

Mapping the low palmitate *fap1* mutation and validation of its effects in soybean oil and agronomic traits in three soybean populations

Andrea J. Cardinal · Rebecca Whetten · Sanbao Wang · Jérôme Auclair · David Hyten · Perry Cregan · Eleni Bachlava · Jason Gillman · Martha Ramirez · Ralph Dewey · Greg Upchurch · Lilian Miranda · Joseph W. Burton

Received: 2 May 2013 / Accepted: 22 September 2013 / Published online: 17 October 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract

Key message *fap1* mutation is caused by a G174A change in *GmKASIII*A that disrupts a donor splice site recognition and creates a GATCTG motif that enhanced its expression.

Abstract Soybean oil with reduced palmitic acid content is desirable to reduce the health risks associated with consumption of this fatty acid. The objectives of this study were: to identify the genomic location of the reduced palmitate *fap1* mutation, determine its molecular basis, estimate the amount of phenotypic variation in fatty acid composition explained by this locus, determine if there are epistatic interactions between the *fap1* and *fap_{nc}* loci and, determine if the *fap1* mutation has pleiotropic effects on seed yield, oil and protein content in three soybean populations. This study detected two major QTL for 16:0

content located in chromosome 5 (*GmFATB1a*, *fap_{nc}*) and chromosome 9 near BARCSOYSSR_09_1707 that explained, with their interaction, 66–94 % of the variation in 16:0 content in the three populations. Sequencing results of a putative candidate gene, *GmKASIII*A, revealed a single unique polymorphism in the germplasm line C1726, which was predicted to disrupt the donor splice site recognition between exon one and intron one and produce a truncated *KASIII*A protein. This G to A change also created the GATCTG motif that enhanced gene expression of the mutated *GmKASIII*A gene. Lines homozygous for the *GmKASIII*A mutation (*fap1*) had a significant reduction in 16:0, 18:0, and oil content; and an increase in unsaturated fatty acids content. There were significant epistatic interactions between *GmKASIII*A (*fap1*) and *fap_{nc}* for 16:0 and oil contents, and seed yield in two populations. In conclusion, the *fap1* phenotype is caused by a single unique SNP in the *GmKASIII*A gene.

Communicated by I. Rajcan.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-013-2204-8) contains supplementary material, which is available to authorized users.

A. J. Cardinal (✉) · R. Whetten · S. Wang · R. Dewey
Department of Crop Science, North Carolina State University,
Raleigh, NC 27695-7620, USA
e-mail: andrea.cardinal@syngenta.com

J. Auclair
La Coop fédérée, Montreal, QC, Canada

D. Hyten · P. Cregan
Soybean Genomics and Improvement Laboratory,
U.S. Department of Agriculture, Agricultural
Research Service, Beltsville, MD 20705, USA

D. Hyten
DuPont Pioneer, P.O. Box 1000, Johnston, IA 50131-0184, USA

Abbreviations

16:0 Palmitic acid
18:0 Stearic acid

E. Bachlava
Monsanto Vegetable Seeds, 37437 State Hwy 16,
Woodland, CA 95695, USA

J. Gillman
Plant Genetics Research Unit, USDA-ARS,
University of Missouri, Room 204, Curtis Hall,
Columbia, MO 65211, USA

M. Ramirez · G. Upchurch · L. Miranda · J. W. Burton
Soybean and Nitrogen Fixation Research Unit,
USDA-ARS, Raleigh, NC 27607, USA

18:1	Oleic acid
18:2	Linoleic acid
18:3	Linolenic acid
BLUP	Best linear unbiased predictor
DAP	Days after planting
FATB	16:0-ACP thioesterase
KASIII	3-Ketoacyl-ACP synthase enzyme III
R2	Flowering date
R8	Maturity date

Introduction

Soybean seed oil quality is determined by the relative composition of fatty acids in triacylglycerols. Palmitic acid (16:0) is the predominant saturated fatty acid in soybean oil. Oil from typical soybean cultivars contains 110–120 g kg⁻¹ palmitic acid (16:0) (Burton et al. 1994b; USDA-ARS, National Genetics Resources Program). Oils with reduced palmitic acid content are desirable to reduce the risks of coronary diseases and, breast, colon, and prostate cancer incidence associated with consumption of this fatty acid (Hu et al. 1997; Henderson 1991).

Soybean oil quality can be improved through genetics. Several lines with reduced palmitic acid concentration were developed by ethyl-methyl sulfonate mutagenesis (C1726, homozygous for the *fap1* allele and ELLP2, homozygous for the *fap** or *fapx* alleles); *N*-nitroso-*N*-methyl urea mutagenesis (A22, homozygous for the *fap3* allele); and X-radiation (J3 homozygous for the *sop1* allele) (Fehr et al. 1991; Stojšin et al. 1998; Wilcox and Cavins 1990; Rahman et al. 1996). In addition, one natural mutation (*fap_{nc}*) was discovered in line N79-2077-11 and its sister line N79-2077-12, both derived from the fifth cycle of a recurrent selection program for high oleic concentration (Burton et al. 1983; Burton et al. 1994b). Among the five alleles conferring low palmitic acid in these lines, three are at independent loci (*fap1*, *fap3*, and *fap**) (Schnebly et al. 1994; Stojšin et al. 1998; Wilcox et al. 1994; Kinoshita et al. 1998; Primomo 2000; Primomo et al. 2002). *fap_{nc}* is allelic to *fap3* (Primomo 2000; Primomo et al. 2002). *sop1* is independent of the *fap1* locus, but its allelic relationship to the other loci is unknown (Kinoshita et al. 1998). In addition to the major genes mentioned above, minor/modifier genes influence the low palmitic acid content in soybean oil (Horejsi et al. 1994; Li et al. 2002; Rebetzke et al. 1998b, Rebetzke et al. 2001; Stojšin et al. 1998).

Several steps in the TAG biosynthetic pathway could affect palmitate concentration. In the case of low palmitic acid mutants, changes in activity or substrate specificity in 3-ketoacyl-ACP synthase II (KAS II), 16:0-ACP thioesterase,

lysophosphatidic acid acyltransferase, glycerol-3-phosphate acyltransferase, or diacylglycerol acyl transferase enzymes could be responsible for their phenotype. Northern analysis of C1726 (*fap1fap1*), N79-2077-12 (*fap_{nc}fap_{nc}*) and the cultivar Dare revealed that *fap_{nc}fap_{nc}* genotypes have reduced accumulation of transcripts corresponding to a 16:0-ACP thioesterase (*GmFATB*) gene and no difference in transcript accumulation for genes encoding 18:1-ACP thioesterase (*FATA*) or 18:0-ACP desaturase activities (Wilson et al. 2001b). No differences in transcript accumulation for *fap1-fap1* genotypes were observed for any of the genes assayed (Wilson et al. 2001b). Southern analysis using the *GmFATB* cDNA as a probe revealed that multiple copies of the 16:0-ACP thioesterase gene are present in the soybean genome and that the *fap_{nc}fap_{nc}* genotype has a deletion in one of those copies (Wilson et al. 2001a). The mutation in N79-2077-12 (*fap_{nc}* allele) is caused by a deletion in one of the copies of the 16:0-ACP thioesterase gene (*GmFATB1a*) and allele-specific markers were developed to be used in marker-assisted selection schemes (Cardinal et al. 2007). The mutation in A22 (*fap3* allele) is caused by a non-conservative amino acid substitution in the *GmFATB1a* gene (Vries et al. 2011).

Correlated changes in stearic acid and unsaturated fatty acids in the soybean oil have been observed in low palmitic acid lines that carry the major genes described above. Lines homozygous for the *GmFATB1a* deletion (homozygous *fap_{nc}* mutants) had decreased stearic acid (18:0) and increased linoleic acid (18:2) in two populations (Cardinal et al. 2007). Similarly, progenies that inherited the major *fap_{nc}* allele from the cross of N87-2122-4 to two high yielding normal palmitate lines had a 6–15 % reduction in 18:0 (Rebetzke et al. 1998b, 2001), a 4–10 % increase in oleic acid (18:1) content in both crosses (Rebetzke et al. 1998a, 2001), and an increase in linolenic acid (18:3) content in one cross (Rebetzke et al. 1998a). No differences in 18:2 were observed between progenies that inherited the major *fap_{nc}* allele and those that inherited the normal allele (Rebetzke et al. 1998a). In addition, when lines homozygous for the *fap1* and *fap3* alleles were compared to normal lines derived from the same population, the low 16:0 lines have reduced 18:0, and increased 18:1, 18:2, and 18:3 contents (Ndzana et al. 1994). Palmitate content in soybean oil was negatively correlated with 18:2 and positively correlated with 18:1 in a population that was segregating for the *fap1* and *fan* loci (Nickell et al. 1991). In summary, homozygous lines carrying *fap_{nc}* or both *fap1* and *fap3* alleles, not only have reduced 16:0 content but also have reduced 18:0 content and the effect on unsaturated fatty acid contents varies among populations.

Also, correlated changes in agronomic traits have been observed in low 16:0 lines that carry the major genes described above. Several studies have shown that lines homozygous for major *fap1*, *fap3* or *fap_{nc}* alleles have

reduced seed yield (Ndzana et al. 1994; Rebetzke et al. 1998a; Cardinal et al. 2008; Cherrak et al. 2003). Reduced 16:0 lines homozygous for the *fap1* and *fap3* alleles had reduced oil content, were taller, and had greater protein content in two out of three populations, when compared to normal lines derived from the same population (Ndzana et al. 1994). Progenies homozygous for the *fap_{nc}* allele had no differences in protein content in all populations studied and an increase in oil content in only one cross (Rebetzke et al. 1998a). Lines homozygous for the *GmFATB1a* deletion (homozygous *fap_{nc}* mutants) had decreased plant height and increased protein content in two populations but had a significantly decreased oil content in only one population (Cardinal et al. 2008). Small but significant positive correlations between palmitate content and seed oil were observed in a population derived from the cross of a low-palmitate, low-linolenate line homozygous for the *fap_{nc}*, *fap1*, and *fan* alleles to the cultivar Anand (Cherrak et al. 2003).

The objectives of this study were to: (1) identify the genomic location of the *fap1* mutation, (2) determine its molecular basis, (3) estimate the amount of phenotypic variation in the 16:0, 18:0, 18:1, 18:2 and 18:3 contents that is explained by this locus, (4) determine if there are epistatic interactions between the *fap1* and *fap_{nc}* loci and (5) determine if the *fap1* mutation has pleiotropic effects on seed yield, oil and protein content; in three adapted populations grown in replicated trials.

Materials and methods

Population development

Three populations, FAE, FAHH00 and FADD00 were developed by single seed descent (Brim 1966). The FAE population consisted of 746 F₅-derived lines from the cross of N98-4445 × Satellite. N98-4445 is a F₃-derived high-oleate (563.1 g kg⁻¹), low-linoleate (26.3 g kg⁻¹); homozygous for *fan* PI 123440) line that is a sister line of the released germplasm N98-4445A developed by the USDA-ARS in Raleigh, NC (Burton et al. 2006). Satellite is a low-palmitate (39.8 g kg⁻¹; homozygous for *fap_{nc}* and *fap1*), low-linolenate cultivar (*fan* PI 123440) (Cardinal et al. 2007; Bachlava et al. 2008a, b, 2009b; Cardinal et al. 2011). FAHH00 (F₄ derived lines) was derived from the cross of a high yielding line, ‘Corsica’ (Kenworthy 1996), and N97-3681-11, a low linolenic (*fan* PI 123440) and low palmitic acid (homozygous for *fap_{nc}* and *fap1*) line. FADD00 (F₄ derived lines) was derived from the cross of a high yielding line, Brim (Burton et al. 1994a), and N97-3708-13, a low linolenic (*fan* PI 123440), low palmitic acid line (homozygous for *fap_{nc}* and *fap1*).

N97-3681-11, N97-3708-13 and the cultivar Satellite are sister lines homozygous for the low-palmitate mutations *fap_{nc}* and *fap1* and they are highly related to Brim (coefficient of parentage = 0.777).

Experimental design

The FAE population was planted in a sets-within-replications experimental design (Hallauer and Miranda 1988) with two replications in each location. The experimental design was described by Bachlava et al. (2008b). Seven hundred nineteen experimental lines were initially divided into five groups according to maturity. Next, experimental lines of each maturity group were randomly assigned to sets. The maturity groups were randomly assigned to columns in each set and the lines of each maturity group were randomly assigned to plots within each column. Lines of the same maturity were planted in each column of each set to facilitate mechanical harvest. Each maturity group was included at least in one column within each set. Maturity groups with most of the segregating lines were allocated to two or three columns within each set. The parental line Satellite was also randomly assigned to one plot within a single set in each replication according to its maturity group. The other parental line, N98-4445, was randomly assigned, based on its maturity group, to one plot within each of the 15 sets in each replication of the experiment. Once the experimental and parental lines were assigned to a set, they remained in the same set in all replications, locations, and years. The order of sets and the lines within each set were randomized in each replication, location and year. Overall, the sets-within-replications design for the FAE population consisted of 15 seven-row by seven-column sets. Each plot consisted of four 4.88 m rows with 0.97 m spacing between rows. Only the two middle rows, end-trimmed to 3.96 m were harvested.

The experimental design used for FADD00 and FAHH00 was described by Cardinal et al. (2007). Briefly, 98 F₄-derived lines and the parents of the FADD00 population and 99 F₄-derived lines and one parent (N97-3681-11) of the FAHH00 population were grown in Plymouth (NC) in 2002 and Caswell and Plymouth (NC) in 2003. Within each environment, the experimental design for each population was a 10 × 10 lattice with two replications.

All yield plots for the three populations were sown at a seeding rate of 30 seed per meter.

Phenotypic evaluation

Flowering and maturity dates were recorded for FAE, FAHH00, and FADD00 at the R2 and R8 reproductive stages, respectively, as days after planting (Fehr and Caviness 1977). Plant height was also measured for FAE,

FAHH00, and FADD00 populations as the distance from the soil to the plant apex in cm. Lodging was measured on a 1–5 scale (scored 1 if plant 90 % erect to 5 if plants 90 % procumbent). Yield and moisture content were evaluated on seed harvested from the two middle rows of each plot after end-trimming and mechanical harvest. Yield was adjusted to 13 % moisture content.

Approximately 10 and 30 g of seed were sub-sampled from each plot for the evaluation of oil content and fatty acid composition in the FAE and FADD00/FAHH00 populations, respectively. Seed oil content was determined by nuclear magnetic resonance (NMR) using a Maran pulsed NMR (Resonance Instruments, Witney, Oxfordshire, UK) in the FAE population. Protein and oil contents (dry weight basis) in FADD00 and FAHH00 populations were analyzed by near infrared reflectance spectroscopy at the USDA North Regional Research Center, Peoria, IL, in 2003. Fatty acid composition was evaluated by gas chromatography, as described by Burkey et al. (2007). Fatty acid content is reported in g kg⁻¹ of total lipids.

Statistical analysis

Statistical analyses of the sets-within-replications designs for FAE were conducted using PROC MIXED in SAS 9.1 (SAS Institute 2004). Environments, replication (environments), sets, sets*environments, sets*replication (environments), lines (set), environments*lines (set), and the error terms were considered random effects. Best linear unbiased predictors (BLUPs) were obtained for all traits as the sum of the overall mean and the random effect for each line of the FAE population (Littell et al. 1996). Degrees of freedom for FAE population were derived with the containment method in order to reduce the computing power required. Analysis of the FAE population was conducted separately for each environment, as well as combined across all the environments for all measured traits. In order to investigate whether maturity date affected the associations between fatty acid traits and molecular markers, maturity date was included as a covariate in the analysis conducted for each environment as well as across environments and BLUPs from the models with the covariate were obtained. Overall mean, range of BLUPs, heritabilities, genetic and phenotypic correlations were previously reported (Bachlava et al. 2008b). The lower bound of the pooled genetic correlations among environments for each trait, that did not account for heterogeneity of genotypic variances among environments were calculated by the ratio of $\sigma_g^2 / (\sigma_g^2 + \sigma_{g \times e}^2)$ (Cooper and DeLacy 1994; Bachlava et al. 2008b).

The statistical analysis of the FADD00 and FAHH00 populations was described by Cardinal et al. (2007, 2008).

The fatty acid data from each population were analyzed as a lattice design across environments with PROC MIXED, SAS 9.1. (SAS Institute 2004). Locations, replications, incomplete blocks, and lines were considered random effects. A BLUP for each line was obtained by adding the overall mean effect to the random effect of each line (Littell et al. 1996).

Genotypic evaluation and linkage mapping

Genomic DNA was extracted from leaf tissue of approximately 10–20 plants for each of 746 lines of the FAE population, collected from plants grown from seed increases grown in Clayton, NC in 2004. DNA isolation was conducted in a 96-well plate format with the Gentra PureGene DNA purification kit (Gentra Systems, Minneapolis, MN). Ninety-eight polymorphic simple sequence repeat (SSR) markers of the consensus linkage map (Creagan et al. 1999; Song et al. 2004) were genotyped across the 20 linkage groups of the soybean genome in 746 lines of the FAE population following the protocol described by Bachlava et al. (2008a). SSR markers very tightly linked to fatty acid candidate genes (*GmFAD2-2C*, *GmFAD2-2D*, *GmFAD6*, *GmFAD2-1B*, *GmAAPT1b*) were developed and genotyped as described by Bachlava et al. (2009a). The same 746 lines were assayed with the *GmFATB1a* allele-specific marker developed to detect homozygote mutant lines carrying the low-palmitate *fap_{nc}* mutation Cardinal et al. (2007). In addition, the parents and 562 F_{5:7} derived lines were genotyped with the GoldenGate 1,536 SNP panel (1536 Universal Soy Linkage Panel 1.0) (Hyten et al. 2010). The simple sequence repeat marker (SSR) BARCSOYSSR_09_1707 (Sat_1707) (Song et al. 2010) located at approximately Gm09:46.3 Mb was genotyped in the same 562 lines to more precisely define the location of a low-palmitate QTL on Gm09. The PCR conditions for this SSR marker were 95 °C for 2 min, 13 cycles of 94 °C for 20 s, a touchdown step of 63 °C for 15 s reducing 1 °C per cycle, and 72 °C for 20 s, then 18 cycles of 94 °C for 20 s, 50 °C for 20 s, and 72 °C for 20 s, and final step of 72 °C for 5 min.

DNA extraction, genotyping protocols and methods for FADD00 and FAHH00 were described by Cardinal et al. (2007). Briefly, an allele-specific marker for the deletion of *GmFATB1a* observed in lines homozygous for the *fap_{nc}* mutation was genotyped in these populations. Both populations were also genotyped with SSR marker BARCSOYSSR_09_1707, to validate its effect using these populations.

Linkage analysis of the FAE population was conducted with MAPMAKER/EXP 3.0 (Lander et al. 1987) according to Cardinal et al. (2001), using a minimum likelihood of odds (LOD) score of 3.0 and a maximum Haldane distance

of 50 cM. Heterozygous genotypes were coded as missing data in order for the F_5 -derived lines to be analyzed as recombinant inbred lines. Loci that could not be uniquely placed on a linkage map or that had no recombination with another locus in a linkage group were not included in the final maps.

Mapping the *fap1* locus and estimating its effect in three populations

QTL analysis was conducted using BLUPs derived from analyses across environments for all traits and 712 experimental lines of the FAE population that had both phenotypic and genotypic data. The marker genotypes were tested for association with palmitate content using single factor analysis, a simple linear regression approach (data not shown). Based on those results, new markers were developed in the vicinity of one major 16:0 QTL on chromosome 9 (LG K) and candidate genes in the region were sequenced.

To determine if the two major QTLs (*fap_{nc}* and *fap₁*) in the FAE population associated with palmitic acid content were also associated with other fatty acids and agronomic traits, an ANOVA was performed with the *GmFATB1a* and *GmKASIIIa* markers, and their interaction as independent variables was analyzed. BLUPs for each fatty acid and agronomic traits for each line were dependent variables in the analysis. The genotypic class means, the difference between means, significance of those differences, and the R-square for the model were obtained using SAS PROC GLM, SAS 9.2. (SAS Institute Inc. 2010).

In the FAHH00 and FADD00 populations an ANOVA was performed with *GmFATB1a* and the BARCSOYSSR_09_1707 marker classes and their interaction as the independent variables in the analysis. Lines that were heterozygous for the BARCSOYSSR_09_1707 marker were considered missing data so that all means could be estimated. The BLUPs for each fatty acid and agronomic traits for each line were dependent variables. The genotypic class means, the difference between means, significance of those differences, and the R-square for the model were obtained using SAS PROC GLM, SAS 9.2. (SAS Institute Inc. 2010).

Sequencing of candidate genes in the vicinity of major QTL on chromosome 9

Genomic DNA from cultivar Century (Wilcox et al. 1980) (WT) and its EMS induced mutant C1726 (*fap1*) was isolated from ~20 mg of mature seed or ~100 mg of leaf tissue, using the DNeasy Plant Mini kit (Qiagen, Inc., Valencia, CA). Gene sequences for target genes near the major QTL on chromosome 9 and potential homologous

genes were downloaded from <http://www.phytozome.net/soybean> and aligned using Vector-NTI (Invitrogen). PCR primers to amplify candidate genes Glyma09g39610 and Glyma09g41380 were designed based on gene sequences downloaded from the Williams 82 whole genome shotgun sequence (<http://www.phytozome.net/soybean>) to be gene specific after alignment with homologous genes using AlignX (Invitrogen). A minimum of two differences per primer were required between the target gene and any gene homologs.

PCR amplification with genomic DNA was performed with 5–50 ng of genomic DNA with Ex taq according to manufacturer's recommendations (Takara, Otsu, Shiga, Japan) in a PTC-200 thermocycler (MJ Research/Bio-Rad, Hercules, CA), under the following conditions: 95 °C for an initial 5 min denaturation, followed by 40 cycles of (95 °C for 30 s, then 60 °C for 30 s, and an extension step at 72 °C for 1 min per kilobase pair of expected product (7 min for Glyma09g39610 and 5 min for Glyma09g41380). PCR products were separated on 1 % agarose gels to ensure appropriate size and were purified using a QIAquick PCR purification kit (Qiagen). Following purification, products were Sanger sequenced at the DNA Core Facility at the University of Missouri-Columbia. Details on primers used for amplification and sequencing are presented in Table 1.

Genotyping the *GmKASIIIa* KASPAR allele-specific marker

A KASPar marker that distinguishes the *GmKASIIIa* allele inherited from Satellite or N98-4445 (Glyma09g41380) was genotyped in the same 562 lines described above. A pre-amplification step was performed to amplify only *GmKASIIIa* (Glyma09g41380) with the following primers (KASIIIa_pcr F 5'-CCGAAGATCATTGTTGGCTTT-3', KASIIIa_pcr R 5'-CCACTGTACCTATCTGTTTATAGCACACTAGT-3') and conditions: 8 µl reaction containing 5–10 ng genomic DNA, 1× NEB standard taq reaction buffer, 234 µM dNTPs, 0.75 units NEB taq polymerase, 3.5 pmol KASIIIa_pcr F, 3.5 pmol KASIIIa_pcr R. The PCR was performed using a touchdown program with an initial 2 min at 95 °C, followed by 13 cycles of 94 °C for 20 s, 63 °C for 15 s (–1 °C/cycle), 72 °C for 20 s; followed by 18 cycles of 94 °C for 20 s, 50 °C for 20 s, 72 °C for 20 s, with a final extension at 72 °C for 5 min. The KASPar protocol to genotype the KASIIIa SNP used the following primers: (KASIIIa_C1726snp_G (wild type): 5'-ga aggtcggagtcacacggattcttccaatctccaataccaga-3', KASIIIa_C1726snp_C3: 5'-CTTGAGATGGCAAACAAAACACTGAGTA-3').

One ul of the KASIIIa pre-amplified product was dried on a PCR plate. The KASPar PCR had a total volume of

Table 1 Gene locus, primers and expected amplicon size of candidate genes located in the vicinity of the major QTL on Chromosome 9

Gene	Forward primer (5' → 3')	Reverse primer (5' → 3')	Amplicon size (bp)
Glyma09g39610 (Thioesterase)	CTCGGCTCTACGCTTTGACT TGGGGCTCATATCATATCCTG CAAGGTCAAATTTAAAAATCCAAA TCACTCCAAGTATCCAATGA AGCGTTTTGCTTTCACCTA TTTGGAACACCATTGTGATGA CAAAGGTGCCTTTTGGCTTA GCAACTGCTTTTGTGCTG CTTTTCTGGTTCTTTCCTCCA TCCCAATTCCTCATTCTTC TCATATATGCCCTAGCTCTTGTC TGCTATGGTTGTTGACAGAGATTT AAGGCTCGTGCTGTGAGTCT GGGGTTAATTTACTGTGAAGCCTA	GCGACAACAA AAGGACCAAT	7,139
Glyma09g41380 (ketoacyl-synthase III)	AATATTCCTTGTGGTGTACATATAC GCCATCTCAAGCATTATTTCTAGG CTGAATCCGAAATAGCCTCT TATAAACCCCGTTCTTGTTCAT CAAACCTCCTTGTCAAATGATTA TGGTAGCAAATGTCAGTTAGCAAT ATTCAAATGAATGGCAAAGAAGTC GAATTGCATGGCATACTTCATAAT ACATCCAACCTTGCACCTTT TGACAGTTAAGCATAGAAAATAGCTG TAAGACAATGGATTTGTTTGCAG TCTTTGAATGATATGAGAAGACA CATGCAATTCCTCAAGTTAACTAAA	GAAACATAAATA GTAAGCCTCGTTTC	4,110

4 µl, which included 2 µl of 2× KASPar reagent (KBioscience, KBS-1004-002), 0.66 pmol of allele-specific primer KASIIIa_C1726snp_G, 0.66 pmol of allele-specific primer KASIIIa_C1726snp_A, 1.65 pmol of common primer KASIIIa_C1726snp_C3, a final concentration of 0.2 mM Tris pH 8.3, and a final concentration of 2.7 mM MgCl₂.

The KASPar PCR was performed using a touchdown program with an initial 15 min at 94 °C, followed by 10 cycles of 94 °C for 20 s, 65 °C for 60 s (−0.8 °C/cycle), followed by 26 cycles of 94 °C for 20 s, 57 °C for 60 s. Results were read on a Roche LightCycler480.

qPCR expression analysis of fatty acid related genes *GmKASIIIa*, *GmFATB1a*, *GmFATB1b*, *GmFATB2a* and *GmFATB2b*

Thirty-five days after the R2 flowering stage, multiple pods of similar size and location within the soybean canopy were harvested from one row plots grown at Clayton, NC in 2010. These pods from Satellite and each of five lines

from the FAE population that had similar flowering date and maturity date were immediately frozen.

These lines were homozygous or heterozygous for the *GmFATB1a* allele (*Fap_{nc}Fap_{nc}* or *Fap_{nc}fap_{nc}* genotypes) but two of those lines were homozygous for the *GmKASIIIa* allele inherited from N98-4445, one line was heterozygous and two lines were homozygous for the *GmKASIIIa* allele inherited from Satellite. Seed obtained from each sample were pooled and ground under liquid nitrogen and total RNA was isolated using the RNeasy Plant mini kit and protocol from Qiagen (Valencia, CA). Traces of genomic DNA were eliminated using on-column DNase digestion (Qiagen). First strand cDNA synthesis was carried out using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems) with MultiScribe reverse transcriptase (RT) and random primers. 2.5 µg of RNA in 12.5 µl of nuclease free water were mixed with the same volume of a 2× RT master mix. The reaction conditions were: (1) 10 min at 25 °C, (2) 120 min at 37 °C, and (3) 5 s at 85 °C. Quantitative RT-PCR assays with SYBR green were performed with the i-Cycler

Table 2 Primers, gene locus, annealing temperature, amplicon and PCR amplification efficiency of fatty acid related genes *GmKASIIIa*, *GmFATB1a*, *GmFATB1b*, *GmFATB2a* and *GmFATB2b* studied in expression analysis in soybean lines, using qRT-PCR

Gene	Forward/reverse primers (5' → 3')	Gene locus	Ta ^a (°C)	Amplicon (bp)	Efficiency ^b (%)	Function	References
<i>GmFATB1a</i>	CAAGTGGACACTTGGGTTTC GCCTCCGTGTTAGCTTATTC	Glyma05g08060	61	136	93.3	Acyl-ACP thioesterase	This study
<i>GmFATB1b</i>	AGTAGACACCTGGGCTTCTG GCCTCCGTGTCAGCTTATTC	Glyma17g12940	62	134	100	Acyl-ACP thioesterase	This study
<i>GmFATB2a</i>	TCTCTGGTGCTGATGTAGGA AAAGGTTTGGGCCTCCATTG	Glyma04g21910	60	121	101	Acyl-ACP thioesterase	This study
<i>GmFATB2b</i>	AGAATGATTTGCCATGGTTG GATCTTCCAATCCCAAAGG	Glyma06g23560	57	197	109	Acyl-ACP thioesterase	This study
<i>GmKASIIIa</i>	ATTTCCAATGACGACCTTTC TGCCTCTACTGCTAAACTTG	Glyma09g41380	57	126	98	3-ketoacyl-ACP synthase	This study
<i>Soy57 actin</i>	5' GAGCTATGAATTGCCTGATGG	Glyma15g005570	58	118	100	Cytoskeletal structures (Reference gene)	Moniz de Sa and Drouin 1996
<i>GmEF1α</i>	5' CGTTTCATGAATTCCAGTAGC CTGTAACAAGATGGATGCCACTAC CAGTCAAGGTTTGTGGACCT	Glyma17g23900	61	176	98	Translation elongation factor (Reference gene)	Aguilar et al. 1991

^a Annealing temperature

^b PCR amplification efficiency

thermocycler (BioRad, Hercules, CA) for *GmFATB1a*, *GmFATB1b*, *GmFATB2a*, *GmFATB2b*, *GmKASIIIa*, *Soy 57 actin*, and *GmEF1α*. PCR reactions were carried out in a total volume of 25 µl, containing 200 nM of each forward and reverse primer (Table 2), 1X iQ SYBR green Supermix (BioRad, Hercules, CA) and template cDNA equivalent to 20 ng of total RNA. As a control, the amplification was also conducted using water instead of nucleic acid template (no template control). The amplification conditions were 95 °C for 3 min to activate the hot-start AmpliTaq Gold DNA Polymerase. Cycling conditions were: 15 s at 94 °C, 30 s at the respective annealing temperature and 30 s at 72 °C, followed by melting curve analysis. Primer sequences and annealing temperatures are shown in Table 2. Threshold cycles (C_t) for each sample were determined using i-Cycler software. PCR reactions were conducted in triplicate for each sample.

The amounts of transcripts were estimated from standard curves obtained using PCR products amplified with the respective primer pairs and applying a normalization factor that was obtained based on the analysis of soybean actin 57 (Moniz de Sa and Drouin 1996) and elongation factor EF1α (Aguilar et al. 1991) and using geNorm software (Vandesompele et al. 2002). To ensure that the correct gene transcript sequences were amplified, amplicons were sequenced at the Iowa State University Biotechnology Center, Ames, IA.

Sequencing of fragment amplified from wildtype and *GmKASIIIa* mutant cDNA

Primers were selected to specifically amplify a fragment from cDNA of KASIIIa on Gm09 (Table 3). Forward primer 5'-AATATTCCCTTGTTGGTTACAT-3' is located in the 5' UTR of Glyma09g41380, and reverse primer 5'-GATTTGTTTTGCAGCCAAGC-3' is located in exon 4 of Glyma09g41380. PCR conditions included an initial extension step of 2 min at 94 °C, followed by 30 cycles of (94 °C for 14 s, 57 °C for 30 s, 72 °C for 1 min), and 5 min at 72 °C. The fragments were excised from a 1 % agarose gel and purified using a Qiaquick Gel Extraction Kit (Qiagen # 28704). The purified fragments were cloned into a pCR2.1 vector using the TA Cloning Kit (Invitrogen #28704), and DNA from these clones was submitted to GeneWiz for sequencing. Sequence analysis was performed using VectorNTI software (Invitrogen).

Northern analysis of *GmKASIIIa*

Pods and unifoliate leaves were harvested from N95-4445A and Satellite to isolate RNA and perform Northern analysis. Pods were harvested 35 days and 50 days after flowering from plots grown at Clayton in 2012 and leaves were harvested from seedlings germinated in the laboratory. Total RNA was isolated using the RNeasy Plant mini

Table 3 Primers, gene locus, and amplicon of fatty acid related gene *GmKASIIIa*, and loading control *GmCONS6* used to test splice site in putative *fap1* (C1726) mutation in Glyma09g41380

Gene	Forward primer (5' → 3')	Reverse primer (5' → 3')	Gene locus	Amplicon (bp)	Notes
<i>GmKASIIIa</i>	ccgccaaagtgtttgct	gcctgaagaactcgtcgc	Glyma09g41380	cDNA 239, DNA 459	exon 1 to exon2
<i>GmCONS6</i> (loading control)	agatagggaaatggtgcaggt	ctaagccaattgcagctctc	Glyma12g05510	cDNA 93 bps	Plant Genome (2008) 1:44–54

kit and protocol from Qiagen (Valencia, CA). Northern blots were performed as described by Sambrook and Russell (2001). A 1,208 bp fragment *GmKASIIIa* probe was amplified from cDNA synthesized for qPCR expression analysis. The probe was labeled using the Random Primed DNA Labeling Kit from Roche and [³²P] dCTP from Perkin Elmer. Hybridization was done overnight at 65 °C using PerfectHyb Plus Hybridization Buffer from Sigma. The autorad was developed after 5 days of exposure.

Results

Linkage map

Linkage analysis was conducted using MAPMAKER/EXP 3.0 with 98 polymorphic SSR markers and 420 polymorphic SNPs in the FAE population, which cover the 20 linkage groups (LGs) of the soybean genome. A total of 451 markers were uniquely placed on the final map of 31 LGs and 3 unlinked loci because some markers were redundant (no recombination) or could not be placed uniquely on the map and due to the lack of polymorphisms among the parents across large genomic areas, linkage could not be declared among all markers on the same chromosome (Supplementary Table 1). Chromosomes 1, 2, 4, 7, 8, 13, 14, 15, 16, 17, and 20 were split into to 2 LGs each. In addition, there were 14 large gaps (26–50 cM) in the final map. The total length of the linkage map was 2,351 cM, equivalent to published consensus maps (Hyten et al. 2010). The order of marker loci within each linkage group was the same as the order presented in Hyten et al. (2010) with a few exceptions: in chromosome 4, BARC-062503-17863 and BARC-042719-08395; in chromosome 5, BARC-021775-04203; in chromosome 9, BARC-061897-17586; in chromosome 10, BARC-038327-10019, in chromosome 13, BARC-064707-18784 and BARC-041141-07916; in chromosome 14, BARC-017589-02630; in chromosome 15, BARC-030171-06819 and BARC-061277-17149; in chromosome 17 BARC-047903-10427; in chromosome 18, BARC-063581-18909 and BARC-064177-18581; in chromosome 19, BARC-064609-18739,

BARC-021733-04193, BARC-035235-07156, and BARC-047496-1294; are all in slightly different local orders (Supplementary Table 1).

Loci were declared to have significant segregation distortion if the Chi-test had a *p* value <0.0016 (0.05/31 LGs) (Supplementary Table 1). Nine markers encompassing a large genomic region (~45 cM) in chromosome 5 are severely distorted. The low-palmitate *fap_{nc}* mutation caused by a deletion of the *GmFATB1a* gene maps to the end of this chromosome (Cardinal et al. 2007; Li et al. 2002; and this study).

There are 9 marker loci on chromosome 6 (~20 cM), 4 markers on chromosome 17 (~7.8 cM), 17 markers on chromosome 18 (~23.6 cM), 15 markers in chromosome 19 (~43.4 cM) that show significant segregation distortion.

Mapping the *fap1* locus

Only BLUPs derived from analyses across environments were utilized for QTL analysis because the lower bound of the pooled genetic correlations among environments for each trait (Cooper and DeLacy 1994), were higher than 0.8 (except for oil content and seed yield, data not shown). Single factor marker analysis for palmitate content detected two major QTL, one in chromosome 5 (*GmFATB1a*) and the other at the end of chromosome 9. Based on those results, BARCSOYSSR_09_1707 (46.31 Mb, Supplemental Table 1) was the closest marker to the QTL on chromosome 9.

In order to identify candidate genes for reduced palmitic acid content due to *fap1*, we examined the Williams 82 reference sequence (Schmutz et al. 2010), corresponding to the genomic location near BARCSOYSSR_09_1707. Two genes were annotated as involved in fatty acid biosynthesis and were chosen as putative candidate genes: Glyma09g39610, annotated as a putative *Acyl-CoA thioesterase* gene and Glyma09g41380, annotated as a putative *ketoacyl-ACP synthase III* gene (KASIII) (<http://www.phytozome.net/soybean>). We PCR amplified and Sanger sequenced both genes using genomic DNA from Century (WT) and its EMS induced mutant C1726 (*fap1*). For Glyma09g39610, assembled gene sequences were identical

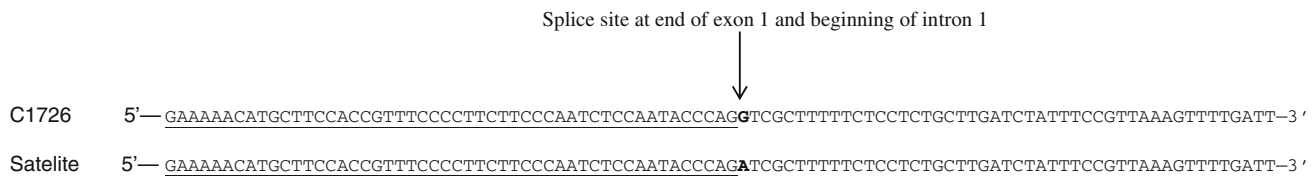


Fig. 1 Region of KASIIIa genomic sequence Glyma09g41380 in Century and in the EMS induced mutant C1726 (*fap1*). A portion of Glyma09g41380 exon 1 is *underlined*, and *an arrow* indicates

position of splice site at junction between exon 1 and intron 1. SNP in C1726 is located at the first base of intron 1

between Century and C1726 (and Williams 82). In contrast, examination of sequencing results for Glyma09g41380 revealed a single unique polymorphism in C1726 (G174A, relative to the start codon) shown in Fig. 1, which was predicted to disrupt the donor splice site recognition between exon one and intron one (Hebsgaard et al. 1996; Long and Deutsch 1999). The resulting KASIIIa protein would have a truncation of the open reading frame of the Williams 82 version of 397 predicted residues to:

MANASGFFTPSPVPHFKVRVKPLNTAIGFSKVVCFGNIEGAEKHASTVSPSSQSPIRSLFLLCLIYFR*. In addition, the G to A change in C1726 creates the GATCTG motif that has been associated with enhanced gene expression when present in the first intron in *Arabidopsis* (Rose et al. 2008).

A KASPar marker was developed for the SNP in *GmKASIIIa* gene and genotyped in the FAE population. *GmKASIIIa* mapped to Chromosome 9 between BARC-059411-15797 and BARCSOYSSR_09_1707 (Sat_1707), 1.8 and 0.8 cM apart, respectively (Supplemental Table 1), coinciding with its location in the Williams 82 genome assembly. In FAE, the *fap_{nc}*, *GmKASIIIa* and their interaction explained 66 % variation in 16:0 content (Table 4). Genotypes homozygous for *fap_{nc}*, and homozygous for the *GmKASIIIa* mutation had 49 g of 16:0 per kg of total lipids. The alleles inherited from Satellite (*fap_{nc}* and *fap1*) also significantly reduced stearic acid content and increased linoleic acid content (Table 4). The *GmKASIIIa* allele inherited from Satellite significantly increased linolenic acid content in this population. To test if *fap_{nc}* and *GmKASIIIa* had an effect on agronomic traits, an ANOVA was performed with *GmFATB1a* and *GmKASIIIa* markers and their interaction as independent variables and BLUPs for each agronomic trait for each line as dependent variables in the analysis. *GmFATB1a* and *GmKASIIIa* and their interaction significantly affected seed yield; lines homozygous for both mutant alleles had the lowest yield in the across environments ANOVA analysis (Table 5). However, when ANOVA analysis was performed for each environment separately (data not shown) only lines homozygous for the *fap_{nc}* mutation significantly reduced yield in 5 out of the 6 environments. The *fap_{nc}* mutation

slightly increases oil content in FAE (Table 5) but when ANOVA is performed in each environment separately (data not shown), the increase is significant in only two environments out of six. However, lines homozygous for the *GmKASIIIa* allele inherited from Satellite (*fap1*) have a highly significant reduction in oil content (Table 5) in the analysis across and within environments, (data not shown). The *fap_{nc}* mutation also significantly reduced plant height in the FAE population (Table 5).

In the FAHH00 and FADD00 populations an ANOVA was performed with *GmFATB1a* and the BARCSOYSSR_09_1707 (Sat_1707, closely linked to *GmKASIIIa*) and their interaction as independent variables and BLUPs for each fatty acid and agronomic traits for each line as dependent variables. The *fap_{nc}*, Sat_1707 (*GmKASIIIa*) and their interaction (in FADD00) explained 86–94 % of the variation for 16:0 content (Table 6). Lines homozygous for the Satellite allele at both loci had on average 44–45 g of 16:0 per kg of total lipids. Both alleles significantly reduced stearic acid content in both populations and the Sat_1707 allele inherited from the *fap1* mutant parent (Satellite) significantly increased linoleic acid content in one population. The Sat_1707 allele from Satellite and its interaction decreased yield in FADD00 (Table 7). The Sat_1707 allele from Satellite and its interaction decreased oil content in both populations (Table 7). Lines homozygous for the Satellite allele at both loci matured significantly later in both populations.

qPCR expression of fatty acid related genes *GmKASIIIa*, *GmFATB1a*, *GmFATB1b*, *GmFATB2a* and *GmFATB2b* in soybean lines from FAE population

In order to test if the SNP that creates a GATCTG motif in the *GmKASIIIa* gene (from C1726) for enhanced gene expression as predicted from studies in *Arabidopsis*, RNA was obtained from seeds harvested 35 days after flowering from five lines of the FAE population of similar flowering date and maturity date and also from the parental line Satellite. cDNA was synthesized to perform quantitative RT-PCR analysis. Lines homozygous or heterozygous for the *GmFATB1a* allele (*Fap_{nc}Fap_{nc}* or *Fap_{nc}fap_{nc}* genotypes) and homozygous for the *GmKASIIIa* allele inherited from Satellite (mutant allele) had 1,247 Attamoles more of

Table 4 Palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid content BLUPs (g kg⁻¹) of seed oil of parental lines and genotypic means of progenies homozygous *fap_{nc}fap_{nc}* and homozygous wild type for KASIII A (AA), homozygous *fap_{nc}fap_{nc}*and homozygous mutant for KASIII A (BB), homozygous or heterozygous *Fap_{nc}*- and homozygous wild type for KASIII A (AA), and homozygous or heterozygous *Fap_{nc}*- and homozygous mutant for KASIII A (BB) in FAE population (Satellite × N98-4445)

Genotype	No. of individuals	Palmitic acid (g kg ⁻¹)	Stearic acid (g kg ⁻¹)	Oleic acid (g kg ⁻¹)	Linoleic acid (g kg ⁻¹)	Linolenic acid (g kg ⁻¹)
<i>fap_{nc}fap_{nc}</i> KASIII A (AA)	76	63.3	36.6	428.1	439.4	32.1
<i>fap_{nc}fap_{nc}</i> KASIII A (BB)	49	49.0	35.5	419.7	460.9	34.0
<i>Fap_{nc}</i> - KASIII A (AA)	157	89.0	37.9	427.5	413.8	32.4
<i>Fap_{nc}</i> - KASIII A (BB)	186	69.8	37.0	432.7	426.6	33.7
<i>fap_{nc}fap_{nc}</i> vs <i>Fap_{nc}</i> - KASIII A (AA vs BB)		-23.2****	-1.4****	n.s.	29.9****	n.s.
Epistatic interaction		16.7***	1*	n.s.	-17.2**	-1.6****
R ² model (%)		66.4	3.5	0.4	6.2	6.2
Satellite		47.7	36.6	311.6	562.4	42.7
N98-4445		89.9	35.9	563.4	384.3	26.3

Difference between genotypic means and their significance and the R² for the full model in Proc GLM are included

n.s. not significant

Significant at *****p* < 0.0001; significant ****p* < 0.001; significant ***p* < 0.01; and significant **p* < 0.05**Table 5** Genotypic means of seed yield, date of flowering (R2), maturity date (R8), oil concentration, and height of lines homozygous *fap_{nc}fap_{nc}* and homozygous wild type for *GmKASIII A* (AA), lines homozygous *fap_{nc}fap_{nc}* and homozygous mutant for *GmKASIII A*(BB), lines homozygous or heterozygous *Fap_{nc}*- and homozygous wild type for *GmKASIII A* (AA), and lines homozygous or heterozygous *Fap_{nc}*- and homozygous mutant for *GmKASIII A* (BB) in FAE population (Satellite × N98-4445)

Genotype	No. of individuals	Yield (kg ha ⁻¹)	R2 (DAP ^a)	R8 (DAP ^a)	Oil (g kg ⁻¹)	Height (cm)
<i>fap_{nc}fap_{nc}</i> KASIII A (AA)	76	1,474	53	130	196.3	89
<i>fap_{nc}fap_{nc}</i> KASIII A (BB)	49	1,384	55	133	188.9	88
<i>Fap_{nc}</i> - KASIII A (AA)	157	1,535	53	130	193.5	99
<i>Fap_{nc}</i> - KASIII A (BB)	186	1,533	53	130	188.9	99
<i>fap_{nc}fap_{nc}</i> vs <i>Fap_{nc}</i> - KASIII A (AA vs BB)		-105****	1*	n.s.	1.4*	-10****
R ² model (%)		45*	n.s.	n.s.	6.0****	n.s.
		5.0	1.1	0.5	14.9	5.0

Differences between genotypic means, their significance and the R² for the full model in Proc GLM are included

n.s. not significant

^a Days after plantingSignificant at *****p* < 0.0001; significant ****p* < 0.001; significant ***p* < 0.01; and significant **p* < 0.05

GmKASIII A RNA per mg of total RNA than lines homozygous for the *GmKASIII A* allele inherited from N98-4445 (wild type) (Table 8), indicating that the G to A change present in the first intron in C1726 and Satellite that creates the GATCTG motif enhanced gene expression. However, this transcript would produce a truncated protein. Sequencing analysis confirmed that the amplified products corresponded to the *GmKASIII A* (Glyma09g41380) homolog as expected.

Satellite (mutant *fap_{nc}* and *GmKASIII A* genotype) had a 10× reduction in the amount of *GmFATB1a* transcript and significant increases of *GmFATB1b*, *GmFATB2a*, and *GmFATB2b* transcripts.

Sequencing of fragment amplified from wildtype and *GmKASIII A* mutant cDNA

A fragment of 579 bp was amplified from cDNA of the wildtype *GmKASIII A* sample, and a fragment of 799 bp was amplified from cDNA of the KASIII A mutant sample, which would result from the presence of the unspliced 220 bp intron 1 in the KASIII A mutant. Alignment of fragment sequences from 3 clones from each genotype ID shows that introns 1, 2, 3 are not present in the wildtype fragment. The mutant fragment did not contain introns 2 and 3 but it contained the entire sequence of intron 1.

Table 6 Palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid content BLUPs (g kg⁻¹) of seed oil of parental lines and genotypic means of progeny homozygous *fap_{nc}fap_{nc}* and homozygous AA for Sat_1707 (AA), lines homozygous *fap_{nc}fap_{nc}* andhomozygous BB for Sat_1707 (BB), lines homozygous or heterozygous *Fap_{nc}-* and homozygous AA for Sat_1707 (AA), and lines homozygous or heterozygous *Fap_{nc}-* and homozygous BB for Sat_1707 (BB) in two F4-derived populations

Genotype	No. of individuals	Palmitic acid (g kg ⁻¹)	Stearic acid (g kg ⁻¹)	Oleic acid (g kg ⁻¹)	Linoleic acid (g kg ⁻¹)	Linolenic acid (g kg ⁻¹)
‘Corsica’ × N97-3681-11 (FAHH00)						
<i>fap_{nc}fap_{nc}</i> Sat_1707 (AA)	6	59.2	29.2	296.4	556.6	54.9
<i>fap_{nc}fap_{nc}</i> Sat_1707 (BB)	15	44.4	27.6	287.5	579.8	55.5
<i>Fap_{nc}-</i> Sat_1707 (AA)	37	97.2	31.3	286.5	536.7	51.5
<i>Fap_{nc}-</i> Sat_1707 (BB)	32	79.8	30.5	298.9	535.6	54.8
<i>fap_{nc}fap_{nc}</i> vs <i>Fap_{nc}-</i> Sat_1707 (AA vs BB)		-36.7****	-2.5****	n.s.	32.1****	n.s.
Epistatic interaction		16.1***	1.2*	n.s.	n.s.	n.s.
R ² model (%)		n.s.	n.s.	n.s.	n.s.	n.s.
N97-3681-11		85.8	25.9	4.2	27.2	2.1
		38.9	26.9	315.4	563.9	44.7
‘Brim’ × N97-3708-13 (FADD00)						
<i>fap_{nc}fap_{nc}</i> Sat_1707 (AA)	16	58.4	31.2	316.5	533.1	56.9
<i>fap_{nc}fap_{nc}</i> Sat_1707 (BB)	7	45.1	29.1	303.2	561.5	57.7
<i>Fap_{nc}-</i> Sat_1707 (AA)	30	101.9	33.4	300.2	513.2	53.2
<i>Fap_{nc}-</i> Sat_1707 (BB)	33	77.1	32.9	295.6	537.6	57.2
<i>fap_{nc}fap_{nc}</i> vs <i>Fap_{nc}-</i> Sat_1707 (AA vs BB)		-37.8****	-3.0****	12.0*	21.9**	n.s.
Epistatic interaction		19.0****	1.3****	n.s.	-26.4****	n.s.
R ² model (%)		-11.5****	1.7***	n.s.	n.s.	n.s.
N97-3708-13		94.0	65.5	12.4	24.3	2.8
BRIM		66.5	31.2	277.9	573.6	52.1
		110.0	34.5	288.9	503.0	68.7

Difference between genotypic means and their significance and the R² for the full model in Proc GLM are included

n.s. not significant

Significant at *****p* < 0.0001; significant ****p* < 0.001; significant ***p* < 0.01; and significant **p* < 0.05

Northern analysis of *GmKASIII*A

In order to prove if the single unique polymorphism in C1726 (G174A, relative to the start codon) and Satelite, disrupted the donor splice site recognition between exon one and intron one, Northern Analysis was performed. Northern blots showed the presence of a larger *GmKAS-III*A mRNA isolated from pods and leaves of Satelite, confirming the presence of intron 1 and stability of the mRNA. The *GmKASIII*A mRNA isolated from pods harvested 50 days after flowering clearly shows a smaller band in N98-4445 that corresponds to the spliced mRNA lacking intron 1 (Fig. 2).

Discussion

In most regions, the linkage map developed had the same order expected from the Williams 82 *Glycine max* genome sequence (Mb) (v 1.01, verified 7/2012), but there are some important discrepancies in chromosomes 4, 5, 8, 9, 11, 14,

15, and 20 (Supplementary Table 1). Because N98-4445 (female parent) was an F3-derived line, it was not completely homozygous. Since the RILs developed in this study were derived from 3 different F₁ plants, there were instances in which a marker was polymorphic in a subset of RILs but not in another subset of RILs derived from a different F₁ plant. In such instances, the marker was scored as missing data in the subset of RILs derived from 1 or 2 F₁ plants that showed no segregation (monomorphic). There were a few marker loci in almost all LGs that showed this pattern of segregation (Supplementary Table 1), however, there were large genomic regions in chromosomes 8, 9, 12, 14, and 17 in which many markers segregated in a subset of RILs. Marker order and QTL (and their effect) observed in those regions would probably be biased due to the small sample of RILs that provided genotypic and phenotypic information.

The severe segregation distortion observed in the vicinity of the low-palmitate *fap_{nc}* mutation had been reported before (Cardinal et al. 2007; Li et al. 2002) and the allele-specific marker designed for this mutation was

Table 7 Genotypic means of seed yield, date of flowering (R2), maturity date (R8), protein and oil concentrations, lodging rating, and height lines homozygous *fap_{nc}fap_{nc}* and homozygous AA for Sat_1707 (AA), lines homozygous *fap_{nc}fap_{nc}* and homozygous BB for Sat_1707 (BB), lines homozygous or heterozygous *Fap_{nc}*- and homozygous AA for Sat_1707 (AA), and lines homozygous or heterozygous *Fap_{nc}*- and homozygous BB for Sat_1707 (BB) in two F₄-derived populations

Genotype	No. of individuals	Yield (kg ha ⁻¹)	R2 (DAP ^a)	R8 (DAP ^a)	Protein (g kg ⁻¹)	Oil (g kg ⁻¹)	Lodging (1-5 ^b)	Height (cm)
‘Corsica’ × N97-3681-11 (FAHH00)								
<i>fap_{nc}fap_{nc}</i> Sat_1707 (AA)	6	2,588	56	136	437.5	199.9	2.6	87
<i>fap_{nc}fap_{nc}</i> Sat_1707 (BB)	15	2,402	61	149	438.5	186.6	2.5	97
<i>Fap_{nc}</i> - Sat_1707 (AA)	37	2,736	58	143	430.0	197.1	2.6	107
<i>Fap_{nc}</i> - Sat_1707 (BB)	32	2,731	57	142	429.8	193.6	2.5	107
<i>fap_{nc}fap_{nc}</i> vs <i>Fap_{nc}</i> - Sat_1707 (AA vs BB)		-238****	n.s.	n.s.	0.8**	n.s.	n.s.	-1.5***
Epistatic interaction		n.s.	n.s.	-6*	n.s.	8.4****	n.s.	n.s.
R ² model (%)		27.0	6.9	11.3	12.7	28.2	4.3	15.1
‘Brim’ × N97-3708-13 (FADD00)								
<i>fap_{nc}fap_{nc}</i> Sat_1707 (AA)	6	2,683	66	149	438.5	192.3	2.5	106
<i>fap_{nc}fap_{nc}</i> Sat_1707 (BB)	15	2,589	67	152	442.5	185.0	2.5	104
<i>Fap_{nc}</i> - Sat_1707 (AA)	37	2,750	67	150	437.8	189.9	2.4	114
<i>Fap_{nc}</i> - Sat_1707 (BB)	32	2,734	68	150	436.9	188.3	2.5	114
<i>fap_{nc}fap_{nc}</i> vs <i>Fap_{nc}</i> - Sat_1707 (AA vs BB)		-106****	n.s.	n.s.	3.2**	n.s.	n.s.	-9****
Epistatic interaction		55***	n.s.	-1.5*	n.s.	4.5****	n.s.	n.s.
R ² model (%)		35.3	5.0	7.2	10.9	25.4	4.4	32.8

Differences between genotypic means and their significance and the R² for the full model in Proc GLM are included

n.s. not significant

^a Days after planting

^b Scored 1 if plants are 90 % erect to 5 if plants are 90 % procumbent

Significant at *****p* < 0.0001; significant ****p* < 0.001; significant ***p* < 0.01; and significant **p* < 0.05

Table 8 RT-PCR results (Attomoles/mg RNA) from developing seeds RNA of 5 F₅-derived lines from FAE population and ‘Satelite’ harvested at 35 days after flowering in Clayton 2010. Assays were done in triplicates (subsamples)

Genotype	N ^a	GmKASIII A	GmFATB1a	GmFATB1b	GmFATB2a	GmFATB2b
<i>fap_{nc}fap_{nc}</i> KASIII A (BB) (Satelite)	1	2,788.5	187.5	568.9	181.4	32.8
<i>Fap_{nc}</i> - KASIII A (H)	1	1,072.6	1,547.5	306.5	60.3	19.2
<i>Fap_{nc}</i> - KASIII A (AA)	2	953.5 (170.5)	1,591.3 (23)	381.5 (55.1)	84.5 (5.7)	16.9 (2.4)
<i>Fap_{nc}</i> - KASIII A (BB)	2	19.06.8 (141.8)	1,425 (141.4)	383.1 (51.6)	91.5 (21.7)	17.5 (4.7)
<i>T</i> test <i>fap_{nc}fap_{nc}</i> vs <i>Fap_{nc}</i> -		n.s.	-1,328.5 (121.7)***	201.8 (55.6)*	99.0 (18.7)**	15.3 (3.1)**
<i>T</i> test KASIII A (AA vs BB) ^c		-1,247.2*(397)	n.s.	n.s.	n.s.	n.s.

n.s. not significant

^a N number of lines with a specific genotype

^b *T* test (Satelite vs Average LSMeans of 5 lines)

^c *T* test includes Satelite as BB genotype. Similar results were obtained when Satelite was excluded in the *T* test

Significant ****p* < 0.001; significant ***p* < 0.01; and significant **p* < 0.05

the most severely distorted. The other markers on the same chromosome (LG5) become less distorted the farther away they map from this locus.

Other chromosomal regions that also showed segregation distortion in this study could possibly be related to

fitness loci. For example, on chromosome 6, in the genomic region where markers displayed segregation distortion, QTL for abnormal seedling, pod maturity and flowering date have been reported (<http://soybeanbreederstoolbox.org>; verified 8/2012). On chromosome 18, in the genomic



Fig. 2 Northern-blot analysis of *GmKASIIIA* mRNA isolated from pods harvested 50 days after flowering in the germplasm line N98-4445 (lane 1) and in the cultivar ‘Satelite’ (lane 2)

region where markers displayed segregation distortion, QTL for flowering time have been reported (<http://soybeanbreederstoolbox.org>; verified 8/2012). On chromosome 19 a seed set QTL was reported in the genomic region near markers with segregation distortion. If QTL for all these traits were segregating in our population, unintended selection against alleles with undesirable effects may have occurred during inbreeding causing the observed segregation distortion. Standard likelihood approaches for linkage mapping are not best suited for markers displaying segregation distortion (Lorieux et al. 1995), therefore the markers order and distance in the genomic regions described above should be taken with caution. It should also be noted that the recombination frequency among markers and map distances are slightly upwardly biased by analyzing the FAE F₅-derived population as a recombinant inbred line population in linkage analysis performed with MAPMAKER/EXP 3.0 (Supplement Table 1).

It was previously shown that the *fap_{nc}* mutation also causes a small yield reduction (Cardinal et al. 2007) and now our results show that the *fap_{nc}* mutation (or the deletion around the locus per se) reduces fitness either of gametes carrying the mutation or of individuals homozygous for the mutation. Interestingly, the genomic location of the *GmFATB1a* gene in Williams 82 whole genome shotgun sequence (<http://www.phytozome.net/soybean>) is at 7.99 Mb on Chromosome 5, downstream from BARC-021573-04148, however in our populations (Cardinal et al. 2007) and in Li et al. (2002), the *fap_{nc}* mutation is mapped to the tip of chromosome 5 upstream of BARC-048987-10780 and Satt684 at 0.27 and 1.8 Mb, respectively. The discrepancy between the genomic location of the wild type *GmFATB1a* and the genetic location of the *fap_{nc}* mutation could be due to biased estimation of map orders and

recombination frequency with standard likelihood methods (Lorieux et al. 1995) or problems in the assembly of the Williams 82 whole genome shotgun sequence (v1.01) in this region.

The results from the sequence analysis of the two candidate genes on chromosome 9 were validated with the gene expression and Northern blot assays, confirming that the *fap1* low-palmitate phenotype is due to a single unique polymorphism in C1726 (G174A, relative to the start codon), which was predicted to disrupt the donor splice site recognition between exon one and intron one and also creates the GATCTG motif. The gene expression assay demonstrated that this motif enhanced gene expression of the *GmKASIIIA* locus, as had been observed before in *Arabidopsis* (Rose et al. 2008). The increases in *GmFATB1b*, *GmFATB2a*, and *GmFATB2b* transcripts observed in Satellite (*fap_{nc}* and *GmKASIIIA* mutations) showed that when the transcription of the most highly transcribed *GmFATB1a* homolog is compromised (Wilson et al. 2001b; Cardinal et al. 2007), the soybean plant up regulates the transcription of the other three *GmFATB* genes to try to maintain a “normal” fatty acid biosynthesis.

We speculated that since the 3-ketoacyl-ACP synthase enzyme III catalyzed the first condensation reaction in the biosynthesis of C₁₆ and C₁₈ fatty acids in developing seeds, a significant reduction in 16:0, 18:0 and total oil content would be observed in lines homozygous for the *GmKASIIIA* mutation. This was confirmed in all three populations and is in agreement with previous studies with the same mutation (Ndzana et al. 1994; Horejsi et al. 1994).

As we indicated previously (Cardinal et al. 2007), the *GmFATB1a* gene, which encodes the palmitate thioesterase in soybean and is deleted in the *fap_{nc}* allele, is putatively active toward stearoyl acyl carrier protein (ACP) substrates and, therefore, the mutations in this gene would reduce both 16:0 and 18:0 contents in the seed oil. This effect was observed in all the populations. The *fap_{nc}* locus also caused a small increase in total oil content.

Because our study includes the FAE population with 712 F₅-derived lines, our estimates of the genetic parameters have excellent precision. Lines homozygous for the *fap_{nc}* mutation showed significantly reduced yield in all of the populations evaluated and this decreased yield was fairly consistent across the environments. A significant reduction in plant height was also observed for this mutation. In contrast, lines homozygous for the *fap1* and *fap3* alleles have increased plant height (Ndzana et al. 1994), and therefore, since *fap3* (amino acid substitution) and *fap_{nc}* (deletion of unknown size) are mutations in the same *GmFATB1a* gene, it is highly likely that the deletion per se may include another gene that is causing the plant height reduction observed in *fap_{nc}* mutants. Yield reductions associated with the *fap1* mutation were of smaller

magnitude and not always consistent across environments. This observation is in agreement with previous studies on the effect of the *fap1* mutation on agronomic traits (Rebetzke et al. 1998a; Cardinal et al. 2008). Whether this small negative effect on yield is due to the mutation per se or to negative linkage drag is something that cannot be conclusively determined in these populations because none of the parents was a high yielding line, comparable to modern cultivars. Studying the agronomic effects of this mutation on a high yielding-high oil genetic background will help to reveal its potential as a source of low palmitate that will not compromise agronomic performance.

Acknowledgments The authors would like to thank William Novitzky at the USDA-ARS Soybean and Nitrogen Fixation Research Unit, Raleigh, NC, for providing training and equipment for the fatty acid analysis; James B. Holland for providing technical expertise in QTL analysis and discussion of results; Martha Ramirez for isolating RNA; and Carol Griffin for performing Northern analysis. This project was funded by the United Soybean Board (Grant # 59-.6645-2-071).

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Standards The experiments performed for this publication comply with the current laws of the United States of America.

References

- Aguilar F, Montandon PE, Stutz E (1991) Two genes encoding the soybean translation elongation factor eEF-1 alpha are transcribed in seedling leaves. *Plant Mol Biol* 17:351–360
- Bachlava E, Dewey RE, Auclair J, Wang S, Burton JW, Cardinal AJ (2008a) Mapping genes encoding microsomal ω -6 desaturase enzymes and their cosegregation with QTL affecting oleate content in soybean. *Crop Sci* 48:640–650
- Bachlava E, Burton JW, Brownie C, Wang S, Auclair J, Cardinal AJ (2008b) Heritability of oleic acid content in soybean seed oil and its genetic correlation with fatty acid and agronomic traits. *Crop Sci* 48:1764–1772
- Bachlava E, Dewey RE, Burton JW, Cardinal AJ (2009a) Mapping candidate genes for oleate biosynthesis and their association with unsaturated fatty acid seed content in soybean. *Mol Breed* 23:337–347
- Bachlava E, Dewey RE, Burton JW, Cardinal AJ (2009b) Mapping and comparison of Quantitative Trait Loci for oleic acid seed content in two segregating soybean populations. *Crop Sci* 49:1–10
- Brim CA (1966) A modified pedigree method of selection in soybean. *Crop Sci* 6:220
- Burkey KO, Booker FL, Pursley WA, Heagle AS (2007) Elevated carbon dioxide and ozone effects on peanut: II. Seed yield and quality. *Crop Sci* 47:1488–1497
- Burton JW, Wilson RF, Brim CA (1983) Recurrent selection in soybean: IV. Selection for increased oleic acid percentage in seed oil. *Crop Sci* 23:744–747
- Burton JW, Carter TE, Huie EB (1994a) Registration of ‘Brim’ soybean. *Crop Sci* 34:301
- Burton JW, Wilson RF, Brim CA (1994b) 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 JW, Wilson RF, Rebetzke GJ, Pantalone VR (2006) Registration of N98-4445A mid-oleic soybean germplasm line. *Crop Sci* 46:1010–1012
- Cardinal AJ, Lee M, Sharopova N, Woodman-Clikeman WL, Long MJ (2001) Genetic mapping and analysis of quantitative trait loci for resistance to stalk tunneling by the European corn borer in maize. *Crop Sci* 41:835–845
- Cardinal AJ, Burton JW, Camacho-Roger AM, Yang JH, Wilson RF, Dewey RE (2007) Molecular analysis of soybean lines with low palmitic acid content in the seed oil. *Crop Sci* 47:304–310
- Cardinal AJ, Dewey RE, Burton JW (2008) Estimating the individual effects of the reduced palmitic acid alleles *fap_{nc}* and *fap1* alleles on agronomic traits in two soybean populations. *Crop Sci* 48:633–639
- Cardinal AJ, Burton JW, Camacho-Roger A, Whetten R, Chappell AS, Bilyeu KD, Auclair J, Dewey RE (2011) Molecular analysis of *GmFAD3A* in two soybean populations segregating for the *fan*, *fap1*, and *fap_{nc}* loci. *Crop Sci* 51:2104–2112
- Cherrak CM, Pantalone VR, Meyer EJ, Ellis DL, Melton SL, West DR, Mount JR (2003) Low-palmitic, low-linolenic soybean development. *J Am Oil Chem Soc* 80:539–543
- Cooper M, DeLacy IH (1994) Relationships among analytical methods used to study genotypic variation and genotype-by-environment interaction in plant breeding multi-environment experiments. *Theor Appl Genet* 88:561–572
- Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung L, Specht JE (1999) An integrated genetic linkage map of the soybean genome. *Crop Sci* 39:1464–1490
- Fehr WR, Caviness CE (1977) Stages of soybean development. Special Report. Agriculture and Home Economics Experiment Station. Iowa State University 80:11
- Fehr WR, Welke GA, Hammond EG, Duvick DN, Cianzio SR (1991) Inheritance of reduced palmitic acid content in seed oil of soybean. *Crop Sci* 31:88–89
- Hallauer AR, Miranda JB (1988) Quantitative genetics in maize breeding, 2nd edn. Iowa State University Press, Ames
- Hebsgaard SM, Korning PG, Tolstrup N, Engelbrecht J, Rouzé P, Brunak S (1996) Splice site prediction in *Arabidopsis thaliana* Pre-mRNA by combining local and global sequence information. *Nucleic Acids Res* 24:3439–3452
- Henderson MM (1991) Correlations between fatty acid intake and cancer incidence. In: Nelson GJ (ed) Health effects of dietary fatty acids. American Oil Chemists’ Society, Champaign, pp 136–149
- Horejsi TF, Fehr WR, Welke GA, Duvick DN, Hammond EG, Cianzio SR (1994) Genetic control of reduced palmitate content in soybean. *Crop Sci* 34:331–334. doi:10.2135/cropsci1994.001183X003400020003x
- Hu FB, Stampfer MJ, Manson JE, Rimm E, Colditz GA, Rosner BA, Hennekens CH, Willett WC (1997) Dietary fat intake and the risk of coronary heart disease in women. *N Engl J Med* 337:1491–1499
- Hyten DL, Choi IY, Song Q, Specht JE, Carter TE, Shoemaker RC, Hwang EY, Matukumalli LK, Cregan PB (2010) A high density integrated genetic linkage map of soybean and the development of a 1,536 Universal Soy Linkage Panel for QTL mapping. *Crop Sci* 50:960–968. doi:10.2135/cropsci2009.06.0360
- Kenworthy WJ (1996) Registration of ‘Corsica’ soybean. *Crop Sci* 36:1078
- Kinoshita T, Rahman SM, Anai T, Takagi Y (1998) Inter-locus relationship between genes controlling palmitic acid contents in soybean mutants. *Breed Sci* 48:377–381

- Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln ES, Newburg L (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174–181
- Li Z, Wilson RF, Rayford WE, Boerma HR (2002) Molecular mapping genes conditioning reduced palmitic acid content in N87-2122-4 soybean. *Crop Sci* 42(2):373–378
- Littell RC, Milliken GA, Stroup WW, Wolfinger RD (1996) SAS[®] system for mixed models. SAS Institute Inc, Cary
- Long M, Deutsch M (1999) Association of intron phases with conservation at splice site sequences and evolution of spliceosomal introns. *Mol Biol Evol* 16:1528–1534
- Lorieux M, Perrier X, Goffinet B, Lanaud C, González de León D (1995) Maximum-likelihood models for mapping genetic markers showing segregation distortion. 2. F2 populations. *Theor Appl Genet* 90:81–89
- Moniz de Sa M, Drouin G (1996) Phylogeny and substitution rates of angiosperm actin genes. *Mol Biol Evol* 13:1198–1212
- Ndzana X, Fehr WR, Welke GA, Hammond EG, Duvick DN, Cianzio SR (1994) Influence of reduced palmitate content on agronomic and seed traits of soybean. *Crop Sci* 34:646–649
- Nickell AD, Wilcox JR, Cavins JF (1991) Genetic relationships between loci controlling palmitic and linolenic acids in soybean. *Crop Sci* 31:1169–1171
- Primomo VS (2000) Inheritance and stability of palmitic acid alleles in soybeans (*Glycine max* L Merr.). MS Thesis, Department of Plant Agriculture, University of Guelph, Guelph, Ontario, Canada
- Primomo 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 SM, Kinoshita T, Anai T, Takagi Y (1996) Genetic analysis of palmitic acid content using two soybean mutants. *Breed Sci* 46:343–347
- Rebetzke GJ, Burton JW, Carter TE, Wilson RF (1998a) Changes in agronomic and seed characteristics with selection for reduced palmitic acid content in soybean. *Crop Sci* 38:297–302
- Rebetzke GJ, Burton JW, Carter TE, Wilson RF (1998b) Genetic variation for modifiers controlling reduced saturated fatty acid content in soybean. *Crop Sci* 38:303–308
- Rebetzke GJ, Pantalone VR, Burton JW, Carter TE, Wilson RF (2001) Genetic background and environmental influence palmitate content of soybean seed oil. *Crop Sci* 41:1731–1736
- Rose AB, Elfersi T, Parra G, Korfa I (2008) Promoter-proximal introns in *Arabidopsis thaliana* are enriched in dispersed signals that elevate gene expression. *Plant Cell* 20:543–551
- Sambrook J, Russell DW (2001) Molecular cloning, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- SAS Institute (2010) SAS/STAT software. Release 9.2. SAS Institute, Cary
- SAS Institute Inc (2004) SAS/STAT software: release 9.1. SAS Institute Inc, Cary
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463:178–183
- Schneibly SR, Fehr WR, Welke GA, Hammond EG, Duvick DN (1994) Inheritance of reduced and elevated palmitate in mutant lines of soybean. *Crop Sci* 34:829–833
- Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delanay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. *Theor Appl Genet* 109:122–128
- Song Q, Gaofeng J, Zhu Y, Grant D, Nelson R, Hwang EY, Hyten DL, Cregan PB (2010) Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR-1.0) in soybean. *Crop Sci* 50:1950–1960
- Stojšin D, Ablett GR, Luzzi BM, Tanner JW (1998) Use of gene substitution values to quantify partial dominance in low palmitic acid soybean. *Crop Sci* 38:1437–1441
- Vandesompele JK, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:1–12. doi:10.1186/gb-2002-3-7-research0034
- Vries BD, Fehr WR, Welke GA, Dewey RE (2011) Molecular characterization of the mutant **fap3**(A22) allele for reduced palmitate concentration in soybean. *Crop Sci* 51:1611–1616
- Wilcox RF, Cavins JF (1990) Registration of C1726 and C1727 soybean germplasm with altered levels of palmitic acid in seed oil of soybean. *Crop Sci* 30:240
- Wilcox JR, Athow KL, Abney TS, Laviolette FA, Richards TL (1980) Registration of century soybean1 (Reg. No. 135). *Crop Sci* 20:415. doi:10.2135/cropsci1980.0011183X002000030043x
- Wilcox JR, Burton JW, Rebetzke GJ, Wilson RF (1994) Transgressive segregation for palmitic acid in seed oil of soybean. *Crop Sci* 34:1248–1250
- Wilson RF, Burton JW, Novitzky WP, Dewey RE (2001a) Current and future innovations in soybean (*Glycine max* L. Merr.) oil composition. *J Oleo Sci* 50:353–358
- Wilson RF, Marquardt TC, Novitzky WP, Burton JW, Wilcox JR, Kinney AJ, Dewey RE (2001b) Metabolic mechanisms associated with alleles governing the 16:0 concentration of soybean oil. *J Am Oil Chem Soc* 78:335–340