Restriction Fragment Length Polymorphism Analysis of Soybean Fatty Acid Content[1]

B.W. Diers[a,*] and **R.C. Shoemaker**[b]
[a]Department of Agronomy, and [b]Departments of Agronomy and Genetics and USDA/ARS/FCR, Iowa State University, Ames, Iowa 50011

The quality of soybean [*Glycine max* (L.) Merr.] oil depends greatly upon its fatty acid composition. The objective of this study was to map quantitative trait loci affecting the relative content of the fatty acids composing soybean seed oil. The mapping was done in a population formed from a cross between a *G. max* experimental line and a *G. soja* plant introduction. Sixty F2-derived lines from this population were genotyped with 243 restriction fragment length polymorphism, five isozyme, one storage protein and three morphological markers. The genotyped lines were grown in a replicated trial, and the seed harvested was analyzed for palmitic, stearic, oleic, linoleic and linolenic fatty acids. Markers significantly ($P > 0.005$) associated with each fatty acid were found. Individual markers were associated with up to 38% of the total variance in specific fatty acids among the F2 lines. The greatest associations between markers and fatty acid content were observed with markers on two major linkage groups.

KEY WORDS: Fatty acids, restriction fragment length polymorphism markers, soybean, quantitative trait loci.

Soybean oil is one of the most important vegetable oils in the world (1). The fatty acid composition of vegetable oil is quite important in the determination of its quality. Oil from soybean seed is composed predominantly of the following five fatty acids: the saturated 16-carbon palmitate and 18-carbon stearate and the unsaturated 18-carbon oleate, linoleate and linolenate, which have one, two and three double bonds, respectively (2). It is thought that fatty acids are synthesized first by the addition of two carbons to palmitate to form stearate, and then by the desaturation of stearate to form oleate, linoleate and finally linolenate (3).

The manipulation of soybean oil quality by varying fatty acid composition has been an important soybean research emphasis (4). Although the fatty acid composition of oil is inherited quantitatively (5), several qualitative mutations result in major changes in the relative percentages of several fatty acids (6,7).

Restriction fragment length polymorphism (RFLP) markers are used extensively to create saturated linkage maps for crop plants. These saturated genetic maps provide the resolution necessary to map quantitative trait loci (QTL) (8). Maps containing hundreds of RFLP markers have been developed (9–11) and used to map QTL in soybean. Keim *et al.* (9,12) mapped QTL for hard seededness, maturity and morphological traits in an interspecific population formed from a cross between the cultivated soybean, *G. max*, and a wild relative and putative ancestor of *G. max*, *G. soja* (13). Diers *et al.* (11) also mapped QTL for seed protein and oil content in this population. We are using the same interspecific soybean population to map QTL associated with fatty acid composition of seed oil in this study.

MATERIALS AND METHODS

A population formed from a cross between the *G. max* experimental line A81-356022 and the *G. soja* accession PI 468916 was used in this study. Sixty F2 plants from this population were grown during the summer of 1987 near Ames, IA. Leaf samples were taken from each plant for DNA extraction and RFLP analysis. The plants were allowed to self-pollinate naturally, and at maturity each plant was harvested and threshed separately to form F2-derived lines.

The population was scored for 252 loci to construct the linkage map and to map the QTL. Two-hundred and twenty-seven loci were clones from a *PstI* genomic library from soybean (14), 16 loci were recombinant DNA clones from other labs, five loci were isozymes, three loci were morphological markers and one locus was a storage protein variant. Keim *et al.* (15) have described DNA extraction, Southern blot, and hybridization procedures used in the population. The software program Mapmaker (16) was used to construct a linkage map of the soybean genome from the F2 data (9,11).

The F2-derived lines in the F3 generation and parents were grown in a randomized complete block design experiment at three locations near Ames, IA, during the summer of 1988. There were two replications per location. The plots were single rows 1.5 m long with 1-m row spacing and a seeding rate of 33 seeds m^{-1}. Location 1 was planted 1 May; location 2, 15 May; and location 3, 29 May. At maturity, each plot was hand-harvested, and the seed was analyzed for fatty acid content on a plot basis.

Fatty acid content was measured on a 10-seed sample from each plot by gas chromatography. The analyses were conducted by E. G. Hammond and D. N. Duvick in the Department of Food Technology at Iowa State University. A hydraulic press at a pressure of approximately 36 kg $(cm^2)^{-1}$ was used to crush the seeds between two aluminum plates. Two five-seed samples from each plot were crushed in grooves 8 mm wide and 1.6 mm deep. The sample was transferred to a test tube and extracted approximately 18 h later with sufficient hexane to cover the seed. Approximately 0.1 mL of the hexane extract was transferred to a 1.5-mL autosampler vial and reacted with 0.5 mL of 1 M sodium methoxide solution in methanol for 30 min at 40°C, with gentle mixing for 10 min. Next, 0.8 mL of water was added, and after 3–5 min, the floating oil was dissolved in 0.5 mL hexane. The hexane layer was transferred to a clean vial and diluted to about 1.1 mL with additional hexane. The samples were analyzed in a Hewlett-Packard (Avondale, PA) 5890 gas chromatograph fitted with flame detectors and 15-M Durabond-23 capillary columns (J&W Scientific, Deerfield, IL), 0.25-mm i.d. and 0.25-μm film thickness, at 200°C. Percentages were calculated from electronically integrated peak areas cor-

[1]Joint contribution of North Central Region, USDA-ARS and Journal Paper no. J-14849 of the Iowa Agriculture and Home Economics Experiment Station, Ames, IA 50011; Project 2974.

*To whom correspondence should be addressed at Department of Crop and Soil Sciences, Plant and Soil Sciences Building, Michigan State University, East Lansing, MI 48824.

rected by factors based upon the number of C-H bonds. The reliability of these correction factors was verified daily by injection of a standard of known composition.

The fatty acid content data were analyzed by standard analysis of variance procedures for a randomized complete-block design model. Broad-sense heritabilities were calculated according to Fehr (17). Each marker-trait combination was analyzed to test for significant associations between markers and traits. The lines were divided into three classes for codominant markers (homozygous for *G. max* alleles, homozygous for *G. soja* alleles, and heterozygous) or into two classes for dominant markers (heterozygous class and homozygous dominant class contrasted with homozygous recessive class). An analysis of variance was used to determine whether a significant difference was present among the marker classes. R^2 values, or the proportions of total variance among lines that is explained by the segregation of markers, were calculated for significant markers.

RESULTS AND DISCUSSION

A RFLP map that encompassed 2,147 centimorgans was constructed using the markers (11). QTLs were then mapped by using these markers together with the quantitative data from the 1988 field trial. The trial included the parents of the population and F2-derived families. The parents had significantly different levels for each fatty acid (Table 1), and there were significant differences among the F2:3 lines for each fatty acid. The heritability for each fatty acid is given in Table 1.

Markers were significantly ($P < 0.005$) associated with each fatty acid (Table 2), and those that explained the most variation were concentrated on 'A' and 'B' linkage groups. Because of the great number of F tests used to find associations between markers and traits in this experiment, there is a great probability that some markers have been falsely declared significant. Such a probability for each trait is approximately 0.72 for $P < 0.005$ but is reduced to 0.025 for a $P < 0.0001$, which is the significance level of those markers most associated with fatty acid content.

The markers on the 'A' and 'B' linkage groups were associated with contrasting effects. The *G. max* alleles at significant loci in the 'A' linkage group were associated with greater content of oleate and lesser content of linoleate and linolenate than were *G. soja* alleles. The *G. max* alleles in the 'B' group, however, were associated with lesser content of oleate and greater content of linoleate than were *G. soja* alleles. Because the alleles in each linkage group had contrasting effects on oleate and linoleate contents, it is likely that the significant loci in groups 'A' and 'B' are linked to genes directly associated with oleate or linoleate synthesis. The effect of these loci on other fatty acids thus would be a secondary consequence of the primary effect. Although several mutants at individual loci that substantially alter the content of fatty acids in seed have been reported (6,7), many genes are involved in fatty acid synthesis (3). The significant markers may be linked to any of these genes.

The RFLP mapping allowed us to map two regions having a major effect on fatty acid composition. The analysis provides information regarding the steps on the fatty acid synthesis pathway that the mapped QTL affect. Analysis does not, however, provide information regarding which gene involved in fatty acid synthesis has been mapped. Further genetic analysis of these regions should provide additional information about the control of seed composition in soybean. As genes involved in fatty acid synthesis are cloned, we hope to map them to determine if they cosegregate QTL associated with fatty acid content. This

TABLE 1

Mean Performance of *G. max* and *G. soja* Parents, and the Broad-Sense Heritabilities of the Population

Trait	Mean performance: *G. max* A81-356022	Mean performance: *G. soja* PI468.916	Heritability
	g(kg oil)$^{-1}$		
Palmitate	102[a]	112	0.85
Stearate	52	41	0.86
Oleate	217	173	0.86
Linoleate	544	528	0.88
Linolenate	84	146	0.86

[a]Values for all traits are significantly different between parents ($P < 0.001$).

TABLE 2

Markers Significantly ($P < 0.005$) Associated with Variation for Seed Traits

Marker	R^2	P > F	Means of genotypic classes: MM[a]	SM	SS	Linkage group
			g (kg oil)$^{-1}$			
Palmitate						
pA-343a[b]	0.24	0.001	112	114	109	M
pA-18	0.19	0.005	115	111	111	J
pK-375	0.18	0.005	111	111	115	Q
Stearate						
pA-233	0.19	0.005	42	44	46	Q
Oleate						
pA-82	0.28	0.0001	176	189	199	B
pA-104	0.26	0.0003	175	188	198	B
pA-170	0.23	0.001	177	188	198	B
pb[c]	0.21	0.001	192	191	175	A
pA-242b	0.20	0.002	192	191	176	A
pA-619	0.19	0.004	179	193	187	J
Linoleate						
pA-82	0.38	0.0001	558	546	534	B
pA-104	0.33	0.0001	559	547	536	B
pA-170	0.30	0.0001	557	547	535	B
pA-242b	0.21	0.002	545	543	557	A
pA-118	0.20	0.006	537	549	550	H
pb[c]	0.20	0.002	544	543	557	A
Linolenate						
pSAC-7a[d]	0.31	0.0005	104	110	116	A
pA-23	0.26	0.0005	106	111	115	A
pA-242b	0.23	0.0008	106	111	115	A
pA-203	0.22	0.002	108	109	116	A
pA-454	0.22	0.001	106	110	115	A
pA-65	0.20	0.003	111	113	105	G
pK-229	0.20	0.003	107	110	116	A

[a]MM designates homozygous *G. max* class; SM, heterozygous class; and SS, homozygous *G. soja*.
[b]Markers labelled pA or pK were developed at Iowa State University from a *Pst*I library.
[c]Morphological marker blunt-sharp pubescence tip (ref. 19).
[d]Actin gene probe provided by Richard Meagher (University of Georgia).

would provide information whether the QTL are actually the fatty acid synthesis genes. Also, interesting regions such as linkage group 'A', which is significant in terms of unsaturated fatty acids and total protein and oil content (11), could be studied by the construction of near isogenic lines (NILs) having different sections of the linkage group. Such an analysis could provide information about whether the regions significant for more than one trait are the result of gene linkage or of pleiotropy. The NILs could then be used to pinpoint and characterize the region in which important genes are located (18).

ACKNOWLEDGMENTS

This work was supported in part by Iowa State Biotechnology Grant No. 480-46-09 and by the Iowa Soybean Promotion Board.

REFERENCES

1. Smith, K.J., and W. Huyser, in *Soybean: Improvement, Production and Uses*, 2nd edn., edited by J.R. Wilcox, The American Society of Agronomy, Madison, 1987, No. 16, pp. 1–23.
2. Wilson, R.F., *Ibid. 16*:643 (1987).
3. Slack, C.R., and J.A. Browse, in *Seed Physiology. Vol. 1. Development*, edited by D.R. Murray, Academic Press Australia, North Ryde, N.S.W, 1984, pp. 209–244.
4. Mounts, T.L., K. Warner, G.R. List, R. Kleiman, W.R. Fehr, E.G. Hammond and J.R. Wilcox, *J. Am. Oil Chem. Soc. 65*:624 (1988).
5. Howell, R.W., C.A. Brim and R.W. Rinne, *Ibid. 49*:30 (1972).
6. Nickell, A.D., J.R. Wilcox and J.F. Cavins, *Crop Sci. 31*:1169 (1991).
7. Graef, G.L., W.R. Fehr, L.A. Miller, E.G. Hammond and S.R. Cianzio, *Ibid. 28*:55 (1988).
8. Lander, E.S., and D. Botstein, *Genetics 121*:185 (1989).
9. Keim, P, B.W. Diers, T. Olson and R.C. Shoemaker, *Ibid. 126*:735 (1990).
10. Tingey, S., S. Sebastian and J.A. Rafalski, in *Agronomy Abstracts*, ASA Madison, p. 180, 1989.
11. Diers, B.W., P. Keim, W.R. Fehr and R.C. Shoemaker, *Theor. Appl. Genet. 83*:608 (1992).
12. Keim, P., B.W. Diers and R.C. Shoemaker, *Ibid. 79*:465 (1990).
13. Hymowitz, T., and R.J. Singh, in *Soybean: Improvement, Production and Uses*, 2nd edn., edited by J.R. Wilcox, The American Society of Agronomy, Madison, 1987, No. 16, pp. 1–23.
14. Keim, P., and R.C. Shoemaker, *Soybean Genet. Newsl. 15*:147 (1988).
15. Keim, P., R.C. Shoemaker and R.G. Palmer, *Theor. Appl. Genet. 77*:786 (1989).
16. Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M.J. Daly, S.E. Lincoln and L. Newburg, *Genomics 1*:174 (1987).
17. Fehr, W.R., *Principles of Cultivar Development*, Macmillan Publishing Co., New York, 1987, p. 98.
18. Paterson, A.H., J.W. DeVerna, B. Lanini and S.D. Tanksley, *Genetics 124*:735 (1990).
19. Palmer, R.G., and T.C. Kilen, in *Agronomy 16*:135 (1987).

[Received June 3, 1992; accepted October 15, 1992]