Theor Appl Genet (2003) 106:615–619
DOI 10.1007/s00122-002-1086-y

M. M. Spencer · V. R. Pantalone · E. J. Meyer
D. Landau-Ellis · D. L. Hyten Jr

Mapping the *Fas* locus controlling stearic acid content in soybean

Received: 26 January 2002 / Accepted: 8 July 2002 / Published online: 22 October 2002

Abstract Increasing the stearic acid content to improve soybean [*Glycine max* (L) Merr] oil quality is a desirable breeding objective for food-processing applications. Although a saturated fatty acid, stearic acid has been shown to reduce total levels of blood cholesterol and offers the potential for the production of solid fat products (such as margarine) without hydrogenation. This would result in the reduction of the level of *trans* fat in food products and alleviate some current health concerns. A segregating F_2 population was developed from the cross between Dare, a normal stearic acid content cultivar, and FAM94-41, a high stearic acid content line. This population was used to assess linkage between the *Fas* locus and simple sequence repeat (SSR) markers. Three SSR markers, Satt070, Satt474 and Satt556, were identified to be associated with stearic acid ($P < 0.0001$, $r^2 > 0.61$). A linkage map consisting of the three SSR markers and the *Fas* locus was then constructed in map order, *Fas*, Satt070, Satt474 and Satt556, with a LOD score of 3.0. Identification of these markers may be useful in molecular marker-assisted breeding programs targeting modifications in soybean fatty acids.

Keywords Simple sequence repeats · *Glycine max* · Genetic linkage map

Introduction

Over the past decades soybean [*Glycine max* (L) Merr] has become an important crop because of its varied usefulness in an array of products, including pharmaceuticals, livestock feed, and industrial and food products. Products containing soybean have, for example, an ability to lower serum cholesterol levels (Kito et al. 1993). Soy products such as tofu, soymilk, tempeh, flour, concentrates, isolates, textured protein and oil are increasingly popular as healthy foods (Liu 1997). Improvements in nutritional value and functional properties of soybean proteins and oils is a major objective in the food-processing industry (Utsumi et al. 1997). This provides breeders with an incentive and a challenge to optimize the composition and concentration of soybean protein and oil.

Soybean oil accounts for about 30% of the world's vegetable oil market (Rebetzke et al. 1998). The average fatty acid content of commercial soybean oil is 11% palmitic, 4% stearic, 24% oleic, 54% linoleic and 7% linolenic (Hui 1996). Soybean genotypes with modified fatty acids, including stearic acid, are useful for certain food and industrial products.

Soybean breeders would benefit by having the potential to increase stearic acid, because approximately 52% of USA soybean oil is utilized for the production of various 'baking fats' such as margarines and shortenings. The solid-fat content of a vegetable oil determines its suitability for use in the production of baking fats. A common industry parameter is the solid-fat index (SFI), which is directly related to the concentration of saturated fatty acids in vegetable oil. The process of hydrogenation increases the SFI of soybean oil in order to make it suitable for use as margarine or for shortening. However, hydrogenation also forms *trans* fatty acids. Public concerns regarding *trans* fatty acids has prompted the U.S. Food and Drug Administration (FDA) to propose a regulation (U.S. FDA 1999) which will require food manufacturers to list the levels of *trans* fatty acids on food labels. Hence, vegetable oil processors are actively seeking to find ways to minimize hydrogenation, while maintaining quality products. A soybean oil, which contains approximately 30% total saturates, can make a suitable *trans*-free margarine through the process of interestification (List et al. 2000). Palmitic acid and

Communicated by C. Möllers

M.M. Spencer (✉) · V.R. Pantalone · E.J. Meyer
D. Landau-Ellis · D.L. Hyten Jr
The University of Tennessee,
Department of Plant Sciences and Landscape Systems,
252 Ellington Plant Science Building, 2431 Center Drive,
Knoxville, TN 37996-4561, USA
e-mail: mspence1@utk.edu
Tel.: (+1-865)-974-0375, Fax: (+1-865)-974-7997

stearic acid are the two predominant saturated fatty acids found in soybean oil. Unlike palmitic acid, stearic acid has been show to either reduce or to have no effect on serum cholesterol levels in humans (Emken 1994; Grundy 1994; Kris-Etherton and Yu 1997). Identification of novel sources of soybean with an enhanced stearic acid content will help to meet the needs of consumers and the oil-processing industry.

Stearic acid concentration in soybean may be genetically altered by mutations at the *Fas* locus. Nearly all the variants currently known have been induced by chemical or X-ray mutagenesis. Five soybean germplasm lines are reported to carry modified stearic acid alleles: *fas*^{a} [A6, (Hammond and Fehr 1983)], *fas*^{b} [FA41545 (Graef et al. 1985a)], *fas* [A81-606085, (Graef et al. 1985b)], st_1 [KK-2 (Rahman et al. 1997)] or st_2 [M25, (Rahman et al. 1997)]. As reported by Graef et al. (1985b), *fas*^{a} (30% stearic acid), *fas*^{b} (15% stearic acid) and *fas* (19% stearic acid) are allelic and represent different mutations in the same gene.

In N-nitroso-N-methyurea (NMU)-treated sunflower (*Helianthus annuus* L.), two additive alleles designated fas_2 and fas_x have been described for determining lower stearic acid concentration (Miller and Vick 1999). In ethylmethanesulfonate (EMS)-treated Arabidopsis two alleles influencing stearic acid concentration have been observed (Lightner et al. 1997). These alleles represent mutations at loci designated *Fab* and are identified as $Fab_{2\text{-}1}$ (20% stearic acid) and $Fab_{2\text{-}2}$ (6% stearic acid). Another property of these alleles in Arabidopsis is to cause up to a 5-fold reduction in plant size. Interestingly, *fas*^{a}, *fas*^{b} and *fas* alleles also severely depressed yield in soybean (Lundeen et al. 1987; Hartman et al. 1997). This decrease in yield may be one of the reasons for the slow progress in the development of acceptable commercial cultivars with higher stearic acid content.

A newly developed soybean line, FAM94-41, carrying a natural mutation, fas_{nc}, has been recently described (Pantalone et al. 2002). FAM94-41 is an agronomically robust high stearic acid line, which may become useful in reversing the apparent yield depression of material developed with other *fas* alleles.

An initial step toward the testing of this hypothesis would be the identification of molecular markers closely associated with stearic concentration. Such markers have not yet been reported in soybean and could enable further marker-assisted breeding for rapid genetic gain. In this report, we identified and mapped SSR markers (Akkaya et al. 1992; Cregan and Quigley 1998; Cregan et al. 1999) linked to the *Fas* locus using bulk segregant analysis (Giovannoni et al. 1991; Michelmore et al. 1991).

Materials and methods

Plant material

The genetic material used consisted of F_2 and $F_{2:3}$ individuals from the cross between the parental lines Dare (normal stearic acid content approximately 4%) and FAM94-41 (high stearic acid content approximately 7%). The cross was made at the University of Tennessee in the summer of 1998. The F_1 was grown at the Knoxville Experiment Station in Knoxville, Tenn. in 1999. F_2 individuals were grown at the Knoxville Experiment Station in Knoxville, Tenn., in 2000.

Gas chromatography

Fatty acid analyses were performed using gas chromatography, as described by Hammond, (1991). Fatty acids were extracted by soaking crushed seed chips with 0.5 ml of a mixture of chloroform:hexane:methanol (8:5:2 v/v/v) for a minimum of 4 h. Extracts were subsequently transferred to a 1.5-ml autosampler vial, and 75 µl of methylating reagent [methoxyde/methanol: petroleum ether: ethyl ether (1: 4: 2 v/v/v/)] were added and the vial tightly capped with a crimper.

Composition of palmitic, stearic, oleic, linoleic and linolenic acids were determined with a Hewlett-Packard model HP 6890 gas chromatograph (Palo Alto, Calif.) equipped with a model 7673 autosampler, a flame ionization detector, and an immobilized 30 × 0.53-mm inner diameter Alltech AT-Silar capillary column with an 0.5 µm fused stationary phase. Operating conditions were as follows: carrier, Helium (20 ml/min); 20:1(v/v) split injection; injection temperature 250 °C; detector temperature 275 °C, and column temperature 230 °C. The RM-1 standard (suitable for measuring soybean oil) was used to calibrate and determine the relative fatty acid content of each of the experimental samples.

Bulked segregant analysis

Chipped F_2 and $F_{2:3}$ seeds from the cross Dare × FAM94-41 were used to determine fatty acid composition. Twelve single seeds from each extreme were identified and germinated. DNA was isolated from young F_2 and $F_{2:3}$ leaves according to Fulton et al. (1995), or using the QIAGEN DNeasy Plant Mini Kit (QIAGEN Inc. Valencia, Calif.). Aliquots of 2 µg of DNA from each individual were bulked together into two groups: with a high and normal content of stearic acid. Each combined mixture was then diluted to a final concentration of 20 ng/µl (Giovannoni et al. 1991; Michelmore et al. 1991). The DNA bulks were screened with 85 random SSR markers distributed throughout the 20 molecular linkage groups, and 16 other SSR markers which mapped to molecular linkage group (MLG) B2 (Cregan et al. 1999).

PCR amplification and analysis

PCR conditions in a final volume of 12 µl containing 20 ng/µl of template DNA, were: 1 µl of × 1 PCR-2 Klentaq buffer (Ab Peptides, Inc. St Louis, Mo.); 0.2 mM of each deoxynucleotide triphosphate (USB Corp., Cleveland, Ohio); 1.0 µM of each forward and reverse SSR primer (SIGMA Genosys, The Woodlands, Tex.); and 0.5 Units of Klen*Taq*1 polymerase (Ab Peptides, Inc. St Louis, Mo.). PCR reactions were carried out in a Hybaid multiblock thermocycler (CLP, San Diego, Calif.) with the following profile: (1) 94 °C for 3 min × 1 cycle; (2) 94 °C for 25 s, 47 °C for 25 s and 72 °C for 1 min × 35 cycles; and finally (3) 72 °C for 5 min × 1 cycle. PCR products were then separated by electrophoresis using a 6% non-denaturing polyacrylamide gel containing ethidium bromide dye for visualization. Loci polymorphic between the bulks were confirmed to correlate with the polymorphisms between the parents, Dare and FAM94-41, and were used to screen F_2 individuals in order to determine segregation ratios and map order.

Statistical analyses and mapping

Stearic acid F_2 genotypic classes were constructed based upon two standard deviations about the parental means (Narvel et al. 2000; Rahman et al. 2001; Primono et al. 2002). Molecular marker F_2 genotypic classes for Satt070, Satt556 and Satt474 were constructed based upon the observed DNA banding pattern of the parents (a random line that expressed the same allele as FAM94-41 for that specific SSR locus was designated as an *FAM94-41 type*; a random line that expressed the same allele as Dare for that specific SSR locus was designated as a *Dare type*; a random line that expressed both parental alleles at that specific SSR locus was designated as a *heterozygous type*. Chi-square analyses were performed to determine goodness of fit with expected 1:2:1 F_2 genotypic class ratios.

Data were analyzed using SAS (SAS Institute Inc., 1998, Cary. N.C.) to establish associations between the trait and the SSR markers. Regression of stearic acid content in the F_2 population was used to establish the linkage between the *Fas* locus and the markers. The Mapmaker program (Lander et al. 1987) was used to develop a linkage map for the *Fas* locus in our population. The most-likely order was determined using a LOD threshold of 3.0.

Results and discussion

Fatty acid analysis of F_2 seeds of the population from the cross between Dare and FAM94-41 demonstrated the existence of three stearic acid groups: one of normal content (stearic acid ≤ 4%), one of high content (stearic acid > 6%), and a third (intermediate) group formed by individuals whose stearic acid content was between 4 and 6% (Fig. 1). These stearic acid classes fit χ^2 goodness of fit criteria for 1:2:1 genotypic class ratios (Table 1). This is consistent with the observation that an elevated stearic acid in FAM94-41 is governed by an alternative allele showing single gene inheritance (Pantalone et al. 2002).

From the total of 101 SSR primers tested, three (satt070, satt474 and satt556) showed strong and consistent polymorphisms between the two extreme bulks of F_2 individuals, indicative of linkage with the *Fas* locus. These three SSR markers showed the same corresponding pattern of polymorphisms between the two parents: for each marker, the banding pattern for the normal stearic acid content bulk corresponded to the pattern in Dare; the banding pattern for the high stearic aid content bulk corresponded to the one in FAM94-41. DNA from bulks of normal or high stearic content in the $F_{2:3}$ individuals also exhibited the corresponding banding patterns, indicative of hertitable transmission of this trait. (Fig. 2).

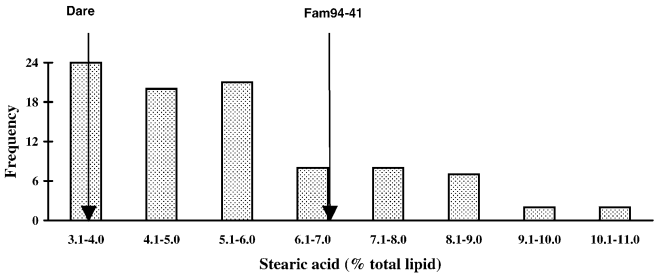

Fig. 1 Frequency distribution of stearic acid concentration of F_2 individuals from the cross between the normal soybean cultivar, Dare, and the elevated stearic acid line, FAM94-41

Results presented in Table 1 suggested a strong association between each of the markers and the *Fas* locus ($r^2 \geq 61\%$). The *P* value for additive genetic effects was highly significant ($P < 0.0001$). Our report is the first to identify SSR markers on MLG B2, which are associated with the *Fas* locus, governing stearic acid concentration.

Rennie and Tanner (1989) previously showed that *Fas* and *Fan* loci are linked at a distance of 21.6 ± 1.7 map units, and also that *Fan* and Idh2 are linked at a distance of 24.6 ± 2.6, 25.2 ± 3.5 and 24.3 ± 3.4, 24.7 ± 3.5 map

Table 1 Markers significantly associated with variation in 18:0 content in an F_2 soybean population of Dare × FAM94-41

Item	18:0 Genotypic class[a]	D[b]	Satt070 Genotypic class[a]	D[b]	Satt556 Genotypic class[a]	D[b]	Satt474 Genotypic class[a]	D[b]
FAM94-41 type (18:0 > 6%)	24	0.04	24	0.23	23	0.04	22	0.41
Heterozygous type (4% < 18:0 ≤ 6%)	41	0.54	39	0.46	36	1.45	39	0.57
Dare type (18:0 ≤ 4%)	27	0.70	24	0.23	29	0.04	24	0.18
Total[c]	92	1.28	87	0.92	88	1.53	88	1.16
R^2			0.67		0.65		0.61	
Additive genetic effect			$P < 0.0001$		$P < 0.0001$		$P < 0.0001$	
Dominance genetic effect			$P < 0.05$		NS		$P < 0.05$	

[a] 18:0 (stearic acid) genotypic classes constructed based upon two standard deviations about the parental means; Satt070, Satt556 and Satt474 genotypic classes constructed based upon observed DNA banding pattern of the parents (a random F_2 individual that expressed the same allele as FAM94-41 for that specific SSR locus was designated as an *FAM94-41 type*; a random F_2 individual that expressed the same allele as Dare for that specific SSR locus was designated as a *Dare type*; a random F_2 individual that expressed both parental alleles at that specific SSR locus was designated as a *heterozygous type*)

[b] D = Squared deviation from expectation of a 1:2:1 F_2 genotypic class ratio

[c] Total number of individuals analyzed reflects missing data for individuals whose bands were not scoreable for a specific marker; total sum of squared deviations which are less than the tabled value of $\chi^2_{(0.05,\ 2\ df)} = 5.99$ indicate that the data fit the expected 1:2:1 F_2 genotypic class ratio

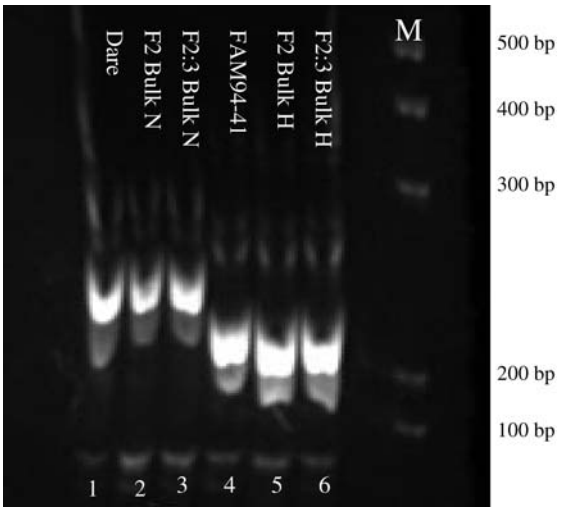

Fig. 2 SSR polymorphism patterns for marker Satt474. *Lane 1* **Dare**; *lanes 2 and 3* $\mathbf{F_2}$ and $\mathbf{F_{2:3}}$ DNA pools for normal stearic acid content. *Lane 4* **FAM94-41**; *lanes 5 and 6* $\mathbf{F_2}$ and $\mathbf{F_{2:3}}$ DNA pools for high stearic acid content. **M** represents a molecular-size standard

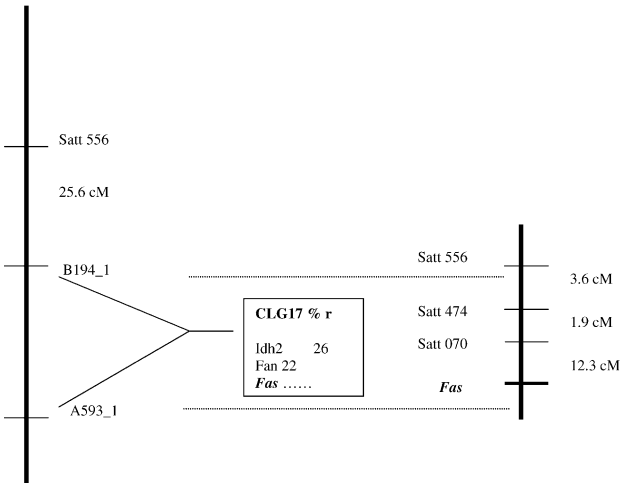

Fig. 3 Molecular map showing the likely position of the *Fas* locus and three SSR markers (Satt070, Satt474 and Satt556) compared to other markers on linkage group B2. The *vertical bar* to the right shows data from this study and represents the positions of the markers and the *Fas* locus at 3.0 LOD using Mapmaker. The *map* on the far left is a simplified map of USDA/Iowa St. University (Cregan et al. 1999), the center is the classical map for the *Fas*, *Fan* and Idh2 loci (Rennie and Tanner 1989). *Dashed lines* represent the presumed location of the *Fas* locus according to data in our study

units for the crosses: Maple Arrow × PI 361088B, A5 × N78-2245, N78-2245 × A5 and N78-2245 × C1640. These authors established that the genes *Fan*, Idh2 and *Fas* were placed on the same classical linkage group 17 and suggested the order *Fas*, *Fan* and Idh2 (Rennie and Tanner 1989).

Brummer et al. (1995) placed the *Fan* locus close to RFLP markers pB194-1 and pB124 on MLG B2, using a segregating population from the cross between C1640 and Zucc. PI 479.450. Since no molecular markers linked to the *Fas* locus have yet been reported, the three SSR markers that we identified are an important finding.

To assess the most-likely position and order of these markers on MLG B2, we used Mapmaker to determine the likelihood estimates for possible linkage. Results showed that Satt070, Satt474 and Satt556 are on the same linkage group as the *Fas* locus, with a LOD score of 3.0. The most-likely order is *Fas*, Satt070, Satt474 and Satt556. The distance between the *Fas* locus and the closest marker, Satt070, was 12.3 cM; Satt474 and Satt556 were found at distances of 14.2 and 17.8 cM, respectively, from the *Fas* locus (Fig. 3). The order between the three SSR markers was consistent with that proposed by the University of Utah (Cregan et al. 1999) and also reported on the latest map from Cregan (personal communication), but differed from the map developed by the University of Nebraska. The observation that all SSR markers located south of Satt070 generated no polymorphisms between the parents (data not shown) did not allow more precise mapping. Nevertheless, it is known that sample size and the number of available markers per linkage group may influence genetic distances and order. Further investigations using other marker types such as RFLP, AFLP or SNPs, for example, may be required to determine a higher density map of the *Fas* locus. These markers would be useful for recognizing high stearic acid lines before flowering and therefore be valuable for determining desirable parents in a molecular-assisted breeding program.

The information in this study establishes a foundation whereby map-based markers can be used to assist selection for stearic acid concentration in soybean. Further research using different populations from more genetically distant parents may provide a more-precise location for the *Fas* locus and other genes involved in soybean fatty acid modification.

Acknowledgements The Tennessee Soybean Promotion Board is gratefully acknowledged for partially supporting this work.

References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131–1139

Brummer EC, Nickell AD, Wilcox Jr, Shoemaker RC (1995) Mapping the Fan locus controlling Linolenic acid content in soybean oil. J Hered 86:245–247

Burton JW, Carter TE Jr, Huie EB (1994) Registration of 'Brim' soybean. Crop Sci 34:301

Cregan PB, Quigley CV (1998) Simple sequence repeat DNA marker analysis. In: Caetano-Anolès G, Gresshoff PM (eds) DNA markers. Protocol, applications and overviews. Wiley-VCH, New York, pp 173–186

Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, Specht JE (1999) An integrated genetic linkage map of the soybean genome. Crop Sci 39:1464–1490

Emken EA (1994) Metabolism of dietary stearic acid relative to other fatty acids in human subjects. Am J Clin Nutr 60:1023S–1028S

Fulton TM, Chungwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and herbaceous plants. Plant Mol Biol 13:207–209

Giovannoni JJ, Wing RA, Ganal MW, Tanksley SD (1991) Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping population. Nucleic Acids Res 19:6553–6558

Graef GL, Hammond EG, Fehr WR (1985a) Inheritance of three stearic acid mutants of soybean. Crop Sci 25:1076–1079

Graef GL, Miller LA, Fehr WR, Hammond EG (1985b) Fatty acid development in a soybean mutant with high stearic acid. J Am Oil Chem Soc 62:773–775

Grundy SM (1994) The influence of stearic acid on cholesterol metabolism relative to other long-chain fatty acids. Am J Clin Nutr 60:986S–990S

Hammond EG (1991) Organization of rapid analysis of lipids in many individual plants. In: Linskins HF, Jackson JF (eds) Essential oils and waxes modern methods of plant analysis. Berlin, pp 321–330

Hammond EG, Fehr WR (1983) Registration of A6 germplasm line of soybean. Crop Sci 23:192–193

Hartmann RB, Fehr WR, Welke GA, Hammond EG, Duvick DN, Cianzo SR (1997) Association of elevated stearate with agronomic and seed traits of soybean. Ibid 37:124–127

Hui YH (1996) Edible oil and fat products: oils and oil seeds. Soybean Oil. In: Bailey's industrial oil and fat products, 5th edn. Wiley, New York, pp 497–601

Kito M, Moriyama S, Kimura Y, Kambara H (1993) Changes in plasma lipid levels in young healthy volunteers by adding an extruder-cooked soy-protein to conventional meals. Biosci Biotech Biochem 57:534–538

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

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage map of experimental and natural populations. Genomics 1:174–181

Lightner J, Lark E, James D, Browse J (1997) Novel mutations affecting leaf stearate content and plant size in Arabidopsis. Theor Appl Genet 94:975–981

List GR, Steidley KR, Neff WE (2000) Commercial spreads formulation, structure and properties. INFORM 11:980–986

Liu KeShun (1997) Soybeans: chemistry, technology and utilization. International Thompson Publishing, New York, pp 34–35

Lundeen PO, Fehr WR, Hammond EG, Cianzo SR (1987) Association of alleles for high stearic acid with agronomic characters of soybean. Crop Sci 27:1102–1105

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Miller JF, Vick BA (1999) Inheritance of reduced stearic and palmitic acid content in sunflower seed oil. Ibid 39:364–367

Narvel JM, Fehr WR, Ininda J, Welke GA, Hammond EG, Duvick DN, Cianzio SR (2000) Inheritance of elevated palmitate in soybean seed oil. Crop Sci 40:635–639

Pantalone VR, Wilson RF, Novitzki WP, Burton JW (2002) Genetic regulation of elevated stearic acid concentration in soybean oil. J Am Oil Chem Soc 79:549–553

Primono VS, Falk DE, Ablett GR, Tanner JW, Rajcan I (2002) Inheritance and interaction of low palmitic and low linolenic soybean. Crop Sci 42:31–36

Rahman S, Kinoshita T, Anai T, Takagi Y (2001) Combining ability in loci for high oleic and low linolenic acids in soybean. Crop Sci 41:26–29

Rahman SM, Takagi Y, Kinoshita T (1997) Genetic control of high stearic acid content in seed oil of two soybean mutants. Theor Appl Genet 95:772–776

Rebetzke GE, Burton JW, Carter TE, Wilson RF (1998) Changes in agronomic and seed characteristics with selection for reduced palmitic acid content in soybean. Crop Sci 38:297–302

Rennie BD, Tanner JW (1989) Mapping a second fatty acid locus to soybean linkage group 17. Crop Sci 29:1081–1083

US Food and Drug Administration (1999) FDA proposes new rules for trans fatty acids in nutrition labeling, nutrient content claims and health claims. Federal Register 12 Nov 1999

USDA, ARS National Genetics Resources Program. Germplasm resources information network online database, National Germplam Resources Laboratory, Beltsville Md. Available: http://www.ars-grin.gov/

Utsumi S, Katsube T, Ishige T, Takaiwa F (1997) Molecular design of soybean glycinins with enhanced food qualities and development of crops producing such glicinins. In: Damodaran (ed) Food proteins and lipids. Lightner et al. 1997. Plenum Press, New York, pp 1–15