

Molecular Mapping and Identification of Soybean Fatty Acid Modifier Quantitative Trait Loci

David L. Hyten^{a,*}, Vincent R. Pantalone^a, Arnold M. Saxton^b,
Michael E. Schmidt^c, and Carl E. Sams^a

^aThe University of Tennessee, Department of Plant Sciences, Knoxville, Tennessee,

^bThe University of Tennessee, Department of Animal Science, Knoxville, Tennessee,

and ^cSouthern Illinois University, Department of Plant, Soil, and General Agriculture, Carbondale, Illinois

ABSTRACT: Altering FA content in soybean [*Glycine max* (L.) Merr.] oil for improved functionality is a research goal of many soybean breeders. Several of the genes that alter palmitic, stearic, oleic, linoleic, and linolenic acids are modifier genes with small effects, causing these FA traits to act as quantitative traits. The objective of this study was to identify modifier FA quantitative trait loci (QTL) in soybean. A recombinant inbred line population was created from two prominent ancestors of currently available U.S. cultivars (Essex and Williams) and grown in five environments. One hundred simple sequence repeat markers spaced throughout the genome were mapped in this population. QTL were found for all five FA traits on the soybean linkage groups C2, D2, D1b, F, K, and L. A single marker interval on linkage group L contained the largest QTL for palmitic ($r^2 = 13.1\%$), oleic ($r^2 = 35.3\%$), linoleic ($r^2 = 50.5\%$), and linolenic acids ($r^2 = 24.8\%$); however, this interval also contained the gene for growth habit (Dt1) and was significantly associated with maturity. Other modifier QTL found in this study may be of use in marker-assisted selection to enable breeders to increase genetic gains for desirable FA composition of soybean.

Paper no. 10921 in *JAOCs* 81, 1115–1118 (December 2004).

KEY WORDS: Composite interval mapping, FA, quantitative trait loci, soybean [*Glycine max* (L.) Merr.].

One of the main uses of soybean is for extracted vegetable oil in food and industrial products. Soybean oil is predominately composed of five FA: palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2), and linolenic acids (18:3) (1). Saturated fat and *trans*-FA (formed during the hydrogenation process of soybean oil) have been found to be linked to heart disease (2,3). The U.S. Food and Drug Administration (FDA) has recently passed a requirement for food processors to include the amount of *trans*-FA present on all food nutrition labels by 2006 (4). This new requirement is increasing pressure on soybean breeders to alter current soybeans to have less saturated fat and less *trans*-isomer producing unsaturated FA through nontransgenic breeding methods. A natural soybean oil that enables manufacture of foods with reduced saturated fat and *trans*-FA will help U.S. soybeans to remain competitive with other oil crop alternatives (1).

*To whom correspondence should be addressed at Soybean Genomics and Improvement Laboratory, USDA-ARS-PSI, Bldg. 006, 10300 Baltimore Ave., Beltsville, MD 20705. E-mail: hytend@ba.ars.usda.gov

Several genes have been discovered that alter FA content considerably (for a review, see Ref. 5). Some genes are naturally occurring, and others were produced through mutation (6). It has been suggested that there are modifying genes along with the major FA genes that affect the different FA levels (5,7). These modifier genes have small effects that cause the FA profiles of populations to act as a quantitative trait. Through selection of these modifier genes, a wide range of FA profiles could be developed by breeders to help meet changing industry demands.

Quantitative traits are controlled by quantitative trait loci (QTL), which are regions of the chromosome that contain single or multiple genes that have small or large effects on the trait. Molecular markers such as simple sequence repeat (SSR) markers and restriction fragment length polymorphism (RFLP) markers have made it easier to map QTL in soybean. Many QTL have been discovered for yield, protein, oil, maturity, height, and other traits, but few modifier QTL have been found for seed FA concentration (8). Methods for QTL mapping also have improved and now include simple linear regression, interval mapping (9), and composite interval mapping (CIM) (10,11). CIM is a QTL mapping procedure that tests an interval on a genetic map for significant QTL while using the most important unlinked markers to control for genetic background. The main advantage of CIM is that it gives a more accurate location and estimated additive effect of the QTL in a population than simple linear regression or interval mapping.

Integration of new FA modifier alleles into current elite cultivars and breeding lines to meet consumer and processor needs will be facilitated with the knowledge of existing FA QTL. The objective of this study was to map soybean seed FA composition modifier QTL in a population that does not segregate for any known major FA genes. The major FA genes often negatively affect agronomic traits such as yield (5) so we utilized a prominent population known to have agronomically important QTL as our mapping population. The prominent population is derived from the cross of Essex by Williams. This cross has contributed to the genetic background of many northern and southern United States cultivars and elite breeding lines (12,13). Lines selected out of populations formed from these parents are responsible for the majority of currently used breeding material, and any modifier

QTL present could immediately be selected through marker-assisted selection in current breeding programs.

MATERIALS AND METHODS

One hundred and thirty-one F_6 -derived recombinant inbred line (RIL) were created at The University of Tennessee (Knoxville Experiment Station (KES), Knoxville, TN) from the cross between the soybean cultivars Essex and Williams. The $F_{6,8}$ RIL were planted in a randomized complete block design with three replications in each of two environments [KES-C field and the West Tennessee Experiment Station (WTES) at Jackson, TN] during the year 2000 and three replications in each of two environments [KES-V field and Harrisburg, IL] with two replications at one environment (WTES) during the year 2001 as described in Hyten *et al.* (14). DNA was isolated from all 131 RIL by taking a single leaf from five plants and combining them into one sample. Then Qiagen Plant Easy DNA Extraction Kit (Qiagen, Hilden, Germany) was used to obtain purified DNA.

Essex and Williams were tested with 568 SSR markers previously placed on the integrated soybean genetic linkage map (15). The sequence information for the soybean SSR primers and current integrated soybean genetic linkage map are publicly available from the USDA internet site http://bldg6.arsusda.gov/~pooley/soy/cregan/soy_map1.html. One hundred polymorphic markers were chosen to achieve a reasonable coverage of all 20 molecular linkage groups. Polymerase chain reaction (PCR) conditions and visualization of PCR products are described in Hyten *et al.* (14).

Seed FA composition was analyzed using GC of the methyl esters. Preparation of seeds consisted of three seeds crushed per replicate per location and soaked overnight in 1.5 mL chloroform/hexane/methanol (8:5:2, by vol). One hundred microliters of the extract was transferred to a 1.5-mL autosampler vial and 75 μ l of methylating reagent [0.5 M sodium methoxide in methanol/petroleum ether/ethyl ether (1:4:2, by vol) was added. FAME were separated from these samples using a Hewlett-Packard model HP 6890 gas chromatograph (Palo Alto, CA) equipped with a model 7673 autosampler, flame-ionization detector, and an immobilized 30 \times 0.53 mm inner diameter, AllTech AT-Silar capillary column with 0.5 μ m fused stationary phase. Operating conditions were: carrier, He (20 mL/min); 20:1 (vol/vol) split injection; injection temperature, 250°C; detector temperature, 275°C, and column temperature, 240°C.

One hundred and thirty-one lines were mapped with 100 SSR markers covering all 20 soybean linkage groups and three phenotypic markers [flower color (*W1*), pubescence (*T*), and growth habit (*Dt1*)]. MAPMAKER/EXP (16,17) was used to estimate genetic distances between SSR markers. A minimum likelihood of odds (LOD) ≥ 3.0 and a maximum distance ≤ 50 centimorgan (cM) were used to test linkages among markers.

QTL analysis was performed using the phenotypic trait values and SSR RIL scores on 131 RIL planted in the field

environments. A one-way ANOVA (18) was initially used to identify candidate QTL. The soybean composite integrated map was used as the genetic map because the map distances are based on five populations (15) and should have less variability associated with the map distances. Simple linear regression and CIM were performed using QTL Cartographer (19,20). CIM was used to identify and map significant QTL where two or more markers were linked and simple linear regression was utilized for QTL detection when a marker was not linked to another marker. One thousand permutation tests were performed on each trait to establish empirical LOD thresholds at the 5% alpha level (21). Broad-sense heritability estimates were computed as

$$h^2 = \sigma_g^2 / [\sigma_g^2 + (\sigma_{ge}^2/e) + (\sigma^2/re)] \quad [1]$$

where h^2 = heritability; σ_g^2 = genotypic variance; σ_{ge}^2 = genotype \times environment variance; σ^2 = error variance; r = number of replications, and e = number of environments (22).

RESULTS AND DISCUSSION

A total of 100 SSR markers plus three phenotypic markers were mapped in the Essex \times Williams RIL population. The resulting genetic linkage map covered 869.4 cM, which is about 34% of the currently known recombination distance of 2524 cM, averaged from the integrated soybean genetic linkage map (15). The Essex \times Williams genetic map did not show any major discrepancies with the integrated soybean genetic linkage map. SSR markers on linkage groups A2, B1, and B2 were not polymorphic except for a few unlinked SSR markers, so only simple linear regression was performed to test for QTL on these linkage groups. Continual placement of markers in the future should help to assure more coverage of the genome so that additional QTL can be detected.

As expected, the population means fell within the normal range for FA composition that is typical of most currently grown soybean cultivars (Table 1). Nevertheless, the diversity of the two parents did lead to a range within the RIL population, which enabled modifier FA QTL to be detectable. Palmitic, stearic, linoleic, and linolenic acids have small ranges for the RIL population, which verifies that only modifier QTL with small additive effects were detected and mapped.

TABLE 1
Mean, Standard Deviation, Range, Heritability, and Parental Means for Soybean Seed FA Concentration in an F_6 -Derived Recombinant Inbred Line Population of Essex \times Williams Soybean Grown in Five Environments

FA	Mean g/kg	Standard deviation g/kg	Range g/kg	Essex g/kg	Williams g/kg	h^2
16:0	110	3.6	102–119	114	106	0.56
18:0	41	3.3	33–52	40	39	0.74
18:1	237	29.9	182–340	215	253	0.71
18:2	536	24.7	453–575	549	534	0.70
18:3	77	6.6	58–90	82	69	0.43

The range for 18:1 was larger, which could be due to a few important modifier genes or several minor modifier genes of the FA pathway or maturity that will cause earlier maturing lines to have higher 18:1 due to higher temperature during seed fill (1).

Minor ($r^2 < 10\%$) and major QTL ($r^2 \geq 10\%$) were found for all five FA (Table 2). Three 16:0 QTL were found on linkage groups D2, K, and L with the allele for decreased 16:0 content in the seed being present in Williams. The interval significant for the 16:0 QTL on LG L was also significant for 18:1, 18:2, and 18:3 and explained the most total variance of these four FA traits (Table 2). The large effect of this interval on LG L could be due to the gene for growth habit (*Dt1*), which is located in the same interval at 89.1 cM on the integrated soybean genetic linkage map (15). The locus for growth habit segregates in the Essex (determinate) by Williams (indeterminate) cross. Rebetzke *et al.* (7) found growth habit and reduced palmitic acid content were not independent and that normal 16:0 lines were frequently associated with an indeterminate growth habit. This would agree with our discovery of a major QTL from the Williams cultivar increasing 16:0 by 1.4 g/kg. A major and a minor QTL for reduced palmitic content were recently mapped to LG A1 and M from the low palmitic line N87-2122-4 (23). In addition Diers and Shoemaker (24) found three major QTL on LG B2 and J. The three 16:0 QTL alleles found in our study do not coincide with any previously found QTL and are new QTL for reducing palmitic acid.

The 18:0 QTL allele from Essex on C2 and L (Table 2) decreased the concentration of 18:0 in the seed. Rebetzke *et al.* (25) observed that maturity could have a significant effect on

TABLE 2
Position and Effects of QTL Significant at the 5% Alpha Level Associated with Soybean Palmitic Acid (16:0), Stearic Acid (18:0), Oleic Acid (18:1), Linoleic Acid (18:2), and Linolenic Acid (18:3) in an Essex x Williams F6 Recombinant Inbred Line Population, Based on QTL Cartographer Analysis

Trait	LG ^a	QTL position ^b (cM)	95% C.I. of QTL position ^c	r^2	LOD ^d	Effect ^e
16:0	D2	28.5	<24.5–46.5	9.6	3.0	1.1
	K	78.6	66.6–>78.6	8.3	3.2	1.1
	L	89.1	70.5–93.1	13.1	5.4	1.4
18:0	C2	121.3	117.8–125.3	13.1	4.7	-1.2
	L	58.2	54.6–64.2	16.8	5.5	-1.4
18:1	D1b	108.9	97.2–>116.9	7.2	3.1	8.1
	L	82.5	68.5–89.1	35.3	15.1	-2.0
18:2	F	93.7	83.7–106.1	10.0	4.2	-7.9
	L	74.5	66.5–88.5	50.5	20.2	17.8
18:3	F	11.4	7.9–17.4	8.1	4.8	-1.9
	L	50.6	42.6–62.2	13.6	4.6	2.6
	L	82.5	72.5–91.1	24.8	8.2	3.7

^aLinkage group in which the quantitative trait loci (QTL) is located based on the integrated soybean genetic linkage map (Ref. 14).

^bEstimated position on the integrated soybean genetic map.

^cC.I., confidence interval.

^dLOD, likelihood of odds.

^eAverage change (g/kg) resulting from substituting one allele of Williams with one allele of Essex.

FA levels. In our study, the population varied from maturity group (MG) III to MG V, and a major maturity gene was mapped on LG C2 ($r^2 > 70\%$), and a minor maturity gene was mapped on LG L ($r^2 < 10\%$) in the same area of the FA QTL. The FA QTL mapped at LG C2 and L could be due to a pleiotropic effect or a consequence of the maturity QTL on the same linkage group. Only one marker was significant for 18:0 and not maturity. SSR marker Satt070 on LG B2 was unlinked to any of the other markers mapped in our population and was significantly associated with 18:0 ($P < 0.0001$, $r^2 = 47\%$). This modifier QTL may influence the major gene *Fas*, which governs 18:0 concentration and is in the same region. Spencer *et al.* (26) detected Satt070 at 1.9 cM from the *Fas* locus having an r^2 of 67% in a population, which contained a natural mutation governing a two-fold increase in 18:0 (6,27). In our study the Essex allele of Satt070 had an additive effect of 4.5 g/kg.

Two new QTL were found for 18:1 on LG D1b and L with the QTL on L having a major effect on 18:1. The interval significant for the QTL on LG L for 18:1 is only significant for 18:1 and 18:3, and its 95% confidence interval does not overlap with *Dt1*. Diers and Shoemaker (24) found the RFLP marker A023_1 on L at 37cM to be a major QTL on LG L for 18:3 with an $r^2 = 26\%$. Linkage group F contained a new QTL for 18:2 and 18:3 although the two QTL for each trait are ~80cM apart, leading to the conclusion that these are separate QTL for each trait.

Several modifier QTL for FA composition of soybean seed were mapped in the RIL population created from an Essex x Williams cross. These QTL should be present in many of the cultivars grown in northern and southern United States. Some of the effects of the QTL could be explained by maturity or growth habit. However, several QTL were found that did not coincide with these factors. With very few studies on modifier QTL, most of the QTL found in this study need to be verified by other researchers and in populations where maturity effects are not present to determine their environmental stability and usefulness for marker-assisted selection.

ACKNOWLEDGMENTS

The authors would like to thank James H. Klein III (Southern Illinois University at Carbondale, IL), Debbie Landau-Ellis (UT-Knoxville, TN), Beth Meyer (UT-Knoxville, TN), and Gordon G. Percell (WTEs) for their technical support and Dr. Perry B. Cregan (USDA-ARS, Beltsville, MD) for his suggestions. This research was supported by the United Soybean Board project #1240 and by the Tennessee Soybean Promotion Board.

REFERENCES

- Wilson, R.W., Seed Composition, in *Soybeans: Improvement, Production, and Uses*, 3rd ed., edited by H.R. Boerma and J.E. Specht, ASA-CSSA-SSSA, Madison, WI, 2004, pp. 621–677.
- Willett, W.C., Diet and Health: What Should We Eat?, *Science* 264:532–537 (1994).
- Beare-Rogers, J., Food Fats and FA in Human Nutrition, in *Development and Processing Vegetable Oils for Human Nutrition*, edited by R. Przybylski and B.E. McDonald, AOCS, Champaign, IL, 1995, pp. 1–7.

4. U.S. Food and Drug Administration, FDA Proposes New Rules for *Trans* FA in Nutrition Labeling, Nutrient Content Claims and Health Claims, *Federal Register*, November 12, 1999.
5. Pantalone, V.R., D.R. Walker, R.E. Dewey, and I. Rajcan, DNA Marker-Assisted Selection for Improvement of Soybean Oil Concentration and Quality, in *Legume Crop Genomics*, edited by R.F. Wilson, H.T. Stalker, and E.C. Brummer, AOCS Press, Champaign, IL, 2004, pp. 283–311.
6. Wilson, R.F., J.W. Burton, V.R. Pantalone, and R.E. Dewey, New Gene Combinations Governing Saturated and Unsaturated FA Composition in Soybean, in *Lipid Biotechnology*, edited by T.M. Kuo and H.W. Gardner, Marcel Dekker Inc., New York, 2002.
7. Rebetzke, G.J., J.W. Burton, T.E. Carter, Jr., and R.F. Wilson, Genetic Variation for Modifiers Controlling Reduced Saturated FA Content in Soybean, *Crop Sci.* 38:303–308 (1998).
8. The USDA-ARS Soybean Genome Database, SoyBase, <http://soybase.org> (accessed July 2004).
9. Lander, E.S., and D. Botstein, Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps, *Genetics* 121:185–199 (1989).
10. Zeng, Z.-B., Theoretical Basis for Separation of Multiple Linked Gene Effects in Mapping Quantitative Trait Loci, *Proc. Natl. Acad. Sci. USA* 90:10972–10976 (1993).
11. Zeng, Z.-B., Precision Mapping of Quantitative Trait Loci, *Genetics* 136:1457–1468 (1994).
12. Sneller, C.H., Pedigree Analysis of Elite Soybean Lines, *Crop Sci.* 34:1515–1522 (1994).
13. Gizlice, Z., T.E. Carter, Jr., T.M. Gerig, and J.W. Burton, Genetic Diversity Patterns in North American Public Soybean Cultivars Based on Coefficient of Parentage, *Ibid.* 36:753–765 (1996).
14. Hyten, D., V. Pantalone, C. Sams, A. Saxton, D. Landau-Ellis, T. Stefaniak, and M. Schmidt, Seed Quality QTL in a Prominent Soybean Population, *Theor. Appl. Genet.* 109:552–561 (2004).
15. Song, Q.J., L.F. Marek, R.C. Shoemaker, K.G. Lark, V.C. Concibido, X. Delannay, J.E. Specht, and P.B. Cregan, A New Integrated Genetic Linkage Map of the Soybean, *Ibid.* 109:122–128 (2004).
16. Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M.J. Daly, S.E. Lincoln, and L. Newburg, MAPMAKER: An Interactive Computer Package for Constructing Primary Genetic Linkage Map of Experimental and Natural Populations, *Genomics* 1:174–181 (1987).
17. Lincoln, S.E., M.J. Daly, and E. Lander, Constructing Genetic Maps with MAPMAKER/EXP 3.0, 3rd edn., Whitehead Institute Technical Report, Cambridge, Massachusetts, 1992.
18. SAS/STAT User's Guide, Version 6, 4th edn., Vol. 1, SAS Institute, Inc., Cary, North Carolina, 1989.
19. Basten, C.J., B.S. Weir, and Z.-B. Zeng, Zmap—A QTL Cartographer, in *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production: Computing Strategies and Software*, edited by C. Smith, J.S. Gavora, B. Benkel, J. Chesnais, W. Fairfull, J.P. Gibson, B.W. Kennedy, and E.B. Burnside, Organizing Committee, 5th World Congress on Genetics Applied to Livestock Production, Guelph, Ontario, Canada, Vol. 22, 1994, pp. 65, 66.
20. Basten, C.J., B.S. Weir, and Z.-B. Zeng, QTL Cartographer, Version 1.15, Department of Statistics, North Carolina State University Raleigh, NC, 2001.
21. Churchill, G., and R. Doerge, Empirical Threshold Values for Quantitative Trait Mapping, *Genetics* 138:963–971 (1994).
22. Nyquist, W.E., Estimation of Heritability and Prediction of Selection Response in Plant Populations, *Crit. ev. Plant Sci.* 10(3):235–322 (1991).
23. Li, Z., R.F. Wilson, W.E. Rayford, and H.R. Boerma, Molecular Mapping Genes Conditioning Reduced Palmitic Acid Content in N86-2122-4 Soybean, *Crop Sci.* 42:373–378 (2002).
24. Diers, B.W., and R.C. Shoemaker, Restriction Fragment Length Polymorphism of Soybean FA Content, *J. Am. Oil Chem. Soc.* 69:1242–1244 (1992).
25. Rebetzke, G.J., V.R. Pantalone, J.W. Burton, T.E. Carter, Jr., and R.F. Wilson, Genetic Background and Environment Influence Palmitate Content of Soybean Seed Oil, *Crop Sci.* 41:1731–1736 (2001).
26. Spencer, M.M., V.R. Pantalone, E.J. Meyer, D. Landau-Ellis, and D.L. Hyten, Jr., Mapping the Fas Locus Controlling Stearic Acid Content in Soybean, *Theor. Appl. Genet.* 106:615–619 (2003).
27. Pantalone, V.R., R.F. Wilson, W.P. Novitzky, and J.W. Burton, Genetic Regulation of Elevated Stearic Acid Concentration in Soybean Oil, *J. Am. Oil Chem. Soc.* 79:549–553 (2002).

[Received August 16, 2004; accepted November 16, 2004]