvist, L.A. 1989. The Chemical Nature of Vegetable Oils. p. 22-37. *In* G. Röbbelen, R.K. Downey, and A. Ashri (ed.) *Oil Crops of the World*. McGraw-Hill Publishing Co. New York, NY, USA.
- Beavis, W.D. 1994. The power and deceit of QTL experiments: Lessons from comparative QTL studies. p. 250-266. *In* Proc. of the 49th Annual Corn and Sorghum Industry Research Conference. ASTA, Washington, D.C., USA.
- Beavis, W.D., O.S. Smith, D. Grant, and R. Fincher. 1994. Identification of quantitative trait loci using a small sample of top-crossed and F₄ progeny of maize. *Crop Sci.* 34:882-896.
- 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.
- Bilyeu, K.D, L. Palavalli, D.A. Sleper, and P.R. Beuselinck. 2005. Mutations in soybean microsomal omega-3 fatty acid desaturase genes reduce linolenic acid concentration in soybean seeds. *Crop Sci.* 45:1830-1836.
- Broman, K. 2002. Introduction to QTL mapping in model organisms. Dept. of Biostatistics. Johns Hopkins University. Available at http://www.biostat.jhsph.edu/~kbroman/teaching/misc/Jax/2002/jax_qtlhandouts.pdf (Verified 25 June, 2004).
- Bromfield, K.R. 1984. Soybean Rust Monograph 2. American Phytopathological Society, St. Paul, MN, USA.
- Bromfield, K.R., and E.E. Hartwig. 1980. Resistance to soybean rust (*Phakopsora pachyrhizi*) and mode of inheritance. *Crop Sci.* 20:254-255.

- Brummer, E.C., G.L. Graef, J. Orf, J.R. Wilcox, and R.C. Shoemaker. 1997. Mapping QTL for seed protein and oil content in eight soybean populations. *Crop Sci.* 37:370-378.
- Bruner, A.C., S. Jung, A.G. Abbott, and G.L. Powell. 2001. The naturally occurring high oleate character in some peanut varieties results from reduced oleoyl-PC desaturase activity from mutation of aspartate 150 to asparagine. *Crop Sci.* 41:522-526.
- Burdon, J.J., and D.R. Marshall. 1981. Evaluation of Australian native species of *Glycine* for resistance to soybean rust. *Plant Disease* 65:44-45.
- Burdon, J.J., and S.S. Speer. 1984. A set of differential *Glycine* hosts for the identification of races or pathotypes of *Phakopsora pachyrhizi* Syd. *Euphytica* 33:891-896.
- 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. 1987. Quantitative genetics: Results relevant to soybean breeding. p. 211- 248. *In* J.R. Wilcox (ed.) *Soybeans: Improvement, production and uses*. 2nd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- Burton, J.W., R.F. Wilson, and C.A. Brim. 1994. Registration of N79-2077-12 and N87-2122-4, two soybean germplasm lines with reduced palmitic acid in seed oil. *Crop Sci.* 34:313.
- Burton, J.W., R.F. Wilson, G.J. Rebetzke, and V.R. Pantalone. 2006. Registration of N98-4445A mid-oleic soybean germplasm line. *Crop Sci.* 46:1010-1011.
- Byrum, J.R., A.J. Kinney, K.L. Stecca, D.J. Grace, and B.W. Diers. 1997. Alteration of the omega-3 fatty acid desaturase gene is associated with reduced linolenic acid in the A5 soybean genotype. *Theor. Appl. Genet.* 94:356-359.
- Carter, T.E., and R.F. Wilson. 1998. Soybean quality for human consumption. p.1-16. *In* A. James (ed.) *Proceedings of the Australian Soybean Conference* 10. CSIRO, Brisbane, Australia.
- Carver, B.F., J.W. Burton, R.F. Wilson, and T.E. Carter. 1986. Cumulative response to various recurrent selection schemes in soybean: oil quality and correlated agronomic traits. *Crop Sci.* 26:853-858.

- 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.
- CFSAN, 2003. Center for Food Safety and Applied Nutrition. Office of Nutritional Products, Labeling, and Dietary Supplements. Available at <http://www.crohns.org/governments/cfsan.htm> (Verified 28 Oct. 2006).
- Cherry, J.H., L. Bishop, P.M. Hasegawa, and H.R. Leffler. 1985. Differences in the fatty acid composition of soybean seed produced in Northern and Southern areas of the USA. *Phytochemistry* 24:237-241.
- Chung, J., H.L. Babka, G.L. Graef, P.E. Staswick, D.J. Lee, P.B. Cregan, R.C. Shoemaker, and J.E. Specht. 2003. The seed protein, oil and yield QTL on soybean linkage group I. *Crop Sci.* 43:1053-1067.
- Crapiste, G.H., M.I. Brevedan, and A.A. Carelli. 1999. Oxidation of sunflower oil during storage. *J. Am. Oil Chem. Soc.* 76:1437-1443.
- Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, T.T. VanToai, D.G. Lohnes, J. 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.
- Diers, B.W., P. Keim, W.R. Fehr, and R.C. Shoemaker. 1992. RFLP analysis of soybean seed protein and oil content. *Theor. Appl. Genet.* 83:608-612.
- Dornbos, D.L., and R.E. Mullen. 1992. Soybean seed protein and oil contents and fatty acid composition adjustments by drought and temperature. *J. Am. Oil Chem. Soc.* 69:228-231.
- Dudley, J.W. 1993. Molecular markers in plant improvements: manipulation of genes affecting quantitative traits. *Crop Sci.* 32:660-668.
- Edwards, M.D., C.W. Stuber, and J.F. Wenderl. 1987. Molecular marker facilitated investigations of quantitative trait loci in maize: I. Numbers, genomic distribution and types of gene action. *Genetics* 116:113-125.

- Edwards, M., and L. Johnson. 1994. RFLPs for rapid recurrent selection. p. 33-40. *In* Proc. of Symposium on Analysis of Molecular Marker Data. Am. Soc. Hort. Sci. and CSSA. Corvallis, OR, USA.
- Erasmus, U. 2000. Fats that heal, fats that kill. Alive Books. Burnaby, BC. Canada. 235 pp.
- Erickson, E.A., J.R. Wilcox, and J.F. Cavins. 1988a. Inheritance of palmitic acid percentages in two soybean mutants. *J. Hered.* 79:465-468.
- Erickson, E.A., J.R. Wilcox, and J.F. Cavins. 1988b. Fatty acid composition of the oil in reciprocal crosses among soybean mutants. *Crop Sci.* 28:644-646.
- Falconer, D.S. 1989. Introduction to quantitative genetics. 3rd edition. Longman Inc., New York, NY, USA.
- FAS, 2002. FOP 07-02, Oilseeds: markets and trade. Foreign Agricultural Service, U.S. Department of Agriculture, Washington, D.C. July 2002.
- Fasoula, V.A., D.K. Harris, and H.R. Boerma. 2004. Validation and designation of quantitative trait loci for seed protein, seed oil, and seed weight from two soybean populations. *Crop Sci.* 44:1218-1225.
- FDA. 1999. Center for Food Safety and Applied Nutrition. A food labeling guide- Appendix A: Definitions of nutrient content claims. Available at <http://vm.cfsan.fda.gov/~dms/flg-6a.html> (Verified 30 Oct 2006).
- FDA. 2004. Food labeling and nutrition. U.S. Food and Drug Administration. Available at <http://www.cfsan.fda.gov/~dms/lab-cat.html#transfat> (Verified 28 Oct 2006).
- Fehr, W.R. 1987. Soybean. p. 533-576. *In* W.R. Fehr (ed.) Principles of Cultivar Development. Vol. 2: Crop Species. Macmillan Publishing Co. Ames, IA, USA.
- Fehr, W.R. 1989. Soybean. p. 283-300. *In* G. Röbbelen, R.K. Downey, and A. Ashri (ed.) Oil crops of the world. McGraw-Hill Publishing Co. New York, USA.
- Fehr, W.R., G.A. Welke, E.G. Hammond, D.N. Duvick, and S.R. Cianzano. 1991. Inheritance of reduced palmitic acid content in seed oil of soybean. *Crop Sci.* 31:88-89.

- 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.
- Gizlice, Z., T.E. Carter, and J.W. Burton. 1994. Genetic base for North American public soybean cultivars released between 1947 and 1988. *Crop Sci.* 34:1143-1151.
- Graef, G.L., W.R. Fehr, L.A. Miller, E.G. Hammond, and S.R. Cianzio. 1988. Inheritance of fatty acid composition in a soybean mutant with low linolenic acid. *Crop Sci.* 28:55-58.
- Graham, G.I., D.W. Wolff, and C.W. Stuber. 1997. Characterization of a yield quantitative trait locus on chromosome five of maize by fine mapping. *Crop Sci.* 37:1601-1610.
- Grau, C.R., A.E. Dorrance, J.S. Bond, and J.S. Russin. 2004. Fungal diseases. p. 679-763. *In* H.R. Boerma, J.E. Specht (ed.) *Soybeans: Improvement, production and uses*. 3rd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- Gunstone, F.D. 1996. *Fatty acid and lipid chemistry*. Blackie Academic and Professional, Glasgow, Scotland.
- Hamada, T., T. Nishiuchi, H. Kodama, M. Nishimura, and K. Iba. 1996. cDNA cloning of a wounding-inducible gene encoding a plastid ω -3 fatty acid desaturase from tobacco. *Plant Cell Physiol.* 37:606-611.
- Hammond, E.G., and W.R. Fehr. 1984. Improving the fatty acid composition of soybean oil. *J. Am. Oil Chem. Soc.* 61:1713-1716.
- Hanson, W.D. 1963. Heritability. p. 125-139. *In* W.D. Hanson and H.F. Robbins (ed.) *Statistical genetics and plant breeding*. Pub. 982. National Academy of Sciences - National Research Council. Washington, DC, USA.
- Hartman, G.L., T.C. Wang, and A.T. Tschanz. 1991. Soybean rust development and the quantitative relationship between rust severity and soybean yield. *Plant Dis.* 75:596-600.
- Hartwig, E.E. 1986. Identification of a fourth major gene conferring resistance to soybean rust. *Crop Sci.* 26:1135-1136.

- Hartwig, E.E., and K.R. Bromfield. 1983. Relationships among three genes conferring specific resistance to soybean rust in soybeans. *Crop Sci.* 23:237-239.
- Harwood, J.L. 1988. Fatty acid metabolism. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 39:101-138.
- Hatje, G. 1989. World Importance of Oil Crops and their Products. p. 1-21. *In* G. Röbbelen, R.K. Downey, and A. Ashri (ed.) *Oil Crops of the World*. McGraw-Hill Publishing Co. New York, NY, USA.
- 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.
- Hayakawa, K., Y.Y. Linko, and P. Linko. 2000. The role of *trans* fatty acids in human nutrition. *European J. Lipid Sci. Tech.* 102:419-425.
- Heppard, E.P., A.J. Kinney, K.L. Stecca, and G.H. Miao. 1996. Developmental and growth temperature regulation of two different microsomal ω -6 desaturase genes in soybean. *Plant Physiol.* 110:311-319.
- Horejsi, T.F., W.R. Fehr, G.A. Welke, D.N. Duvick, E.G. Hammond, and S.R. Cianzio. 1994. Genetic control of reduced palmitate content in soybean. *Crop Sci.* 34:331-334.
- Howell, R.W., C.A. Brim, and R.W. Rinne. 1972. The plant geneticist's contribution toward changing lipid and amino acid composition of soybeans. *J. Am. Oil Chem. Soc.* 49:30-32.
- Hu, F.B., M.J. Stampfer, J.E. Manson, E. Rimm, G.A. Colditz, B.A. Rosner, C.H. Hennekens, and W.C. Willet. 1997. Dietary fat intake and the risk of coronary heart disease in women. *N. Engl. J. Med.* 337:1491-1499.
- Hui, Y.H. 1996. Soybean oil. p. 497-601. *In* Y.H. Hui (ed.) *Bailey's industrial oil and fat products*. Vol. 2. *Edible oils and fat products: Oils and oilseeds*, 5th edition. John Wiley and Sons. New York, NY, USA.
- Hymowitz, T., R.G. Palmer, and H.H. Hadley. 1972. Seed weight, protein, oil and fatty acid relationships within the genus *Glycine*. *Trop. Agric.* 49:245-250.
- Hymowitz, T., and R.J. Singh. 1987. Taxonomy and Speciation. p. 23-48. *In* J.R. Wilcox (ed.) *Soybeans: Improvement, production and uses*. 2nd edition. ASA, CSSA, and SSSA. Madison, WI, USA.

- Isard, S.A., C. Main, S. Redlin, and J.M. Russo. 2004. Weather-based assessment of soybean rust threat to North America final report to APHIS. USDA-ARS, and Department of Crop and Soil Sciences. University of Illinois, Urbana, IL. Available at http://www.aphis.usda.gov/ppq/ep/soybean_rust/ (Verified 28 Oct 2006).
- Jung, S., G. Powell, K. Moore, and A. Abbott. 2000. The high oleate trait in the cultivated peanut (*Arachis hypogaea* L.) II: Molecular basis and genetics of the trait. *Mol. Gen. Genet.* 263:806-811.
- Khan, A.R., D.A. Emery, and J.A. Singleton. 1974. Refractive index as basis for assessing fatty acid composition in segregating populations derived from intra-specific crosses of cultivated peanut. *Crop Sci.* 14:464-468.
- Kinney, A.J. 1996. Development of genetically engineered soybean oils for food applications. *J. Food Lip.* 3:273-292.
- Kinney, A.J., and S. Knowlton. 1997. Designer oils: The high oleic soybean. p. 193-213. *In* S. Harander and S. Roller (ed.) Genetic engineering for food industry: A strategy for food quality improvement. Blackie Academic, London.
- Kinney, A.J. 1998. Manipulating the flux through plant biosynthetic pathways. *Curr. Opin. Plant Biol.* 1:173-178.
- Kirsch, C., K. Hahlbrock, and I.E. Somssich. 1997. Rapid and transient induction of a parsley microsomal Δ^{12} -fatty acid desaturase mRNA by fungal elicitor. *Plant Physiol.* 115:283-289.
- Krupinsky, J.M., K.L. Bailey, M.P. McMullen, B.D. Gossen, and T.K. Turkington. 2002. Managing plant disease risk in diversified cropping systems. *Agron. J.* 94:198-209.
- Kuchler, F., M. Duffy, R.D. Shrum, and W.M. Dowler. 1984. Potential economic consequences of the entry of an exotic fungal pest: The case of soybean rust. *Phytopathology* 74:916-920.
- Lander, E.S., and D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199.
- Lea, P.J., and R.C. Leegood (ed.). 1999. Plant biochemistry and molecular biology. 2nd edition. John Wiley and Sons. New York, NY, USA.

- Lee, S.H., M.A. Bailey, M.A.R. Mian, T.E. Carter, D.A. Ashley, R.S. Hussey, W.A. Parrott, and H.R. Boerma. 1996a. Molecular markers associated with soybean plant height, lodging, and maturity across locations. *Crop Sci.* 36:728-735.
- Lee, S.H., M.A. Bailey, M.A.R. Mian, T.E. Carter, Jr., E.R. Shipe, D.A. Ashley, W.A. Parrott, R.S. Hussey, and H.R. Boerma. 1996b. RFLP loci associated with soybean seed protein and oil content across populations and locations. *Theor. Appl. Genet.* 93:649-657.
- Lichtenstein, A.H., L.M. Ausman, S.M. Jalbert, and E.J. Schaefer. 1999. Effects of different forms of dietary hydrogenated fats on serum lipoprotein cholesterol levels. *N. Engl. J. Med.* 340:1933-1940.
- Livingston, M., R. Johansson, S. Daberkow, M. Roberts, M. Ash, and V. Breneman. 2004. Economic and policy implications of wind-borne entry of Asian soybean rust into the United States. Electronic Outlook Report from the Economic Research Service, OCS-04D-02. Available at <http://www.ers.usda.gov/publications/OCS/Apr04/OCS04D02/> (Verified 29 Aug 2004).
- López, Y., H.L. Nadaf, O.D. Smith, J.P. Connell, A.S. Reddy, and A.K. Fritz. 2000. Isolation and characterization of the Δ^{12} -fatty acid desaturase in peanut (*Arachis hypogaea* L.) and search for polymorphisms for the high oleate trait in Spanish market-type lines. *Theor. Appl. Genet.* 101:1131-1138.
- Mansur, L.M., J.H. Orf, K. Chase, T. Jarvik, P.B. Cregan, and K.G. Lark. 1996. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. *Crop Sci.* 36:1327-1336.
- Mazur, B., E. Krebbers, and S. Tingey. 1999. Gene discovery and product development for grain quality traits. *Science* 285:372-375.
- McConn, M., R.A. Creelman, E. Bell, J.E. Mullet, and J. Browse. 1997. Jasmonate is essential for insect defense in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* 94:5473-5477.
- McLean, R.J., and D.E. Byth. 1980. Inheritance of resistance to rust *Phakopsora pachyrhizi* in soybeans. *Aust. J. Agric. Res.* 31:951-956.

- Melching, J.S., W.M. Dowler, D.L. Koogle, and M.H. Royer. 1989. Effects of duration, frequency, and temperature on leaf wetness periods on soybean rust. *Plant Dis.* 73:117-122.
- Mensink, R.P., E.H. Temme, and G. Hornstra. 1994. Dietary saturated and trans fatty acids and lipoprotein metabolism. *Ann. of Med.* 26:461-464.
- Mercer, L.C., J.C. Wynne, and C.T. Young. 1990. Inheritance of fatty acid content in peanut oil. *Peanut Sci.* 17:17-21.
- Miller, J.F., D.C. Zimmerman, and B.A. Vick. 1987. Genetic control of high oleic acid content in sunflower oil. *Crop Sci.* 27:923-926.
- Miles, M.R., R.D. Frederick, and G.L. Hartman. 2003. Soybean rust: is the US soybean crop at risk? American Phytopathological Society, APSnet. Available at www.apsnet.org/online/feature/rust/top.asp (Verified 28 Oct 2006).
- Miquel, M.F., and J.A. Browse. 1994. High-oleate oilseeds fail to develop at low temperature. *Plant Physiol.* 106:421-427.
- Mounts, T.L., K.A. Warner, G.R. List, J.P. Fredrich, and S. Koritala. 1978. Flavor and oxidative stability of hydrogenated and unhydrogenated soybean oils: Effects of antioxidants. *J. Am. Oil Chem. Soc.* 55:345-349.
- Mounts, T.L., K. Warner, G.R. List, R. Kleiman, W.R. Fehr, E.G. Hammond, and J.R. Wilcox. 1988. Effect of altered fatty acid composition on soybean oil stability. *J. Am. Oil Chem. Soc.* 65: 624-628.
- Murphy, D.J. 1999. Production of novel oils in plants. *Curr. Opin. Biotech.* 10:175-180.
- Neidleman, S.L. 1987. Effects of temperature on lipid unsaturation. *Biotechnol. Genet. Eng. Rev.* 5:245-268.
- Njiti, V.N., M.A. Shenaut, R.J. Suttner, M.E. Schmidt, and P.T. Gibson. 1998. Relationship between soybean sudden death syndrome (SDS) and yield components in F₆-derived lines. *Crop Sci.* 38:673-678.
- O'Bryne, D.J., D.A. Knauff, and R.B. Shireman. 1997. Low fat-monounsaturated rich diets containing high-oleic peanuts improve serum lipoprotein profiles. *Lipids* 32:687-695.

- Ogle, H.J., D.E. Byth, and R.J. McLean. 1979. Effect of rust (*Phakopsora pachyrhizi*) on soybean yield and quality in south-eastern Queensland. *Aust. J. Agric. Res.* 30:883-893.
- Ohlrogge, J.B. 1994. Design of new plant products: Engineering of fatty acid metabolism. *Plant Physiol.* 104:821-826.
- Ohlrogge, J.B., and J. Browse. 1995. Lipid biosynthesis. *Plant Cell* 7:957-970.
- Ohlrogge, J.B., J. Browse, and C. Somerville. 1991. The genetics of plant lipids. *Biochem. Biophys. Acta* 1082:1-26.
- 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.
- Orf, J.H., K. Chase, F.R. Adler, L.M. Mansur, and K.G. Lark. 1999. Genetics of soybean agronomic traits: II. Interactions between yield quantitative trait loci in soybean. *Crop Sci.* 39:1652-1657.
- Pantalone, V.R., R.F. Wilson, W.P. Novitzky, and J.W. Burton. 2002. Genetic regulation of elevated stearic acid concentration in soybean oil. *J. Am. Oil Chem. Soc.* 79:549-553.
- Paterson, A.H., J.W. DeVerna, B. Lanini, and S.D. Tanksley. 1990. Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes in an interspecies cross of tomato. *Genetics* 124:735-742.
- Paterson, A.H., S. Damon, J.D. Hewitt, D. Zamir, H.D. Rabinowitch, S.E. Lincoln, E.S. Lander, and S.D. Tanksley. 1991. Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. *Genetics* 127:181-197.
- Paterson, A.H. 1998. Prospects for cloning the genetic determinants of QTLs. p. 289-293. *In* A.H. Paterson (ed.) *Molecular dissection of complex traits*. CRC Press. Boca Raton, FL, USA.
- Primomo, V.S., D.E. Falk, G.R. Ablett, J.W. Tanner, and I. Rajcan. 2002. Genotype $\times$ environment interactions, stability and agronomic performance of soybean with altered fatty acid profiles. *Crop Sci.* 42:37-44.

- Pryde, E.H., and J.A. Rothfus. 1989. Industrial and nonfood uses of vegetable oils. p. 87-117. *In* G. Röbbelen, R.K. Downey, and A. Ashri (ed.) Oil Crops of the World. McGraw-Hill Publishing Co. New York, NY, USA.
- Rahman, S.M., Y. Takagi, K. Kubota, K. Miyamoto, and T. Kawakita. 1994. High oleic acid mutant in soybean induced by X-ray irradiation. *Biosci. Biotech. Biochem.* 58:1070-1072.
- Rahman, S.M., Y. Takagi. 1997. Inheritance of reduced linolenic acid content in soybean seed oil. *Theor. Appl. Genet.* 94:299-302.
- Rahman S.M., Y. Takagi, and T. Kinoshita. 1997. Genetic control of high stearic acid content in seed oil of two soybean mutants. *Theor. Appl. Genet.* 95:772-776.
- 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.
- Rahman, S.M., T. Kinoshita, T. Anai, and Y. Takagi. 2001. Combining ability in loci for high oleic and low linolenic acids in soybean. *Crop Sci.* 41:26-29.
- Rahman, S.M., T. Anai, T. Kinoshita, and Y. Takagi. 2003. A novel soybean germplasm with elevated saturated fatty acids. *Crop Sci.* 43:527-531.
- Rebetzke, G.J., J.W. Burton, T.E. Carter, and R.F. Wilson. 1998. Genetic variation for modifiers controlling reduced saturated fatty acid content in soybean. *Crop Sci.* 38:303-308.
- Rennie, B.D., and J.W. Tanner. 1989. Fatty acid composition of oil from soybean seeds grown at extreme temperatures. *J. Am. Oil Chem. Soc.* 66:1622-1624.
- Ross, A.J., W.R. Fehr, G.A. Welke, and S.R. Cianzio. 2000. Agronomic and seed traits of 1%-linolenate soybean genotypes. *Crop Sci.* 40:383-386.
- Rytter, J.L., W.M. Dowler, and K.R. Bromfield. 1984. Additional alternative hosts of *Phakopsora pachyrhizi*, causal agent of soybean rust. *Plant Dis.* 68:818-819.
- Schnebly, S.R., and W.R. Fehr. 1993. Effect of years and planting dates on fatty acid composition of soybean genotypes. *Crop Sci.* 33:716-719.
- Schneider, R.W., C.A. Hollier, H.K. Whitan, M.E. Palm, J.M. McKenny, J.R. Hernández, L. Levy, and R. Devries-Paterson. 2005. First report of soybean

- rust caused by *Phakopsora pachyrhizi* in the continental United States. Plant Dis. 89:774.
- Schuppert, G.F. 2004. Molecular mechanisms underlying the high oleic acid phenotype in sunflower. Ph.D. diss. Oregon State Univ.
- Sinclair J.B., and G.L. Hartman. 1996. Proceedings of the Soybean Rust Workshop, Urbana, IL. Natl. Soybean Res. Lab. Publ. 1. Urbana-Champaign, IL, USA.
- Sinclair J.B., and G.L. Hartman. 1999. Soybean Rust. p. 25-26. In G.L. Hartman, J.B. Sinclair, and J.C. Rupe (ed.) Compendium of soybean diseases. 4th edition. APS Press. St. Paul, MN, USA.
- Sommerfeld, M. 1983. *Trans* unsaturated fatty acids in natural products and processed foods. Prog. Lipid Res. 22:221-233.
- Somerville, C., J. Browse, J.G. Jaworski, and J.B. Ohlrogge. 2000. Lipids. p. 456-527. In B.B. Buchanan, W. Gruissem, and R.L. Jones (ed.) Biochemistry and molecular biology of plants. American Society of Plant Physiologists. Rockville, MD, USA.
- 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.
- Staub, J.E., F.C. Serquen, and M. Gupta. 1996. Genetic markers, map construction, and their application in plant breeding. HortSci. 31:729-741.
- Stojšin, 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.
- Stuber, C.W., and M.D. Edwards. 1986. Genotypic selection for improvement of quantitative traits in corn using molecular marker loci. Proc. 41st Annual Corn and Sorghum Industry Res. Conf., American Seed Trade Assoc. Washington, DC, USA. 41:70-83.
- Stuber, C.W., M. Polacco, and M.L. Senior. 1999. Synergy of empirical breeding, marker-assisted selection, and genomics to increase crop yield potential. Crop Sci. 39:1571-1583.
- Sugimoto, T., K. Tanaka, M. Monma, Y. Kawamura, and K. Saio. 1989. Phosphoenolpyruvate carboxylase level in soybean seed highly correlates to its contents of protein and lipid. Agric. Biol. Chem. 53:885-887.

- Tai, Y.P., and C.T. Young. 1975. Genetic studies of peanut proteins and oils. *J. Am. Oil Chem. Soc.* 52:377-385.
- 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.
- Tanksley, S.D., N.D. Young, A.H. Patterson, and M.W. Bonierbale. 1989. RFLP mapping in plant breeding: new tools for an old science. *BioTechnology* 7:257-264.
- Thelen, J.J., and J.B. Ohlrogge. 2002. Metabolic engineering of fatty acid biosynthesis in plants. *Metab. Eng.* 4:12-21.
- Thies, W. 1970. Chloroplast development and biogenesis of linolenic acid in ripening cotyledons of rapeseed. p. 348-356. *In Proc. Int. Conf. Sci. Technol. Marketing of rapeseed and rapeseed products.* Rapeseed Assn. Ottawa, Canada.
- Töpfer, R., and N. Martini. 1994. Molecular cloning of cDNAs or genes encoding proteins involved in *de novo* fatty acid biosynthesis in plants. *J. Plant Physiol.* 143:416-425.
- Töpfer, R., N. Martini, and J. Schell. 1995. Modification of plant lipid synthesis. *Science* 268:681-686.
- USDA, ARS, National Genetic Resources Program. 2004. *Germplasm Resources Information Network - (GRIN)*. [Online Database] National Germplasm Resources Laboratory, Beltsville, MD. Available at <http://www.ars-grin.gov/cgi-bin/npgs/html/obvalue.pl?51068> (Verified 30 Oct 2006).
- USDA-ARS. 2004. ASA alerted to discovery of soybean rust north of the equator. Available at <http://southeastfarmpress.com/news/82304Soybean-rust/> (Verified 30 Oct 2006).
- USDA. 2006. Pest information platform for extension and education. National soybean rust commentary. Available at <http://www.sbrusa.net/> (Verified 30 Oct 2006).
- Vakili, N.G., and K.R. Bromfield. 1976. *Phakopsora* rust on soybean and other legumes in Puerto Rico. *Plant Dis. Rep.* 60:995-999.
- Vakili, N.G. 1979. Field survey of endemic leguminous hosts of *Phakopsora pachyrhizi* in Puerto Rico. *Plant Dis. Rep.* 63:931-935.

- Vance, V., and H. Vaucheret. 2001. RNA silencing in plants: defense and counterdefense. *Science* 292:2277-2280.
- Vles, R.O., and J.J. Gottenbos. 1989. Nutritional characteristics and food uses of vegetable oils. p. 63-85. *In* G. Robbelen, R.K. Downey and A. Ashri (ed.) *Oil crops of the world*. McGraw-Hill Publishing Co. New York, USA.
- White, H.B., F.W. Quackenbush, and A.H. Probst. 1961. Occurrence and inheritance of linolenic and linoleic acids in soybean seeds. *J. Am. Oil Chem. Soc.* 38:113-117.
- WHO. 1982. Prevention of coronary heart disease. WHO Technical Report Series No. 678. World Health Organization. Geneva.
- 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.R. Cavins. 1987. Gene symbol assigned for linolenic acid mutant in soybean. *J. Hered.* 78:410.
- Wilcox, J.R. 2004. World distribution and trade of soybean. p. 1-13. *In* H.R. Boerma, and J.E. Specht (ed.) *Soybeans: Improvement, production and uses*. 3rd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- Willet, W.C. 1994. Diet and health: what should we eat? *Science* 264: 532-537.
- Wilson, R.F., H.H. Weissinger, J.A. Buck, and G.D. Faulkner. 1980. Involvement of phospholipids in polyunsaturated fatty acid synthesis in developing soybean cotyledons. *Plant Physiol.* 66:545-549.
- 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. 1987. Seed Metabolism. p. 643-686. *In* J.R. Wilcox (ed.) *Soybeans: Improvement, production and uses*. 2nd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- 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 (ed.) *Lipid biotechnology*. Marcel Dekker, Inc. IL, USA.

- Wilson, R.F. 2004. Seed composition. p. 621-669. *In* H.R. Boerma and J.E. Specht (ed.) Soybeans: Improvement, production and uses. 3rd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- Wolf, R.B., J.F. Cavins, R. Kleiman, and L.T. Black. 1989. Effect of temperature on soybean seed constituents: oil, protein, moisture, fatty acids, amino acids and sugars. *J. Am. Oil Chem. Soc.* 59:230-232.
- Worthington, R.E., and R.O. Hammons. 1971. Genotypic variation in fatty acid composition and stability of *Arachis hypogaea* L. oil. *Oleagineux* 26:695-700.
- Yadav, N.S. 1995. Genetic modification of soybean oil quality. p.165-188. *In* D.P. S. Verma and R.C. Shoemaker (ed.) Soybean biotechnology. CAB Int., London, England.
- Yang, X.B., M.H. Royer, A.T. Tschantz, and B.Y. Tsia. 1990. Analysis and quantification of soybean rust epidemics from seventy-three sequential planting experiments. *Phytopathology* 80:1421-1427.
- Yamamoto, K.T., H. Mori, and H. Imaseki. 1992. Novel mRNA sequences induced by indole-3-acetic acid in sections of elongating hypocotyls of mung bean (*Vigna radiata*). *Plant Cell Physiol.* 33:13-20.
- Yamamoto, T., Y. Kuboki, S.Y. Lin, T. Sasaki, and M. Yano. 1998. Fine mapping of quantitative trait loci *Hd-1*, *Hd-2* and *Hd-3*, controlling heading date of rice, as single Mendelian factors. *Theor. Appl. Genet.* 97:37-44.
- Zhu, Y.L., Q.J. Song, D.L. Hyten, C.P. Van Tassell, L.K. Matukumalli, D.R. Grimm, S.M. Hyatt, E.W. Fickus, N.D. Young, and P.B. Cregan. 2003. Single-nucleotide polymorphisms in soybean. *Genetics* 163:1123-1134.

Table 2.1. Fatty acids most commonly found in plants (Appelqvist, 1989; Somerville et al., 2000).

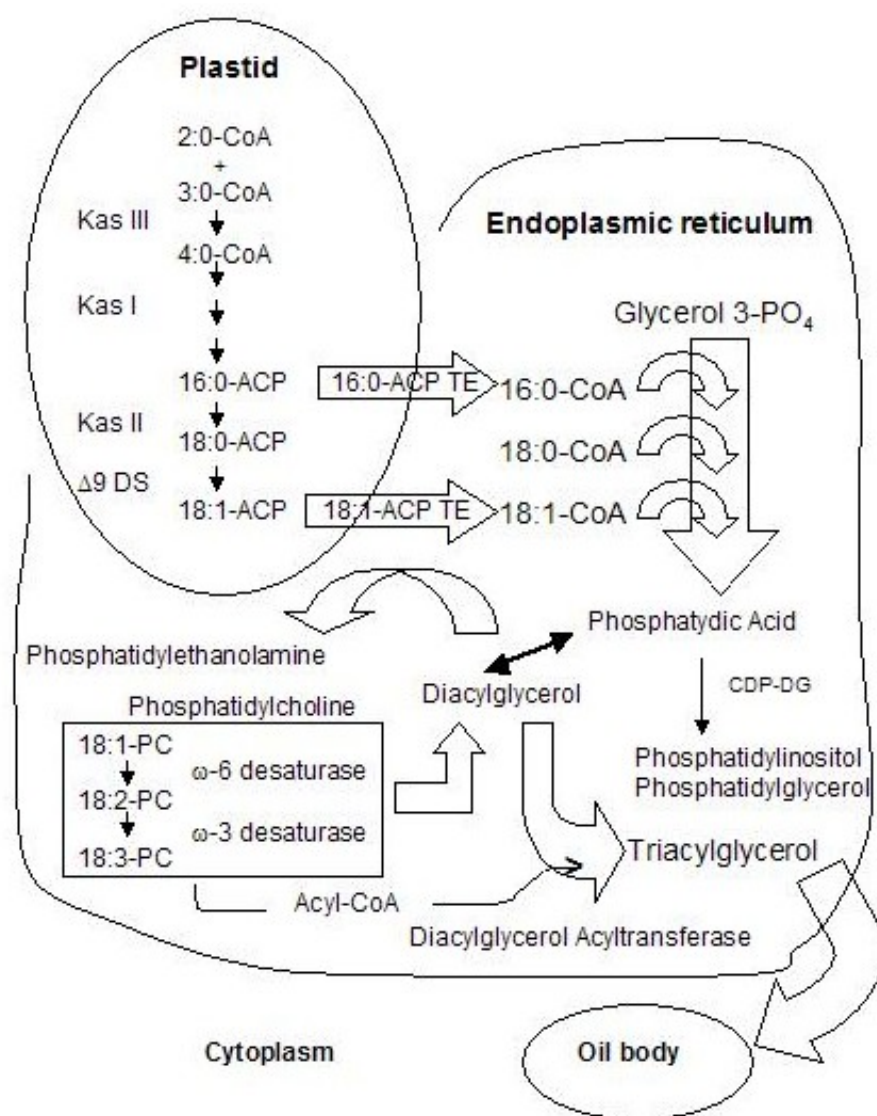
	Traditional Name	Symbol [†]	IUPAC [‡]
Saturated	Lauric acid	12:0	Tetradecanoic acid
	Palmitic acid	16:0	Hexadecanoic acid
	Stearic acid	18:0	Octadecanoic acid
	Arachidic acid	20:0	Eicosanoic acid
	Lignoceric acid	24:0	Tetracosanoic acid
Unsaturated	Oleic acid	18:1	9-octadecenoic acid
	Linoleic acid	18:2	9,12-octadecenoic acid
	Linolenic acid	18:3	9,12,15-octadecenoic acid
	Erucic acid	22:1	13-docosenoic acid

[†] First number refers to the number of C atoms; second number is the number of double bonds. [‡] IUPAC (International Union of Pure and Applied Chemistry).

Table 2.2. Fatty acid composition of different types of oil (Erasmus, 2000; Töpfer et al., 1995).

Oil type	Palmitic (16:0)	Stearic (18:0)	Oleic (18:1)	Linoleic (18:2)	Linolenic (18:3)
	g kg ⁻¹	g kg ⁻¹	g kg ⁻¹	g kg ⁻¹	g kg ⁻¹
Canola	41	18	630	200	86
Olive	1	160	760	80	1
Soybean	90	60	260	500	70
Sunflower	5	120	230	650	1

Figure 2.1. Fatty acid synthesis and glycerolipid synthetic pathways in soybean (Wilson, 2004).



CHAPTER 3
MOLECULAR MAPPING AND CONFIRMATION OF QTL ASSOCIATED WITH
OLEIC ACID CONTENT IN N00-3350 SOYBEAN¹

¹ Maria J. Monteros, Joseph H. Burton, and H. Roger Boerma. To be submitted to Molecular Breeding.

ABSTRACT

The fatty acid composition of soybean [*Glycine max* (L.) Merr.] seed affects the flavor, nutritional value, and stability of the oil. Oleic acid is one of the major fatty acids in soybean, with most current cultivars possessing 160 to 280 g kg⁻¹. Increasing oleic acid content in soybean oil would decrease the total saturated fatty acid content and reduce the need for hydrogenation, a process that creates unhealthy *trans* fatty acids. Several soybean genotypes with increased oleic acid content have been developed. The objective of this study was to map and confirm the areas of the soybean genome associated with oleic acid content from N00-3350 (~583 g kg⁻¹ oleic acid) using simple sequence repeat (SSR) markers. A F_{2:3} population of 259 lines from the cross of G99-G725 × N00-3350 was used as a mapping population, and a F_{2:3} population of 231 lines from the cross of G99-G3438 × N00-3350 was used as a confirmation population. Based on single factor analysis of variance (ANOVA), interval mapping and composite interval mapping (CIM), six QTLs for oleic acid content were found on linkage groups (LG) A1 (Satt211, R² = 4%), LG-D2 (Satt389, R² = 6%), LG-G (Satt394, R² = 13%), LG-G (Satt191, R² = 7%), LG-L (Satt418, R² = 9%), and LG-L (Satt561, R² = 25%) in the G99-G725 × N00-3350 population. All six QTLs for oleic acid were confirmed in the G99-3438 × N00-3350 population. At all of the identified oleic acid QTLs, the N00-3350 allele increased the oleic acid content. We propose the designation *cqOle2-1*, *cqOle2-2*, *cqOle2-3*, *cqOle2-4*, *cqOle2-5*, and *cqOle2-6* for the oleic acid QTL that have been identified and confirmed. The identification of SSR markers closely linked to the oleic acid QTLs will facilitate the use of marker-assisted selection (MAS) in soybean breeding programs to increase the oleic acid content in soybean seed.

INTRODUCTION

The most common fatty acids found in the seeds of most plants, including soybean, belong to a small group of C16 and C18 fatty acids that include palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2), and linolenic (18:3) (Somerville et al., 2000). The fatty acid composition of vegetable oils is variable, and depends on the cultivar (i.e. genetic factors), and environment during the growing season (Vles and Gottenbos, 1989). The content and relative proportions of each fatty acid are important factors because they affect the flavor, stability, and nutritional value of the oil (Mensink et al., 1994). Soybean oil alone accounts for approximately 27% of the world's total edible oil production (FAS, 2002; Carter and Wilson, 1998). Global soybean oil consumption has increased at a steady rate of about 1 MMT (million metric tons) per year since 1994 (Wilson, 2004).

Soybean oil used for human consumption is subject to U.S. Food and Drug Administration (FDA) guidelines for health claims made on labels. FDA labeling regulations, in accordance with the Nutritional Labeling and Education Act of 1990 require that a "low-saturated" vegetable oil have less than 7% total saturated fatty acids (USFDA, 1999). Elite soybean cultivars produce, on average, 110 g kg⁻¹ palmitic, 40 g kg⁻¹ stearic, 180 to 240 g kg⁻¹ oleic, 540 g kg⁻¹ linoleic, and 80 g kg⁻¹ linolenic acids (Wilcox et al., 1984; Diers et al., 1992; Schnebly and Fehr, 1993; Hui, 1996). Thus, soybean oil contains about 15% saturated fat, which is higher than those of both canola (*Brassica rapa* L.) and sunflower (*Helianthus annuus* L.) oils (Wilson et al., 2002). Soybean oil also contains a relatively high level of linoleic acid. The major limitation with vegetable oils that contain high concentrations of polyunsaturated fatty acids, like linoleic and linolenic acid, is flavor stability (Wilson 1987). Both of these fatty acids may be oxidized, leading to undesirable odors and flavors (Crapiste et al., 1999; Mounts et al., 1988). Oleic acid, a monounsaturated fatty acid, is less susceptible to oxidation during storage and frying (Miller et al., 1987; Mercer et al., 1990).

The process of hydrogenation is currently used to improve the oxidative stability of soybean oil, which increases the shelf life of fats and foods prepared with it. However, this apparent solution to the stability issue may give rise to another

problem. The bonds of unsaturated fatty acids in crude vegetable oils are predominately in a *cis* configuration. During hydrogenation, some double bonds may be rearranged into a *trans* configuration, where the hydrogen atoms end up on different sides of the chain (Wilson, 2004). These *trans* fatty acids can cause undesirable health effects, including elevated blood levels of low-density lipoproteins (LDL), and an increase in the risk for coronary heart disease (CHD) (Willet, 1994; Hu et al., 1997; Lichtenstein et al., 1999; Mazur et al., 1999). These findings have prompted the FDA to require that, in addition to the saturated fat content, information on the amount of *trans* fatty acids must be included on the Nutrition Facts panel of a product's label (FDA, 2004). The effective date of this labeling mandate was 1 Jan. 2006 (CFSAN, 2003). A viable alternative to increase the oxidative stability of the oil without hydrogenation would be to produce soybean oil with a more favorable fatty acid composition (i.e., 500 to 550 g kg⁻¹ oleic acid and less than 30 g kg⁻¹ linolenic acid) (Wilson et al., 2002).

Consumer and food manufacturer interest in obtaining vegetable oil with high oleic acid and low polyunsaturated fatty acid content is increasing because the oil maintains its quality for a longer period of time (O'Bryne et al., 1997; Rahman et al., 2001). Additionally, consumer demand for healthier foods has created a potential market to develop specialty crops with less saturated fat (Darroch et al., 2002). Consumption of oil with a high oleic acid content has been shown to have positive health-related effects in that it can lower serum cholesterol levels, particularly LDLs, and it may reduce the incidence of CHD (O'Bryne et al., 1997; Thelen and Ohlrogge, 2002). High oleic acid soybean oil can benefit consumers in that they will obtain oil with a lower saturated fat content and an increase in oil quality (Hayakawa et al., 2000; Darroch et al., 2002). Manufacturers will also benefit from this high oleic acid oil because it is naturally more heat-stable than commodity-grade soybean oil, and therefore needs no hydrogenation for use in cooking or processing to extend the shelf life of foods (Darroch et al., 2002). Therefore, the needs for soybean oil with a lower saturated fat content, and for food with low levels of *trans* fats, have made the production of high oleic acid soybean cultivars important breeding objectives in the USA. Accomplishing this objective will allow soybean oil to remain attractive to food manufacturers (Töpfer et al., 1995; Wilson et al., 2002; Wilson, 2004). Ultimately,

the value of high oleic acid soybean oil will be determined by the relative prices of the products using this oil, consumer preference, and the perceived benefits of high oleic acid soybean oil (Giannakas and Yiannaka, 2004).

Incorporation of a unique fatty acid composition into commercial cultivars would likely be enhanced by knowledge of the inheritance of the fatty acid composition in soybean. Oleic acid and polyunsaturated fatty acid contents in soybean are quantitatively inherited (White et al., 1961; Burton et al., 1983; Carver et al., 1987). Success in drastically altering the fatty acid composition in plants has been achieved with no obvious ill effects to the plant (Hammond and Fehr, 1984). Yield trials of soybean lines with increased oleic acid content have shown that higher oleic acid content does not negatively affect yield (Carver et al., 1986; Diers and Shoemaker, 1992).

Soybean oil use has been affected by the identification of mutations and naturally occurring variations affecting seed fatty acid synthesis (Palmer et al., 2004). In 1975, the use of mutation breeding with ethyl methane sulfonate (EMS) or sodium azide was implemented, and changes in linolenic acid content were obtained (Hammond and Fehr, 1984). The line N97-3363-4 has recessive fatty acid desaturase alleles and contains about 60% oleic acid content. Selection for higher oleic acid content has been used to indirectly reduce linolenic acid. The level of linolenic acid in the line N78-2245 was reduced from 90 g kg⁻¹ to 6 g kg⁻¹, by selecting for oleic acid, which increased from 220 g kg⁻¹ to 420 g kg⁻¹. Ancestors of the parental line N79-2473 were selected for lower linolenic acid (Wilson et al., 2002). The lines N78-2245, N79-2077, N87-2122-4, and C1726 were selected for low palmitic acid content (Burton et al., 1994). The breeding line C1726 was developed by mutagenesis from the cultivar Century, and has reduced palmitic acid levels (Wilcox et al., 1980; Erickson et al., 1988; Wilcox and Cavins, 1990). All of these lines selected for lower palmitic acid and lower linolenic acid content were used as parents in developing N98-4445A. The line N98-4445A belongs to Maturity Group (MG) IV and contains 500 to 600 g kg⁻¹ oleic acid content (Burton et al., 2006). N00-3350 originated as a single plant from N98-4445A and possesses a similar oleic acid content (Fig. 3.1).

Advances in DNA marker technology, including the development of SSR markers and the development of an integrated soybean genetic linkage map, have facilitated the genetic mapping of quantitative traits in soybean (Cregan et al., 1999). The most recent integrated genetic map for *G. max* contains over 1,000 mapped simple sequence repeat (SSR) markers (Song et al., 2004). Twelve to 29 additional markers per linkage group were added to the earlier map of Cregan et al. (1999). The objectives of this study were to map and confirm the areas of the soybean genome that are associated with oleic acid content from N00-3350 (550 g kg⁻¹ oleic acid) using SSR markers.

MATERIALS AND METHODS

Plant material

An F₂ population consisting of 259 plants derived from the cross of G99-G725 (~206 g kg⁻¹ 18:1) × N00-3350 (~620 g kg⁻¹ 18:1) was developed for use as a mapping population. G99-G725 is a glyphosate-resistant backcross conversion of Boggs (Boerma et al., 2000). Boggs is a MG VI cultivar with white flowers that was derived from an F₅ plant from the cross G81-152 × 'Coker 6738'. G81-152 was derived from the cross D74-7741 × 'Coker 237'. D74-7741 is a MG VI breeding line selected from the cross of 'Forrest' × D70-3001. Twelve seeds per line were planted in the greenhouse on 16 Oct. 2001.

A total of 231 F₂ plants from the cross G99-G3438 (~185 g kg⁻¹ 18:1) × N00-3350, were used as a confirmation population. G99-G3438 is a glyphosate-resistant backcross conversion of Benning (Boerma et al., 1997). Benning is a MG VII cultivar with purple flowers derived from an F₄ plant from the cross 'Hutcheson' × 'Coker6738'. Twelve seeds per line were planted in the greenhouse on 17 June 2003. Three entries from each of the parents were included as checks.

On 7 May 2004, the 259 F_{2:3} lines from G99-G725 × N00-3350 and the 231 F_{2:3} lines from G99-G3438 × N00-3350 were planted in 1-m row plots in Isabela, Puerto Rico. Three entries of each of the parents were planted in each experiment as checks. Plots were harvested individually on 2 Aug. 2004. In addition, 12 seeds each of 231 F_{2:3} lines from the cross of G99-G3438 × N00-3350 were planted 25

May 2004 in 0.76-m x 0.45-m hills at the Univ. of Georgia Plant Sciences Farm, near Athens, GA. Hills were thinned to eight plants on 10 June 2004. Due to limited seed, a single replication was planted in the two locations. Individual plants from within some hills from G99-G3438 $\times$ N00-3350 grown in Athens matured at different times. For these hills, the early-maturing and late-maturing plants were harvested independently to create sub-samples from each entry. Fatty acid analysis was determined separately for each early and late maturing sub-sample from each line. In these cases, a random sample of six seeds from within each subgroup was planted in the greenhouse on 10 Aug 2005. Trifoliolate leaves for each sub-sample were collected for DNA extraction and SSR genotyping.

Phenotypic data

For the initial fatty acid determination of the G99-G725 $\times$ N00-3350 mapping population F_{2:3} seed were analyzed for fatty acid composition. If there were 35 F_{2:3} seeds or more, a bulked random sample of 10 seeds from each plant was sent to USDA-ARS laboratories in Peoria, IL, and Raleigh, NC, for fatty acid analysis. If 18 to 34 seeds were available, 10 ½-seed chips were sent to each laboratory. If fewer than 18 seeds were available, ¼-seed chips of each seed were sent to each laboratory. Fatty acid content was determined using gas chromatography (Hewlett Packard Model 5890/6890) to evaluate methyl esters. Phenotypic data for each fatty acid were tested for a normal distribution using SAS® (PROC UNIVARIATE PLOT) (SAS Institute, Cary, NC). The fatty acid data of each entry used in the analysis is the average from the Peoria and Raleigh laboratories.

For the G99-G3438 $\times$ N00-3350 confirmation population, if > 49 seeds were available, 15 seeds were sent to each laboratory for fatty acid analysis. If 40 to 50 seeds per line were available, then 12 seeds per line were sent to each laboratory. If 29 to 39 seeds per line were available, then 10 seeds per line were sent to each laboratory. If 22 to 28 seeds were available, then 8 seeds per line were sent to each laboratory. The fatty acid content of the 67 lines from G99-G3438 $\times$ N00-3350 that differed in maturity when planted in the field in Athens was determined separately for each sub-group. Most subgroups consisted of either four plants with early maturity and four plants with late maturity, three plants early and five plants late, or vice

versa. In two cases the early subgroup consisted of six plants and the late subgroup consisted of two plants.

SSR analysis

Leaf tissue from each F₂ plant in the two populations was collected, lyophilized, and macerated. DNA was extracted using a modified CTAB (hexadecyltrimethylammonium acid) protocol (Keim et al., 1988), and re-suspended in TE buffer. PCR reactions were similar to the protocol by Li et al., (2001), with some modifications. The 10- μ L reaction mix contained 2 μ L of 50 ng μ L⁻¹ template DNA, 1.0X PCR buffer, 2.5 mM MgCl₂, 100 μ M of each dNTP, 0.2 μ M each of the forward and reverse primers, and 0.5 unit of Promega *Taq* DNA polymerase (Madison, WI). Primers were labeled with the fluorescent dyes 6-FAM, NED, or HEX (PE-ABI, Foster City, CA). A 384-well or a 96-well GENE AMP PCR System 9700 thermal cycler (Applied Biosystems, Foster City, CA) was used for DNA amplification.

Pooled PCR products (3-4 μ L) were combined with 2 μ L deionized formamide, 0.75 μ L loading buffer, and 0.2 μ L Genescan ROX-500 internal size standard. The mixture was denatured at 95°C for 4 min., and 1-2 μ L were loaded into each of 96 lanes on 12-cm acrylamide:bisacrylamide (19:1) gels, using KLOEHN micro syringes (Kloehn Ltd., Las Vegas, NV). DNA amplicons were run on gels using an ABI Prism 377 DNA sequencer at 750 V for 1.5 to 2 hours, with 1X TBE buffer. PE ABI 377 DNA sequencer collection software was used to collect the marker data (ABI, Foster City). GeneScan® Version 3.0 was used to analyze the DNA amplicons. Gels were scored with Genotyper® Version 2.5 (ABI, Foster City), and manually verified.

A genome-wide screen of DNA markers was conducted using evenly spaced markers from all 20 linkage groups (Cregan et al., 1999). A total of 350 SSR markers covering all of the linkage groups were tested for polymorphism in the G99-G725 $\times$ N00-3350 population. A total of 180 SSR markers (51.4%) was polymorphic between the two parents. Markers evaluated for the confirmation population using 231 lines from the cross of G99-G3438 $\times$ N00-3350 were selected based on

significant associations with oleic acid content obtained from the mapping population. The early and late maturing sub-samples from G99-G3438 × N00-3350 were genotyped separately.

Data analyses and mapping

The correlation coefficient between the oleic acid data from the F_{2:3} seed from G99-G725 × N00-3350 obtained from both laboratories and the single factor analysis of variance (ANOVA) for the marker analysis were calculated using SAS® Version 8 (SAS Institute, Cary, NC). For the oleic and linoleic acid content from Athens and Puerto Rico from the cross of G99-G725 × N00-3350, genotypes and environments were considered random effects and the genotype by environment interaction (G × E) was used as an error term. The variance-component heritability was calculated as $h^2 = \sigma^2_G / (\sigma^2_{GE/2} + \sigma^2_G)$; where σ^2_G = genotypic variance component and σ^2_{GE} = genotype by environment variance component (Fehr, 1987).

Single factor analysis of variance (SF-ANOVA) was used to determine the significance of SSR genotypic class means using the general linear model (PROC GLM) SAS® Version 8 (SAS Institute, Cary, NC). Single linkage group multiple regression analysis (SLG-MR), and multiple linkage group regression analysis (MLG-MR) using the STEPWISE selection criteria was used to identify the significant markers in the model associated with oleic acid content at the 5% significance level. Linkage maps were initially constructed with MapManager QTX b20 (Manly et al., 2001) using the Kosambi mapping function (Kosambi, 1944). Oleic acid content and marker data were analyzed to determine the presence and estimate the positions of QTL using interval mapping with MapManager QTXb 20, and composite interval mapping (CIM) with QTL Cartographer V2.0 (Wang et al., 2005). For QTL detection, one thousand permutations were used to establish the minimum logarithm of odds (LOD) score. The CIM options and parameters used were similar to those described by Chung et al. 2003. A multiple regression model using SAS® was also used for the two-factor analysis of variance to detect epistatic interactions between all pairs of significant markers.

RESULTS AND DISCUSSION

G99-G725 × N00-3350 mapping population

The correlation coefficient between the oleic acid percentages for the $F_{2:3}$ plants when grown in the greenhouse in Athens obtained from the laboratories in Peoria and Raleigh was 0.94 ($p < 0.001$). The fatty acid data are reported as an average for the data from the two laboratories (Table 3.1). The mean oleic acid content for the G99-G725 parent from Athens and Isabela was 206 g kg^{-1} and for the N00-3350 parent was 583 g kg^{-1} . The range of oleic acid content of the 259 $F_{2:3}$ lines in the mapping population was 228 to 613 g kg^{-1} and the mean was 388 g kg^{-1} . The variance-components heritability for oleic acid content was 0.71, and for linoleic acid content was 0.54. The oleic heritability estimate is higher than a previous report from Hawkins et al. (1983) of 0.50 to 0.58 likely due to a wider range in oleic acid content of the lines evaluated in the present study.

Based on SF-ANOVA, significant markers associated with oleic acid content were found on LG's A1, D2, G, and L (Table 3.2). Linkage maps with additional markers on LG-A1, D2, G, and L were constructed (Fig 3.2). The order of the markers in these linkage groups based on our data is in close agreement with that of the integrated soybean genetic linkage map (Cregan et al., 1999; Song et al., 2004).

On the basis of SF-ANOVA, five markers on LG-A1 near Satt200 were significantly ($p < 0.01$) associated with oleic acid content. These markers were located in the 85 to 96 cM region on the USDA/ARS soybean consensus map. Using SLG-MR, only marker Satt211 was retained in the model (Table 3.2). Interval mapping shows the QTL likelihood plot near Satt200 and Satt599 (Fig. 3.2). Using the markers found significant by SF-ANOVA on LG-D2, the SLG-MR retained Satt389, which explained 4% of the variation in the oleic acid content. The LOD peak for oleic acid content based on interval mapping was located near Satt389 (Fig 3.2b).

On LG-G, eight SSR markers were found to be associated with the oleic acid content based on SF-ANOVA ($p < 0.001$) (Table 3.2). When these markers were placed in a SLG-MR equation, the markers Satt394 (43.4 cM), Satt594 (52.9 cM), and Satt303 (53.4 cM) had the largest overall effect with an R^2 value of 13%. When

LG-G was evaluated by interval mapping, three peaks were identified (Fig. 3.2c). One peak was near Satt235, one at Satt394, and one near Satt191.

Based on SF-ANOVA, seven markers on LG-L were associated with the variation in oleic acid content ($p < 0.001$). Of these seven markers, only Satt418 and Satt156 were retained in the LG-L SLG-MR (Table 3.2). Using interval mapping, three QTLs associated with oleic acid were identified (Fig. 3.2d). One QTL was located near Satt418, the second QTL on LG-L was approximately 40.5 cM away near Satt561, and a third QTL was identified between Satt418 and Satt561. Using SF-ANOVA, Satt418 explains 9% of the variation, Satt561 explains 25% of the variation, and Satt156 explains 15% of the variation (Table 3.2).

Composite interval mapping (CIM) was used to determine whether the significant effects at several linked markers or intervals were independently conditioning oleic acid content. A limited number of background markers were identified via the forward/backward stepwise regression option of QTL Cartographer V 1.3 using conservative probability thresholds ($P_{in} = 0.01$; $P_{out} = 0.01$). A 1-cM window parameter was chosen to exclude from the background marker group any marker located within 1 cM of the two markers flanking any interval being tested for a putative QTL peak (Chung et al., 2003). The CIM of oleic acid QTL on LG-L identified Satt561 as a background marker and the LOD peak at Satt418 surpassed the significance threshold determined by 1000 permutation tests, indicating the presence of two independent QTLs for oleic acid content on this LG (Fig. 3.3b). The CIM on LG-G revealed a similar pattern to that on LG-L. The oleic acid QTL near both Satt394 and Satt191 exceeded the LOD threshold. However, using CIM the QTL near Satt235 did not exceed the significance threshold (Fig. 3.3a).

A STEPWISE multiple regression analysis across the four linkage groups with significant SSR markers associated with oleic acid using 259 lines from G99-G725 $\times$ N00-3350 indicated that Satt211 ($R^2 = 8\%$) on LG-A1, Satt389 ($R^2 = 4\%$) on LG-D2, Satt235 ($R^2 = 8\%$) and Satt191 ($R^2 = 4\%$) on LG-G, and Satt418 ($R^2 = 3\%$), and Satt561 ($R^2 = 3\%$) on LG-L contribute to the oleic acid content, and together explain 30% of the variation in oleic acid content (Table 3.2). To evaluate possible epistatic interactions, all pairs of significant markers were tested for interaction using a two-

factor ANOVA (SAS Institute). None of the two-marker interactions were significant ($p \leq 0.01$).

When all six putative independent QTLs associated with oleic acid content identified using SF-ANOVA, interval mapping, and CIM were considered individually, the N00-3350-derived allele increased oleic acid content at each QTL (Table 3.2). When the six markers were homozygous for the N00-3350 allele, the predicted mean for oleic acid content increased by 356 g kg^{-1} compared to those having the G99-G725 allele. In most cases, except for the QTL near Satt418, the relative oleic acid content between the homozygous and the heterozygous genotypes indicated additive gene action for the alleles at these QTLs (Table 3.3). For the QTL near Satt418, the N00-3350 allele for increased oleic acid was recessive. Although this work explained a considerable amount of genetic variation, additional variation in oleic acid content due to effects unaccounted for by the markers, epistasis, and the genotype by environment interaction remains.

In sunflower, Schuppert (2004) showed that the *Ol₁* mutation associated with FAD2-1 is necessary but not sufficient to produce the high oleic acid phenotype, presumably because additional QTLs segregate in some of the genetic backgrounds evaluated. Additionally when a segregating population was evaluated, the data indicated that the oleic acid phenotype was caused by the main effect and the interaction of several genes. When a biosynthetic pathway is involved it is not unlikely that the availability of a substrate and the activity of genes upstream in the pathway may affect the accumulation of individual components downstream in the pathway. These findings provide evidence for the complexity of the oleic acid phenotype in the N00-3350 soybean line.

Several of the oleic content QTLs identified in this study are in the same genomic region as previously reported oil QTLs. Oil content QTLs have been reported on LG-A1 in the interval delimited by Satt174 (88.6 cM) to B170_1 (94.9 cM) using two different populations (Brummer et al., 1997; Orf et al., 1999; Mansur et al., 1996; Specht et al., 2001). On LG-D2, an oil content QTL was reported near Satt082 (87.2 cM) (Hyten et al., 2004b). QTLs for oil content have also been reported on LG-G (Brummer et al., 1997; Lee et al., 1996). Five different studies have reported oil content QTLs on LG-L and one of them is closely associated with

Satt166 (66.5 cM, $R^2 = 8\%$) (Diers et al., 1992; Lee et al., 1996; Hyten et al., 2004b, Mansur et al., 1996; Orf et al., 1999).

The mid-oleic QTLs identified in this study were also associated with the variation in linoleic acid content (Table 3.4). Oleic acid and linoleic acid were found to be negatively correlated, with a correlation coefficient of $r = -0.91$ ($p < 0.001$). Previous studies have similarly shown oleic acid and linoleic acid levels to be negatively correlated (Howell et al., 1972; Burton et al., 1983). The additive effects for the N00-3350 alleles for linoleic acid content were negative, indicating that the N00-3350 allele contributed to a decrease in linoleic acid content, and as previously shown, an increase in oleic acid content. Biochemical evidence indicates that linoleic and linolenic acid are produced by the consecutive desaturation of oleic acid (Howell et al., 1972; Wilson et al., 1981). Therefore, a reduction in linoleic acid content is likely when its precursor, oleic acid, is increased in the seed.

None of the markers associated with oleic acid on LG-A1, D2, G, and L explained any of the variation in the linolenic acid content. However, Satt318 on LG-B2 explained 12% of the variation in linolenic acid content, $p < 0.001$ (data not shown). At this QTL, N00-3350 is contributing an allele that decreases the amount of linolenic acid in soybean seeds. These results are consistent with a previous study where the *fan* locus controlling reduced linolenic acid was mapped to this region of LG-B2 (Brummer et al., 1995).

SSR markers on LG-A1, D2, G, and L are also significant for palmitic acid content (Table 3.4). A QTL near Satt211 on LG-A1 explains 4% of the variation in palmitic acid content. A study conducted by Li et al. 2002 indicated that N87-2122-4 contributed an allele for a major QTL for palmitic acid on LG-A1, which explained on average 34% of the variation in palmitic acid content. This QTL was located near the top of LG-A1, which is more than 90 cM from the oleic acid QTL. An association between palmitic acid and Satt166 on LG-L was also found in a cross of Cook x C1726 (David Hulburt, personal communication). In the same study, Satt458 on LG-D2 was also found to be associated with palmitic acid content.

G99-G3438 × N00-3350 confirmation population

The mean oleic acid content in the Puerto Rico environment was 186 g kg⁻¹ for the G99-G3438 parent and 668 g kg⁻¹ for the N00-3350 parent (Table 3.5). The range for oleic acid content in the 231 F_{2:3} lines from G99-G3438 × N00-3350 was 212 to 635 g kg⁻¹, and their progeny mean was 409 g kg⁻¹.

Data from the SF-ANOVA analysis of the 231 F_{2:3} lines from G99-G3438 × N00-3350 indicated that five of the QTLs associated with oleic acid content in the G99-G725 × N00-3350 mapping population were also significant ($p = 0.05$) in this population (Table 3.6). The putative oleic QTL located on LG-A1 (Satt211), two on LG-G (Satt394 and Satt191), and one on LG-L (Satt561), were significant when oleic acid content was measured in the Puerto Rican environment. However, the amount of variation in oleic acid content explained by these QTL was lower than in the mapping population (Table 3.2). The QTL on LG-D2 near Satt226 was not significant in this population.

In the G99-G3438 × N00-3350 population, segregation for maturity date across and within lines was observed in the field in Athens GA. For the 67 lines containing plants with more than a 7-day range in maturity, the plants were segregated into early and late sub-samples. Fatty acid determination for these sub-samples was done separately for the 67 lines. The analysis of maturity date sub-samples included only the data from Athens, since only limited differences in maturity were observed in Puerto Rico. The sub-samples from these lines were genotyped separately with SSR markers on linkage groups where putative oleic acid QTLs had been identified. In this case, the oleic acid QTL near Satt561 was significant across all the lines regardless of whether the data from the early or the late maturing sub-samples were considered in the analysis (Table 3.7). In this case, Satt397, which is 9.9 cM from Satt389 on the consensus map, suggests that this putative QTL on LG-D2 is significant. Consistent with the mapping population, at all of these QTLs, the N00-3350 alleles were associated with increased oleic acid content.

In the SF-ANOVA analysis, Satt418 on LG-L was not significant either with or without the lines that had different maturity dates (Table 3.6, Table 3.7). Satt153 on

LG-O was significant in the sub-samples from G99-G3438 $\times$ N00-3350. However, Satt153 also explained 9% of the variation observed in maturity dates. The parents used in the mapping and confirmation population belong to different maturity groups. Benning, the recurrent parent of G99-G3438 is a Maturity Group VII cultivar and Boggs, the recurrent parent of G99-G725 is a Maturity Group VI cultivar. In a SF-ANOVA analysis with SSR markers on linkage groups previously associated with oleic acid content, Satt561 was significantly associated with maturity date ($p < 0.001$). However, Satt418 was not significantly associated with maturity date (data not shown). Satt501, which is located 16 cM from Satt418 on the consensus map, is significantly associated with oleic acid content in this population.

The results from genotyping G99-G725 and G99-G3438 with SSR markers on LG-D2 and LG-L reveal different allele sizes (data not shown). These results suggest that differences in genetic background, in addition to the differences in maturity date, may be affecting the oleic acid QTLs. Studies in *Brassica* reported significant effects of genetic background on the content of major target fatty acids (Tang and Scarth, 2004). Studies with soybean near isogenic lines (NILs) in which background differences are minimized could be useful in evaluating the individual and combined effects of the oleic acid QTLs identified.

Due to the size of the mapping population there was sufficient statistical power to be able to detect QTLs with relatively small effects, which would have gone undetected using a smaller sample size or whose effects would have likely been overestimated. The oleic acid QTLs near Satt211 on LG-A1, Satt389 on LG-D2, Satt394 and Satt191 on LG-G, and Satt418 and Satt561 on LG-L were confirmed across different environments and in two independent populations. Overall, six putative QTLs for oleic acid content were identified and confirmed in an independent population. These results indicate that considerable progress can be made through selection of the N00-3350 alleles at the identified oleic acid QTLs to obtain soybean oil with a higher oleic acid content.

The designation Ole1-1 through Ole1-6 have been assigned in Soybase to oleic acid QTLs identified from a cross between A81356022 $\times$ PI468916 (Diers and Shoemaker, 1992). The designation *cq* for “confirmed QTL” has been proposed as

a way to allow breeders and researchers to recognize these QTLs as having been mapped and confirmed in a population derived by independent meiotic events (Fasoula et al., 2004). Knowledge of which QTLs have been confirmed can be used by breeders and researchers to prioritize the QTLs they want to incorporate in their programs. Therefore, we propose to designate the confirmed oleic acid QTLs on LG-A1 near Satt211 as *cqOle2-1*, on LG-D2 near Satt389 as *cqOle2-2*, on LG-G near Satt394 as *cqOle2-3*, on LG-G associated with Satt191 as *cqOle2-4*, on LG-L near Satt418 as *cqOle2-5*, and on LG-L near Satt561 as *cqOle2-6*.

The quantitative nature of the oleic acid trait makes phenotypic selection of the mid-oleic acid content a challenging task. Molecular markers can be used to identify the number, location, and contribution of QTL that affect oleic acid content in soybean oil. One type of breeding strategy to incorporate all oleic acid QTLs would involve conducting two separate backcross populations to introgress the oleic acid QTLs from N00-3350. In each backcross population three of the oleic QTLs would be introgressed. At the final step in the program, the two backcross populations would be crossed and MAS used to select a line with all six oleic QTLs. Increasing the number of transferred segments also increases the risk of inadvertently introgressing undesirable agronomic trait alleles at loci on linked donor chromosomal segments (Stuber et al., 1999). Therefore, the two QTLs on LG-G and LG-L should be selected for in different backcross populations to maximize the genetic contribution of the elite parent. This strategy would reduce the potential for linkage drag by introgressing a smaller genomic region from N00-3350. Although no QTLs conditioning traits of agronomic importance, such as seed shattering or lodging, are reported on LG's where oleic acid QTL have been found, some germination problems and potential shattering have been associated with the N00-3350 line (data not shown).

The presence of six QTLs for oleic acid content may reduce the effectiveness of MAS when compared to a trait determined by relatively few QTLs with major effects. However, as new fatty acid gene-based SNP markers become available, it may be possible to increase the mapping resolution of the oleic acid QTL identified in this study and allow an increased throughput in future MAS applications.

ACKNOWLEDGEMENTS

We would like to thank Donna Thomas and Bill Novitsky for the fatty acid determination of the lines evaluated. We appreciate the assistance of Francisco Fernández from Monsanto in growing the plants in Puerto Rico. Funding for this research was obtained from the United Soybean Board and the Georgia Agricultural Experiment Research Stations.

REFERENCES

- Boerma, H.R., R.S. Hussey, D.V. Phillips, E.D. Wood, G.B. Rowan, and S.L. Finnerty. 1997. Registration of 'Benning' Soybean. *Crop Sci.* 37:1982.
- Boerma, H.R., R.S. Hussey, D.V. Phillips, E.D. Wood, G.B. Rowan, and S.L. Finnerty. 2000. Registration of 'Boggs' Soybean. *Crop Sci.* 40:294-295.
- 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.
- Brummer, E.C., G.L. Graef, J. Orf, J.R. Wilcox, and R.C. Shoemaker. 1997. Mapping QTL for seed protein and oil content in eight soybean populations. *Crop Sci.* 37:370-378.
- 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.* 46:1010-1011.
- Burton, J.W., R.F. Wilson, and C.A. Brim. 1983. Recurrent selection in soybeans. IV. Selection for increased oleic acid percentage in seed oil. *Crop Sci.* 23:744-747.
- Burton, J.W., R.F. Wilson, and C.A. Brim. 1994. Registration of N79-2077-12 and N87-2122-4, two soybean germplasm lines with reduced palmitic acid in seed oil. *Crop Sci.* 34:313.
- Carter, T.E., and R.F. Wilson. 1998. Soybean quality for human consumption. p.1-16. *In* A. James (ed.) *Proceedings of the Australian Soybean Conference 10*. CSIRO, Brisbane, Australia.
- Carver, B.F., J.W. Burton, R.F. Wilson, and T.E. Carter. 1986. Cumulative response to various recurrent selection schemes in soybean: oil quality and correlated agronomic traits. *Crop Sci.* 26:853-858.
- 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.
- CFSAN, 2003. Center for Food Safety and Applied Nutrition. Office of Nutritional Products, Labeling, and Dietary Supplements. Available at <http://www.crohns.org/governments/cfsan.htm> (Verified 28 Aug 2004).

- Chung, J., H.L. Babka, G.L. Graef, P.E. Staswick, D.J. Lee, P.B. Cregan, R.C. Shoemaker, and J.E. Specht. 2003. The seed protein, oil, and yield QTL on soybean linkage group I. *Crop Sci.* 43:1053-1067.
- Crapiste, G.H., M.I. Brevedan, and A.A. Carelli. 1999. Oxidation of sunflower oil during storage. *J. Am. Oil Chem. Soc.* 76:1437-1443.
- 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, J. Chung, and J.E. Specht. 1999. An integrated genetic linkage map of the soybean genome. *Crop Sci.* 39:1464-1490.
- Darroch, M.A., J.T. Akridge, and M.D. Boehlje. 2002. Capturing value in the supply chain: the case of high oleic acid soybeans. *Inter. Food Agribus. Man. Rev.* 5:87-103.
- 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.
- Diers, B.W., P. Keim, W.R. Fehr, and R.C. Shoemaker. 1992. RFLP analysis of soybean seed protein and oil content. *Theor. Appl. Genet.* 83:608-612.
- Erickson, E.A., J.R. Wilcox, and J.F. Cavins. 1988. Inheritance of palmitic acid percentages in two soybean mutants. *J. Hered.* 79:465-468.
- FAS, 2002. FOP 07-02, Oilseeds: Markets and Trade. Foreign Agricultural Service, U.S. Department of Agriculture, Washington, D. C. July 2002.
- Fasoula, V.A, D.K. Harris, and H.R. Boerma. 2004. Validation and designation of quantitative trait loci for seed protein, seed oil, and seed weight from two soybean populations. *Crop Sci.* 44:1218-1225.
- FDA. 2004. Food labeling and nutrition. U.S. Food and Drug Administration. Available at <http://www.cfsan.fda.gov/~dms/lab-cat.html#transfat> (Verified 28 Oct 2006).
- Fehr, W.R. 1987. Heritability. p. 95-105. *In* W.R. Fehr (ed.) Principles of cultivar development. Vol. 1: Crop species. Macmillan Publishing Co. Ames, IA, USA.
- Giannakas, K., and A. Yiannaka. 2004. The market potential of a new high-oleic soybean oil: An ex ante analysis. *AgBioForum* 7:101-112.

- Hammond, G., and W.R. Fehr. 1984. Improving the fatty acid composition of soybean oil. *J. Am. Oil Chem. Soc.* 61:1713-1716.
- 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.
- Hayakawa, K., Y.Y. Linko, and P. Linko. 2000. The role of trans fatty acids in human nutrition. *European J. of Lipid Sci. and Tech.* 102:419-425.
- Howell, R.W., C.A. Brim, and R.W. Rinne. 1972. The plant geneticist's contribution toward changing lipid and amino acid composition of soybeans. *J. Am. Oil Chem. Soc.* 49:30-32.
- Hu, F.B., M.J. Stampfer, J.E. Manson, E. Rimm, G.A. Colditz, B.A. Rosner, C.H. Hennekens, and W.C. Willett. 1997. Dietary fat intake and the risk of coronary heart disease in women. *N. Engl. J. Med.* 337:1491-1499.
- Hui, Y.H. 1996. Soybean oil. p. 497-601. *In* Y.H. Hui (ed.) *Bailey's industrial oil and fat products*. Vol. 2. *Edible oils and fat products: Oils and oilseeds*, 5th edition. John Wiley and Sons. New York, NY, USA.
- Hyten, D.L., V.R. Pantalone, A.M. Saxton, M.E. Schmidt, and C.E. Sams. 2004a. Molecular mapping and identification of soybean fatty acid modifier quantitative trait loci. *J. Am. Oil Chem. Soc.* 81:1115-1118.
- Hyten, D.L., V.R. Pantalone, C.E. Sams, A.M. Saxton, D. Landau-Ellis, T.R. Stefanik, and M.E. Schmidt. 2004b. Seed quality QTL in a prominent soybean population. *Theor. Appl. Genet.* 109:552-561.
- Keim, P., T.C. Olson, and R.C. Shoemaker. 1988. A rapid protocol for isolating soybean DNA. *Soybean Genet. News.* 15:150-152.
- Kosambi, D.D. 1944. The estimation of map distances from recombination values. *Ann. Eugen.* 12:172-175.
- Lee, S.H., M.A. Bailey, M.A.R. Mian, T.E. Carter, Jr., E.R. Shipe, D.A. Ashley, W.A. Parrott, R.S. Hussey, and H.R. Boerma. 1996. RFLP loci associated with soybean seed protein and oil content across populations and locations. *Theor. Appl. Genet.* 93:649-657.
- Li, Z., L. Jakkula, R.S. Hussey, J.P. Tamulonis, and H.R. Boerma. 2001. SSR mapping and confirmation of QTL from PI96354 conditioning soybean

- resistance to southern root-knot nematode. *Theor. Appl. Genet.* 103:1167-1173.
- Li, Z., R.F. Wilson, W.E. Rayford, and H.R. Boerma. 2002. Molecular mapping genes conditioning reduced palmitic acid content in N87-2122-4 soybean. *Crop Sci.* 42:373-378.
- Lichtenstein, A.H., L.M. Ausman, S.M. Jalbert, and E.J. Schaefer. 1999. Effects of different forms of dietary hydrogenated fats on serum lipoprotein cholesterol levels. *N. Engl. J. Med.* 340:1933-1940.
- Manly, K.F., R.H. Cudmore Jr, and J.M. Meer. 2001. Map Manager QTX, cross-platform software for genetic mapping. *Mammalian Genome* 12:930-932.
- Mansur, L.M., J.H. Orf, K. Chase, T. Jarvik, P.B. Cregan, K.G. Lark. 1996. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. *Crop Sci.* 36:1327-1336.
- Mazur, B., E. Krebbers, and S. Tingey. 1999. Gene discovery and product development for grain quality traits. *Science* 285:372-375.
- Mensink, R.P., E.H. Temme, and G. Hornstra. 1994. Dietary saturated and *trans* fatty acids and lipoprotein metabolism. *Ann. of Med.* 26:461-464.
- Mercer, L.C., J.C. Wynne, and C.T. Young. 1990. Inheritance of fatty acid content in peanut oil. *Peanut Sci.* 17:17-21.
- Miller, J.F., D.C. Zimmerman, and B.A. Vick. 1987. Genetic control of high oleic acid content in sunflower oil. *Crop Sci.* 27:923-926.
- Mounts, T.L., K. Warner, G.R. List, R. Kleiman, W.R. Fehr, E.G. Hammond, and J.R. Wilcox. 1988. Effect of altered fatty acid composition on soybean oil stability. *J. Am. Oil Chem. Soc.* 65:624-628.
- O'Bryne, D.J., D.A. Knauff, and R.B. Shireman. 1997. Low fat-monounsaturated rich diets containing high-oleic peanuts improve serum lipoprotein profiles. *Lipids* 32:687-695.
- 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.
- Palmer, R.G., T.W. Pfeiffer, G.R. Buss, and T.C. Kilen. 2004. Qualitative genetics. p. 137-233. *In* H.R. Boerma and J.E. Specht (ed.) *Soybeans: Improvement*,

- production, and uses. 3rd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- Rahman, S.M., T. Kinoshita, T. Anai, and Y. Takagi. 2001. Combining ability in loci for high oleic and low linolenic acids in soybean. *Crop Sci.* 41:26-29.
- SAS® Institute. 2001. SAS/STAT User's guide. Version 8. SAS Inst., Inc. Cary, NC.
- Schnebly, S.R., and W.R. Fehr. 1993. Effect of years and planting dates on fatty acid composition of soybean genotypes. *Crop Sci.* 33:716-719.
- Schuppert, G.F. 2004. Molecular mechanisms underlying the high oleic acid phenotype in sunflower. Ph.D. diss. Oregon State Univ.
- Somerville, C., J. Browse, J.G. Jaworski, and J.B. Ohlrogge. 2000. p. 456-527. Lipids. *In* B.B. Buchanan, W. Gruissem, and R.L. Jones (ed.) *Biochemistry and molecular biology of plants*. American Society of Plant Physiologists. Rockville, MD, USA.
- 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.
- Soybase, The Soybean Breeder's Toolbox. 2006. [Online Database] Available at <http://soybase.org/> (Verified 28 Oct 2006).
- Specht, J.E., K. Chase, M. Macrander, G.L. Graef, J. Chung, J.P. Markwell, M. Germann, H.H. Orf, and K.G. Lark. 2001. Soybean response to water: A QTL analysis of drought tolerance. *Crop Sci.* 41:493-509.
- Stuber, C.W., M. Polacco, and M.L. Senior. 1999. Synergy of empirical breeding, marker-assisted selection, and genomics to increase crop yield potential. *Crop Sci.* 39:1571-1583.
- Tang, J. and R. Scarth. 2004. Effect of genetic background on the target fatty acids of acyl-ACP thioesterase transgenes in *Brassica napus*. *Plant Breeding* 123:254-261.
- Thelen, J.J., and J.B. Ohlrogge. 2002. Metabolic engineering of fatty acid biosynthesis in plants. *Metab. Eng.* 4:12-21.
- Töpfer, R., N. Martini, and J. Schell. 1995. Modification of plant lipid synthesis. *Science* 268:681-686.

- USFDA. 1999. Center for Food Safety and Applied Nutrition. A food labeling guide- Appendix A: Definitions of nutrient content claims. Available at <http://vm.cfsan.fda.gov/~dms/flg-6a.html> (Verified 28 Oct 2006).
- Vles, R.O., and J.J. Gottenbos. 1989. Nutritional characteristics and food uses of vegetable oils. p. 63-85. *In* G. Robbelen, R.K. Downey and A. Ashri. (ed.) Oil Crops of the World. McGraw-Hill Publishing Co. New York, USA.
- Wang, S., C.J. Basten, and Z.B. Zeng. 2005. Windows QTL Cartographer 2.5. Department of Statistics, NC State Univ. Raleigh, NC. Available at <http://statgen.ncsu.edu/qtlcart/> (Verified 28 Oct 2006).
- White, H.B., F.W. Quackenbush, and A.H. Probst. 1961. Occurrence and inheritance of linolenic and linoleic acids in soybean seeds. *J. Am. Oil Chem. Soc.* 38:113-117.
- Wilcox, J.R., K.L. Athow, T.S. Abney, F.A. Laviolette, and T.L. Richards. 1980. Registration of 'Century' soybean. *Crop Sci.* 20:415.
- 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.R. Cavins. 1990. Registration of C1726 and C1727 soybean germplasm with altered levels of palmitic acid. *Crop Sci.* 30:240.
- Willet, W.C. 1994. Diet and health: what should we eat? *Science* 264: 532-537.
- 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. 1987. Seed Metabolism. p. 643-686. *In* J.R. Wilcox (ed.) Soybeans: Improvement, production, and uses. 2nd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- 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 (ed.) Lipid biotechnology. Marcel Dekker, Inc. IL, USA.
- Wilson, R.F. 2004. Seed composition. p. 621-669. *In* H.R. Boerma and J.E. Specht (ed.) Soybeans: Improvement, production, and uses. 3rd edition. ASA, CSSA, and SSSA. Madison, WI, USA.

Table 3.1. Mean fatty acid content of the parents and mean range of progeny from the G99-G725 × N00-3350 population grown in an Athens, GA greenhouse and in the field at Isabela, PR. The values reported for the parental lines are an average of 16 samples per entry (two locations with eight samples per location).

Line (s)	Palmitic	Stearic	Oleic	Linoleic	Linolenic
	----- g kg ⁻¹ -----				
G99-G725	113	27	206	568	87
N00-3350	86	35	583	280	36
F _{2:3} range	81.5 – 124.5	19.5 - 39	228 - 613	230 - 579	27 - 109
F _{2:3} mean	101	28	388	432	51
LSD _(0.05)	3.3	2.8	29.7	24.2	4.8

Table 3.2. SSR markers associated with the oleic acid content in 259 lines from G99-G725 × N00-3350.

LG	cM	Marker	SF-ANOVA [‡]		Single LG MR [§]	Multiple LG MR
			R ² [¶]	2a [‡]	Partial R ²	Partial R ²
			%	-- g kg ⁻¹ --	%	%
A1	85.6	Satt599	6**	53.3		
A1	92.9	Satt200	5**	44.2		
A1	93.2	Satt236	4**	39.6		
A1	95.2	Satt225	4**	43.2		
A1	96.0	Satt211	4**	42.3	4**	8***
D2	79.2	Satt389	6***	56.2	4***	4**
D2	84.6	Satt311	7***	56.1		
D2	85.1	Satt226	8***	42.5		
G	21.9	Satt235	8***	61.8	8***	8***
G	43.4	Satt394	13***	76.8		
G	47.3	Satt501	10***	57.8		
G	52.9	Satt594	13***	54.0		
G	53.4	Satt303	13***	62.5		
G	76.8	Satt288	8***	44.5		
G	80.4	Satt612	7***	41.9		
G	96.6	Satt191	7***	47.1		4***
L	30.2	Satt143	7**	46.3		
L	30.9	Satt418	9***	55.4	3**	3**
L	34.5	Satt313	8***	52.7		
L	56.1	Satt156	15***	72.6	7***	
L	66.5	Satt166	8***	71.0		
L	70.4	Satt527	10***	62.3		
L	71.4	Satt561	25***	78.8		3*
TOTAL						30.0

[‡]SF-ANOVA = single factor analysis of variance. [§]MR = multiple regression analysis including significant markers within each linkage group and across linkage groups. [¶]R² = % of the total trait variance explained by the genotype at a marker locus. [‡]2a = the difference in oleic acid content at a SSR marker homozygous for the N00-3350 allele - homozygous for the G99-G725 allele. ** p < 0.01, *** p < 0.001.

Table 3.3. Mean oleic acid content for SSR markers associated with putative oleic acid QTL in 259 lines from G99-G725 × N00-3350.

Oleic					
QTL	LG	Marker	GG [‡]	GN	NN
			-- g kg ⁻¹ --	-- g kg ⁻¹ --	-- g kg ⁻¹ --
1	A1	Satt211	365.6 (58) [§]	384.6 (122)	409.0 (71)
2	D2	Satt389	357.9 (49)	383.4 (121)	412.0 (64)
3	G	Satt394	353.6 (56)	383.9 (132)	430.5 (54)
4	G	Satt191	351.4 (62)	382.9 (118)	401.1 (55)
5	L	Satt418	371.2 (50)	377.7 (148)	427.6 (57)
6	L	Satt561	356.4 (67)	386.3 (134)	436.0 (52)

[‡]GG = Homozygous for the allele from G99-G725; NN = homozygous for the allele from N00-3350; GN = Heterozygous for the alleles from G99-G725 and N00-3350; [§]number of F_{2:3} lines in each class is shown in parenthesis.

Table 3.4. SF-ANOVA for markers associated with soybean fatty acid content in G99-G725 × N00-3350.

LG	SSR	cM	Palmitic		Stearic		Linoleic		Linolenic	
			R ² [¶]	2a [‡]	R ²	2a	R ²	2a	R ²	2a
			%	g kg ⁻¹	%	g kg ⁻¹	%	g kg ⁻¹	%	g kg ⁻¹
A1	Satt599	85.6	1*	2		ns	5***	-41		ns
A1	Satt225	95.2	2*	3		ns	6***	-44		ns
A1	Satt211	96.0	4***	4		ns	5***	-42		ns
D2	Satt311	84.6	2*	-3		ns	6***	-46		ns
D2	Satt226	85.1	ns [§]			ns	5***	-41		ns
G	Satt235	21.9	3**	-4		ns	4***	-43		ns
G	Satt394	43.4	6***	-5		ns	9***	-57		ns
G	Satt191	96.6	4***	-4		ns	6***	-43		ns
L	Satt418	30.9	2*	-3		ns	5***	-45		ns
L	Satt561	71.4	4***	-4		ns	8***	-55		ns

[¶]R² = % of the total trait variance explained by the genotype at a marker locus. [‡]2a = the difference in oleic acid content at a SSR marker homozygous for the N00-3350 allele - homozygous for the G99-G725 allele. * p < 0.05, ** p < 0.01, *** p < 0.001. [§]ns = indicates a non-significant marker association.

Table 3.5. Mean fatty acid content of parents and the mean range of 231 progeny lines from the G99-G3438 × N00-3350 population used for confirmation. The values reported for the parental lines are an average of 16 samples per entry (two locations with eight samples per location).

Line(s)	Palmitic	Stearic	Oleic	Linoleic	Linolenic
----- g kg ⁻¹ -----					
<u>Athens, GA 2004</u>					
G99-G3438	120	38	185	565	92
N00-3350	83	32	631	231	23
F _{2:3} range	79 - 128	25 - 54	226 - 649	207 - 574	27 - 74
F _{2:3} mean	103	34	395	424	44
LSD _(0.05)	2.8	2.3	23.6	19.5	2.6
<u>Isabela, PR 2004</u>					
G99-G3438	124	32	186	588	69
N00-3350	85	28	668	200	20
F _{2:3} range	81 – 131	20 – 37	212 – 635	216 – 590	21 – 67
F _{2:3} mean	108	26	409	419	38
LSD _(0.05)	2.7	1.7	8.8	7.3	1.7

Table 3.6. Marker regression analysis for oleic acid content using 231 lines from G99-G3438 × N00-3350.

QTL	LG	cM	Marker	Puerto Rico 2004	
				R ² [¶]	2a [‡]
				%	g kg ⁻¹
1	A1	85.6	Satt599	2*	37
	A1	96.0	Satt211	3**	45
2	D2	85.1	Satt226	ns [§]	
3	G	43.3	Satt394	3*	43
4	G	96.6	Satt191	2*	41
5	L	30.9	Satt418	ns	
	L	47.3	Satt501	7***	70
6	L	71.4	Satt561	2**	36
	L	106.4	Satt513	ns	

[¶]R² = % of the total trait variance explained by the genotype at a marker locus. [‡]2a = the difference in oleic acid content at a SSR marker homozygous for the N00-3350 allele - homozygous for the G99-G725 allele. * p < 0.05, ** p < 0.01, *** p < 0.001. [§]ns = indicates a non-significant marker association.

Table 3.7. Marker associations with oleic acid content from the sub-samples from G99-G3438 × N00-3350 in Athens. The early and late sub-samples were separately genotyped and analyzed for fatty acids.

LG	cM	Marker	163 lines without E and L		163 lines + 67 early lines		163 lines + 67 late lines	
			R ² [¶]	2a [‡]	R ²	2a	R ²	2a
			%	g kg ⁻¹	%	g kg ⁻¹	%	g kg ⁻¹
A1	85.6	Satt599	ns [§]		ns		ns	
	96.0	Satt211	ns		ns		ns	
D2	69.3	Satt397	2*	39	ns		ns	
	85.1	Satt226	ns		ns		ns	
	84.6	Satt311	5**	46	4*	43	3*	42
	93.7	Satt301	ns		ns		ns	
G	21.9	Satt235	ns		ns		ns	
	43.4	Satt394	ns		ns		ns	
	96.6	Satt191	ns		ns		ns	
L	30.9	Satt418	ns		ns		ns	
	66.5	Satt166	4**	43	5***	45	6***	52
	70.4	Satt527	ns		ns		ns	
	71.4	Satt561	4*	44	7***	54	7***	58
	106.4	Satt513	3*	47	4**	44	12***	74
	115.1	Sat_245	ns		2*	29	2*	28
O	118.1	Satt153	7**	88	5**	49	9***	63

[¶]R² = % of the total trait variance explained by the genotype at a marker locus. [‡]2a = the difference in oleic acid content at a SSR marker homozygous for the N00-3350 allele - homozygous for the G99-G725 allele. * p < 0.05, ** p < 0.01, *** p < 0.001. [§]ns = indicates a non-significant marker association.

Figure 3.1. Pedigree of N00-3350.

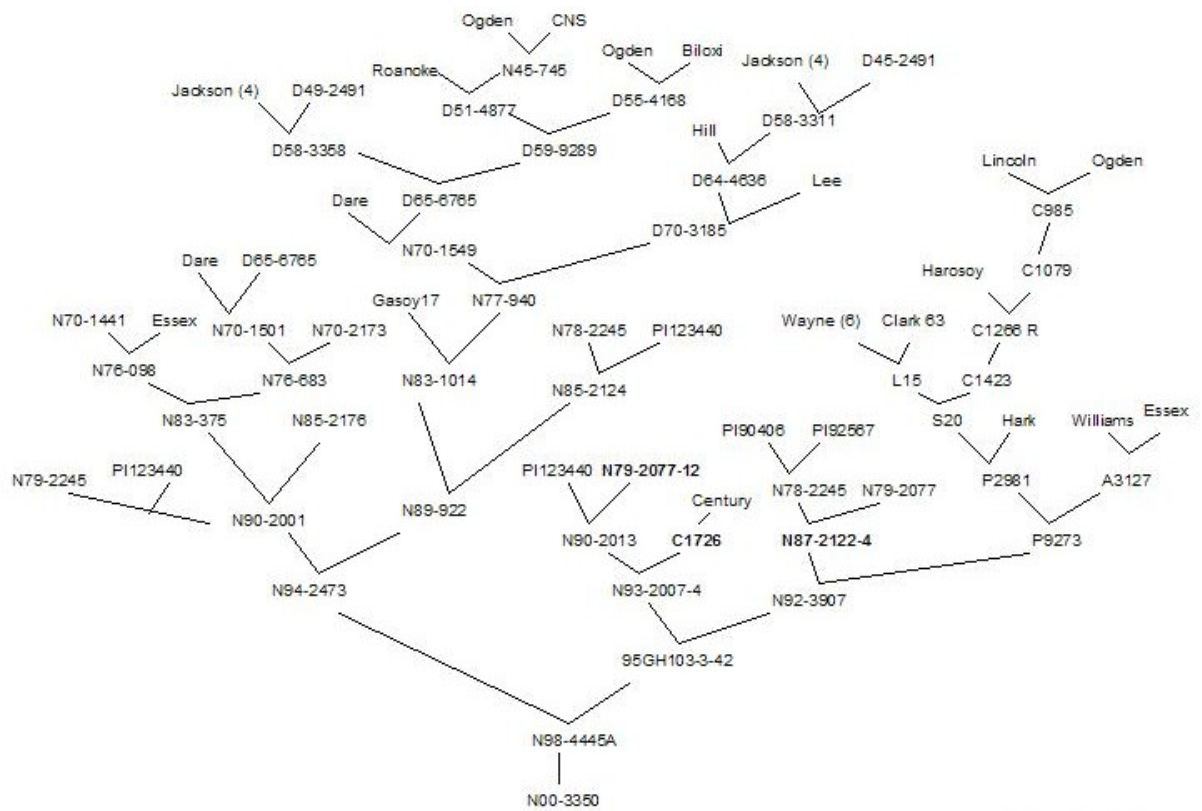


Figure 3.2. QTL likelihood plots from interval mapping for oleic acid QTL using 259 lines from the G99-G725 $\times$ N00-3350 population. For each linkage group (LG), the permutation-derived ($n = 1000$ per trait) LOD score significance criteria are indicated by a vertical dotted line at 3.0.

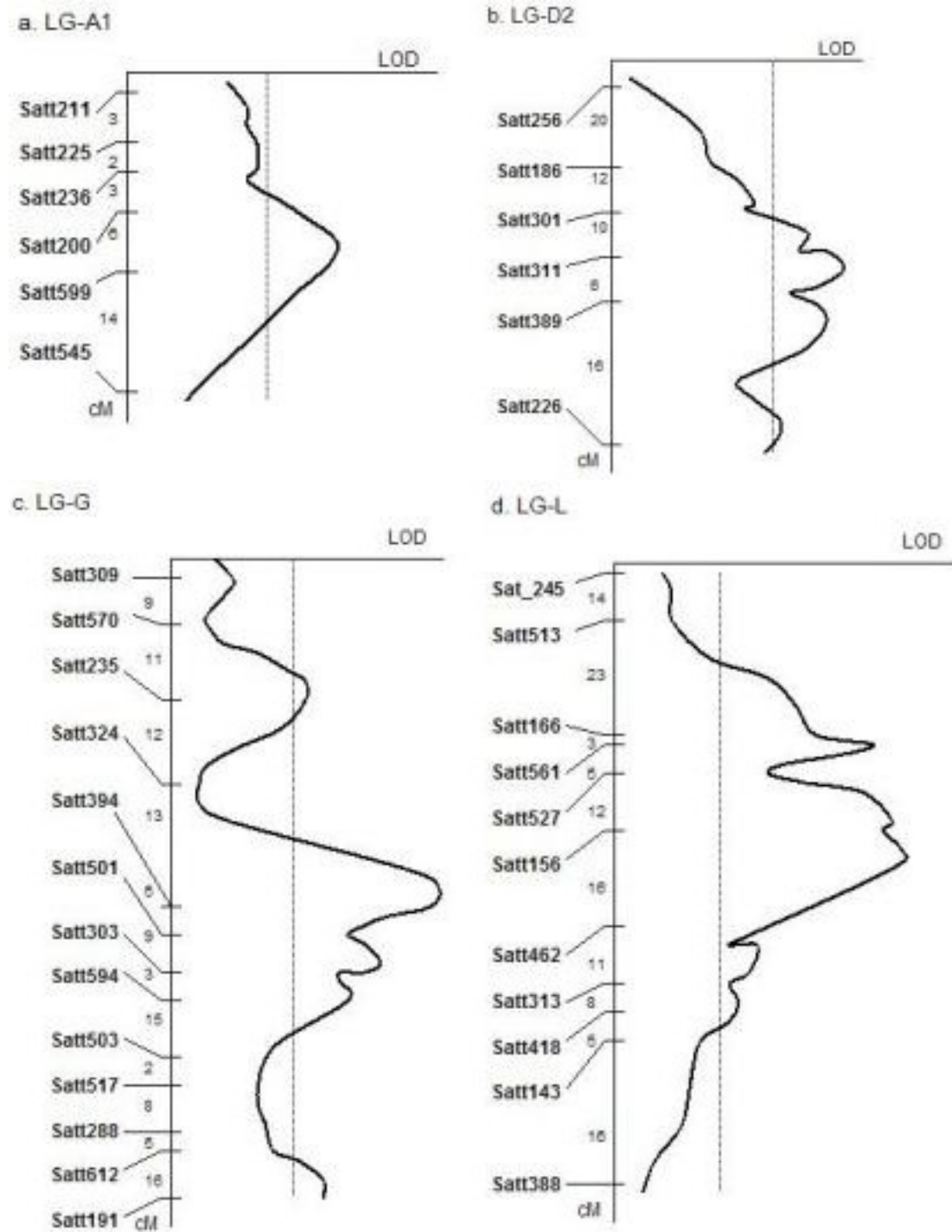
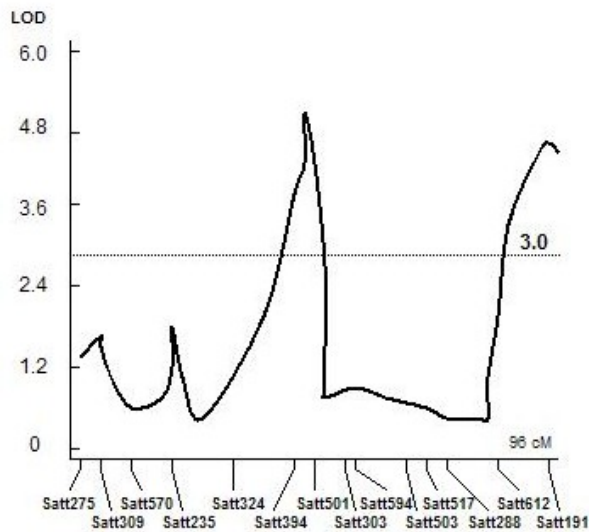
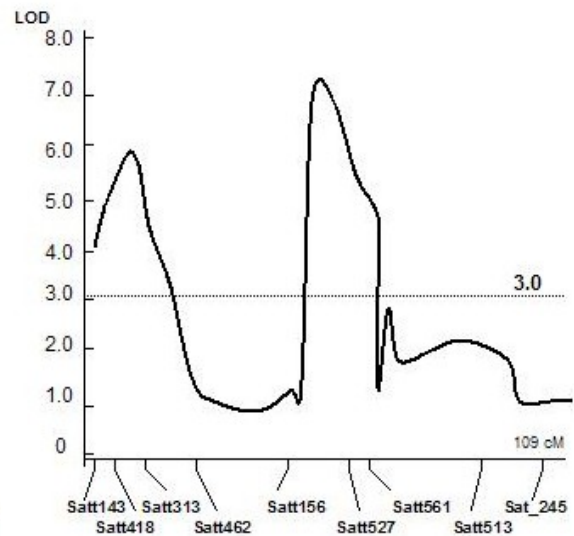


Figure 3.3. Composite interval mapping for oleic acid QTL using 259 lines from G99-G725 $\times$ N00-3350 and the combined oleic acid content. The dashed line shows the LOD plot when background markers were used. a. LG-G. b. LG-L. The permutation-derived ($n = 1000$ per trait) LOD significance criteria are indicated by a dashed horizontal line at 3.0.

a. LG-G



b. LG-L



CHAPTER 4
DISCOVERY AND MAPPING OF SEQUENCE-BASED MARKERS
ASSOCIATED WITH OLEIC ACID CONTENT IN N00-3350 SOYBEAN¹

¹ Maria J. Monteros, Perry B. Cregan, and H. Roger Boerma. To be submitted to Molecular Breeding.

ABSTRACT

Soybean oil, *Glycine max* (L.) Merr, is the major oilseed produced and consumed in the world today. An increase in the amount of oleic acid in soybean oil would decrease its total saturated fatty acid content and reduce the need for hydrogenation, a process that creates unhealthy *trans* fatty acids. Oleic acid-QTLs from N00-3350 soybean ($\sim 550 \text{ g kg}^{-1}$ oleic acid) mapped to linkage groups (LG) A1, D2, G, and L. The objectives of this study were to develop and map sequence-based molecular markers from fatty acid pathway genes and sequence-tagged sites in regions of the soybean genome previously associated with oleic acid QTLs. Single-strand conformation polymorphisms (SSCP) in polymerase chain reaction (PCR) products are used to evaluate the polymorphisms in the targeted regions. An $F_{2:3}$ population of 259 individuals from the cross of G99-G725 ($\sim 206 \text{ g kg}^{-1}$ 18:1) $\times$ N00-3350 ($\sim 583 \text{ g kg}^{-1}$ 18:1) was used to map the location of the markers developed. Using this approach four gene sequence-based markers and twelve sequence-tagged markers have been mapped to linkage groups previously associated with oleic acid QTLs. SNP markers associated with soybean fatty acid genes were also mapped. These markers can be used by breeders in MAS to incorporate desirable alleles for fatty acid content into elite soybean cultivars. SSCP markers can provide additional information for the development of breeder-friendly markers.

INTRODUCTION

Soybean oil is the major oilseed produced and consumed in the world and accounts for approximately 35% of the world's total oilseed production (Wilcox, 2004). Soybean oil predominately consists of palmitic, stearic, oleic, linoleic, and linolenic fatty acids (Cahoon, 2003). Vegetable oils, like that from soybean, contain high concentrations of linoleic and linolenic acid are susceptible to the oxidation process, which affects the flavor stability of the oil (Wilson 1987). The process of hydrogenation increases the oxidative stability of the oil, but also forms *trans* fatty acids. The intake of *trans* fatty acids has been correlated with detrimental health effects (Lichtenstein et al., 1999). These findings have prompted the FDA to require that the amount of *trans* fatty acids be listed on the Nutrition Facts panel of a product's label (U.S. FDA, 2004). An increase in the monounsaturated oleic acid would provide the oxidative stability without the need for hydrogenation (Mercer et al., 1990).

Although fatty acid composition in various plant organs of soybean and cultured cells from those organs may be different, the biochemical mechanism for fatty acid synthesis is highly conserved in plant tissues (Harwood, 1988). The fatty acid biosynthetic pathway in plants has been characterized and many of the genes underlying plant lipid synthesis have been cloned and sequenced (Somerville et al., 2000; Töpfer and Martini, 1994) (Fig. 4.1). Key members of the pathway include acetyl-CoA carboxylase (ACCase), ketoacyl-ACP synthase (KAS I, II and III), stearyl-ACP desaturase (FAB2), fatty acid desaturase (FAD2 and FAD3) (Somerville et al., 2000).

In soybean, two stearyl-ACP desaturase (SACPD) genes, designated SACPDA and SACPDB have been identified. The gene expression or enzyme activity of SACPD may affect the levels of both stearic and oleic acid in soybean seeds (Byfield et al., 2006). Heppard et al., (1996) reported the isolation of two different cDNA sequences that encode omega-6 desaturases in soybean. Soybean possesses three Fad3 genes, *GmFad3A*, *GmFad3B*, and *GmFad3C*. Mutations in two of the three soybean Fad3 genes were

accompanied by an increase in linoleic acid content and a reduction in linolenic acid levels (Bilyeu et al., 2005). The availability of gene sequences can be used to develop genetic markers for these genes (Slabaugh et al., 1997).

Researchers at the USDA-ARS in Raleigh, NC developed the line N98-4445A, the seed of which contains 500 to 600 g kg⁻¹ oleic acid content and can be used to increase the oleic acid content in soybean cultivars (Burton et al., 2006). N00-3350 is a single plant selection from the mid-oleic acid line N98-4445A. The average oleic acid content for N00-3350 when grown in Athens and Puerto Rico was 627 g kg⁻¹ oleic acid content (Chapter 3).

One of the goals of the dissection of complex traits is to identify the possible genes involved in the trait. The identification of quantitative trait loci (QTLs) that affect a trait is a first step to identifying the underlying genes controlling the trait of interest. Six oleic acid QTLs were identified in a F_{2:3} population of 259 lines from the cross of G99-G725 × N00-3350, and confirmed in a F_{2:3} population of 231 lines from the cross of G99-3438 × N00-3350. The confirmed oleic acid-QTLs are on LG-A1 near Satt211, on LG-D2 near Satt389, on LG-G near Satt394 and Satt191, and on LG-L near Satt418 and Satt561 (Table 3.4, Chapter 3). To be able to manipulate a trait more precisely, it is desirable to understand the structure and function of the genes involved in the expression of the trait (Stuber et al., 1999). One technique for identifying genes underlying QTLs is to utilize a candidate gene approach (Faris et al., 1999).

Use of the candidate gene approach utilizes information regarding the biochemical pathway involved in expression of the trait and the candidate gene co-segregation with a previously identified QTL for the trait. EST sequencing projects in soybean have provided a wealth of sequence information that allows for rapid access to many of soybean gene sequences (Shoemaker et al., 2002). Annotations from genes in the fatty acid synthesis pathway from the model organism *Arabidopsis* provide information that can

be used to determine the functions of these sequences (Arondel et al., 1992; Iba et al., 1993; Lemieux et al., 1990).

Methods used to map sequenced genes have different costs, requirements for quantity and quality of DNA, complexity, and resolution (Rafalski and Tingey, 1993). Single strand conformation polymorphism (SSCP) is a widely used method based on the electrophoretic detection of conformational changes in single-stranded DNA molecules resulting from point mutations or other forms of small nucleotide changes (Xie et al., 2002). This technique has been used extensively in human genetics to detect mutations within genes, and was able to detect single-point mutations (Hayashi, 1992). Single-strand conformation polymorphism analysis is considered a sensitive, relatively inexpensive, and rapid method to detect sequence variation (Sekiya, 1993). SSCP markers have been used for gene mapping (Slabaugh et al., 1997).

Single-nucleotide polymorphisms (SNPs) are single base differences between homologous DNA fragments plus small insertions and deletions (indels) (Zhu et al., 2003). Fourteen soybean germplasm accessions that were estimated to have contributed 80.5% of the allelic diversity present in North American soybean cultivars were used to determine the SNP frequency in coding and non-coding soybean DNA sequence (Gizlice et al., 1994; Zhu et al., 2003). In general, nucleotide diversity was higher in random non-coding genomic sequences obtained from BAC clones and SSR flanking regions than in genomic DNA associated with genes. It was estimated that in approximately 18 kb of soybean DNA, the frequency of SNPs is 3.4 per kb (Zhu et al., 2003). SNPs may occur within the coding region or outside the coding regions, but most are in the non-coding regions. Mutations within the coding regions may affect protein function or result in a neutral substitution that may not affect protein function. Alternatively, non-coding SNPs may alter the regulation of gene expression. For example, a SNP in the promoter region may decrease the activity for sequence-specific DNA binding proteins (Shastry, 2004). The availability of a large number of SSR loci can be used

as a resource of sequence-tagged-sites (STSs) from which SNPs can be discovered (Cregan et al., 1999b). These SSR or RFLP markers can be used to identify BAC clones from which sequence can be obtained to discover SNPs at defined positions in the genome (Cregan, 2000).

An understanding of the likely genes involved in the fatty acid synthesis pathway is an important component in the development of soybean lines with a stable expression of higher oleic acid in seeds. Additionally, this will permit a better characterization of the fatty acid synthesis pathway in soybean, and ultimately the development of markers well suited for use in marker-assisted selection (MAS) for the improvement of oleic acid content. The objectives of this research were to develop and map sequence-based markers from genes in the fatty acid biosynthetic pathway and genomic DNA in regions of the soybean genome where putative oleic acid-QTLs have been identified in N00-3350.

MATERIALS AND METHODS

An $F_{2:3}$ population consisting of 259 plants derived from the cross of G99-G725 ($\sim 206 \text{ g kg}^{-1}$ 18:1) $\times$ N00-3350 ($\sim 583 \text{ g kg}^{-1}$ 18:1) was used as a mapping population. G99-G725 is a glyphosate-resistant version of Boggs (Boerma et al., 2000). The $F_{2:3}$ lines were planted and harvested in Athens and Puerto Rico as previously described in Chapter 3. DNA was isolated as described in Keim et al. (1988).

A bulked random sample of 10 seeds from each plant was sent to USDA-ARS laboratories in Peoria, IL, and Raleigh, NC, for fatty acid analysis. If there were ≥ 35 $F_{2:3}$ seeds, a bulked random sample of 10 seeds from each plant was sent to USDA-ARS laboratories in Peoria, IL, and Raleigh, NC, for fatty acid analysis. If 18 to 34 seeds were available, 10 $\frac{1}{2}$ -seed chips were sent to each laboratory. If fewer than 18 seeds were available, $\frac{1}{4}$ -seed chips of each seed were sent to each laboratory. Fatty acid content was determined using gas chromatography (Hewlett Packard Model 5890/6890) to

evaluate methyl esters. The fatty acid data of each entry used in the analysis is the average from the Peoria and Raleigh laboratories.

Molecular primers evaluated

Data from the soybean lipid metabolic pathway deposited in The Institute for Genomic Research (TIGR) database (available at http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=soybean) was used to obtain sequences for genes involved in the lipid biosynthesis pathway from GenBank (NCBI, Bethesda, MD, USA) (Table 4.1). UniGene sets for soybean fatty acid gene sequences were obtained from NCBI (available at <http://www.ncbi.nlm.nih.gov/UniGene/UGOrg.cgi?TAXID=3847>) (Table 4.2). Each UniGene entry is a set of transcript sequences that appear to come from the same transcription locus (gene or expressed pseudogene). SequencherTM was used to construct a consensus sequence from all of the EST sequences available for each UniGene set. This consensus sequence was used to develop sequence-based primers using the Oligo Lite V 6.0 program (Molecular Biology Insights, Inc., Cascade, CO, USA). Primers for the soybean Fad2, Fad3, and SACP genes have been previously reported (Heppard et al., 1996; Bilyeu et al., 2003; Bilyeu et al., 2005; Byfield et al., 2006). Fad2 primers from olive, *Olea europaea* cv. Picual, were obtained from Hernández et al. (2005). Primers from sunflower, *Helianthus annuus*, fatty acid genes were obtained from Schuppert (2004). Additional fatty acid gene sequence-based primers were obtained from Dr. Perry Cregan (USDA-ARS, Bethesda, MD) (Table A.3.1).

BAC contigs of interest were identified using anchored SSR markers associated with oleic acid QTLs in N00-3350 soybean (Soybean Breeder's Toolbox, 2006). BAC-end sequences were obtained from GenBank (<http://www.ncbi.nlm.nih.gov/>) and used as templates for primer development (Table 4.3). The software OligoLite V 6.0 program (Molecular Biology Insights, Inc., Cascade, CO, USA) was also used to design primers using as template sequence from the physical map (Rychlik et al., 1990). Additional

primers that amplified soybean genome sequences on LG-A1, D2, G, and L were also obtained from Perry Cregan (USDA/ARS, Beltsville, MD).

For each gene or genomic sequence tested, the parents G99-G725 and N00-3350, were used to optimize PCR conditions and identify primer pairs that produced a single product from genomic DNA. The 10- μ L PCR reaction mix contained 2 μ L of 50 ng μ L⁻¹ template DNA, 2 μ L primer pairs (2.5 μ M each), 5 μ L Epicentre MasterAmpTM PCR B buffer, and 0.5 unit of Promega *Taq* DNA polymerase (Madison, WI). The optimal annealing temperature varied with each primer pair and ranged from 48 to 60°C (Appendix 3). PCR product amplification was verified using 2% agarose gels stained with ethidium bromide. PCR products were subjected to single strand conformation polymorphism (SSCP) analysis according to Slabaugh et al. (1997). A 3- μ L sample of each PCR product was added to 9 μ L denaturing solution (95% formamide, 0.01 M NaOH, 0.05% xylene cyanol, 0.05% bromophenol blue), heated to 94°C for 2 min, then chilled in a ice-water slurry. Samples of 3-5 μ L were run on 0.5 x MDE gels (Cambrex Bio Science, Rockland, Rockland, ME, USA) (1 mm thick x 50 cm wide x 22 cm high) using 0.6 x TBE running buffer. One of the glass plates was treated with γ -methacryloxypropyltrimethoxysilane (Sigma Chemical Co., St. Louis, MO, USA), so that the gel would remain attached to the glass plate during staining. Gels were run on a DASG-400-50 polyacrylamide gel apparatus (C.B.S. Scientific Co., Del Mar, CA, USA) at room temperature at 7.0 watts at constant power for 14 to 20 h, depending on the size of the fragments. SSCP gels were silver stained and manually scored as described in Sanguinetti et al. (1994). The segregating population of 259 F_{2:3} lines from the cross of G99-G725 $\times$ N00-3350 was screened with the identified polymorphic markers.

Sequencing

Single bands from the parental lines were excised from SSCP gels, and added to 100 μ L of TE buffer. DNA was eluted for 60 min at 65°C with shaking. Two μ L of the buffer-DNA solution were added to a second PCR

reaction (20 μ l) and re-amplified using the same primers. PCR products were digested with shrimp alkaline phosphatase, SAP (0.1 U μ l⁻¹, USB Corporation, Cleveland, Ohio USA) and *ExoI* nuclease (0.02 U μ l⁻¹, USB Corporation, Cleveland, Ohio USA) according to the manufacturer's protocol. The forward and reverse primers were used to directly sequence the PCR products in both directions in separate sequencing reactions. Each sequencing reaction consisted of 2 μ l of BigDye Terminator Cycle Sequencing Ready Reaction mix from the v 3.0 Kit (Applied Biosystems, Foster City, CA), 1.0 μ M of each primer, 1% DMSO, 2 μ l of 5x sequencing buffer (supplied with the sequencing kit), and 4 μ l of PCR product in a total volume of 10 μ l. Cycle sequencing conditions were as recommended by the kit manufacturer using annealing temperatures optimized for each primer pair. Sequencing reactions were purified using the MultiScreen Filtration System (Millipore Corporation, Bedford, MA) and Sephadex G50 Superfine (Sigma-Aldrich, St. Louis, MO) per the manufacturer's protocol. Sequence analysis was carried out on a Perkin Elmer 3730 XL capillary DNA Analyzer (Applied Biosystems, Foster City, CA). Sequences were aligned using ClustalW (Thompson et al., 1994), and potential SNPs were identified by visually inspecting the alignments considering the quality of the sequence. The sequence quality was based on the chromatograms, which were visualized using Chromas 2.31 software (Technelysium Pty Ltd).

SNP Assays

The SNaPshotTM SBE procedure uses an oligonucleotide probe that anneals adjacent to the SNP of interest, and is followed by an extension step with a fluorescently-labelled dideoxy terminator (ABI PRISM®, 2000). The single-base extension (SBE) capture probes used to interrogate each potential SNP site were designed to terminate one nucleotide 5'- downstream of the SNP location. The probes also included a slightly modified 21-nucleotide ZipCode sequence from Iannone et al. (2000), at the 5' end and a site-specific sequence at the 3' end adjacent to the SNP location. The two

parents (G99-G725 and N00-3350) were used to evaluate the potential SNP following the protocol in ABI Prism® SNaPshot™ Multiplex Kit (Applied Biosystems, Foster City, CA). Reactions were run on 12-cm gels using an ABI Prism 377 DNA sequencer at 750 V for 1.5 to 2 hours, with 1X TBE buffer. Each nucleotide has a unique color label, and therefore a different color product observed from each parent confirms the SNP. For SNP determination of the F_{2:3} lines, an SBE assay was developed for the Luminex100 flow cytometry platform (Luminex Corp., Austin, TX). This platform has high accuracy and efficiency at differentiating between the two homozygous and the heterozygous classes (Lee et al., 2004). The PCR reaction consisted of 30 ng template DNA, 0.5 µM of each primer, 1.5 mM MgCl₂, 200 µM of each dNTP, 0.5 µl 10X AccuPrimeTaq® buffer, and 0.1 µl of AccuPrimeTaq® DNA polymerase (Invitrogen Corp., Carlsbad, CA) in a 5 µl reaction. PCR conditions consisted of an initial denaturing step at 94°C for 2 min, followed by 30 cycles of denaturing at 94°C for 30 seconds, annealing at 46°C for 30 seconds, and extension for 68°C for 1 min. PCR reactions were run in a PTC-225 Peltier Thermal Cycler (MJ Research, Inc., Watertown, MA) using an annealing temperature optimized for each primer pair.

A total of 5 µl PCR product was treated with 1 unit each of SAP and *ExoI* to degrade excess primers and dNTPs. The reaction mixture was incubated at 37°C for 1 hour, followed by 15 min at 80°C to inactivate the enzymes. Two reactions were set up, differing only in the biotin-labeled ddNTP added, which was determined based on the nucleotide of the parental lines at the specific SNP. A second extension reaction containing 2.5 µl aliquot of the enzyme-treated PCR products, 0.5 µl 10X Promega buffer, 0.064 units of Thermo Sequenase (USB Corp., Cleveland, OH), 3 mM MgCl₂, 0.12 µM SBE capture probe primer, 0.4 µM allele-specific biotin-labeled ddNTP determined by the parental dNTP's, and 0.4 µM of each of the other three non-labeled ddNTPs. The same SNaPshot™ SBE capture probes were used for this assay.

A total of 5.0×10^6 carboxylated LabMAP microspheres (MiraiBio, Alameda, CA, USA) per assay was centrifuged into a pellet, and re-suspended in 50 μL of 0.1 M 2-(N-morpholino) ethanesulfonic acid (MES) buffer (pH = 4.5). A total of 1 nm of amino-substituted cZipCode oligonucleotide (1 μL of a 1 mM solution) was added to the suspension. Additionally, a 2.5- μL aliquot of 1-ethyl-3-(3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) solution at 10 mg mL^{-1} was added to the same tube and incubated in the dark at room temperature. After 30 min, another 2.5 μL of fresh EDC solution were added, and again incubated for 30 min in the dark at room temperature. One ml of 0.02% Tween 20, and then 1 ml of 0.1% sodium dodecyl sulphate (SDS) solution were used to wash the microspheres. The microspheres were re-suspended in 100 μL of 0.1 M MES (pH = 4.5). Microspheres coupled with ZipCodes were stored in the dark at 4°C until ready to use.

SBE products were precipitated using 75% ethanol to a final concentration of 60% ethanol, incubated in the dark at room temperature for 30 minutes, centrifuged to obtain a pellet, and dried prior to hybridization. In a 50 μL total reaction volume, 1X TMAC (3 M tetramethylammonium chloride, 50 mM Tris-HCl, pH = 8.0, 4 mM EDTA, pH 8.0, 0.1% Sarkosyl), and 3,000 microspheres coupled with the ZipCode, were denatured at 90°C for 10 minutes. The reaction was hybridized for 30 min at 50°C to 55°C, based on the T_m of the capture probe, and labeled with 200 ng streptavidin in 10 μL of 1X TMAC at 55°C for 5 minutes. Fluorescence of the microspheres and the samples were analyzed using a Luminex 100 cytometer, a Luminex XY plate reader, and Luminex analysis software from MiraiBio Inc. (Alameda, CA). The fluorescence of the microspheres was measured and converted to a mean fluorescence intensity (MFI) value using a minimum reading of 100 microspheres in a 50 μL sample.

Linkage maps were constructed with MapManager QTX b20 (Manly et al., 2001) using the Kosambi (1944) mapping function. Single factor analysis

of variance (SF-ANOVA) was used to determine the significance of marker genotypic class means for oleic acid content using the general linear model (PROC GLM) SAS® Version 8 (SAS Institute, Cary, NC).

RESULTS AND DISCUSSION

A total of 123 markers derived from soybean fatty acid gene sequences was tested for polymorphisms in the G99-G725 and N00-3350 parents. Additionally, 74 markers developed from BACs anchored to the physical map by SSRs were tested (Table 4.3). The SSCP assay was initially used to detect polymorphisms and screen markers. As sequence information from the target amplification sites became available, SBE assays were developed and tested using SNaPshot™ assays to confirm the existence of a SNP at the predicted site. As expected, polymorphism rate between the mapping parents was lower from SSCP markers developed from gene sequences when compared to markers developed from genomic sequences from BACs anchored to the physical map (5.4% vs. 24.0%, respectively).

The order of the SSR markers on the linkage maps obtained from genotyping the 259 F_{2:3} individuals from G99-G725 × N00-3350 on LG-A1, LG-D2, LG-G, and LG-L where oleic acid QTLs have been found generally corresponds with the consensus soybean genetic linkage map (Cregan et al., 1999a; Song et al., 2004). These SSR markers were used as anchors to determine the relative position on the soybean genetic linkage map of the SSCP and SNP markers developed.

Markers developed from soybean fatty acid gene sequences mapped on LG-D2 and LG-G. On LG-D2, the SNP primer SNP16289, obtained from a soybean Fad2 sequence, mapped to the same location as the previously identified oleic acid QTL and based on SF-ANOVA explains 10% of the variation in oleic acid (Table 4.4). The oleic acid content in lines homozygous for the N00-3350 allele at SNP16289 averaged 62.8 g kg⁻¹ more oleic acid content than those homozygous for the G99-G725 allele. The Fad2 genes are targets for improving the oleic content in plants because they encode a

desaturase that controls most of the polyunsaturated lipid synthesis in plant cells and converts oleic acid to linoleic acid (Somerville et al., 2000; López et al., 2000; Schuppert, 2004).

On LG-G we mapped six SSCP markers, four of which (U6041_6, G1, accA_6, and accB_2) were developed from fatty acid gene sequences (Fig. 4.2). U6041_6 mapped within 14 cM of Satt394, which has been associated with an oleic acid QTL (Chapter 3). This SSCP marker was developed from a soybean KasIII (ketoacyl-ACP synthase) sequence, which converts the three-carbon chain in malonyl-ACP to a four-carbon chain (Somerville et al., 2000). An additional marker developed from genomic BAC-end DNA, QTLG_1, mapped 3.9 cM from U6041_6. A QTL near Satt394 also explained 6% of the variation in palmitic acid (Chapter 3), suggesting that this might be an important step in determining the amount of carbon flux through the pathway.

The importance of ACCase in the control of fatty acid synthesis in seeds was confirmed when developing seeds from plants transgenic for biotin carboxyl carrier protein in which ACCase activity was reduced by up to 65% produced mature seeds with a significant decrease in their fatty acid content (Thelen and Ohlrogge, 2002). The SSCP marker accB_2 was mapped near a predicted oleic acid QTL on LG-G (Chapter 3). The biotin carboxyl carrier protein subunit is one of three functional components of ACCase, which catalyzes the first step in the fatty acid biosynthetic pathway. The marker accA_6, was developed from the carboxyl transferase subunit sequence from ACCase, and was mapped 9.2 cM from Satt199 (Fig. 4.2). A previous study by Tajuddin et al. (2003) mapped one of the components of the ACCase to a position 3.3 cM from Satt199 on LG-G. Mapping at least two fatty acid gene sequence-based markers to LG-G provides additional support for multiple oleic acid-QTLs on this linkage group.

The SSR marker Satt211 positioned at 96.0 cM on LG-A1 of the soybean consensus map has been associated with an oleic acid QTL (Chapter 3). The sequence-tagged SSCP markers, QTLA-1 and A1-14639, were mapped to this LG (Fig. 4.2). An oleic acid QTL on LG-A1 in the 88.3 –

93.5 cM region was identified in a previous study from a cross between A81356022 × PI468916 (Soybean Breeder's Toolbox, 2006).

Five new markers have been mapped to LG-L. Four of the SSCP markers mapped to LG-L are near a mapped and confirmed oleic acid QTL identified by Satt561 (Fig. 4.2). The marker 13835, developed from a sequence-tagged site on LG-L explains 7% of the variation in oleic acid content (Table 4.4). At this point we have not mapped any fatty acid gene-sequence markers to LG-A1 or LG-L.

QTLs for oleic acid content have been found on LG-G and LG-L. Shoemaker et al. (1996) reported the existence of duplicated regions in the soybean genome between LG-G and LG-K, as well as LG-K and LG-L, suggesting possible duplications of fatty acid genes. The organization of the soybean genome is consistent with a polyploid origin and a possible additional round of genome duplication (Shoemaker et al., 1996). Many genes following tetraploidization events undergo mutation to eliminate or alter their function (Pickett and Meeks-Wagner, 1995). It is possible that natural selection factors have helped to maintain the function of redundant QTLs for agronomic traits, such as plant height and oil content (Shoemaker et al., 1996).

In addition to the Fad2-based SNP16289 that maps to LG-D2, SSCP and SNP markers developed from Fad2 gene sequences have also been mapped to LG-D1b, and O (Table 4.5; Fig. 4.3). In the G99-G3438 × N00-3350 confirmation population, a population-specific oleic acid QTL was identified on LG-O near Satt153 that explains 9% of the variation in oleic acid content ($p < 0.01$) (Chapter 3). The marker SNP15783, developed from Fad2 sequence mapped 0.8 cM from Satt153 (Fig. 4.3c). SNP15155 developed from Fad3 was also mapped to LG-O. The KAS sequence-based marker U6041_3 was mapped to LG-K.

The lack of extensive sequence variation in coding regions of the enzymes in the fatty acid pathway may limit the usefulness of the available EST sequence data available, suggesting that post-transcriptional

modifications, including gene silencing, and other interactions with the genetic background, may play a role in the variation in oleic acid content observed in N00-3350 soybean. However, we have been able to use SSR markers as sequence-tagged-sites to develop and map new markers to linkage groups in the soybean genome associated with oleic acid content.

Although markers developed using sequence-tagged-sites map close to previously identified oleic acid QTLs, they provide no information about the genes underlying the effects of these QTLs. Using a candidate gene approach, we have been able to identify fatty acid gene-based markers and map them to the same regions as oleic acid QTLs previously identified on LG-A1, D2, and G. However, we were unable to map any markers to LG-L using this approach. Although the metabolic pathways leading to unsaturated fatty acid synthesis are known, the regulation of these pathways is poorly understood (Heppard et al., 1996). Studies in *Arabidopsis* and *Brassica* indicate that the expression of the ACCase and fatty acid synthase genes is controlled by the activity of a global regulator (Ke et al., 2000; O'Hara et al., 2002). Therefore, it is possible that a regulatory element or a putative fatty acid gene that has not been mapped is present on LG-L.

Another consideration of the candidate gene approach is that sequence conservation of the genes in the fatty acid biosynthetic pathway has made it difficult to identify polymorphic markers between the two mapping parents for all the genes in the fatty acid pathway. Previous studies have indicated that there is a higher level of sequence variation in non-coding regions, making them good regions in which to search for SNPs (Van et al., 2004). This strategy could be used to identify candidate linkage groups to locate fatty acid QTLs in soybean compared to a random genomic scan using molecular markers.

Overall, we have mapped gene-based markers that could potentially be used to predict putative genes likely responsible for the previously mapped and confirmed oleic acid QTLs. However, a genetic linkage between candidate genes and the QTLs for oleic acid does not definitely demonstrate

a causal relationship. Incorporating information from fine-mapping experiments and the development of NILs for a given QTL would reinforce a possible causal relationship (Francia et al., 2005). Another strategy is to sequence the parental lines and look for sequence variation that is predicted to have a functional consequence (Borevitz and Chory, 2004). Gene expression studies comparing the parental lines for differences in gene expression for these candidate genes could provide relevance to the sequence variations observed. A strategy using only the lines from the mapping population with extreme phenotypes could facilitate the identification of these expression differences (Borevitz and Chory, 2004).

Ultimately, the putative candidate genes need to be functionally tested. One approach to do this is identify a null mutation, reintroduce the alternate allele into reciprocal QTL lines or null mutant backgrounds using transgenic approaches to show that each allele has a significantly different effect on the phenotype (Borevitz and Chory, 2004). This approach has been used successfully in tomato (Frary et al., 2000). Other ways to confirm a QTL-gene correspondence is to use gene replacement or RNAi to silence the candidate gene and determine if there is an effect on the phenotype. Although some gene-based markers map to previously identified oleic acid QTLs, complementation tests using genetic transformation to insert these candidate genes and determine the change in oleic acid content will provide definite evidence for function.

The availability of gene-based molecular markers should greatly enhance the successful incorporation of multiple alleles for genes at independent loci to increase the oleic acid content of soybean seed. Additionally, although functional verification will be necessary, these markers provide a framework of candidate genes that are putatively responsible for the oleic acid phenotype. Future studies will include development of breeder-friendly and cost-effective SNP markers amenable to high throughput in MAS applications to increase oleic acid content in soybean seed using these SSCP sequence-based primers.

ACKNOWLEDGEMENTS

We would like to thank Dr. David Hyten (USDA-ARS, Beltsville, MD) for providing the fatty acid sequence-based and physical-map based primer sequences on LG-A1, D2, G, and L, and Dr. Randy Shoemaker for the identification of the BACs.

REFERENCES

- ABI PRISM®, 2000. SNaPshot™ Multiplex Kit: Protocol. Applied Biosystems. Foster City, CA.
- Aghoram, K., R.F. Wilson, J.W. Burton, R.E. Dewey. A mutation in a 3-ketoacyl-ACP synthase II gene associated with elevated palmitic acid levels in soybean seeds. *Crop Sci.* 46:2453-2459.
- 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.
- 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.
- Bilyeu, K.D, L. Palavalli, D.A. Sleper, and P.R. Beuselinck. 2005. Mutations in soybean microsomal omega-3 fatty acid desaturase genes reduce linolenic acid concentration in soybean seeds. *Crop Sci.* 45:1830-1836.
- Boerma, H.R., R.S. Hussey, D.V. Phillips, E.D. Wood, G.B. Rowan, and S.L. Finnerty. 2000. Registration of 'Boggs' Soybean. *Crop Sci.* 40:294-295.
- Borevitz, J.O., and J. Chory. 2004. Genomics tools for QTL analysis and gene discovery. *Curr. Opin. Plant Bio.* 7:132-136.
- 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.* 46:1010-1011.
- Byfield, G.E., H. Xue, and R.G. Upchurch. 2006. Two genes from soybean encoding soluble delta 9 stearoyl-ACP desaturases. *Crop Sci.* 46:840-846.
- Cahoon, E.B. 2003. Genetic enhancement of soybean oil for industrial uses: Prospects and challenges. *AgBioForum* 6:11-13.
- 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, J. Chung, and J.E. Specht.

- 1999a. An integrated genetic linkage map of the soybean genome. *Crop Sci.* 39:1464-1490.
- Cregan, P.B., J. Mudge, E.W. Fickus, L.F. Marek, D. Danesh, R. Denny, R.C. Shoemaker, B.F. Mathews, T. Jarvik, and N.D. Young. 1999b. Targeted isolation of simple sequence repeat markers through the use of bacterial artificial chromosomes. *Theor. Appl. Genet.* 98:919-928.
- Cregan, P.B. 2000. DNA markers, maps, and technologies. World Soybean Research Conference. VI. p.46.
- Faris, J.D., W.L. Li, D.J. Liu, P.D. Chen, and B.S. Gill. 1999. Candidate gene analysis of quantitative disease resistance in wheat. *Theor. Appl. Genet.* 98:219-225.
- Frary, A., T.C. Nesbitt, S. Grandillo, E. Van der Knaap, B. Cong, J. Liu, J. Meller, R. Elber, K.B. Alpert, S.D. Tanksley. 2000. *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science* 289:85-88.
- Gizlice, Z., T.E. Carter, and J.W. Burton. 1994. Genetic base for North American public soybean cultivars released between 1947 and 1988. *Crop Sci.* 34:1143-1151.
- Harwood, J.L. 1988. Fatty Acid Metabolism. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 39:101-138.
- Hayashi, K. 1992. PCR-SSCP: A method for detection of mutations. *Genet. Anal. Tech. Appl.* 3:73-79.
- Heppard, E.P., A.J. Kinney, K.L. Stecca, and G.H. Miao. 1996. Developmental and growth temperature regulation of two different microsomal ω -6 desaturase genes in soybean. *Plant Physiol.* 110:311-319.
- Hernández, M.L., M. Mancha, and J.M. Martínez-Rivas. 2005. Molecular cloning and characterization of genes encoding two microsomal oleate desaturases (FAD2) from olive. *Phytochemistry* 66:1417-1426.
- Iannone, M.A., J.D. Taylor, J. Chen, M.S. Li, P. Rivers, K.A. Slentz-Kesler, and M.P. Weiner. 2000. Multiplexed single nucleotide polymorphism

- genotyping by oligonucleotide ligation and flow cytometry. *Cytometry*. 39:131-140.
- Iba, K., S. Gibson, T. Nishiuchi, T. Fuse, M. Nishimura, V. Arondel, S. Hugly, and S. Somerville. 1993. A gene encoding a chloroplast ω -3 fatty acid desaturase complements alterations in fatty acid desaturation and chloroplast copy number in the *fad7* mutant of *Arabidopsis thaliana*. *J. Biol. Chem.* 268:24099-24105.
- Ke, J., T.N. Wen, B.J. Nikolau, E.S. Wurtele. 2000. Coordinate regulation of the nuclear and plastidic genes coding for the subunits of the heteromeric acetyl-coenzyme A carboxylase. *Plant Physiol.* 122:1057-1071.
- Keim, P., T.C. Olson, and R.C. Shoemaker. 1988. A rapid protocol for isolating soybean DNA. *Soybean Genet. News.* 15:150-152.
- Kosambi, D.D. 1944. The estimation of map distances from recombination values. *Ann. Eugen.* 12:172-175.
- Lee, S.H., D.R. Walker, P.B. Cregan, and H.R. Boerma. 2004. Comparison of four flow cytometric SNP detection assays and their use in plant improvement. *Theor. Appl. Genet.* 110:167-174.
- Lemieux, B., M. Miquel, C. Somerville, and J. Browse. 1990. Mutants of *Arabidopsis* with alterations in seed lipid fatty acid composition. *Theor. Appl. Genet.* 80:234-240.
- López, Y., H.L. Nadaf, O.D. Smith, J.P. Connell, A.S. Reddy, and A.K. Fritz. 2000. Isolation and characterization of the Δ^{12} -fatty acid desaturase in peanut (*Arachis hypogaea* L.) and search for polymorphisms for the high oleate trait in Spanish market-type lines. *Theor. Appl. Genet.* 101:1131-1138.
- Manly, K.F., R.H. Cudmore Jr, and J.M. Meer. 2001. Map Manager QTX, cross-platform software for genetic mapping. *Mammalian Genome* 12:930-932.

- O'Hara, P., A.R. Slabas, and T. Fawcett. 2002. Fatty acid and lipid biosynthetic genes are expressed at constant molar ratios but different absolute levels during embryogenesis. *Plant Physiol.* 129:310-320.
- Rafalski, J.A. and S.V. Tingey. 1993. Genetic diagnostics in plant breeding: RAPDs, microsatellites, and machines. *Trends Genet.* 9:275-280.
- Ross, A.J., W.R. Fehr, G.A. Welke, and S.R. Cianzio. 2000. Agronomic and seed traits of 1%-linolenate soybean genotypes. *Crop Sci.* 40:383-386.
- Rychlik, W., W.J. Spencer, and R.E. Rhoads. 1990. Optimization of the annealing temperature for DNA amplification *in vitro*. *Nucleic Acids Res.* 18:6409-6412.
- Sanguinetti, C.J., E. Dias Neto, and A.J.G. Simpson. 1994. Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *Biotechniques* 17:914-921.
- SAS® Institute. 2001. SAS/STAT User's guide. Version 8. SAS Inst., Inc. Cary, NC.
- Sekiya, T. 1993. Detection of mutant sequences by single-strand conformation polymorphism analysis. *Mutat. Res.* 288:79-83.
- Shastry, B.S. 2004. Role of SNP / haplotype map in gene discovery and drug development: an overview. *Drug Devel. Res.* 62:143-150.
- Schuppert, G.F. 2004. Molecular mechanisms underlying the high oleic acid phenotype in sunflower. Ph.D. diss. Oregon State Univ.
- Shoemaker, R.C., K. Polzin, J. Labate, J. Specht, E.C. Brummer, T. Olson, N. Young, V. Concibido, J. Wilcox, J.P. Tamulonis, G. Kochert, and H.R. Boerma. 1996. Genome duplication in soybean (*Glycine* subgenus *soja*). *Genetics* 144:329-338.
- Shoemaker, R.C, P. Keim, L. Vodkin, E. Retzel, S.W. Clifton, R. Waterson, D. Smoller, V. Coryell, A. Khanna, J. Erpelding, X. 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.
- Slabaugh, M.B., G.M. Huestis, J. Leonard, J.L. Holloway, C. Rosato, V. Hongtrakul, N. Martini, R. Töpfer, M. Voetz, J. Schell, and S.J. Knapp. 1997. Sequence-based genetic markers for genes and gene families: single-strand conformational polymorphisms for the fatty acid synthesis genes of *Cuphea*. *Theor. Appl. Genet.* 94:400-408.
- Somerville, C., J. Browse, J.G. Jaworski, and J.B. Ohlrogge. 2000. p. 456-527. Lipids. *In* B.B. Buchanan, W. Gruissem, and R.L. Jones (ed.) *Biochemistry and molecular biology of plants*. American Society of Plant Physiologists. Rockville, MD, USA.
- 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.
- Stuber, C.W., M. Polacco, and M.L. Senior. 1999. Synergy of empirical breeding, marker-assisted selection, and genomics to increase crop yield potential. *Crop Sci.* 39:1571-1583.
- Tajuddin, T., S. Watanabe, N. Yamanaka, and K. Harada. 2003. Analysis of quantitative trait loci for protein and lipid contents in soybean seeds using recombinant inbred lines. *Breeding Sci.* 53:133-140.
- Thelen, J.J., and J.B. Ohlrogge. 2002. Both antisense and sense expression of biotin carboxyl carrier protein isoform 2 inactivates the plastid acetyl-coenzyme A carboxylase in *Arabidopsis thaliana*. *Plant J.* 32:419-431.
- Thompson, J.D., D.G. Higgins, and T.J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673-4680.
- Töpfer, R., and N. Martini. 1994. Molecular cloning of cDNAs or genes encoding proteins involved in *de novo* fatty acid biosynthesis in plants. *J. Plant Physiol.* 143:416-425.

- U.S. FDA. 2004. Food labeling and nutrition. U. S. Food and Drug Administration. Available at <http://www.cfsan.fda.gov/~dms/lab-cat.html#transfat> (Verified 28 Oct 2006).
- Van, K., E.Y. Hwang, M.Y. Kim, Y.I. Cho, P.B. Cregan, and S.H. Lee. 2004. Discovery of single nucleotide polymorphisms in soybean using primers designed from ESTs. *Euphytica* 139:147-157.
- Wilcox, J.R. 2004. World distribution and trade of soybean. p. 1-13. *In* H.R. Boerma and J.E. Specht (ed.) Soybeans: Improvement, production, and uses. 3rd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- Wilson, R.F. 1987. Seed Metabolism. p. 643-686. *In* J.R. Wilcox (ed.) Soybeans: Improvement, production and uses. 2nd edition. ASA, CSSA, and SSSA. Madison, WI, USA.
- Xie, J., T.C. Wehner, and M.A. Conkling. 2002. PCR-based single-strand conformation polymorphism (SSCP) analysis to clone nine aquaporin genes in cucumber. *J. Amer. Soc. Hort. Sci.* 127:925-930.
- Zhu, Y.L., Q.J. Song, D.L. Hyten, C.P. Van Tassell, L.K. Matukumalli, D.R. Grimm, S.M. Hyatt, E.W. Fickus, N.D. Young, and P.B. Cregan. 2003. Single-nucleotide polymorphisms in soybean. *Genetics* 163:1123-1134.

Table 4.1. *G. max* sequences from genes in the fatty acid biosynthetic pathway.

GenBank No.	Acronym	Description
AF164511	accA-2	carboxyl transferase alpha subunit
GMU40979	accA-2	alfa-carboxyltransferase precursor
AF165158	accA-3	carboxyl transferase alpha subunit
AF165159	accA-3	carboxyl transferase alpha subunit
L42814	ACCase	acetyl coA carboxylase A
AF164510	accA-1	carboxyl transferase alpha subunit
AF162283	accB-1	acetyl-CoA carboxylase
U40666	accB-1	biotin carboxyl carrier protein precursor
AF271071	accB-2	acetyl-CoA carboxylase
AF271796	accB-2	biotin carboxyl carrier protein subunit
AF163149	accC-2	acetyl-CoA carboxylase
AF007100	accC-2	biotin carboxylase precursor
AF068249	accC-3	biotin carboxylase precursor
AF163150	accC-3	acetyl-CoA carboxylase
U26948	accD	beta-carboxyltransferase subunit
AF243182	KASI	beta-ketoacyl-ACP synthetase I
AF243183	KASI-2	beta-ketoacyl-ACP synthetase I-2
AF244518	KASII	beta-ketoacyl-ACP synthetase 2
AY907523	KASII-A	3-keto-acyl-ACP synthase II-A
AY907522	KASII-B	3-keto-acyl-ACP synthase II-B
AF260565	KASIII	beta-ketoacyl-acyl synthase III
AY885234	SACPDA	stearoyl-acyl carrier protein desaturase A
AY885233	SACPDB	stearoyl-acyl carrier protein desaturase B
AY611472	Fad2	FAD2-1 mRNA
AJ271842	Fad2	FAD2 gene for intron 1
L43921	Fad2	FAD2-2 mRNA
L43920	Fad2	FAD2-1 mRNA
L22964	Fad3	Fad3 mRNA
AY204710	Fad3A	ω -3 fatty acid desaturase mRNA
AY204711	Fad3B	ω -3 fatty acid desaturase mRNA
AY204712	Fad3C	ω -3 fatty acid desaturase mRNA

Table 4.2. *G. max* UniGene sets from candidates in the fatty acid biosynthetic pathway.

UniGene	Acronym	Markers tested
		-- no. --
U51	Desaturase	5
U1839	Fad2_1	3
U5046	Fad2_2	9
U6041	KAS	7
U6256	SACPD	9
U8460	Acyl-CoA	2
U8463	FA desaturase	2
U8476	Fas2	6
U16819	accB-2	2
U18367	Fad2_1	9

Table 4.3. Oligonucleotides developed from sequence-tagged-sites in linkage groups from the soybean genome associated with oleic acid QTL.

LG	SSR marker	GenBank No.
A1	Satt211	BH126417
	Satt200	BH126407
	Satt225	BH126426
D2	Satt226	ISQ62010
		UM143E12
		ISO65P14
G	Satt570	AZ254187
	Satt303	BH610222
	Satt324	AQ989195
		AQ989298
	Satt235	UM140J16
		UM091M12
L	Satt418	B45M04
	Satt143	AZ254187
		AZ254185

Table 4.4. SF-ANOVA marker associations with oleic acid content from G99-G725 × N00-3350.

Marker	Source	LG	R ² [¶]	2a [‡] g kg ⁻¹
QTL_A1_1	BAC	A1	2*	23.2
A1-114639	BAC	A1	5***	44.9
SNP16289	Fad2	D2	10***	62.8
D2-32349	BAC	D2	7***	56.2
G-10857	BAC	G	ns [§]	
U6041_6	KasIII	G	ns	
QTL_G1	BAC	G	3**	41.7
accA_6	accA	G	ns	
G1	accase	G	7***	50.7
accB_2	accB	G	5**	36.9
17027	BAC	L	5***	50.3
13835	BAC	L	7***	58.7
L-44913	BAC	L	4**	49.3
21575	BAC	L	3**	48.82
L-35235	BAC	L	2*	29.3

[¶]R² = % of the total trait variance explained by the genotype at a marker locus. [‡]2a = the difference in oleic acid content at a SSR marker homozygous for the N00-3350 allele - homozygous for the G99-G725 allele. * p < 0.05, ** p < 0.01, *** p < 0.001.

[§]ns = indicates a non-significant marker association.

Table 4.5. SSCP and SNP markers from fatty acid gene-based sequences and the linkage group in which they were mapped.

Marker	Gene ID	LG
SNP12447	FatB2 [¶]	C1
SNP12453	FatB2	C1
SNP12507	Fad3	B1
SNP15809	Fad2 [‡]	D1b
SNP15231	Fad2	D1b
U6041_3	KasIII	K
SNP15783	Fad2	O
SNP15155	Fad3	O

[¶]Primers developed from FatB2 sequence from GenBank No. AW201449 and AW201449. [‡] Primers developed from Fad2 sequences GenBank No. L43920 and L43921 (Heppard et al., 1996).

Figure 4.1. Outline of the fatty acid biosynthetic pathway. ACCase: acetyl-CoA carboxylase; KasIII: condensation enzyme, together with KasI and KasII are part of the fatty acid synthase (FAS) complex; ACP: acyl carrier protein; Fab2: stearoyl-ACP desaturase; FatA: fatty acid thioesterase A; Fad2: fatty acid desaturase 2, Fad3: fatty acid desaturase 3; ER: endoplasmic reticulum (Adapted from Somerville, 2000; Aghoram et al., 2006).

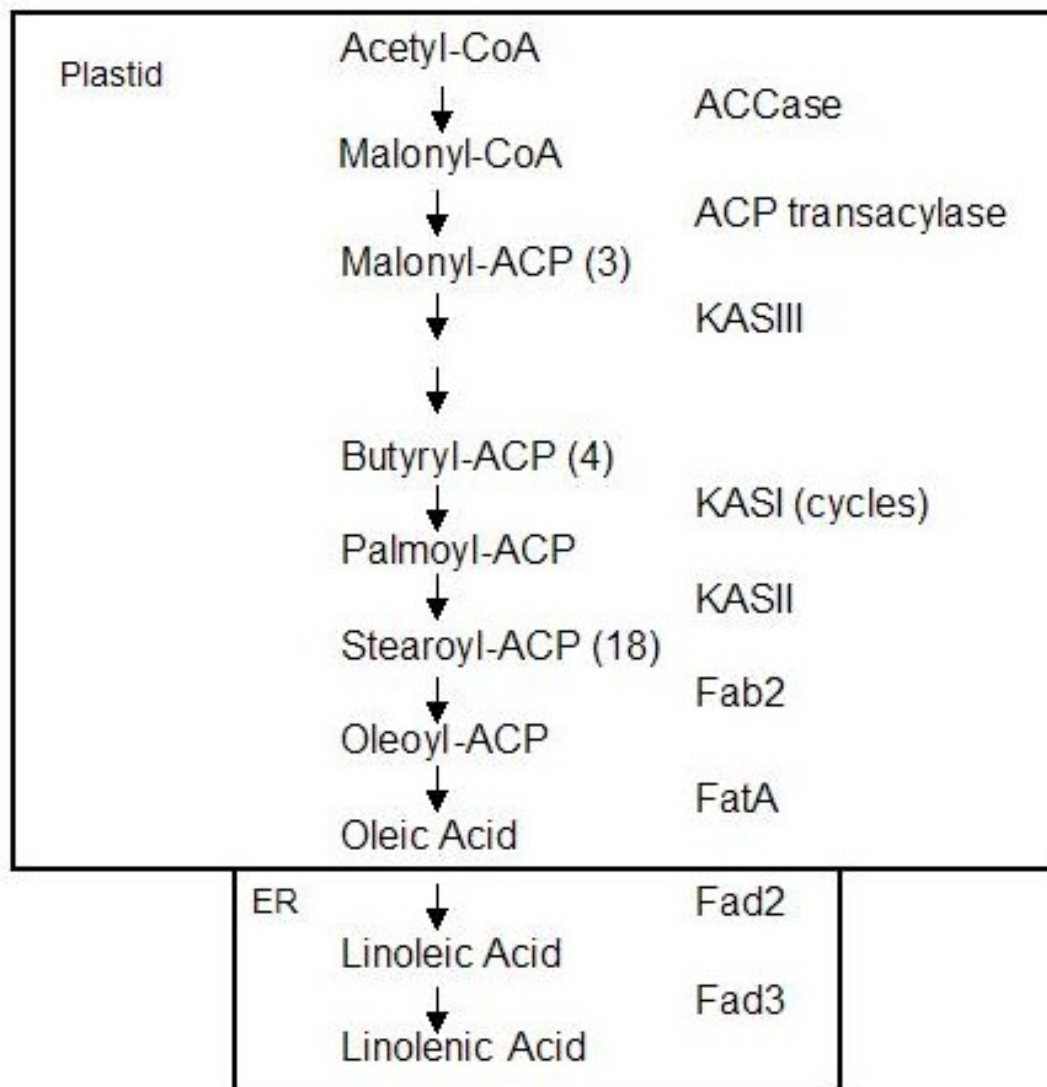


Figure 4.2. Genetic linkage map of the G99-G725 $\times$ N00-3350 $F_{2:3}$ population in linkage groups with oleic acid QTLs. The arrow indicates the most likely QTL position.

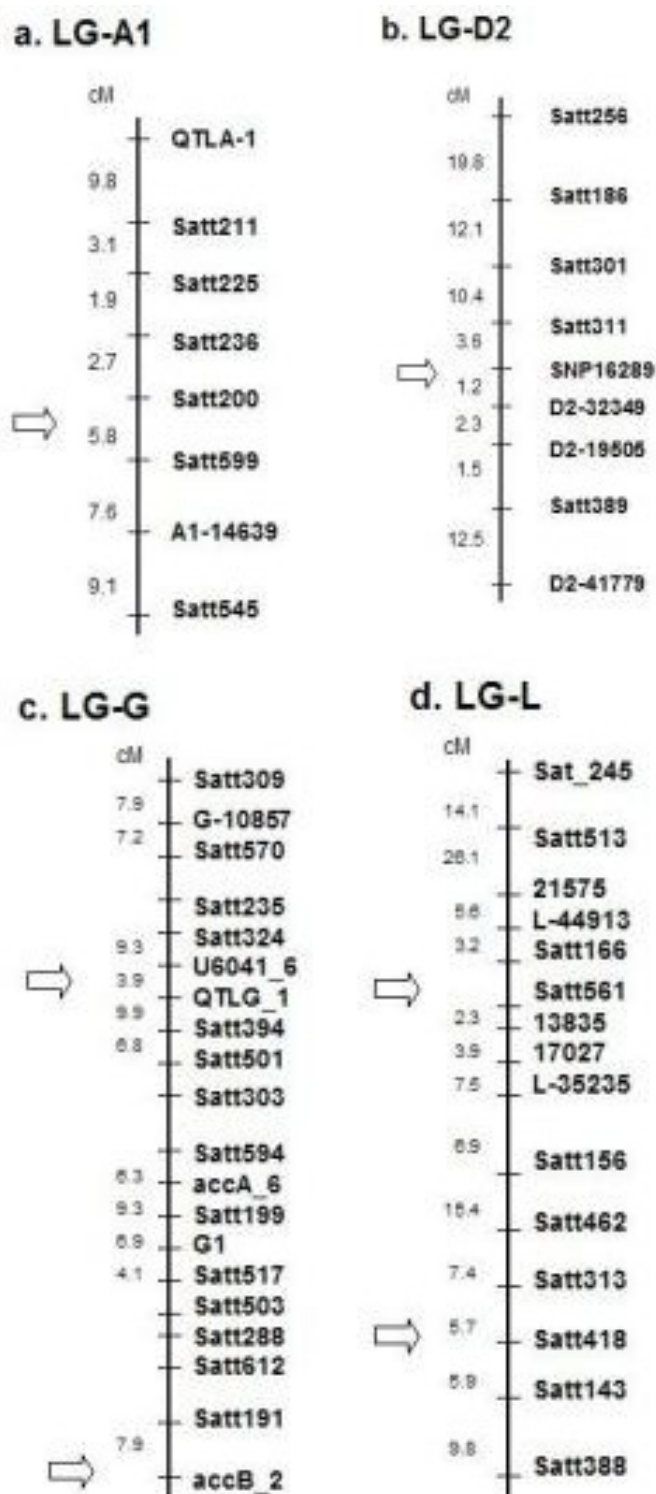
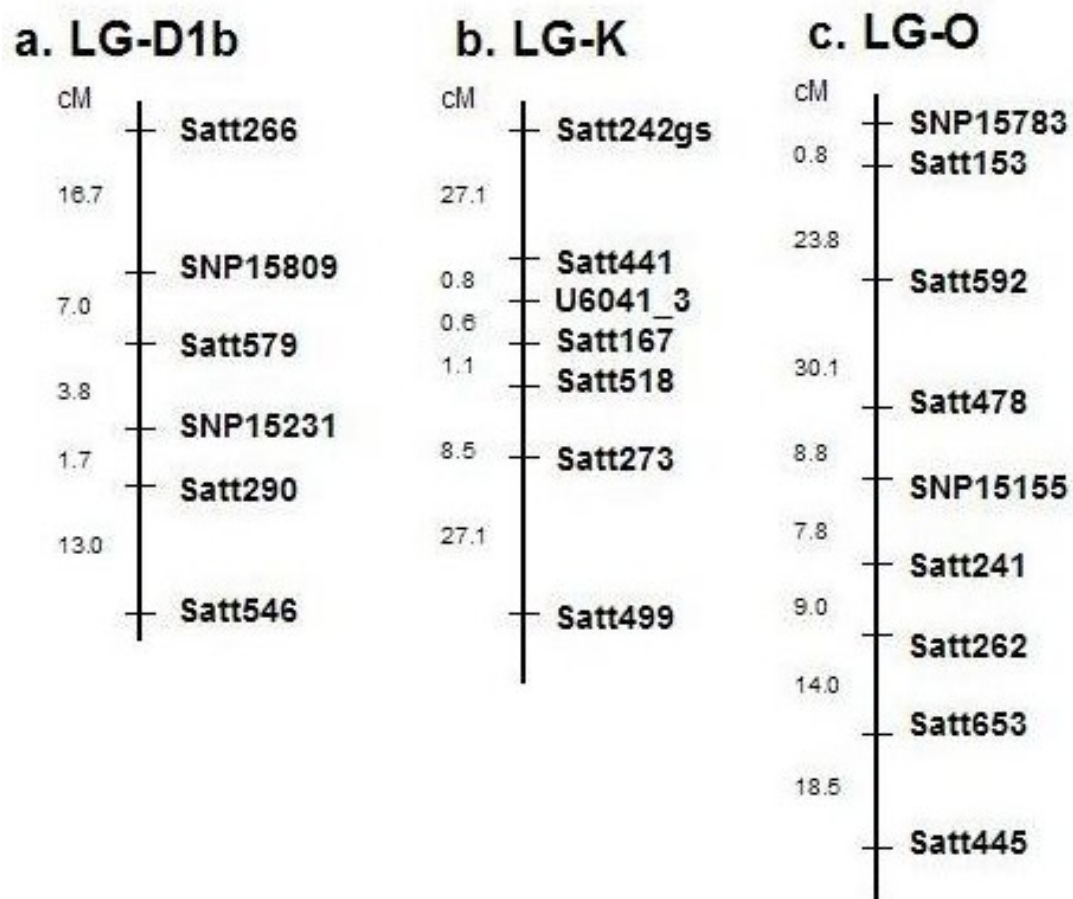


Figure 4.3. Genetic linkage map for LG-D1b, LG-K, and LG-O of the G99-G725 × N00-3350 F_{2:3} population.



CHAPTER 5
MAPPING AND CONFIRMATION OF THE ‘HYUUGA’ RED-BROWN
LESION RESISTANCE GENE FOR ASIAN SOYBEAN RUST¹

¹Maria J. Monteros, Ali M. Missaoui, Daniel V. Phillips, David R. Walker, and H. Roger Boerma. Submitted to Crop Science.

ABSTRACT

Asian soybean rust (ASR), caused by *Phakopsora pachyrhizi*, is a widespread disease of soybean [*Glycine max* (L.) Merr.] with the potential to cause serious economic losses. The objective of this study was to genetically map red-brown lesion type resistance from the cultivar Hyuuga. A population of 117 RILs from the cross of Dillon (tan lesion) x Hyuuga (red-brown lesion, RB) was rated for ASR lesion type in the field and inoculated with *P. pachyrhizi* in the greenhouse. The RB resistance gene mapped between Satt460 and Satt307 on linkage group (LG) C2. When field severity and lesion density in the greenhouse were mapped, the *Rpp?*(Hyuuga) locus explained 22% and 15% of the variation, respectively ($p < 0.0001$). The RB lesion type was associated with lower severity, fewer lesions, and reduced sporulation when compared to the tan lesion type. A population from the cross of Benning x Hyuuga was screened with SSR markers in the region on LG-C2 flanked by Satt134 and Satt460. Genotype at these markers was used to predict lesion type when the plants were exposed to *P. pachyrhizi*. All the lines selected for the Hyuuga markers in this interval had the RB lesion type and they averaged approximately 50% fewer lesions compared to lines with tan lesions. Sporulation was only detected in 6% of the RB lines compared with 100% of the tan lines. Marker-assisted selection can be used to develop soybean cultivars with the *Rpp?*(Hyuuga) gene for resistance to ASR.

INTRODUCTION

Plant rusts, caused by Basidiomycetes of the order Uredinales, are some of the most destructive plant diseases (Agrios, 1997). *Phakopsora pachyrhizi* and *P. meibomia* are the causal agents of rust in soybean, *Glycine max* (L.) Merr (Sinclair and Hartman, 1999). *P. meibomia* causes the 'American' form of soybean rust. This species is native to South America, and has been found on wild and cultivated legumes from Puerto Rico to southern Brazil (Vakili, 1979). *P. pachyrhizi* is the causal agent of 'Asian' soybean rust, and is native to the traditional growing areas in the Orient. Asian soybean rust (ASR) is considered among the top 25 of the 100 most destructive exotic pests in the world (Ogle et al., 1979).

P. pachyrhizi can infect and spread from many non-soybean wild and cultivated hosts, including many garden legumes (Vakili and Bromfield, 1976). The host range for soybean rust includes more than 90 legume species, including cowpea (*Vigna unguiculata*), and kudzu (*Pueraria lobata*) (Rytter et al., 1984; Sinclair and Hartman, 1999). Soybean is susceptible to *P. pachyrhizi* at any stage of its development, but symptoms are more likely to appear after a plant enters the reproductive stages (Melching et al., 1989). Yield losses from 13 to 80% caused by soybean rust in commercial soybean fields have been reported (Ogle et al., 1979; Yang et al., 1990; Yang et al., 1991; Sinclair and Hartman, 1996). Reduction in yield results from the production of fewer pods, fewer seed in each pod, and reduced seed weight (Melching et al., 1989; Sinclair and Hartman, 1999).

Water-soaked lesions are the first symptoms of a rust infection. These increase in size and become chlorotic as the disease progresses (Sinclair and Hartman, 1999). The color of the lesions may be grayish brown, tan to dark brown, or reddish brown depending on the virulence of the pathogen, the host genotype, the interaction of pathogen and host genotypes, and the age of the lesion. The three types of infection described by Bromfield and Hartwig (1980) and Bromfield (1984) on soybean inoculated with *P. pachyrhizi* are tan lesions with many uredinia and abundant sporulation, RB lesions with few

uredinia and abundant sporulation, and an immune reaction with a lack of visible infection. Tan lesions indicate a compatible interaction and a susceptible reaction (Sinclair and Hartman, 1999; Miles et al., 2003). The RB reaction has been associated with resistance conditioned by three of the four known single resistance genes (Hartman et al, 2004).

PI200492 ('Komata') has a single dominant allele (*Rpp*) that confers resistance to an Australian Asian rust isolate (McLean and Byth, 1980). When PI230970, which exhibited resistance to several isolates of *P. pachyrhizi*, was crossed with a susceptible cultivar, segregation ratios suggested a single dominant gene for resistance (Bromfield and Hartwig, 1980). PI462312 ('Ankur') has a single dominant allele for resistance (Singh and Thapliyal, 1977). Intercrossing the three sources of resistance and inoculating them with two rust isolates revealed that the dominant alleles for resistance in PI200492, PI230970, and PI462312 were at different loci (Hartwig and Bromfield, 1983). Hartwig and Bromfield (1983) suggested that the *Rpp* symbol for the resistance gene from PI200492 be changed to *Rpp1*, and assigned *Rpp2* and *Rpp3* to the resistance genes in PI230970 and PI462312, respectively. Later studies showed that the cultivar Bing Nan (PI459025) from China had a single dominant resistance allele (*Rpp4*) at a locus different from the other three resistance alleles (Hartwig, 1986). At least nine races of *P. pachyrhizi* have been described (Burdon and Speer, 1984; Sinclair and Hartman, 1999). New races of the pathogen have been able to overcome some of these resistance genes, particularly *Rpp1* and *Rpp3* (Godoy, 2005).

Over 95% of the soybean cultivars for which ASR resistance has been assessed have been found to be highly susceptible (Burdon and Marshall, 1981). Currently available soybean cultivars are susceptible to at least some races of ASR (Burdon and Speer, 1984; Sinclair and Hartman, 1999). Although plant resistance that is race-specific to rust has been identified (Bromfield and Hartwig, 1980; Hartwig and Bromfield, 1983; Hartwig, 1986), no cultivars have been developed that have an acceptable level of resistance

to all strains of *P. pachyrhizi*. A lack of resistance to the virulent races of *P. pachyrhizi* demonstrates the vulnerability of the soybean crop to this pathogen. Although some tolerance to *P. pachyrhizi* has been identified, no U.S. soybean cultivars have been reported to possess tolerance to soybean rust (Hartman et al, 1991).

Prior to 2004, ASR was not present in the continental USA. Therefore, assessment of the effects of soybean rust on U.S. soybean cultivars was performed only in Bio Safety Level 3 containment facilities or in countries where ASR was already established. In November 2004, the disease was first reported in the USA in plots at a Louisiana State Univ. research station near Baton Rouge (Schneider et al., 2005). After its initial detection, it was found on soybean in Alabama, Arkansas, Florida, Georgia, Mississippi, Missouri, South Carolina, and North Carolina. The objectives of the current research were to map gene(s) conditioning the RB-lesion type resistance to ASR from the Japanese cultivar Hyuuga using endemic southeastern U.S. isolates of *P. pachyrhizi* and confirm the genomic location of this gene in an independent population.

MATERIALS AND METHODS

ASR rating in Attapulcus, GA

Hyuuga is a Japanese cultivar that was found to produce RB lesions when it was tested along with U.S. public soybean cultivars in Maturity Groups (MG) VI to VIII and parents of Univ. of Georgia mapping populations in greenhouse screenings in Londrina, Brazil, early in 2005. Hyuuga is a Maturity Group VII cultivar that has partial resistance to bacterial pustule, a foliar disease caused by *Xanthomonas campestris* pv. *glycines* (GRIN, 2005). It had previously been crossed with 'Dillon' at the Univ. of Georgia to create a QTL mapping population. Dillon is a Maturity Group VI cultivar derived from an F₄ plant selection from the cross 'Centennial' × 'Young' (Shipe et al., 1997). It is resistant to bacterial pustule and susceptible to ASR (GRIN, 2005).

A population of 117 F_{5:6} recombinant inbred lines (RILs) from the Dillon × Hyuuga cross was planted 3 Sept 2005, at the Univ. of Georgia's Attapulgus Research and Education Center, located in the southwest corner of Georgia, near the Georgia-Florida border. Each RIL was planted in a 2.44 m row at a seeding rate of 33 seeds m⁻¹, and a row spacing of 0.91 m. Each entry was planted in two replications in a randomized complete block design. Three entries each of Dillon and Hyuuga were included as resistant and susceptible checks in each replication. After planting, mobile lighting towers were used to extend the photoperiod to 24 h, until 1 Oct 2005. The ASR screening nursery was planted late in the summer to allow flowering to be induced simultaneously across a wide range of maturity groups being evaluated in plots adjacent to this study, and to maximize the opportunity for a *P. pachyrhizi* epidemic to develop. To reduce the probability of bacterial diseases becoming established, primarily bacterial pustule and bacterial blight (*Pseudomonas savastanoi* pv. *glycinea*), all plots were sprayed with "Bac-Master" agricultural streptomycin at a rate of 8 oz 100 gal⁻¹ (100 ppm) on a weekly basis starting 21 Sept. 2005.

Beginning 6 Oct 2005, suspensions of *P. pachyrhizi* spores, concentration unknown, were added to the spray tank with the streptomycin solution to promote uniform infection. The spores were obtained from ASR-infected soybean plants collected at the research station. Plants were inoculated with *P. pachyrhizi* spores on 6 Oct, 17 Oct, 24 Oct, 31 Oct, 9 Nov, and 18 Nov 2005. On 17 Nov 2005, 10 leaflets from the mid to lower canopy of individual plants within each plot were collected, placed in plastic bags, and transported to the laboratory on ice. Most plants were at the R5 to R6 stage of development at the time (Fehr and Caviness, 1977). Leaflets were observed under 10X magnification and scored based on the type of lesion: tan, RB, or mixed. A score of 1 was assigned if only tan lesions were present, 3 if all lesions were RB, and a 2 was given if both types of lesions were observed.

In addition to the lesion type data, ASR severity ratings were also recorded on a plot-basis on 1 Dec 2005 using the scale 1 = dark green canopy, 2 = some yellowing on the bottom leaves, 3 = significant yellowing on the bottom leaves, 4 = significant yellowing at the bottom of the canopy and some yellowing in the middle of the canopy, 5 = significant yellowing at the bottom and middle of the canopy, and some yellowing in the top of the canopy. Two separate ratings of each plot were made for each replication.

Greenhouse studies

Three replications of the same 117 RILs from the Dillon $\times$ Hyuuga population tested in the field at Attapulcus were planted in the greenhouse in a randomized complete block design, which included three entries of each parent. The experimental unit was a 10.2-cm pot. Six seeds of each genotype were planted in each pot, and the pots were arranged in a multiple-pot tray. Each tray contained 14 pots of the RILs and one pot of the ASR susceptible cultivar Cobb. The first replication of this study was planted in the greenhouse at the Univ. of Georgia's Griffin campus on 3 Feb 2006. The second replication was planted 17 Feb, and the third replication was planted on 6 March 2006 at the same location. A second population of 92 F_{6:7} lines from the cross of Benning $\times$ Hyuuga was planted 1 Feb 2006 in a Univ. of Georgia greenhouse in Athens, GA, and was used to confirm the location of the rust lesion type locus that had been mapped using the Dillon $\times$ Hyuuga data. Benning is a Maturity Group VII cultivar derived from an F₄ plant selection from the cross 'Hutcheson' $\times$ 'Coker 6738'. It is resistant to bacterial pustule (Boerma et al., 1997).

Trifoliolate leaves from the Benning $\times$ Hyuuga population were collected, freeze-dried, and used for DNA extraction and genotyping with SSR markers. DNA was extracted and lines from the Benning $\times$ Hyuuga were screened with SSR markers. A total of 16 lines homozygous for the Benning allele and 16 lines homozygous for the Hyuuga allele at four SSR markers (Satt134, Satt489, Satt100, and Satt460) associated with lesion type were

selected. Two replications of the 32 selected lines and two entries per replication of each of the parents were planted in a randomized complete block design. Six seeds of each entry were planted in a greenhouse in Griffin, GA, on 6 March 2006 in 10.2 cm pots that fit into a 15-pot tray.

Inoculation with *P. pachyrhizi* in the greenhouse

The spores used for inoculation in the greenhouse were collected from field-grown soybean plants and surrounding kudzu plants near Athens, Attapulcus, Griffin, and Eatonton, GA, during the summer 2005. Additionally, leaves with *P. pachyrhizi* spores from susceptible plants grown in the Griffin greenhouse were collected and stored in plastic bags overnight. The inoculum was prepared from a combination of field and greenhouse-collected spores using 1.2 L of sterile water and 0.04% Tween-20. A funnel with cheesecloth was used to filter the suspension of water, surfactant, and spores. The concentration of spores used for inoculation was approximately 7.5×10^4 spores ml⁻¹. Plants were inoculated with the spore suspension using an atomizer. The first replication of RILs from Dillon × Hyuuga was inoculated on 1 March 2006. The second replication was inoculated 6 March 2006, and the third replication was inoculated 24 March 2006. The selected lines from Benning × Hyuuga were inoculated on 24 March 2006. After inoculation, the plants were placed in a humidity chamber in the dark for 24 h.

The first and second replications of lines from the cross of Dillon × Hyuuga were rated for ASR on 23 March 2006. The third replication of Dillon × Hyuuga, and the two replications of Benning × Hyuuga experiment were rated on 10 April 2006. For each pot, the leaflet with the most ASR lesions from each of two of the most severe ASR-infected plants from the pot was harvested, and inspected under 10X magnification to determine lesion type. The number of lesions in a 6.45-cm² leaf area delimited by a small plastic frame was determined. The presence or absence of *P. pachyrhizi* spores was also recorded.

DNA fingerprinting

Leaf tissue for DNA extraction was collected from both the 117 RILs of Dillon × Hyuuga and the 92 RILs of Benning × Hyuuga plants grown in the greenhouse, lyophilized, and macerated. DNA was extracted using a modified CTAB (hexadecyltrimethylammonium acid) protocol (Keim et al., 1988), and re-suspended in Tris-EDTA (TE) buffer pH 8.0. PCR reactions were similar to the protocol described by Li et al., (2001), with some modifications. For the 117 RILs, a genome-wide scan was conducted using 138 polymorphic SSR primers from the 20 linkage groups of soybean. The 10-μL reaction mix contained 2 μL of 50 ng/μL DNA template, 1.0X PCR buffer, 2.5mM MgCl₂, 100 μM of each dNTP, 0.2 μM each of the forward and reverse primers, and 0.5 unit of *Taq* DNA polymerase (Promega, Madison, WI). Primers were labeled with the fluorescent dyes 6-FAM, NED, or HEX (PE-ABI, Foster City, CA). A 384-well GENE AMP PCR System 9700 (Applied Biosystems, Foster City, CA) was used for DNA amplification. Pooled PCR products (3-4 μL) were combined with 2 μL deionized formamide, 0.75 μL loading buffer, and 0.2 μL Genescan ROX-500 internal size standard (ABI, Foster City, CA). The mixture was denatured at 95°C for 4 min., and 1-2 μL were loaded in each of 96 lanes on 12-cm well-to-read distance acrylamide:bisacrylamide (19:1) gels, using KLOEHN micro syringes (Kloehn Ltd., Las Vegas, NV). DNA amplicons were run on 12-cm gels using an ABI Prism 377 DNA sequencer at 750 V for 1.5 to 2 hours, with 1X TBE buffer. Genescan sequencer software (ABI, Foster City) was used to collect the marker data.

Data analyses

Map Manager QTXb20 (Manly et al., 2001) was used with the Kosambi (1944) mapping function and the recombinant inbred option to create the genetic map, and determine the linkage of SSR markers and lesion type when evaluated as a qualitative trait. Marker segregation distortion identified seven lines from Dillon × Hyuuga that were possibly a mix, since they had a

higher percent of both parental bands than what was expected for F₆-derived lines. SAS® (SAS Institute, Cary, NC) PROC GLM was used to detect statistically significant differences between the average lesion number for each lesion type. Map Manager QTXb20 was also used for the marker regression analysis and the interval mapping of severity and lesion number. Permutation tests were performed to determine the significance threshold for QTL analysis.

RESULTS AND DISCUSSION

In the summer of 2005, the 117 Dillon (tan lesion type) $\times$ Hyuuga (RB lesion type) RILs mapping population was phenotyped for lesion type. The type of lesion for each RIL observed in the field at Attapulgus, GA was almost always identical in the two replications, except when mixed lesion types were found. In those cases, one replication showed a mixture and the other replication was rated as having either the tan or RB lesion type. The lines planted were F_6 -derived and therefore the percent of heterogeneous lines derived from a heterozygous F_6 plant is expected to be around 3.12%. RILs from the Dillon $\times$ Hyuuga cross that had an eight percent or higher average level of heterogeneity across the 138 SSR markers were excluded from the analysis because they may have resulted from a seed mixture. The lines with a lesion type that was inconsistent among replications were not included in the final mapping. The segregation ratio for lesion type among the RILs used in the final analysis was 59 tan, 8 mixed, and 35 RB (Table 5.1). These RILs had been previously selected for having maturity similar to that of Dillon, which could explain the higher number of individuals with tan lesions. When lesion type was mapped as a qualitative trait using phenotypic data from the field, a locus associated with lesion type mapped close to Satt460 on LG-C2. Previous mapping in this population had identified a major pod maturity QTL near Satt460 (data not shown).

When the Dillon $\times$ Hyuuga RILs were rated in the greenhouse, the segregation ratio for lesion type was 61 tan, 3 mixed, and 37 RB. Using the lesion type data from the greenhouse, the lesion type locus also mapped to LG-C2 in a 3.2-cM interval between Satt307 and Satt460 (Fig 5.1). The linkage map for the SSR markers on LG-C2 was in close agreement to that previously reported by Song et al. (2004) for this region of the linkage group. On the consensus map, these markers are 3.5 cM apart. Although Satt100 and Satt134 are inverted in order on the Dillon $\times$ Hyuuga map compared to the consensus map, they are estimated to be only 0.5 cM apart on the latter.

Lesion type data from the greenhouse were obtained by inspecting all the plants within a pot. The plants were re-evaluated whenever inconsistencies between the three replications were found. Inconsistencies occurred in six of the lines, but

these were all reconciled upon re-evaluation. Nine lines had lesion types that were inconsistent between the field and the greenhouse. Five lines had a mixture of lesion types in the field, but only RB lesions in the greenhouse. Two lines had mixed lesions in the field, but only tan lesions in the greenhouse. Two lines had RB lesions in the field and mixed lesions in the greenhouse. A re-examination of leaflets from the field was not possible, but we believe that some entries in the field may have been misidentified as having a particular lesion type, and that the greenhouse data for these were more reliable. Unlike the observations of two different lesion types expressed on the same plant reported by Bonde et al. (2006), we only observed different lesion types on different plants in the same pot, not on the same plant.

Brogin et al. (2004), using 113 F_{2:4} lines from a cross between FT-2 (RB lesion type) and the susceptible cultivar, Davis, found that a locus associated with lesion type mapped to LG-C2. Their lesion type gene mapped to the same region of LG-C2, although the interval distances in their map were longer, possibly resulting from incorrect marker or lesion type scores in some of the lines. At this point we are unable to assign a definite gene symbol to the lesion-type resistance allele that has been mapped to LG-C2. In accordance with the guidelines for the assignment of gene symbols from the Soybean Genetics Committee (Soybean Genetics Newsletter, 1997), we propose the temporary designation *Rpp?*(Hyuuga). Before a definite gene symbol can be assigned, it will be necessary to cross Hyuuga with sources of known *Rpp* genes and inoculate progenies with different ASR isolates to determine whether resistance genes from the parents are inherited independently. The *Rpp1* resistance gene has been mapped to a different linkage group in the soybean genome (P.B. Cregan, personal communication).

In our greenhouse evaluations, the three entries of Dillon averaged 4.11 lesions cm⁻², and the three entries of Hyuuga had an average of 2.25 lesions cm⁻². The 61 tan lesion RILs averaged 3.36 lesions cm⁻², and the 37 RILs with an RB lesion type averaged 2.50 lesions cm⁻² (Table 5.1). Additionally, lines from the Dillon × Hyuuga cross with an RB lesion type were less likely to show sporulation than lines with a tan lesion type. In the 61 lines with tan lesions, sporulation was detected on all but one line (i.e., 98.4% of the lines with tan lesions were sporulating). In contrast, among the 37 lines with RB lesions, sporulation was observed on only three of the lines (91.9% of

lines with an RB-lesion type showed no sporulation). In a previous greenhouse study, Hartman (1995) showed that plants with RB reactions, tended to have longer latent periods, a smaller number of pustules over time, and smaller lesions compared with plants that had a tan reaction type.

In the field evaluations, the three entries of Dillon had a mean severity rating of 2.9, and the mean severity rating for Hyuuga was 2.3 (Table 5.1). The average rating for RILs with tan lesions was 3.1, and for RILs with RB lesions the average rating was 2.4. When RB and tan lesion types were compared, both the severity ratings in the field and the number of lesions in the greenhouse were significantly different ($p = 0.05$). The ASR severity ratings and the number of lesions showed a normal distribution (data not shown). When severity and lesion number were mapped as quantitative traits, the *Rpp?*(Hyuuga) locus explained 22% ($p < 0.0001$) of the variation in the severity ratings at Attapulugus and 15% ($p < 0.0001$) of the variation in lesion number observed in the greenhouse (Fig 5.2, Table 5.2). In both cases the LOD score peaks for severity and number of lesions are at the *Rpp?*(Hyuuga) locus. The significance threshold determined by a permutation test was $\text{LOD} = 1.8$. The difference in severity ratings in the field for RILs homozygous for the Hyuuga allele at the *Rpp?*(Hyuuga) locus vs. those homozygous for the Dillon allele (2a) was 0.86. The difference in lesion number for RILs homozygous for Hyuuga vs. homozygous for Dillon is $0.89 \text{ lesions cm}^{-2}$ (Table 5.2). These results indicate that lesion type has an effect on the severity of the ASR infection in the field and the number of lesions observed in the greenhouse. In both environments, the RB lesion type reduced the disease development. The correlation coefficient between lesion number in the greenhouse and canopy severity in the field was 0.32 ($p < 0.001$), which suggests that the RB lesion lines which overall developed fewer lesions in the greenhouse were also among the lines with a lower severity rating in the field (Table 5.2). This provides some evidence for the practical value of the RB lesion type resistance.

Using an independent RIL population with Hyuuga as a common parent, 32 $F_{5:6}$ lines from a Benning $\times$ Hyuuga cross were selected based on their marker genotype in a 4-cM region between Satt460 and Satt134 on LG-C2. Sixteen lines homozygous for Benning alleles and 16 lines homozygous for Hyuuga alleles were screened for their reaction to ASR in the greenhouse. We expected that all 16 lines homozygous for

Benning alleles would have a tan lesion type (16:0:0 tan, mixed, and RB, respectively), and the lines homozygous for Hyuuga alleles would have an RB lesion type (0:0:16). In all lines except for one, which had mostly tan lesion types and only one plant with the RB lesion type, those homozygous for the Benning alleles had tan lesions, and those homozygous for the Hyuuga alleles had the RB lesion type. Thus, with the exception of the line that had a mixture of lesion types within the replications, the marker data correctly predicted the type of lesion. Therefore, markers Satt460 and Satt307 can be used to select for the RB lesion phenotype inherited from Hyuuga.

Similar to the results obtained from the population derived from Dillon as a susceptible parent, lines selected from the cross of Benning $\times$ Hyuuga that had an RB lesion type had fewer lesions on average than the lines that had a tan lesion type. The two entries of Benning had an average of 6.90 lesions cm^{-2} , and the two entries of Hyuuga had an average of 2.48 lesions cm^{-2} . The RILs with the RB lesion type averaged 3.35 lesions cm^{-2} compared to 6.09 lesions cm^{-2} for the RILs with tan lesions (Table 5.1). Additionally, on RILs with a tan lesion type, sporulation was always detected, whereas when RILs had RB lesions, sporulation was visible on only 6% of the lines. Our results indicate that RB lesion type is associated with a lower average number of lesions, and limited sporulation. Bonde et al. (2006) observed similar results where RB reaction types had a lower amount of sporulation compared to those that had a tan reaction type. Evaluation of inheritance of the RB lesion type in crosses with two different susceptible parents suggests that this type of resistance is likely to be effective in different genetic backgrounds.

The ASR resistance locus associated with the RB lesion type maps to a region of LG-C2 that has been associated with resistance to fungal diseases, nematodes and insects. In a cross between the soybean cultivars 'Forrest' $\times$ 'Essex', marker Satt371 on LG-C2 explained 12% of the variation in susceptibility to sudden death syndrome (SDS) caused by the soil fungus *Fusarium solani* f. sp. *glycines* (Iqbal et al., 2001; Rupe and Hartman, 1999). In a cross between 'Douglas' $\times$ 'Pyramid', Satt307 identified a QTL for resistance to SDS (Njiti et al., 2002). A soybean cyst nematode (*Heterodera glycines* Ichinohe) resistance-QTL has been identified near Satt100 (Wang et al., 2001), and a QTL for insect resistance has also been reported on LG-C2 (Rector et al., 1999).

The resistance associated with the RB type of lesion reported here has been effective with the *P. pachyrhizi* isolates currently present in both the U.S. and Brazil. Soybean breeders can use the SSR markers on LG-C2 to select lines with the Hyuuga RB-lesion type and its associated decreased number of lesions, lower ASR sporulation, and reduced ASR severity. Marker-assisted selection (MAS) can thus be used by researchers to select for the RB-lesion type even when ASR is not present in their area, and without the need to use bio-containment facilities for phenotypic assays. In addition, markers would allow *Rpp?*(Hyuuga) to be pyramided with other genes conditioning resistance to the same races of ASR. Incorporating the Hyuuga type of resistance could potentially be used to reduce the vulnerability of U.S. soybean to ASR, and to reduce the dependency on fungicides for the management of ASR. The use of MAS could enable soybean breeders to select for the RB lesion-type resistance in early generations of soybean crosses, and to speed the initial development of ASR resistant soybean cultivars adapted to the various production regions of the USA.

ACKNOWLEDGEMENTS

This research was supported by grants from the American Seed Trade Association, the United Soybean Board, the Tinker Foundation, and funds allocated to the Georgia Experiment Stations. We would like to thank Dr. Leones Alves de Almeida and Dr. Jose Tadashi Yorinori from Embrapa-Soja in Londrina, Brazil for their technical assistance.

REFERENCES

- Agrios, G.N. 1997. Plant pathology. 4th ed. Academic Press, London, UK.
- Boerma, H.R., R.S. Hussey, D.V. Phillips, E.D. Wood, G.B. Rowan, and S.L. Finnerty. 1997. Registration of 'Benning' Soybean. Crop Sci. 37:1982.
- Bonde, M.R., Nester, S.E., C.N. Austin, C.L. Stone, and R.D. Frederick. 2006. Evaluation of virulence of *Phakopsora pachyrhizi* and *P. meibomia* isolates. Plant Dis. 90:708-716.
- Brogini, R.L., C.A.A. Arias, N.A. Vello, J.F.F. de Toledo, A.E. Pipolo, L.L. Catelli, and S.R.R. Marin. 2004. Molecular mapping of a gene conferring resistance to soybean rust. Poster presented at the VII World Soybean Research Conference, Foz do Iguassu, PR, Brazil.
- Bromfield, K.R. 1984. Soybean Rust Monogr. No. 11. American Phytopathological Society, St. Paul, MN.
- Bromfield, K.R., and E.E. Hartwig. 1980. Resistance to soybean rust and mode of inheritance. Crop Sci. 20:254-255.
- Burdon, J.J., and D.R. Marshall. 1981. Evaluation of Australian native species of *Glycine* for resistance to soybean rust. Plant Dis. 65:44-45.
- Burdon, J.J., and S.S. Speer. 1984. A set of differential hosts for the identification of pathotypes of *Phakopsora pachyrhizi* Syd. Euphytica 33:891-896.
- Fehr, W.R., and C.E. Caviness. 1977. Stages of soybean development. Special Report 80. Cooperative Extension Service. Iowa State University. Ames, Iowa. 11 pp.
- Godoy, C.V. 2005. Asian Soybean Rust. Anti-Rust Consortium. Embrapa-Soja. Londrina, PR, Brazil. 1 Jan – 4 Feb, 2005.
- Hartman, G.L., T.C. Wang, and A.T. Tschanz. 1991. Soybean rust development and the quantitative relationship between rust severity and soybean yield. Plant Dis. 75:596-600.
- Hartman, G.L. 1995. Highlights of soybean rust research at the Asian Vegetable Research and Development Center. p. 19-28. In J.B. Sinclair and G.L. Hartman (ed.) Soybean Rust Workshop, August 9-11, 1995. College of Agriculture, Consumer, and Environmental Sciences, National Soybean Research Laboratory Publication Number 1, Urbana, Illinois.

- Hartman, G.L., M.R. Bonde, M.R. Miles, and R.D. Frederick. 2004. Variation of *Phakopsora pachyrhizi* isolates in soybean. p. 440-446. In F. Moscardi, C.B. Hoffman-Campo, O. Ferreira Saraiva, P.R. Galerani, F.C. Kryzanowski, and M.C. Carrao-Panizzi (ed.) Proceedings of VII World Soybean Research Conference, IV International Soybean Processing and Utilization Conference. Embrapa Soja.
- Hartwig, E.E. 1986. Identification of a fourth major gene conferring resistance to soybean rust. *Crop Sci.* 26:1135-1136.
- Hartwig, E.E., and K.R. Bromfield. 1983. Relationships among three genes conferring specific resistance to rust in soybeans. *Crop Sci.* 23:237-239.
- Iqbal, M.J., K. Meksem, V.N. Njiti, M.A. Kassem, and D.A. Lightfoot. 2001. Microsatellite markers identify three additional quantitative trait loci for resistance to soybean sudden-death syndrome (SDS) in Essex × Forrest RILs. *Theor. Appl. Genet.* 102:187-192.
- Keim, P., T.C. Olson, and R.C. Shoemaker. 1988. A rapid protocol for isolating soybean DNA. *Soybean Genet. News.* 15:150-152.
- Kosambi, D.D. 1944. The estimation of map distances from recombination values. *Ann. Eugen.* 12:172-175.
- Li, Z., L. Jakkula, R.S. Hussey, J.P. Tamulonis, and H.R. Boerma. 2001. SSR mapping and confirmation of the QTL from PI96354 conditioning soybean resistance to southern root-knot nematode. *Theor. Appl. Genet.* 103:1167-1173.
- Manly, K.F., R.H. Cudmore Jr., and J.M. Meer. 2001. Map Manager QTX, cross-platform software for genetic mapping. *Mammalian Genome* 12:930-932.
- McLean, R.J., and D.E. Byth. 1980. Inheritance of resistance to rust (*Phakopsora pachyrhizi*) in soybeans. *Aust. J. Agric. Res.* 31:951-956.
- Melching, J.S., W.M. Dowler, D.L. Koogle, and M.H. Royer. 1989. Effects of duration, frequency, and temperature of leaf wetness periods on soybean rust. *Plant Dis.* 73:117-122.
- Miles, M.R., R.D. Frederick, and G.L. Hartman. 2003. Soybean rust: is the US soybean crop at risk? American Phytopathological Society, APSnet. Available at www.apsnet.org/online/feature/rust/top.asp (Verified 28 Aug 2004).

- Njiti V.N., K. Meksem, M.J. Iqbal, J.E. Johnson, M.A. Kassem, K.F. Zobrist, V.Y. Kilo, and D.A. Lightfoot. 2002. Common loci underlie field resistance to soybean sudden death syndrome in Forrest, Pyramid, Essex, and Douglas. *Theor. Appl. Genet.* 104:294-300.
- Ogle, H.J., D.E. Byth, and R.J. McLean. 1979. Effect of rust (*Phakopsora pachyrhizi*) on soybean yield and quality in South-eastern Queensland. *Aust. J. Agric. Res.* 30:883-893.
- Rector, B.G., J.N. All, W.A. Parrott, and H.R. Boerma. 1999. Quantitative trait loci for antixenosis resistance to corn earworm in soybean. *Crop Sci.* 39:531-538.
- Rupe, J.C., and G.L. Hartman. 1999. Sudden Death Syndrome. p. 37-39. *In* Hartman, G., J. Sinclair, and J. Rupe (ed.) *Compendium of soybean diseases*. 4th ed. American Phytopathological Society. St. Paul, MN.
- Rytter, J.L., W.M. Dowler, and K.R. Bromfield. 1984. Additional alternative hosts of *Phakopsora pachyrhizi*, causal agent of soybean rust. *Plant Dis.* 68:818-819.
- Schneider, R.W., C.A. Hollier, H.K. Whitan, M.E. Palm, J.M. McKenny, J.R. Hernández, L. Levy, and R. Devries-Paterson. 2005. First report of soybean rust caused by *Phakopsora pachyrhizi* in the continental United States. *Plant Dis.* 89:774.
- Shipe, E.R., J.D. Mueller, S.A. Lewis, P.F. Williams Jr., and J.P. Tomkins. 1997. Registration of 'Dillon' Soybean. *Crop Sci.* 37:1983.
- Sinclair J.B., and G.L. Hartman. 1996. Proceedings of the Soybean Rust Workshop, Urbana, IL. Natl. Soybean Res. Lab. Publ. 1. Urbana-Champaign, IL, USA.
- Sinclair, J.B. and G.L. Hartman. 1999. Soybean Rust. p. 25-26. *In* G.L. Hartman, J. Sinclair, and J. Rupe (ed.) *Compendium of soybean diseases*. 4th ed. American Phytopathological Society. St. Paul, MN.
- Singh, B.B., and P.N. Thapliyal, 1977. Breeding for resistance to soybean rust in India. p. 62-65. *In* R.E. Ford and J.B. Sinclair (ed.) *Rust of soybeans, the problem and research needs*. Int. Agric. Publ. INTSOY Ser. 12. Univ. of Illinois, Urbana, IL, USA.
- 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.* 109:121-128.

- Soybean Genetics Newsletter. 1997. Rules for genetic symbols. 24:19-23.
- USDA, ARS, National Genetic Resources Program. 2005. *Germplasm Resources Information Network - (GRIN)*. [Online Database] National Germplasm Resources Laboratory, Beltsville, Maryland. Available: <http://www.ars-grin.gov/cgi-bin/npgs/acc/display.pl?1443590> (Verified 28 October 2006).
- Vakili, N.G. 1979. Field survey of endemic leguminous hosts of *Phakopsora pachyrhizi* in Puerto Rico. Plant Dis. Rep. 63:931-935.
- Vakili, N.G, and K.R. Bromfield. 1976. *Phakopsora* rust on soybean and other legumes in Puerto Rico. Plant Dis. Rep. 60:995-999.
- Wang, D., P.R. Arelli, R.C. Shoemaker, and B.W. Diers. 2001. Loci underlying resistance to Race 3 of soybean cyst nematode in *Glycine soja* Plant Introduction 468916. Theor. Appl. Genet. 103:561-566.
- Yang, X.B., M.H. Royer, A.T. Tschantz, and B.Y. Tsia. 1990. Analysis and quantification of soybean rust epidemics from seventy-three sequential planting experiments. Phytopathology 80:1421-1427.
- Yang, X.B., A.T. Tschanz, W.M. Dowler, and T.C. Wang. 1991. Development of yield loss models in relation to reduction of components of soybean infected with *Phakopsora pachyrhizi*. Phytopathology 81:1420-1426.

Table 5.1. Field and greenhouse evaluations for type of lesion, severity, and average number of lesions of RILs from Dillon × Hyuuga and Benning × Hyuuga.

Parents/lesion type	Field (2005)			Greenhouse (2006)				
	Lines	Severity		Lines	Lesions cm ⁻²	Sporulation		
	No.	Range	Mean	No.	Range	Mean [‡]	+	-
		----- score [†] -----			----- no. -----		---- % of lines ----	
Dillon × Hyuuga								
Dillon (Tan)	3	2.5 – 3.5	2.9	3	0.70 – 11.01	4.11 ^x	100	0
Hyuuga (RB)	3	1.8 – 3.0	2.3	3	0.46 – 5.97	2.25 ^y	0	100
Tan lesion RILs	59	1.5 – 5.0	3.1 ^a	61	1.47 – 6.82	3.36 ^a	98.4	1.6 ^a
RB lesion RILs	35	1.0 – 4.3	2.4 ^b	37	0.85 – 4.70	2.50 ^b	8.1	91.9 ^b
Benning × Hyuuga								
Benning (Tan)	-	-	-	4	5.43 – 8.53	6.90 ^x	100	0
Hyuuga (RB)	-	-	-	4	1.47 – 4.50	2.48 ^y	0	100
Tan lesion RILs	-	-	-	15	5.40 – 6.98	6.09 ^a	100	0 ^a
RB lesion RILs	-	-	-	16	1.94 – 5.61	3.35 ^b	6.2	93.8 ^b

[†] Score = 1 (dark green canopy) to 5 (significant yellowing at bottom and middle of canopy and some yellowing at the top of the canopy).

[‡] Means followed by different letters are significantly different based on a t-test at $\alpha = 0.05$. Differences in the number of lines in the field and in the greenhouse are due to poor germination of seeds of a given line either in the field or in the greenhouse or inconsistencies in lesion type between the replications.

Table 5.2. QTL mapping of ASR severity from the field and lesion number in the greenhouse using RILs from Dillon x Hyuuga.

Marker	Severity (2005)			Lesion No. (2006)		
	R ²	P-value	2a [†]	R ²	P-value	2a [†]
			Score [‡]			No.
Satt460	21	< 0.0001	0.86	15	< 0.00001	0.89
<i>Rpp?</i> (Hyuuga)	22	< 0.0001	0.86	15	< 0.00001	0.89
Satt307	14	< 0.0001	0.68	14	< 0.0001	0.86

[†] 2a = difference in severity or number of lesions (cm⁻²) at a marker homozygous for the Dillon allele minus homozygous for the Hyuuga allele.

[‡] Score = 1 (dark green canopy) to 5 (significant yellowing at bottom and middle of canopy and some yellowing at the top of the canopy).

Figure 5.1. Genetic linkage map of a region of soybean LG-C2 containing a locus associated with the type of lesion (tan, red-brown, or mixed) caused by Asian soybean rust. Interval distances and the estimated location of a resistance gene conditioning red-brown lesions [*Rpp?*(Hyuuga)] are based on segregation in 117 RILs derived from a cross of Dillon (tan lesions) × Hyuuga (red-brown lesions). a. Soybean consensus C2 linkage map (Song et al., 2004), b. Mapping of lesion type on LG-C2 using greenhouse phenotypes. The map was generated using Kosambi's mapping function. The values to the left of the maps are cM distances.

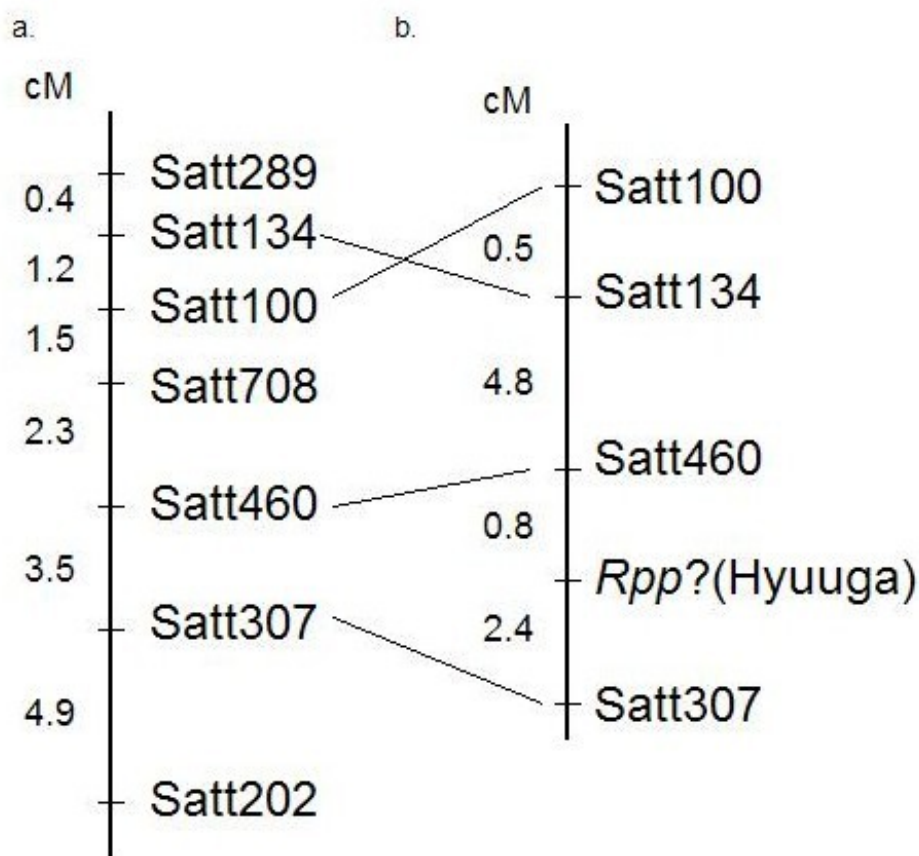
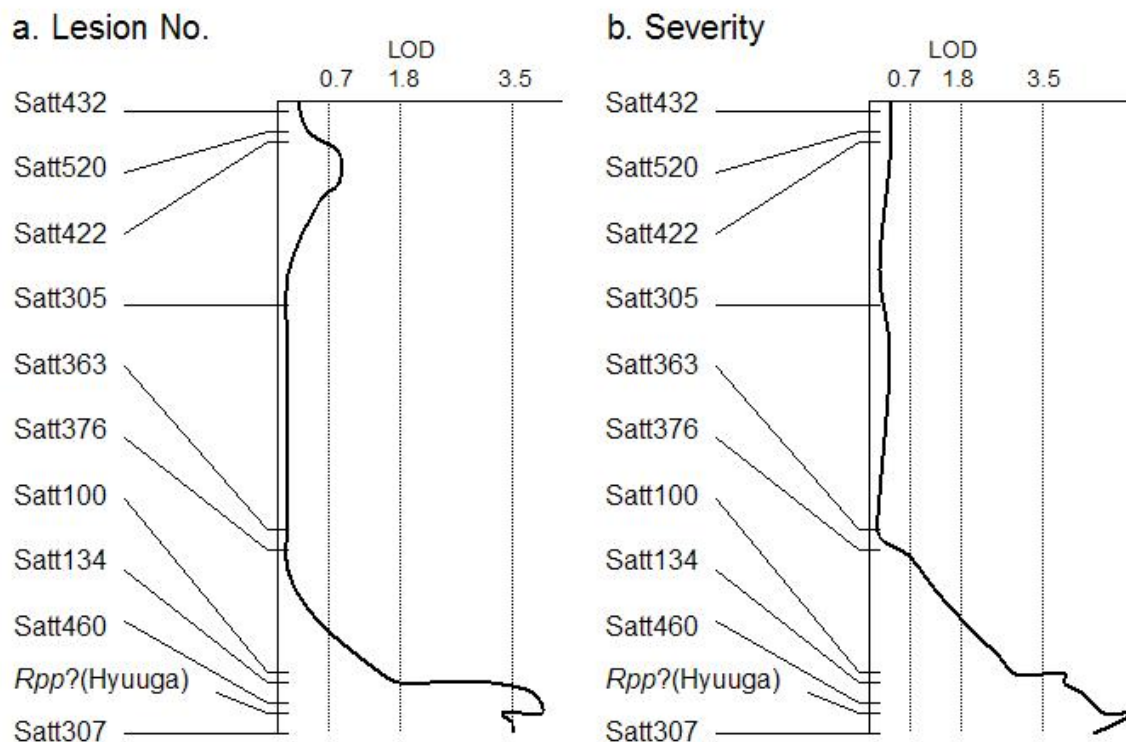


Figure 5.2. QTL likelihood plots for ASR severity and lesion number from the Dillon × Hyuuga RILs. a. Association with lesion number from Griffin greenhouse (LOD = 3.5, and R^2 of *Rpp?*(Hyuuga) = 15% of the phenotypic variation explained), b. Association with severity ratings from Attapulugus, GA (LOD = 3.5, and R^2 of *Rpp?*(Hyuuga) = 22% of the phenotypic variation explained). The significance threshold is indicated by a line at LOD = 2.0.



CHAPTER 6

SUMMARY

Soybean is grown for its protein and oil content, not only as livestock feed, but also in many human food products and industrial applications, making it a major commodity in the world today. Soybean is grown commercially in more than 35 countries. Although yield is still an important component in soybean breeding, improving the chemical composition of soybean oil and incorporating resistance to Asian soybean rust have also become critical targets for improvement. Consumer health concerns and the capability to use soybean oil for biodiesel applications have increased consumer and breeder interest in modifying the fatty acid composition of soybean oil. An increase in the oleic acid content would reduce the need to process the oil, reducing costs and the production of *trans* fatty acids, which have negative health effects. Beginning January 2006, the FDA required that all processed foods contain the *trans* fatty acid content in the nutrition facts panel label.

Six QTLs for oleic acid content have been identified and confirmed in the soybean line N00-3350. Sequence-based SSCP markers developed from genes in the fatty acid biosynthetic pathway and from sequence-tagged-sites near the oleic acid QTLs have allowed us to increase the mapping resolution and provide a resource that can be used to develop additional SNP markers in these regions of the soybean genome. Knowledge of the location of oleic acid-QTLs in soybean and the existence of additive gene action in these QTLs can be used in soybean breeding programs to increase the oleic acid content in the seed. Although incorporating multiple alleles from N00-3350 to achieve this goal could be challenging, molecular markers linked to them are currently being used in soybean breeding programs in the USA to achieve this goal. These markers enable breeders to more effectively select lines for advancement and further testing based on their genotype, reducing the need to continuously phenotype, which is costly, time-consuming and sensitive to

environmental conditions. The identified SSR markers and newly developed SSCP and SNP markers associated with oleic acid content can be used to incorporate QTLs for increased oleic acid content into cultivars that possess high yields and good agronomic performance. The availability of an elite soybean cultivar with the mid-oleic acid phenotype would allow soybean to provide highly cost competitive healthy oil for human consumption.

Crop diseases may be managed using crop rotation, chemical application, or planting resistant cultivars. My research has identified a potentially novel resistance gene to Asian soybean rust in the Japanese cultivar Hyuuga. This research found the presence of the *Rpp?*Hyuuga allele locus produced a red-brown (RB) resistance reaction when exposed to the Asian soybean rust present in Brazil and in the southeastern USA. Data from two different crosses with Hyuuga also showed that the presence of the red-brown lesion type affected the presence or absence of sporulation, which impacts the spread of the rust spores to adjacent plants and fields. The initial mapping studies conducted using phenotypic evaluations from both field and greenhouse experiments with Dillon x Hyuuga recombinant inbred lines placed the *Rpp?*Hyuuga locus in a 3.5-cM region on LG-C2 between Satt307 and Satt460. Additional SSR and SNP markers in this region, and the availability of recombinant lines in this interval allowed mapping and verifying the location of the *Rpp?*(Hyuuga) locus to 2.0 cM between markers Satt460 and Satt079. This increase in mapping resolution is of critical importance for the reduction of linkage drag, and in this case, could be the foundation for eventually cloning this resistance gene.

The SNP marker BARC-10457-640 distinguishes the susceptible parents Dillon and Benning, from Hyuuga. After screening the known sources of ASR resistance genes with six different SNP markers in this region of LG-C2, five haplotypes have been identified in this region. These can potentially be used to determine whether accessions with potential resistance identified through a phenotypic screen possess the *Rpp?*(Hyuuga) gene, or if they represent a novel source of rust resistance.

The availability of linkage maps, the addition of SNP markers to the soybean consensus linkage map, and new detection technologies will greatly accelerate the pace at which QTLs of interest can be identified in mapping populations. Additionally, integrating these new technologies with existing breeding programs will allow the transfer and pyramiding of these QTL/genes at an accelerated pace. Molecular markers linked to the ASR resistance genes can be used to identify and track resistance genes throughout the breeding process. Another advantage of molecular markers linked to disease resistance genes is that marker data can be used to select the best lines for advancement even without the presence of the pathogen. Overall, the availability of molecular markers associated with qualitative and quantitative traits of interest will enhance the effectiveness of soybean breeding programs in incorporating these traits. Future directions in this research area should focus on incorporating the positive alleles for traits of interest including an increase in oleic acid content and rust resistance, and development and use of breeder-friendly molecular markers to reduce the time and cost associated with developing soybean cultivars with desirable agronomic performance, value-added traits, and disease resistance.

APPENDIX 1

PLANT INTRODUCTIONS (PI'S) WITH MID-OLEIC ACID CONTENT

Four PI accessions from Japan were identified in the National Plant Germplasm section of the Genetic Resources Information Network (GRIN) database as having oleic acid contents higher than 44% (Table A.1.1). A total of four pots per entry each with ten seeds, including the parents G99-G725, G99-G3438 and N00-3350, were planted in a greenhouse in Athens 11 Jan. 2005. Trifoliolates for each entry were harvested 3 Feb. 2005. At maturity, seed was harvested independently for each pot 26 April 2005. A bulk sample of 6 seeds per entry was sent for fatty acid analysis to the USDA laboratories in Illinois and North Carolina. Fatty acid content was determined using gas chromatography (Hewlett Packard Model 5890/6890) to evaluate methyl esters. Researchers at the Univ. of Missouri-Columbia re-evaluated some of the PI's with potential for mid-oleic acid content from the greenhouse studies in the field during the summer 2005 (Table A.1.1) (David A. Sleper, personal communication). PI 417360 and PI 506582 had 668 and 594 g kg⁻¹ oleic acid when evaluated in the greenhouse and 415 and 423 g kg⁻¹ oleic acid when evaluated in the field.

Table A.1.1. Fatty acid content for PI's and checks grown in the greenhouse. The fatty acid data presented are the means from the four pots per entry.

		GRIN	Greenhouse 2005	Field 2005
	MG [†]	Oleic acid (g kg ⁻¹)	Oleic acid (g kg ⁻¹)	Oleic acid (g kg ⁻¹)
PI 417360	V	503	668	415
PI 506582	V	445	594	423
PI 549055	I	446	212	na
PI 549057 B	I	449	176	na
G99-G725	VI	225	223	na
G99-G3438	VII	173	184	na
N00-3350	IV	na [‡]	648	na

[†]MG: Maturity Group; [‡] na: indicates that the data is not available.

The SSR amplicon sizes of the PI's and the mapping parents for markers on linkage groups were oleic acid QTL have been found were determined (Chapter 3; Table A.1.2). These can be used as markers to screen segregating populations and eventually map the QTL for oleic acid content from these PI's. This would allow the identification of the number, location, and magnitude of the effect of these QTL, as well as determine whether they are located on the same LG where oleic acid QTL have been previously found. The incorporation of the alleles for oleic acid from these PI's in breeding programs could potentially increase the oleic acid content of soybean seed. PI417360 and PI506582 have already been incorporated in soybean breeding efforts in the southeastern US.

Table A.1.2. Plant Introduction SSR marker amplicon sizes on LG-A1, D2, G, and L.

LG	Marker	PI417360	PI506582	G99-G725	N00-3350
		----- bp [†] -----			
A1	Satt200	250	232	226	229
A1	Satt211	113	113	113	107/111
A1	Satt236	226	235	216	226
A1	Satt511	258	266	250	258
A1	Satt599	185	173	185	173
D2	Satt256	239	239	242	239
D2	Satt301	229	198	245	229
G	Satt191	209	206	191	209
G	Satt199	159	159/170	170	159
G	Satt275	191/198	206	198	230
G	Satt324	246	245	236	246
G	Satt394	272	272	288	272
G	Satt503	268	268	247	265
L	Satt166	258	215	215	261
L	Satt418	255	240	240	234
L	Satt143	300	300	300	297
L	Satt561	249	m [‡]	242	249

[†] bp = base pair; [‡] m = indicates a missing value.

APPENDIX 2

FINE MAPPING A RESISTANCE GENE TO ASIAN SOYBEAN RUST FROM THE CULTIVAR HYUUGA

INTRODUCTION

Hyuuga has a potential novel resistance gene for Asian soybean rust (ASR) caused by the fungus *Phakopsora pachyrhizi*. Using a RIL population from the cross of Dillon × Hyuuga, the *Rpp?*Hyuuga locus was mapped to LG-C2 to a 3.5-cM interval between the SSR markers Satt460 and Satt307. These SSR markers were used to screen 92 individuals from an F_{6:7} population from Benning × Hyuuga, and 16 lines homozygous for Benning alleles and 16 lines homozygous for Hyuuga alleles at these two markers were screened in a greenhouse in Griffin for their reaction to ASR. Based on the results presented in Chapter 5 the marker data correctly predicted the ASR red-brown (RB) or resistance reaction in all lines.

PI200492 ('Komata') has a single dominant gene (*Rpp1*) that confers resistance to an Australian Asian rust isolate (McLean and Byth, 1980; Bromfield and Hartwig, 1980). PI230970 ('Ankur') has *Rpp2* (Singh and Thapliyal, 1977) and PI462312 has the *Rpp3* gene for resistance to ASR (Hartwig and Bromfield, 1983). The cultivar Bing Nan (PI459025) from China possesses a single dominant resistance allele (*Rpp4*) at a locus different from the other three resistance alleles (Hartwig, 1986).

On average, soybean has about 1 SNP per 274 bp (Zhu et al., 2003). A SNP haplotype refers to a distinct combination of SNPs that are tightly linked in a region of a chromosome and have also been described as blocks of DNA that tend to be inherited as entire units from a parent to its progeny (Shastri, 2004). SNPs that differentiate one haplotype from another are potentially useful as markers linked to QTLs or genes. Individuals that are susceptible to a disease and have a shared haplotype can be grouped together, and that haplotype can be used to describe other individuals that will likely be susceptible to a disease. The extent of linkage disequilibrium (LD), or non-random association of alleles at

different loci, can also be used to determine how effectively these haplotypes can be used to categorize susceptible and resistant groups. In contrast with maize, where LD decays within 1500 bp (Remington et al., 2001), data from soybean suggests that LD will be more extensive and significantly decays at distances greater than 2.5 cM, which is roughly equivalent to 1.0 – 1.5 Mbp (Flint-Garcia et al., 2003; Zhu et al., 2003).

OBJECTIVES

The objectives of this research were to increase the resolution in mapping the location of the *Rpp*?Hyuuga locus on LG-C2 and to develop molecular markers associated with the resistance locus that can be used for marker-assisted selection.

MATERIALS AND METHODS

Mapping SNP markers

Dillon and Hyuuga were evaluated for polymorphism using four SNP markers (Table A.2.1) previously mapped between Satt307 and Satt460 on LG-C2 of the consensus soybean linkage map using PCR amplification (Song et al., 2004; David Hyten, personal communication). The PCR products were subjected to single strand conformational polymorphism (SSCP) analysis according to Slabaugh et al. (1997). The polymorphic marker BARC-010457-00640 was used to screen a population of 117 F_{5:6} recombinant inbred lines (RILs) from the Dillon × Hyuuga cross and mapped within the 3.5-cM interval flanked by Satt460 and Satt307 (Fig. A.2.1). Additionally, the SSR markers Satt079 and Staga001 were also mapped between Satt307 and Satt460.

Selection of lines from Benning × Hyuuga

A population of 92 F_{6:7} lines from the cross of Benning × Hyuuga was screened with seven SSR markers on LG-C2 flanked by Satt134 (112.8 cM) and Satt357 (151.9 cM) as described in Chapter 5. A total of 30 lines from this population representing various marker combinations available within this interval were selected to evaluate their reaction to ASR in the greenhouse (Fig. A.2.3).

The lines 11322 and 10700 were included in the assay because they were homozygous for the Hyuuga alleles and Benning alleles, respectively, in the region between Satt460 and Satt307 and their reaction to ASR had been previously determined in March of 2006. A line representing each marker genotype class (Fig. A.2.2), available from a population of RILs from the cross of Dillon x Hyuuga was also included in these greenhouse evaluations.

Phenotypic screening

Three replications of the 30 selected lines from the Benning × Hyuuga population, 11 entries of the Dillon × Hyuuga population, and two entries of each of the parents were planted in a randomized complete block design. Six seeds of each entry were planted in a 10.2-cm pot and the pots were arranged in a 15-pot tray. Each tray contained 14 pots of the lines and one pot of the ASR susceptible cultivar Cobb. The experimental unit was a 10.2-cm pot. All three replications were planted in the greenhouse at the Univ. of Georgia's Griffin campus on 6 Oct. 2006. Plants from Benning × Hyuuga were inoculated with the ASR spore suspension on 23 Oct. 2006, according to the procedure described in Chapter Five. Data on lesion type, lesion number, and visible sporulation was obtained 9 Nov. 2006.

Molecular mapping of the BARC-010457-00640

Map Manager QTXb20 (Manly et al., 2001) was used with the Kosambi (1944) mapping function and the recombinant inbred option to create the genetic map of the Dillon × Hyuuga population including data for the SNP marker BARC-010457-00640 (Fig. A.2.1 and Fig. A.2.2).

Sequencing

PCR products amplified using BARC-010457-00640 were digested with shrimp alkaline phosphatase, SAP (0.1 U μl^{-1}) and *ExoI* nuclease (0.02 U μl^{-1}) according to the manufacturer's protocol. The forward and reverse primers were used to directly sequence the PCR products in both directions in separate sequencing reactions. Each sequencing reaction consisted of 2 μl of BigDye Terminator Cycle Sequencing Ready Reaction mix from the v 3.0 Kit (Applied

Biosystems, Foster City, CA), 1.0 μ M of each primer, 1% DMSO, 2 μ l of 5x sequencing buffer (supplied with the sequencing Kit), and 4 μ l of PCR product in a total volume of 10 μ l. Cycle sequencing conditions were as recommended by the kit manufacturer using annealing temperatures optimized for each primer pair. Sequencing reactions were purified using the MultiScreen Filtration System (Millipore Corporation, Bedford, MA) using Sephadex G50 Superfine (Sigma-Aldrich, St. Louis, MO) per the manufacturer's protocol. Sequencing reactions were evaluated on a Perkin Elmer 3730 XL capillary DNA Analyzer (Applied Biosystems, Foster City, CA). Sequences were aligned using ClustalW, and potential SNPs were identified by visually inspecting the alignments considering the quality of the sequence. The sequence quality was based on the chromatograms, which were visualized using Chromas 2.31 software (Technelysium Pty Ltd).

Confirmation of SNP's using SNaPshot

The SNaPshotTM minisequencing assay was used as a control to test the capture probes which were specifically designed to target potential SNP sites. SNP genotyping for the parents (Dillon, Benning, and Hyuuga) as well as the U.S. cultivar Lee and the Brazilian cultivar FT-2, was performed following the protocol in ABI Prism® SNaPshotTM Multiplex Kit (Applied Biosystems, Foster City, CA). The SNaPshot procedure using single base extension requires an oligonucleotide probe that anneals adjacent to the SNP of interest, and is followed by an extension step with a fluorescently-labeled dideoxy terminator. The capture probes were designed with a modified 21-nucleotide ZipCode sequence from Iannone et al. (2000), at the 5' end and a site-specific sequence at the 3' end adjacent to the SNP location. Reactions were run on 12-cm gels using an ABI Prism 377 DNA sequencer at 750 V for 1.5 to 2 hours, with 1X TBE buffer. Each nucleotide has a unique fluorescent label, and therefore a different color product from each parent observed confirms the SNP.

RESULTS AND DISCUSSION

Four potential SNP markers were identified in the interval between Satt460 and Satt307 (Table A.2.1). Polymorphism between the susceptible parents Dillon and Benning, and the resistant parent, Hyuuga using SSCP gels was only detected for the SNP marker BARC-010457-00640. The marker BARC-010457-00640 was used to screen 100 RILs from Dillon $\times$ Hyuuga and mapped between Satt460 and Satt307 (Fig. A.2.1). The identification of recombination events combined with the ASR reaction data for these 100 RILs indicated the *Rpp?*Hyuuga locus is likely located between Satt079 and Staga001 (Fig. A.2.2). Satt079 is located 0.1 cM from Satt460 and 1.9 cM from Staga001. The Dillon $\times$ Hyuuga RILs 341 and 248 that are homozygous from Dillon in the interval between Satt460 and Staga001 have a tan or susceptible ARS phenotype. RILs 323 and 33 are homozygous for Dillon alleles at Satt460, but homozygous for Hyuuga alleles at the SSR marker Satt079 have a RB or resistant phenotype. These data allowed us to reduce the original 3.5-cM interval between Satt460 and Satt307 in which *Rpp?*(Hyuuga) locus is located by 1.5 cM. The lines from Dillon $\times$ Hyuuga included in Fig. A.2.2 that are heterozygous on LG-C2 have been screened with SSR markers in other linkage groups to rule out the possibility of a seed mixture.

Lines from Benning $\times$ Hyuuga for which phenotypic data was available also indicate that the *Rpp?*(Hyuuga) locus is closer to Satt460 than Satt307. A line homozygous for Hyuuga alleles at Satt460 had a RB lesion type even if it was homozygous for Benning alleles at Satt307 (data not shown). This line also had no visible sporulation. Additional lines from the cross of Benning $\times$ Hyuuga have been genotyped with the SNP marker BARC-010457-00640 and the additional SSR markers in the region. Lines with recombination events in the target region on LG-C2 have been identified (Fig. A.2.3). The reaction to ASR of each of those lines, combined with the marker data confirms the results from the Dillon $\times$ Hyuuga population indicating that the *Rpp?*Hyuuga locus is located near Satt079.

The SNaPshot™ assays allowed confirmation of the nucleotide composition at multiple potential SNP sites of the mapping parents and the four reported sources of resistance genes to ASR. Using this approach we determined the SNP genotypes for six SNP markers on LG-C2 near the *Rpp?*(Hyuuga) locus. Out of the six SNPs evaluated, only the marker 10457-640 distinguished the susceptible parents Dillon (T) and Benning (T), from Hyuuga (A) (Table A.2.2).

When the original sources of resistance and Hyuuga were evaluated in Brazil, PI200492 (*Rpp1*) and PI462312 (*Rpp3*) had a tan lesions (Table A.2.3). PI230970, PI459025A, and Hyuuga all had a RB reaction type. Additionally, a study using SSR markers to evaluate progeny of a cross between the cultivar ‘Williams 82’ and the BC₅ Williams 82 isoline L85-2378 with *Rpp1* identified LG-G as a likely location for the *Rpp1* gene (David Hyten, personal communication). A F_{2:3} population of 108 lines from the cross of Williams 82 × PI462312 (the source of the *Rpp3* resistance) is currently being screened in Ft. Detrick, MD with rust spores from the race India 73-1 (IN73-1) (Perry Cregan, personal communication). These findings indicate that Hyuuga likely has a different resistance gene than *Rpp1* and *Rpp3*. An F_{2:3} mapping population of 130 individuals developed from a cross between PI230970 (*Rpp2*) and the susceptible Brazilian line BRS 184, mapped the location of this gene to the soybean LG-J (Carlos Arias Arrabal, personal communication). A similar study using 80 F₂ individuals from a cross between PI459025 (*Rpp4*) and the same susceptible cultivar mapped the *Rpp4* locus to LG-G (Carlos Arias Arrabal, personal communication).

Brogin et al. (2004), using 113 F_{2:4} lines from a cross between FT-2 (PI 628932; RB lesion type) and the susceptible cultivar Davis, found that a locus associated with lesion type mapped to LG-C2. Although the pedigree for FT-2 does not include any of the previously reported sources of resistance (Fig. A.2.4), when evaluated in Ft. Detrick, MD, for ASR it had a mix of lesion types (USDA, 2006). FT-2 is currently producing a tan lesion type in Brazil, indicating that at least one race of the pathogen has been able to overcome this resistance gene (Carlos Arias Arrabal, personal communication). However, when evaluated in

the Griffin greenhouse assays, FT-2 had a RB lesion type perhaps reflecting the existence of a different ASR strain in the U.S. (Table A.2.3). The SNP markers on LG-C2 that can be used to distinguish between Hyuuga and FT-2 are 10459-641, 10459-643, and 28441-5873 (Table A.2.2). The SNPs designated 641 and 643 are located in the target region amplified by the same primer pair, 10459. Therefore, although the rust resistance from FT-2 has been mapped to the same linkage group, these SNPs on LG-C2 closely linked to *Rpp?*(Hyuuga) indicate that FT-2 has a different haplotype than Hyuuga. This provides additional support that Hyuuga is unique when compared with FT-2.

The proposed markers on LG-C2 are within a 0.96-cM region (Table A.2.1 and Table A.2.2) and could potentially be used as a selection tool for *Rpp?*(Hyuuga). The cultivars Dillon and Benning have the same haplotype when considering the four polymorphic SNP markers, which differs from that of Hyuuga only at the 10457-640 SNP. All four of the previously reported sources of rust resistance genes have a different haplotype from each other in this region (Table A.2.2). However, PI459025A, which has the *Rpp4* gene, has the same haplotype as *Rpp?*(Hyuuga). Overall, from the cultivars and PI's screened we have identified five haplotypes using four SNPs. Zhu et al. (2003) reported that among loci with two or more SNPs there was an average of 3.1 haplotypes. The haplotypes on LG-C2 near the *Rpp?*Hyuuga locus can be used as a tool to screen potentially novel sources of rust resistance genes. At this point, we have evidence to indicate that Hyuuga has a potentially novel source of resistance to ASR that is different from the sources that have been given *Rpp* designation.

Plant disease resistance often results from the presence of a specific resistance (R gene) in the plant, and a corresponding avirulence (avr) gene in the pathogen (Flor, 1956). Studies in wheat (*Triticum aestivum* L.), and sunflower (*Helianthus annuus*) indicate that combining many sources of resistance may be a viable strategy to obtain more durable rust resistance (McIntosh and Brown, 1997; Lawson et al., 1998). Molecular markers associated with the *Rpp?*(Hyuuga) rust resistance and the other sources of rust resistance (*Rpp1* to *Rpp4*) can be used for marker-assisted selection when pyramiding multiple resistance genes into single cultivars.

ACKNOWLEDGEMENTS

We would like to thank Dr. David Hyten and Dr. Perry Cregan (USDA-ARS, Beltsville, MD) for providing the SNP primer sequences in LG-C2 and Dr. Bo-Keun Ha (Univ. of Georgia, Athens, GA) for evaluating the parents using the SNP markers with the SNaPshotTM technique.

REFERENCES

- Brogini, R.L., C.A.A. Arias, N.A. Vello, J.F.F. de Toledo, A.E. Pipolo, L.L. Catelli, and S.R.R. Marin. 2004. Molecular mapping of a gene conferring resistance to soybean rust. Poster presented at the VII World Soybean Research Conference, Foz do Iguassu, PR, Brazil.
- Bromfield, K.R., and E.E. Hartwig. 1980. Resistance to soybean rust (*Phakopsora pachyrhizi*) and mode of inheritance. *Crop Sci.* 20:254-255.
- Flint-Garcia, S.A., J.M. Thornsberry, and E.S. Buckler. 2003. Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54:357-374.
- Flor, H.H. 1956. The complementary genic systems in flax and flax rust. *Adv. Genet.* 8:29-54.
- Hartwig, E.E., and K.R. Bromfield. 1983. Relationships among three genes conferring specific resistance to rust in soybeans. *Crop Sci.* 23:237-239.
- Hartwig, E.E. 1986. Identification of a fourth major gene conferring resistance to soybean rust. *Crop Sci.* 26:1135-1136.
- Iannone, M.A., J.D. Taylor, J. Chen, M.S. Li, P. Rivers, K.A. Slentz-Kesler, and M.P. Weiner. 2000. Multiplexed single nucleotide polymorphism genotyping by oligonucleotide ligation and flow cytometry. *Cytometry.* 39:131-140.
- Kosambi, D.D. 1944. The estimation of map distances from recombination values. *Ann. Eugen.* 12:172-175.
- Lawson, W.R., K.C. Goulter, R.J. Henry, G.A. Kong, and J.K. Kochman. 1998. Marker-assisted selection for two rust resistance genes in sunflower. *Molecular Breeding* 4:227-234.
- McIntosh, R.A., and G.N. Brown. 1997. Anticipatory breeding for resistance to rust diseases in wheat. *Annu. Rev. Phytopathol.* 35:311-326.
- McLean, R.J., and D.E. Byth. 1980. Inheritance of resistance to rust (*Phakopsora pachyrhizi*) in soybeans. *Aust. J. Agric. Res.* 31:951-956.
- Manly, K.F., R.H. Cudmore Jr., and J.M. Meer. 2001. Map Manager QTX, cross-platform software for genetic mapping. *Mammalian Genome* 12:930-932.

- Monteros, M.J., A.M. Missaoui, D.V. Phillips, D.R. Walker, and H.R. Boerma. 2007. Mapping and confirmation of the 'Huyuuga' red-brown lesion resistance gene for Asian Soybean Rust. *Crop Sci.* In Press.
- Remington, D.L., J.M. Thornsberry, Y. Matsuoka, L.M. Wilson, S. R. Whitt, J. Doebley, S. Kresovich, M.M. Goodman, and E.S. Buckler. 2001. Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci. USA.* 98:11479-11484.
- Shastry, B.S. 2004. Role of SNP / haplotype map in gene discovery and drug development: an overview. *Drug Development Res.* 62:143-150.
- Slabaugh, M.B., G.M. Huestis, J. Leonard, J.L. Holloway, C. Rosato, V. Hongtrakul, N. Martini, R. Töpfer, M. Voetz, J. Schell, and S.J. Knapp. 1997. Sequence-based genetic markers for genes and gene families: single-strand conformational polymorphisms for the fatty acid synthesis genes of *Cuphea*. *Theor. Appl. Genet.* 94:400-408.
- Singh, B.B., and P.N. Thapliyal, 1977. Breeding for resistance to soybean rust in India. p. 62-65. *In* R.E. Ford and J.B. Sinclair (ed.) *Rust of soybeans, the problem and research needs.* Int. Agric. Publ. INTSOY Ser. 12. Univ. of Illinois, Urbana, IL, USA.
- Song, Q.J., L.F. Marek, R.C. Shoemaker, K.G. Lark, V.C. Concibido, X. Delannay, J.E. Specht, and P.B. Cregan. 2004. A new integrated genetic linkage map of the soybean. *Theor. Appl. Genet.* 109:122-128.
- USDA, ARS, National Genetic Resources Program. 2006. *Germplasm Resources Information Network - (GRIN)*. [Online Database] National Germplasm Resources Laboratory, Beltsville, MD. Available at <http://www.ars-grin.gov/> (Verified 28 Oct 2006).
- Zhu, Y.L., Q.J. Song, D.L. Hyten, C.P. Van Tassell, L.K. Matukumalli, D.R. Grimm, S.M. Hyatt, E.W. Fickus, N.D. Young, and P.B. Cregan. 2003. Single-nucleotide polymorphisms in soybean. *Genetics* 163:1123-1134.

Table A.2.1. SNP markers tested in the region between Satt307 and Satt460 on LG-C2.

Sequence ID	Locus name	LG	Position (cM) [†]	Marker type	GenBank No.
	<i>Satt460</i>	C2	117.76	SSR	
28441	BARC-028441	C2	118.95	SNP	BE658320
41743	BARC-041743	C2	118.95	SNP	BE821157
10459	BARC-010459	C2	119.09	SNP	AF014502
	<i>Staga001</i>	C2	119.85	SSR	
10457	BARC-010457	C2	119.91	SNP	AB030490
	<i>Satt307</i>	C2	121.27	SSR	

[†] cM = Positions are based on the consensus soybean linkage map (Song et al. 2004; David Hyten, personal communication).

Table A.2.2. SNP genotypes of mapping parents and sources of Asian soybean rust resistance.

Accession	R gene	BARC 028441- 5871	BARC 028441- 5873	BARC 041743 - 8075	BARC 010459 - 641	BARC 010459- 643	BARC 010457- 640	Haplotype
Dillon		A	G	G	G	T	T	1
Benning		A	G	G	G	T	T	1
Lee		A	A	G	G	T	A	2
Hyuuga	<i>Rpp?</i> (Hyuuga)	A	G	G	G	T	A	3
FT-2	<i>Rpp?</i>	A	A	G	T	C	A	4
PI 200492	<i>Rpp1</i>	A	G	G	T	C	A	5
PI 230970	<i>Rpp2</i>	A	A	G	G	T	A	2
PI 462312	<i>Rpp3</i>	A	A	G	T	C	A	4
PI 459025A	<i>Rpp4</i>	A	G	G	G	T	A	3

Table A.2.3. Asian soybean rust reaction of Hyuuga and previously reported sources of resistance.

Name	R gene	Lesion type		
		Brazil greenhouse	Attapulgu field	Griffin greenhouse
PI200492	<i>Rpp1</i>	Tan	Tan	Tan
PI230970	<i>Rpp2</i>	RB	RB	RB
PI462312	<i>Rpp3</i>	Tan	na [†]	RB
PI459025A	<i>Rpp4</i>	RB	na	RB
Hyuuga	<i>Rpp?</i> (Hyuuga)	RB	RB	RB
FT-2	<i>Rpp?</i>	No lesions	na	RB

[†] na indicates that phenotypic data for these lines at the given location is not available.

Figure A.2.1. Molecular mapping of the SNP marker BARC-010457-00640. The soybean consensus linkage map is on the right and our map is on the left.

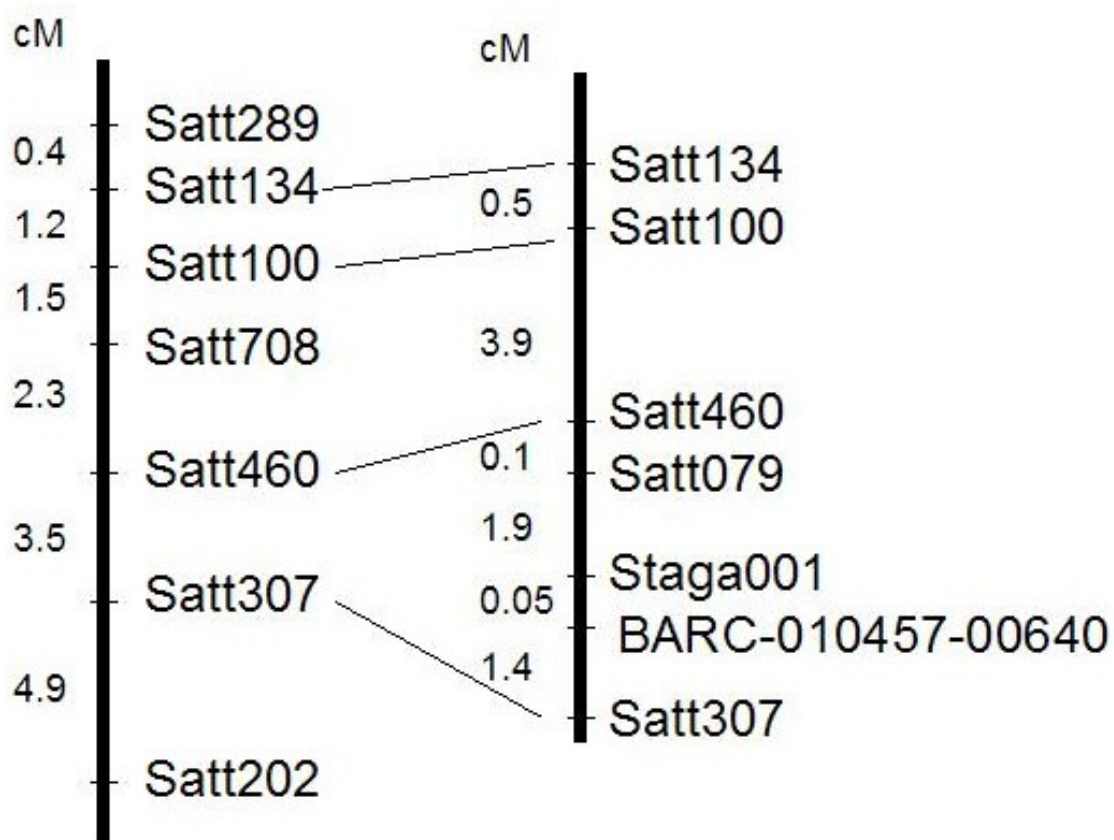


Figure A.2.2. Graphical genotypes of Dillon × Hyuuga RILs. *Phakopsora pachyrhizi* lesion type (LT) and presence (+) or absence (-) of sporulation was obtained from the greenhouse conducted in Feb. 2006. The dashed line indicates the original interval between Satt460 and Satt307 in which *Rpp?*(Hyuuga) was mapped (Monteros et al., 2007). The dotted line indicates the most likely position of the *Rpp?*(Hyuuga) locus.

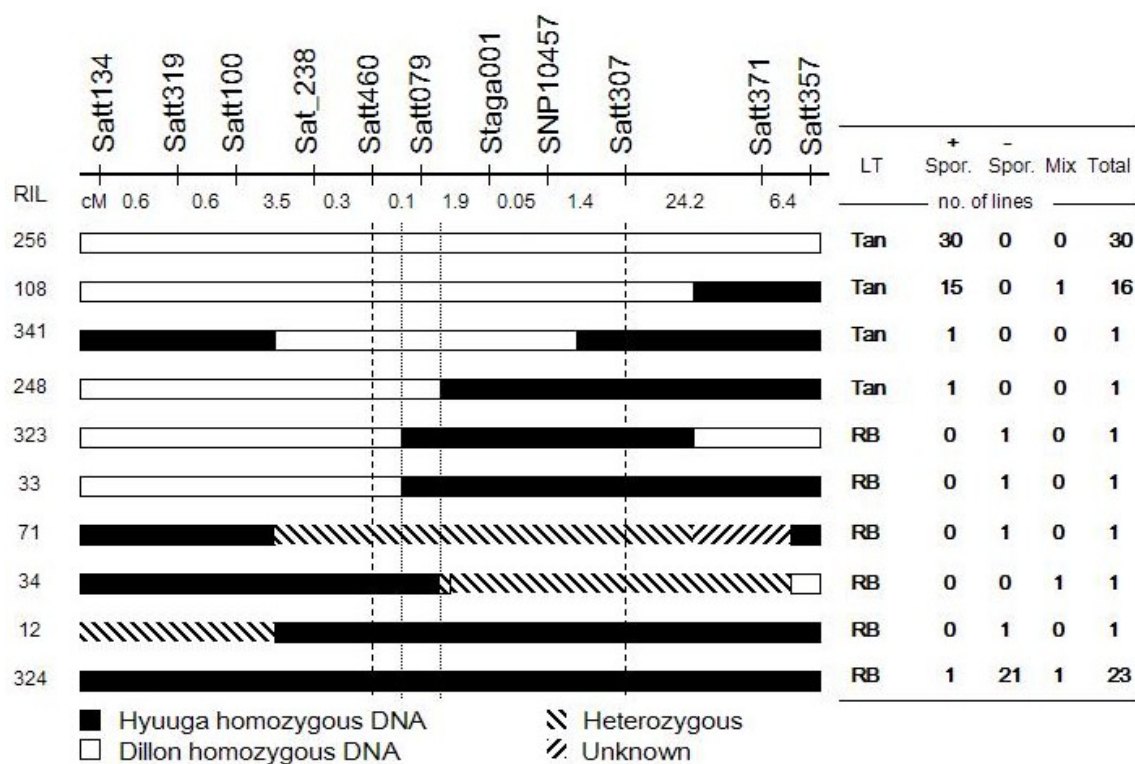


Figure A.2.3. Graphical genotypes of Benning × Hyuuga RILs. *Phakopsora pachyrhizi* lesion type (LT) and presence (+) or absence (-) of sporulation from plants evaluated in a greenhouse in Feb. 2006. The dashed line indicates the original interval between Satt460 and Satt307 in which *Rpp?*(Hyuuga) was mapped. The dotted line indicates the most likely position of the *Rpp?*(Hyuuga) locus.

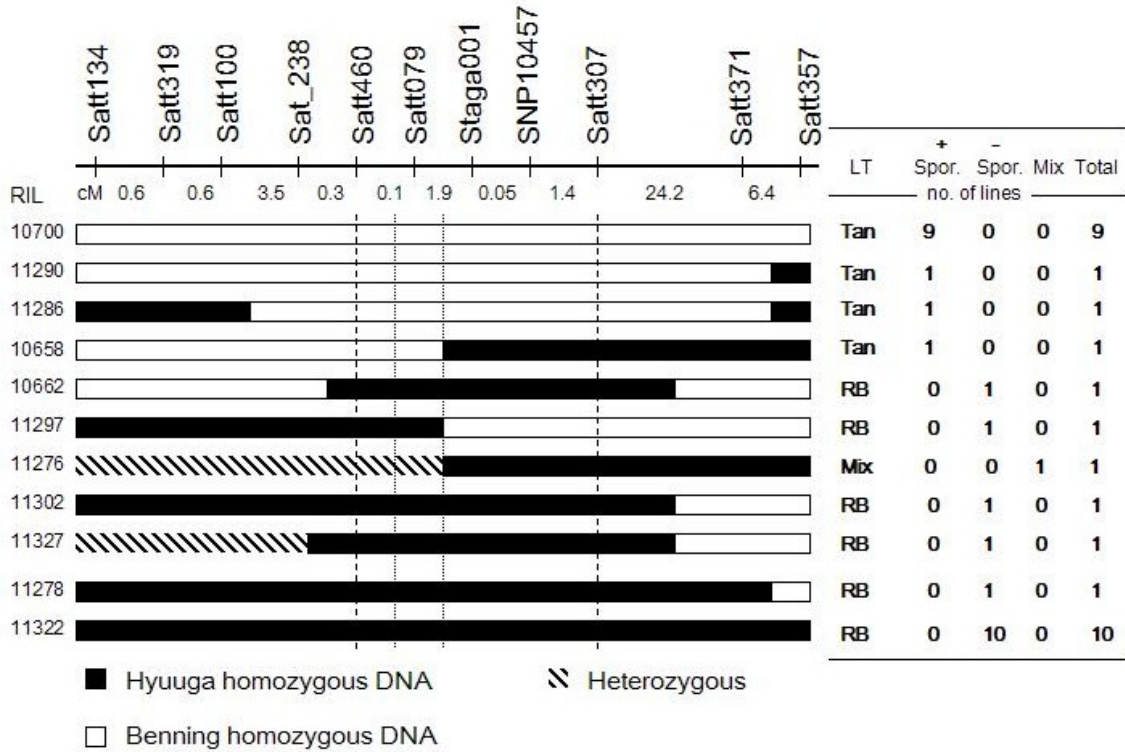
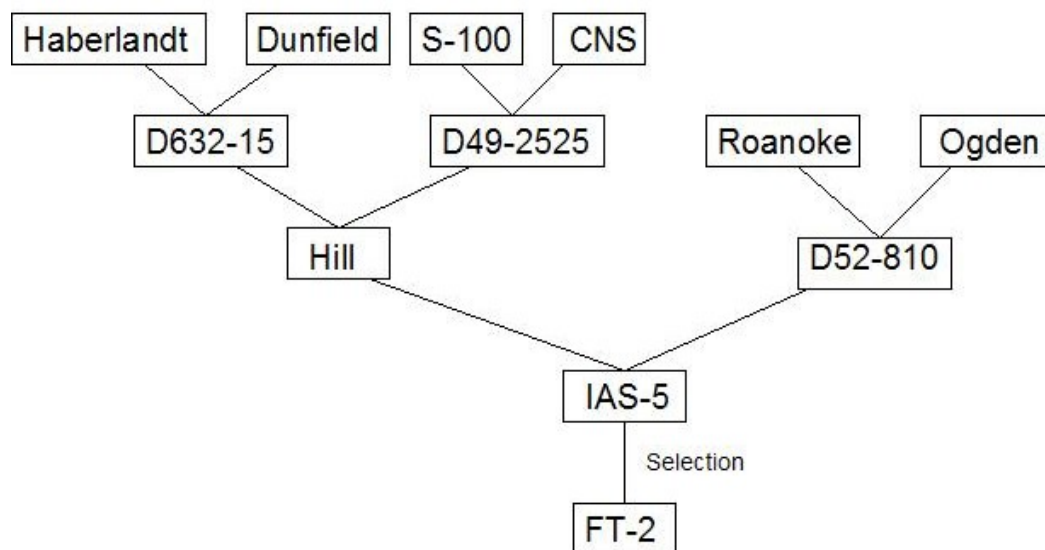


Figure A.2.4. Pedigree of the Brazilian line FT-2.



APPENDIX 3

OLIGONUCLEOTIDES FOR OLEIC ACID QTL

Table A.3.1. Oligonucleotide sequences for SNP markers associated with fatty acid gene sequences.

Primer	5' to 3' primer sequence	Temp annealing
12447F	GCGATTGAGATATGGGCATTA	56
12447R	GCGATAAAAGAAACCACACAAG	
12453F	GCGAGTGTGCCGATTTAC	54
12453R	GCGTGTTTCAATTTACTGTGT	
12507F	GCGTAATATAATGCTTTGAGTG	54
12507R	GCGTTCGTTATTGAGAGTTT	
14049F	GCGAGAGGATAAGTCATAAGTG	54
14049R	GCCCCAATTTGTCTGTGTAATC	
14101F	GCCCGATTCTAATTTTGTGGATG	54
14101R	GCCCATTTAGGTCCATAAGTC	
15155F	GCCCCGAAAAATAGCCTTG GTTG	56
15155R	GCCCGTAAATCCTCCTACTC	
15231F	GCGGCTGCTAAGTTTGAATGTGTAAG	54
15231R	GCGTCCCATGTCGGTTATATTC	
15633F	GCGGGAAACTATAAATAGGGTTCTC	54
15633R	GGGCAATTAAGCATAGGATTCATC	
15783F	GCGGCTATATGTCATAAAGATAAC	54
15783R	GCGGGACGTTGTAATAAAGTTGTG	
15809F	GCGGCAAATAAATAGAGTTTTTC	48
15809R	GCGGGAGAGACACGTCTAATTGAG	
15829F	GCCCCTGTTGCCTTTAGAGGACTAC	56
15829R	GCCCGAACTTTGAATTTTCATTTTC	
16289F	GGGATGGTATCACTGTAAAGAG	54
16289R	GCGGGAATAAAAAGAATTACTCAAG	

Table A.3.2. Oligonucleotide sequences for SSCP markers from genes in the fatty acid biosynthetic pathway.

Primer	5' to 3' primer sequence	Temp annealing
U6041_6F	CTGCAACAGCGAAGAACATGG	54
U6041_6R	ATGAGGACTGCTTGGGAGGAA	
accA_6F	ATAGCCGAAGGTTCAAGGTAA	54
accA_6R	ACTGCCCAACATTGTCCGT	
accB_2F	ACTCTCTCCGCTTCTCTCCTA	54
accB_2R	AATAAGTGCCCAACATAAAGGT	
QTLA_1F	TCTTTTGTAAGTCTGGCTAA	54
QTLA_1R	TCAACGAACTGAAGAGGGTCA	
QTLG_1F	TACACTACAAAAATATGTGGC	50
QTLG_1R	AAATATCAGCTTGCTGC	
G1F	TCAGATTCCACTGCGTTTGT	48
G1R	TGATTACATTTGGGAGGATGA	
17027F	GCGTTGGCAAGGTCTTTATCAT	54
17027R	GCGTGGTTCTGTGACTGAGAGTATTG	
21575F	CGAGAGATCAAAATGTCGTCATAAC	54
21575R	GCAAATACAACCAGGACAAGATGAT	
13835F	GGGTTGGCAAGGTCTTTATCATAC	54
13835R	GCGTTGGTTCTGTGACTGAGAGTA	
A1-14639F	GCGGAACGGAAAAATAAGATAT	58
A1-14639R	GGCTTGGAGGTCCTTGTGACAC	
G-10857F	GCGTGAGAATAAAATCCTGACAAC	54
G-10857R	GCGAACCAACTACAAAATATTG	
L-44913F	TCAACGAAAGTTCCTTAAGTCAAA	58
L-44913R	CCGCGACGACAACAACACTC	
L-35235F	CTGAAATGTTGAAAGAGGATGAG	58
L-35235R	GGCCTAGGTAGAAGATTTGTTGT	
D2-32349F	AAGGGAATGTTCAATTCTCTGGGA	58
D2-32349R	TTGTATTGCCAAGTCTCGCCAAA	
D2-19505F	GCGTTTGGAAGAGATTTTTTGTG	60
D2-19505R	GCGGGTGCTTTGATGACATTCTATTTG	
D2-41779F	AGAAGCAATATCATGAACAGGAA	58
D2-41779R	CAATTGACAACCACTAGGACTGT	

Table A.3.3. Sequences from genes in the fatty acid biosynthetic pathway from other species. At = *Arabidopsis thaliana*, Bj = *Brassica juncea*, Ha = *Helianthus annuus*, Oe = *Olea europaea*.

Source	GenBank No.	Acronym	Description
At	NM_179808	Fad3	omega-3 fatty acid desaturase, ER (At2g29980) mRNA
At	NM_128552	Fad3	omega-3 fatty acid desaturase, ER (At2g29980) mRNA
At	D26508	Fad3	Fad3 gene for microsomal omega-3 fatty acid
Bj	AJ278479	FatA	FatA gene for Acyl-ACP thioesterase, exons 1-3
Bj	AJ294419	FatA	FatA gene for oleoyl hydrolase, exons 1-6
Ha	AY805125	FatA	FATA-2_acyl-acp thioesterase, partial
Ha	AF036565	FatB	FatB1 mRNA, complete cds
Ha	AJ242915	FatB	FatB thioesterase strain CAS-5
Ha	AJ242916	FatB	FatB gene strain CAS-12
Ha	AY803019	FatB	FatB gene_acyl ACP thioesterase
Ha	AY805143	KasI	KasI-2
Ha	AY805139	KasII	KasII mRNA, partial sequence
Oe	AY733077	Fad2_2	Fad2_2 mRNA complete
Oe	AY733076	Fad2_1	FAD2-1 mRNA complete

APPENDIX 4

LIST OF ABBREVIATIONS

ACCase	Acetyl-CoA carboxylase
ACP	Acyl carrier protein
ANOVA	Analysis of variance
ASR	Asian soybean rust
BAC	Bacterial artificial chromosome
Bp	Base pair
CHD	Coronary heart disease
CIM	Composite interval mapping
EMS	Ethylmethanesulfonate
EST	Expressed sequence tag
ExoI	Exonuclease I
FAD	Fatty acid desaturase
FAS	Fatty acid synthase
FAT	Fatty acid thioesterase
FDA	Food and drug administration
GRIN	Genetic resources information network
HDL	High-density lipoprotein
KAS	Ketoacyl-ACP synthase
LDL	Low-density lipoprotein
LG	Linkage group
MAS	Marker-assisted selection
Mg	Megagrams
MG	Maturity group
MMT	Million metric tons
NIL	Near isogenic lines
PI	Plant introduction
PTGS	Post transcriptional gene silencing
PUFA	Polyunsaturated fatty acid
QTL	Quantitative trait loci
RB	Red-brown

RFLP	Restriction fragment length polymorphism
RIL	Recombinant inbred line
SAP	Shrimp alkaline phosphatase
SBE	Single base-chain extension
SNP	Single nucleotide polymorphism
SSCP	Single strand conformational polymorphism
SSR	Simple sequence repeat
USDA	US Department of agriculture
WHO	World health organization
YAC	Yeast artificial chromosome