Genotyping-by-Sequencing-Based Investigation of the Genetic Architecture Responsible for a ~Sevenfold Increase in Soybean Seed Stearic Acid

Crystal B. Heim* and Jason D. Gillman*,†,1
*Division of Plant Sciences, University of Missouri, Columbia, Missouri 65211 and †USDA-ARS, Columbia, Missouri, 65211

ABSTRACT Soybean oil is highly unsaturated but oxidatively unstable, rendering it nonideal for food applications. Until recently, the majority of soybean oil underwent partial chemical hydrogenation, which produces *trans* fats as an unavoidable consequence. Dietary intake of *trans* fats and most saturated fats are conclusively linked to negative impacts on cholesterol levels and cardiovascular health. Two major soybean oil breeding targets are: (1) to reduce or eliminate the need for chemical hydrogenation, and (2) to replace the functional properties of partially hydrogenated soybean oil. One potential solution is the elevation of seed stearic acid, a saturated fat which has no negative impacts on cardiovascular health, from 3 to 4% in typical cultivars to > 20% of the seed oil. We performed QTL analysis of a population developed by crossing two mutant lines, one with a missense mutation affecting a *stearoyl-acyl-carrier protein desaturase* gene resulting in ~11% seed stearic acid crossed to another mutant, A6, which has 24–28% seed stearic acid. Genotyping-by-sequencing (GBS)-based QTL mapping identified 21 minor and major effect QTL for six seed oil related traits and plant height. The inheritance of a large genomic deletion affecting chromosome 14 is the basis for largest effect QTL, resulting in ~18% seed stearic acid. This deletion contains *SACPD-C* and another gene(s); loss of both genes boosts seed stearic acid levels to ≥ 18%. Unfortunately, this genomic deletion has been shown in previous studies to be inextricably correlated with reduced seed yield. Our results will help inform and guide ongoing breeding efforts to improve soybean oil oxidative stability.

KEYWORDS
Glycine max
soybean
fatty acid composition
genotyping-by-sequencing
QTL mapping

Soybean [*Glycine max* (L.) Merr.] is the source of the most widely consumed edible oil in the US, representing 55% of total oil consumption (http://soystats.com/2015-soystats/). Soybean oil and its derivatives are ubiquitous ingredients in packaged foods, especially in cookies and other snack foods (Van Camp *et al.* 2012). Soybean oil from a "typical" cultivar is comprised of five main fatty acids: palmitic, stearic, oleic, linoleic, and linolenic (11, 3–4, 25, 52, and 8%, respectively) (Fehr 2007). Due to the high proportion of polyunsaturated fatty acids, which are oxidatively unstable (Frankel 1991), the majority of soybean oil is chemically hydrogenated in order to increase stability and shelf life. Unfortunately, the process of hydrogenation produces *trans* fats, which have been tied to increased risk of heart disease and elevated low-density lipoprotein (LDL) levels in blood serum (Hunter *et al.* 2009). In 2003, the Food and Drug Administration (FDA) mandated labeling of packaged foods containing *trans* fats in amounts exceeding 0.5 g per serving (Van Camp *et al.* 2012), culminating in the removal of the "Generally Regarded As Safe" status for partially hydrogenated foods, due to the presence of *trans* fats. (http://www.fda.gov/Food/IngredientsPackagingLabeling/FoodAdditivesIngredients/ucm449162.htm accessed 08-24-2016).

For these reasons, a major soybean breeding target has been modification of soybean oil composition, so as to reduce or eliminate the need for chemical hydrogenation. Achievement of this goal would potentially translate to decreased *trans* fat intake for consumers, as well as fewer steps in processing for manufacturers. Biotechnological and traditional breeding have independently increased the oleic acid content of soybean oil and thereby increased oxidative stability (Bilyeu *et al.* 2011; Buhr *et al.* 2002; Pham *et al.* 2010). Soybean oil high in oleic acid (> 70%) has been shown to be more chemically stable than typical soybean oil (Warner and Gupta 2005). However, both conventional and high oleic acid

doi: 10.1534/g3.116.035741
Manuscript received September 19, 2016; accepted for publication November 15, 2016; published Early Online November 18, 2016.
Supplemental material is available online at www.g3journal.org/lookup/suppl/doi:10.1534/g3.116.035741/-/DC1.
[1]Corresponding author: 205 Curtis Hall, University of Missouri, Columbia, MO 65211. E-mail: Jason.Gillman@ars.usda.gov

Table 1 Details on final GBS linkage map for A6 × 194D cross

Chromosome	Marker #	Length (cM)	Average Spacing (cM)	Maximum Spacing (cM)
1	123	67.5	0.6	15.5
2	190	131.8	0.7	13.1
3	115	125.7	1.1	73.7
4	197	136.0	0.7	52.5
5	145	117.8	0.8	19.2
6	233	141.1	0.6	16.4
7	61	66.0	1.1	18.0
8	95	162.7	1.7	51.0
9	196	62.6	0.3	6.9
10	219	122.6	0.6	12.7
11	38	111.9	3.0	65.6
12	49	125.5	2.6	62.0
13	180	126.2	0.7	33.0
14	151	85.9	0.6	30.3
15	203	107.4	0.5	25.9
16	102	118.8	1.2	62.0
17	165	45.3	0.3	8.1
18	190	126.0	0.7	49.8
19	191	105.4	0.6	27.8
20	134	115.2	0.9	16.9
Overall	2977	2201.4	0.7	73.7

Table 2 Statistical analysis of seven phenotypic traits in soybean

Trait	Source of Variation	MS	F		h^2
Plant height	Location	16050.15	84.85	***	0.54
	Line	387.27	2.05	***	
	Location × line	208.22	1.10		
	Rep	874.84	4.62	**	
	Error	189.16			
Total seed oil	Location	146.55	88.55	***	0.28
	Line	3.07	1.86	***	
	Location × line	2.57	1.55	***	
	Rep	1.97	1.19		
	Error	1.66			
Palmitic	Location	12.83	34.25	***	0.89
	Line	2.39	6.38	***	
	Location × line	0.33	0.89		
	Rep	0.17	0.46		
	Error	0.38			
Stearic	Location	300.94	48.50	***	0.88
	Line	58.78	9.47	***	
	Location × line	9.37	1.51	***	
	Rep	0.84	0.14		
	Error	6.20			
Oleic	Location	1278.44	216.52	***	0.81
	Line	42.49	7.20	***	
	Location × line	9.95	1.69	***	
	Rep	15.35	2.60		
	Error	5.90			
Linoleic	Location	925.85	112.96	***	0.73
	Line	53.09	6.48	***	
	Location × line	16.93	2.07	***	
	Rep	10.57	1.29		
	Error	8.20			
Linolenic	Location	337.56	666.36	***	0.72
	Line	2.27	4.48	***	
	Location × line	0.79	1.55	***	
	Rep	1.31	2.58		
	Error	0.51			

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. MS, mean sum of squares; F, value; h^2, broad sense heritability.

soybean oil have very low melting points, and for many applications fully hydrogenated oil is more suitable (Ribeiro *et al.* 2009). For instance, edible oil high in saturated fats is desirable for solid fat baking applications (Renzyaeva 2013; Tarancón *et al.* 2013).

Interestingly, new research has shown that stearic acid (C18:0), a long-chain saturated fatty acid, does not have the same cholesterolemic effects in humans as shorter acyl-chain saturated fats (Hunter *et al.* 2009; Yu *et al.* 1995); it is effectively "heart-neutral." Unfortunately, soybean oil generally contains only ~3–4% stearic acid (Wilson 2004), and > 20% must be achieved to meet current market demands (K. Whiting, Smith-Bucklin/United Soybean Board, personal communication).

Elevated stearic acid is an apparently rare trait; out of 21,849 USDA-GRIN soybean collection entries that have been evaluated for oil composition, only a single entry has stearic acid > 9% (https://npgsweb.ars-grin.gov/gringlobal/descriptors.aspx, accessed 08-19-2016). It is known that soybean seed stearic acid content is determined largely by alterations in the *Stearoyl-Acyl Carrier Protein Desaturase-C* (*SACPD-C*) gene (Glyma14g27990). *SACPD-C* is a soluble enzyme that selectively desaturates stearic acid precursors (C18:0) to oleic acid precursors (C18:1) (Zhang *et al.* 2008). It has been previously reported that *SACPD-C* is a major determinant of seed stearic acid levels (Ruddle *et al.* 2013a,b, 2012; Gillman *et al.* 2014).

To date, there is only one known naturally occurring source of elevated stearic acid, FAM94-41, which contains a missense mutation in *SACPD-C* and ~13% seed stearic acid (Zhang *et al.* 2008). All other lines with elevated stearic acid are due to mutagenesis (Bolon *et al.* 2011; Gillman *et al.* 2014; Rahman *et al.* 1995, 1997). The majority of lines contain 8–12.5% seed stearic acid, but one notable exception is mutant line A6, which has been reported to have arisen by sodium azide-induced mutagenesis (Hammond and Fehr 1983), but which contains several large genomic deletions more consistent with radiation-induced mutagenesis. A6 seeds exhibit a dramatic increase in seed stearic acid levels (~24–28%, as compared to the wild-type 4%). A6 bears several large genomic deletions, including a large portion of chromosome 14 (~6,221,000 bp) corresponding to ~1/8 of the chromosome (Gillman *et al.* 2014). Regrettably, A6 is also extremely agronomically deficient, exhibiting poor germination, low seed yield, reduced seed quality, extreme early maturity (MG 0), and short plant stature when alleles for elevated stearic acid are present (Lundeen *et al.* 1987; Hammond and Fehr 1983). Some of the defects are likely due to one (or more) of the significant genomic deletions (Gillman *et al.* 2014).

We identified multiple *SACPD-C* deletion and missense mutation lines from multiple genetic backgrounds and, except for line A6, loss of functional *SACPD-C per se* can only elevate seed stearic acid levels to ~13%, regardless of mutation or deletion (Gillman *et al.* 2014). The identification of multiple mutants with genetically controlled variation in elevated stearic acid content presented an opportunity to investigate the genetic basis for the extremely elevated stearic acid content (~24%) in line A6. To that end, a recombinant inbred line (RIL) population was developed by crossing a line with ~11% stearic acid (194D) to a line with ~24–28% stearic acid (A6). We then employed a GBS method (Elshire *et al.* 2011), which is routine in maize, but remains less utilized in soybean. GBS is a simple, repeatable, and robust method for identifying single nucleotide polymorphisms (SNPs) via sequencing of reduced representation libraries produced with methylation-sensitive restriction enzymes, which bias against highly repetitive and gene-poor genomic regions (Poland *et al.* 2012).

A dense genetic linkage map was constructed from 2977 high-quality GBS markers. Using phenotypic data from 173 RILs over three

Table 3 Descriptive statistics of traits of the RIL populations and the parents

		RILs			Parents		
	Trait	Average[a]	Range	SD[b]	A6		194D
BR 2014	Height (cm)	49.31	22.50–87.50	11.13	46.25	*	64.17
	Total seed oil (%)	18.63	15.04–21.44	1.13	15.84		19.81
	Palmitic (% seed oil)	8.98	7.14–11.16	0.67	8.70		9.62
	Stearic (% seed oil)	15.12	7.83–26.11	3.44	23.88	*	9.97
	Oleic (% seed oil)	21	15.17–33.29	3.38	18.57		18.05
	Linoleic (% seed oil)	48.06	33.09–55.40	3.83	42.60	**	54.57
	Linolenic (% seed oil)	6.85	4.47–9.37	0.86	6.25	**	8.05
2014 SF	Height (cm)	63.82	31.67–86.67	11.01	40		68.75
	Total seed oil (%)	17.3	14.16–19.57	1.15	11.85		18.59
	Palmitic (% seed oil)	9.44	8.42–10.75	0.51	9.13		9.77
	Stearic (% seed oil)	14.34	9.66–20.99	2.87	22.32	*	11.77
	Oleic (% seed oil)	18	15.13–23.43	1.75	16.92		18.67
	Linoleic (% seed oil)	49.9	44.46–54.42	2.43	43.51	*	48.92
	Linolenic (% seed oil)	8.32	6.82–10.30	0.69	8.13		8.12
RB 2015	Height (cm)	56.8	38.00–91.11	9.42	60.75		78.94
	Total seed oil (%)	17.79	14.91–19.47	0.73	18.55		17.02
	Palmitic (% seed oil)	9.19	7.73–11.30	0.68	8.76		9.09
	Stearic (% seed oil)	15.91	6.03–24.76	3.35	20.85	**	12.67
	Oleic (% seed oil)	21.69	17.24–33.56	2.84	18.76	**	22.19
	Linoleic (% seed oil)	46.76	34.77–53.35	3.51	44.88	*	49.13
	Linolenic (% seed oil)	6.45	4.634–7.87	0.58	6.75		6.91

* Parent values are significantly different at $P < 0.05$; ** parent values are significantly different at $P < 0.01$. RIL, recombinant inbred line; BR, Bradford Research Center; SF, South Farm; RB, Hinkson Field.
[a] Full details on individual RILs are in Supplemental Materials.
[b] SD, standard deviation.

environments, 21 QTL were identified for seven traits in soybean, including four QTL for seed stearic acid content.

MATERIALS AND METHODS

Plant material and population development

Two RIL populations were developed, derived from the cross between A6 and 194D (designated population 2), and the reciprocal cross between 194D and A6 (designated population 3). Population 2 consisted of 100 individuals in 2014 and 2015, and population 3 consisted of 95 individuals in 2014 and 94 individuals in 2015. Walter Fehr generously provided seeds from the high stearic acid line, A6. 194D is a midstearic line (~13%) that has a V211E mutation affecting *SACPD-C* identified by Gillman *et al.* (2014) from an EMS-induced TILLiNG population created by Kristin Bilyeu, described previously (Cooper *et al.* 2008).

Both populations and their parent lines were grown in three distinct locations: (1) South Farm in 2014 (Columbia, MO, Latitude 38.908189, Longitude −92.278693, Mexico silt loam soil); (2) Bradford Research Center in 2014 (Columbia, MO, Latitude 38.893990, Longitude 92.196405, Mexico silt loam soil); and (3) Hinkson Field in 2015 (Columbia, MO, Latitude 38.928015, Longitude −92.351425, Haymond silt loam soil). Both populations were grown in single-row plots in randomized complete block design with three replicates over all locations in both years, except at South Farm in 2014, where population 2 was not grown due to severe weather constraints and seed availability during planting season. A small number of lines were found to be unrelated to A6 or 194D via GBS markers and were excluded; the final population size was 173 RILs.

Quantification of seed traits

The $F_{4:5}$ and the $F_{5:6}$ populations were harvested at physiological maturity and stored in a ventilated grain bin until threshing. After threshing, seed were moved to a temperature and humidity controlled facility (Sears plant growth facility, University of Missouri, 4° and 39% relative humidity) for a minimum of four weeks prior to subsequent analyses, in order to minimize differences in seed moisture content between samples. Samples were allowed to warm in a ~25° lab overnight and then total seed oil content (percent w/w) was estimated using NMR spectroscopy (La *et al.* 2014). A benchtop Oxford MQC NMR machine (Oxford Instruments, Oxford, UK) was used to analyze 5 g seed samples. Total seed oil content is expressed as the percentage of total weight of the seeds. As part of an independent project, seed from 200 RILs produced in SF 2014 and RB 2015 were scanned on a NIRS monochromator model FOSS 6500 (FOSS North America, Eden Prairie, MN) using the transport quarter cup (dimension 97 × 55 mm). The NIRS reflectance (R) spectra were collected at 2 nm intervals in the NIRS region of 400–2500 nm at room temperature. Calibrations previously developed by FOSS were used to estimate moisture content. In 200 samples from multiple plots stored in this manner, seed moisture content was found to be minimally variable (7.7 ± 0.4% SD).

Seed samples were also analyzed for seed fatty acid composition using the derivatized lipid gas chromatography method (Bilyeu *et al.* 2011). In 2014, two samples of three seeds each per plot were analyzed, and in 2015, four or five individual seeds per plot were analyzed. Seeds were crushed then extracted in 1 ml chloroform-hexane-methanol (8:5:2, v/v/v) overnight. Derivatization of 150 µl extracted oils was done with 75 µl methylating reagent (0.25 M methanolic sodium methoxide-petroleum ether-ethyl ether, 1:5:2, v/v/v/). Seed samples were then diluted with hexane to 1 ml. An Agilent (Palo Alto, CA) series 6890 capillary gas chromatograph fitted with a flame ionization detector (275°) was used with an AT-Silar capillary column (Alltech Associates, Deerfield, IL). Standard fatty acid mixtures (Animal and Vegetable Oil Reference Mixture 6, AOACS) were used as reference standards.

Quantification of field traits

Plant height was measured in centimeters on three separate plants within each plot (three plots per genotype) when 95% of the pods showed mature pod color.

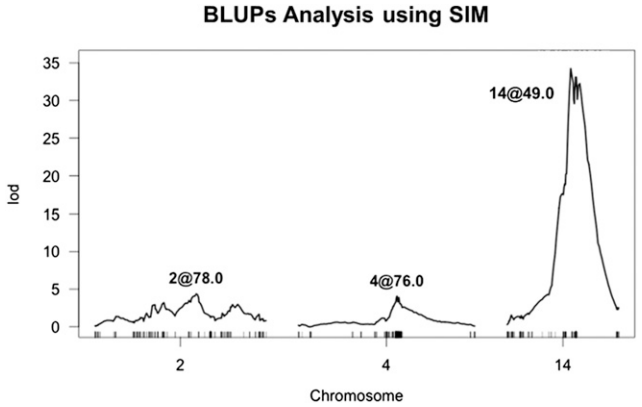

Figure 1 QTL results for the seed stearic acid trait using BLUPs and the SIM method. BLUPs, best linear unbiased predictors; LOD, logarithm of the odds; QTL, quantitative trait loci; SIM, standard interval mapping.

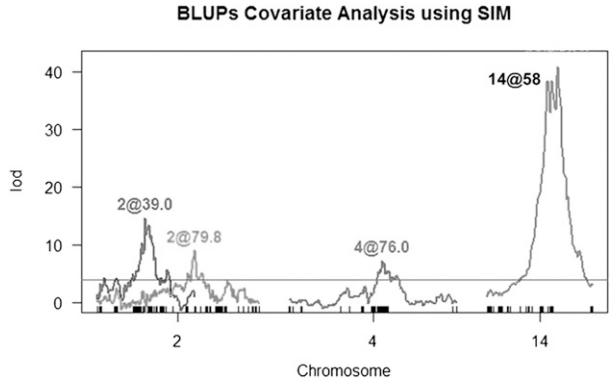

Figure 2 QTL results using BLUPs and the most significant marker associated with stearic acid as a covariate in SIM. LOD threshold is shown in red at LOD = 4. BLUPs, best linear unbiased predictors; LOD, logarithm of the odds; QTL, quantitative trait loci; SIM, standard interval mapping.

Statistical analyses

ANOVA was performed using the "anova" and "aov" functions in R (R Foundation for Statistical Computing, Vienna, Austria 2008) to determine if there were significant differences between any locations or blocks. Pearson's correlation coefficients between phenotypes were determined using PROC CORR in SAS (SAS Institute Inc., 2002).

Best linear unbiased predictors (BLUPs) were calculated using the lme4 package in R (R Foundation for Statistical Computing). All effects were considered random. Heritability was calculated in the broad-sense (h^2) as follows:

$$h^2 = \sigma_g^2 \Big/ \left\{ \sigma_g^2 + \left(\sigma_{ge}^2 \Big/ n \right) + \left(\sigma^2 / rn \right) \right\}$$

where h^2 = heritability, σ_g^2 = genotypic variance among RILs, σ_{ge}^2 = genotype × environment (location), σ^2 = error variance, r = number of reps, and n = number of environments.

DNA isolation and gGBS

DNA was isolated from ~40 mg of lyophilized leaf tissue using the DNeasy Plant Mini kit (QIAGEN, Valencia, CA) according to the manufacturer's instructions. DNA samples from 190 lines from populations 2 and 3 were submitted to the Institute for Genomic Diversity (IGD) at Cornell University. GBS libraries were prepared as previously described (Elshire *et al.* 2011; Swarts *et al.* 2014) using the *ApeKI* enzyme, DNA ligase, and appropriate Illumina adapters. The IGD staff performed the library construction, read mapping, and downstream SNP calls. A total of 455,924,779 reads were produced, 428,671,333 of which were used for downstream analysis. A total of 61.5% of the reads were successfully aligned to unique positions in the *G. max* Wm82.a2.v1 Williams 82 reference sequence (Schmutz *et al.* 2010, http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Gmax) using the BWA 0.7.8-r455 program (Li and Durbin 2009). SNPs were called using the TASSEL 5.0 pipeline (Bradbury *et al.* 2007) resulting in 33,728,018 tags for all 190 samples.

A total of 27,672 SNPs were identified, with an average SNP coverage of 10.98-fold. SNPs were filtered using TASSEL 5.0 to exclude those with > 20% missing data, and to restrict allele frequency to between 0.2 and 0.8, bringing the SNP count down to 5423. Lastly, a chi-square test was performed to remove any severely distorted markers (at $F_{4:5}$), resulting in a final count of 2977 high-quality markers for mapping. Parental allele assignment and imputation were performed using the Window LD function of FsFHap (Swarts *et al.* 2014) incorporated into TASSEL 5.0.

Linkage map construction

Linkage map construction was performed in R/qtl (R Foundation for Statistical Computing) using the 2977 polymorphic, high-quality SNP markers obtained from GBS (Supplemental Material, Figure S1, File S1, File S2, and Table 1). Genetic distances between markers were determined using the est.map function, with an estimated genotyping error rate of 0.01. Iterative evaluation of chromosomes with excessive map distances (> 200 cM) was performed using droponemarker and est. map functions. Finally, all chromosomes were evaluated for correct marker order using the ripple function, using five markers at a time. No better order was identified than that assigned by the position in the Wm82.a2.v1 assembly.

Covariate analysis and QTL mapping

BLUP phenotypes were used for a one-dimensional QTL scan using the SIM method to determine the most significant marker associated with seed stearic acid for use in covariate analysis. Marker c14.loc58 (chromosome 14 @ 58 cM, Gm14:42206409) was identified as significantly associated with stearic acid and was subsequently designated as an interactive covariate using the pull.geno and cbind functions in R/qtl. Further analysis using this interactive covariate was performed using BLUPs values and the SIM method through the "scanone" function.

QTL analysis was performed using the qtl package in R (Broman and Sen 2009). Because R/qtl assumes an F_2, the cross file was first converted into an F_4 using the BC_SF_T tool (Shannon *et al.* 2013). Conditional genotypic probabilities were calculated using an error probability rate of 0.01, due to the inherent likelihood of genotyping errors of the GBS method (Elshire *et al.* 2011). Single QTL analysis was run using the Standard Interval mapping method as described in Broman and Sen (2009). Significance thresholds were calculated using 1000 permutation tests for each phenotype, but the extremely large effects from the *sacpd-c* deletion resulted in substantially elevated LOD scores for oil related traits; in general, the only QTL above threshold was the *sacpd-c* deletion. As a result, we set a LOD significance threshold of LOD > 4. Interactions between QTL and additional QTL were examined using the "addint" and "addqtl" functions. Final model fit was determined by dropping one QTL at a time using the "fitqtl" function. QTL were refined using the "refineqtl" function and fit to obtain a final model.

Table 4 SIM and SIM + covariate analysis for stearic acid-related QTL using BLUPs

Trait	Method	Chromosome	SNP Marker Nucleotide	Position	LOD	R^2	1.5 Interval
Stearic acid	SIM	2	Gm2:15552879	78.00	4.38	4.4	44–111
		4	Gm4:18312993	76.00	4.08	4.1	72–83
		14	Gm14:42206409	49.00	34.25	53.2	48–54
Stearic acid	SIM + Cov	2	Gm2:5946912	39.00	14.63	13.75	38–42
		2	Gm2:15552879	79.79	9.05	7.9	78–80
		4	Gm4:18312993	75.90	7.23	6.1	73–78
		14	c14.loc58	58.00	40.84	56.8	57–59

LOD, logarithm of the odds; SIM, SIM, standard interval mapping; Cov, covariate analysis.

Confidence intervals for QTL locations are presented as 1.5-LOD intervals. QTL were evaluated for overlap with selected literature (listed in results) and historical QTL using genetic/physical map tools available online (http://soybase.org/search/qtllist_by_symbol.php, accessed 01-01-2016 through 05-30-2016). The effect of each QTL was estimated in R/QTL using averaged phenotypic data over the three study locations via the effectplot command, following sim.geno with 1000 draws and an error probability of 0.01.

Data availability

The authors state that all data necessary for confirming the conclusions presented in the article are fully represented either within the article or within Supplemental Materials.

RESULTS

Phenotypic distributions

Analysis of variance showed that there were no significant differences between population 2 and population 3, so we conclude that there is no significant cytoplasmic effect on any of the traits evaluated in this study. Full results of the statistical analysis are presented in Table 2. There was significant genetic variation among RILs for all traits (Table 2), as well as a significant location effect for all traits ($P < 0.05$). The two parents, A6 and 194D, showed significant differences in all traits examined except oleic acid ($P < 0.05$). There were highly significant line and environment effects for all traits ($P < 0.001$) and a genotype × environment interaction was significant for all traits except palmitic acid and plant height ($P < 0.01$). Heritability estimates were moderately high for the five main fatty acid species, the highest value being $h^2 = 0.89$ for palmitic acid.

Phenotypic distributions and parent values are shown in Figure S2 and Table 3. All traits were approximately normally distributed. Transgressive segregation was noted for all traits of interest, particularly fatty acid traits, which suggested the contribution of multiple genes and/or multiple genes of small effect.

Correlations between traits

Significant correlations ($P < 0.01$) were found between several traits with the strongest being between the five fatty acid species ($P < 0.0001$). Pearson's correlation coefficients and associated P-values are presented in Table S1. Significant coefficients between traits ranged between 0.10 and −0.81. The majority of significant correlations were detected in multiple environments for each trait. Stearic acid was significantly associated with all fatty acid traits, including total seed oil, in at least one environment.

Results using BLUPs

In order to attempt to lower the residual phenotypic variance even further, we chose to use the most significant marker on chromosome 14 as a covariate in all subsequent analyses. Using the same BLUPs as described above, we ran the SIM method in order to pick out the most significant marker associated with stearic acid on chromosome 14, (Figure 1). Once this marker was determined, the SIM was rerun with marker c14.loc58 as an interactive covariate. Using this method, the LOD scores and R^2 values were the highest, and an additional QTL was identified on chromosome 2 that no other method had identified previously (Figure 2). The 1.5-LOD intervals were also much narrower using this method than with SIM alone (the largest interval was only 5 cM wide for the QTL on chromosome 4) (Table 4).

QTL analysis

A total of 21 QTL were detected for seven traits on 9 of the 20 chromosomes. We observed colocation of QTL for several traits. The same significant region on chromosome 14 was detected for all oil traits examined (total seed oil, palmitic, stearic, oleic, linoleic, linolenic, and linolenic; see *Discussion*). The same QTL was detected on chromosome 2 for palmitic, stearic, and linolenic acids. A region on chromosome 4 was detected for stearic, oleic, linoleic, and linolenic acids. QTL were considered to be colocated when there was significant overlap between the 1.5-LOD intervals for each QTL. Phenotypic variation explained by each QTL ranged from 6.12% (*q.4s*) to 56.76% (*q.14s*) (Table 5). A6 alleles contributed to an increase in phenotype for 12 out of 21 QTL, while 194D alleles contributed to an increase in phenotype for 11 out of 21 QTL.

Total seed oil QTL

Two QTL were detected for total seed oil, one of minor effect on chromosome 9, accounting for 14.43% of the phenotypic variation, and one major effect QTL on chromosome 14, which accounted for 20.58% of the variation. For both loci, individuals homozygous for the 194D allele showed the highest total seed oil content. Mansur *et al.* (1993) detected an association between markers on chromosome 9 and total seed oil whose QTL intervals overlapped with *q.9oi*. There are several QTL published for total seed oil on chromosome 14 within 10 cM of *q.14oi* on the soybean consensus map (Chen *et al.* 2007; Csanadi *et al.* 2001; Eskandari *et al.* 2012; Liang *et al.* 2010; Qi *et al.* 2011).

Palmitic acid (C16:0) QTL

Three minor QTL were detected for palmitic acid. QTL on chromosome 2, 5, and 10 each accounted for 10–13% of the phenotypic variation. For all three loci, homozygosity for the 194D allele increased palmitic acid content. *Q.2p* mapped to the same genomic region as fatty acid desaturase gene *FAD3-B* (Glyma02g39230) (Bilyeu *et al.* 2003; Gillman and Bilyeu 2012). Two QTL for palmitic acid have been published in this region (Panthee *et al.* 2006; Reinprecht *et al.* 2006). Two QTL for palmitic acid have also been published in the same region of chromosome 5 as *q.5p* (Li *et al.* 2002; Wang *et al.* 2012). *Q.5p* mapped to the

Table 5 Summary of QTL results for seven traits in soybean A6 × 194D cross

Trait	QTL	Chr	SNP Marker Nucleotide	Position (cM)	1.5 Interval	LOD	R^2 (%)	QTL Effect
Palmitic acid	q.2p	2	Gm2:7138451	44.58	38–128	5.55	10.49	+0.36%
	q.5p	5	Gm5:1420686	2.48	0–8	6.79	13.06	+0.44%
	q.14p	14	Gm14:12506615	44.91	42–58	6.89	13.26	+0.44%
Stearic acid	q.2.1s	2	Gm2:5946912	39	38–42	14.63	13.75	−1.16%
	q.2.2s	2	Gm2:15552879	79.79	78–80	9.05	7.86	−0.52%
	q.4s	4	Gm4:18312993	75.99	73–78	7.23	6.12	+1.24%
	q.14s	14	Gm14:42206409	58	57–59	40.84	56.76	−5.00%
Oleic acid	q.4o	4	Gm4:19496796	76.38	74–80	6.57	12.17	+1.37%
	q.14o	14	Gm14:32290452	51	49–54	13.7	28.05	+2.9%
Linoleic acid	q.4e	4	Gm4:32351007	78.88	73–84	6.74	14.75	−2.31%
	q.14e	14	Gm14:34918500	54	45–63	4.48	9.49	+2.02%
Linolenic acid	q.2n	2	Gm2:15552879	79.79	78–81	7.05	10.4	+0.33%
	q.4n	4	Gm4:27332180	78.12	76–84	7.04	10.39	−0.21%
	q.8n	8	Gm8:44116750	150	148–152	6.93	10.2	−0.15%
	q.10n	10	Gm10:45310798	89.68	81–98	8.13	12.16	−0.42%
	q.13n	13	Gm13:41141355	118.19	108–125	6.48	9.48	−0.17%
	q.14n	14	Gm14:32372635	51.27	44–53	5.52	7.97	−0.25%
Plant height	q.10h	10	Gm10:44639359	90	88–94	9.79	20.9	−6.54 cm
	q.19h	19	Gm19:49398020	100.51	91–104	5.2	10.42	+5.03 cm
Oil	q.9oi	9	Gm9:6801513	16	15–17	6.85	14.43	+0.28%
	q.14oi	14	Gm14:42206409	62	49–68	9.43	20.58	+0.71%

"QTL Effect" represents the impact of converting homozygous A6 alleles to homozygous 194D alleles. QTL, quantitative trait loci; Chr, chromosome; SNP, single nucleotide polymorphism; LOD, logarithm of the odds.

same region as *FATB1-A* (Glyma05g26110), a keto-acyl ACP synthase gene (Cardinal *et al.* 2007). The map location of *q.14p* corresponded to *SACPD-C* (Glyma14g27990), a *stearoyl-acyl carrier protein desaturase* gene. No positional confirmation exists for this particular region; however, two QTL have been reported that are within 10 cM (Diers and Shoemaker 1992).

Stearic acid (C18:0) QTL

Four QTL were detected for seed stearic acid content; three minor effect QTL, two on chromosome 2, and one on chromosome 4, as well as one major effect QTL on chromosome 14. For the QTL on chromosomes 2 and 14 (Figure 3), homozygosity for the A6 allele contributed to elevated stearic acid content. *Q.14s* had an overwhelming major effect, which contributed 56.76% of the total variation in seed stearic acid content. *Q.14s* also mapped to the same genomic region as *SACPD-C*, which is also the same result we saw when examining palmitic acid. The two QTL identified on chromosome 2 accounted for 13.75 and 7.86% of the variation in seed stearic acid content. QTL associated with stearic acid have been previously reported on chromosome 2 within 5 cM of *q.2.1s* (Li *et al.* 2011; Xie *et al.* 2011). On chromosome 4, the 194D allele contributed to elevated stearic acid. *Q.2s* also interacted with *q.14s*, which contributed another 9.97% of the phenotypic variation. *Q.4s* may be a novel QTL, as no marker associations for seed stearic acid content have been previously reported on chromosome 4.

Oleic and linoleic acid (C18:1, C18:2) QTL

QTL for oleic acid content were detected on chromosomes 4 and 14, which accounted for 12.17 and 28.05% of the phenotypic variation, respectively. For both QTL, homozygosity for the 194D allele contributed to higher oleic acid levels. Nearly identical QTL for linoleic acid content were detected on chromosomes 4 and 14, which accounted for 14.75 and 9.49% of the phenotypic variation, respectively. On chromosome 4, individuals homozygous for the A6 allele had the highest linoleic acid content, while on chromosome 14, individuals homozygous for the 194D had the highest linoleic acid content. As with palmitic and stearic acid, the chromosome 14 QTL (*q.14o* and *q.14e*) aligned with *SACPD-C* deletion from A6. QTL for seed oleic as well as linoleic acid content have been identified in previous studies within 10 cM of *q.14o* and *q.14e* chromosome 14 (Diers and Shoemaker 1992; Kim *et al.* 2010). *Q.4o* and *q.4e* on chromosome 4 were effectively identical to *q.4s* for seed stearic acid and, like stearic acid, no QTL for oleic or linoleic acid have been previously published on this chromosome.

Linolenic acid (C18:3) QTL

Six QTL were detected for linolenic acid content on chromosomes 2, 4, 8, 10, 13, and 14, more than were detected for any other trait. All six QTL had minor effects, ranging from 7.97 to 12.16% of the phenotypic variation explained (60.6% total). Two of the QTL, *q.8n* and *q.13n* interacted epistatically (Table 6), and accounted for an additional 7.45% of the variation in linolenic acid content. Though all the QTL were of minor effect, we saw some positional overlap with published QTL for seed linolenic acid. Li *et al.* (2011) identified an association between Satt537 and seed linolenic acid content on chromosome 2 at nearly the same map position as *q.2n* (within 5 cM). In the same study, a QTL was identified on chromosome 4, however it was not in the same region as *q.4n*. One QTL for linolenic acid has been published on chromosome 8 within 5 cM of *q.8n* (Bachlava *et al.* 2009). A QTL on chromosome was reported (Shibata *et al.* 2008) in an adjacent region to *q.10n* on chromosome 10, not likely associated with *FAD2-1a*, as our interval was > 20 cM away on the genetic map. Lastly, a single QTL was previously reported on chromosome 13, however it was on the opposite side of the chromosome as *q.13n* (Hyten *et al.* 2004). Several markers have been reported to be associated with alterations in seed linolenic acid on chromosome 14, however they were all greater than 10 cM away from *SACPD-C* based on the consensus map (Bachlava *et al.* 2009; Reinprecht *et al.* 2006; Spencer *et al.* 2004; Xie *et al.* 2011).

Plant height QTL

QTL of major and minor effect for plant height were detected on chromosomes 10 and 19, respectively. *Q.10h* accounted for 20.90% of

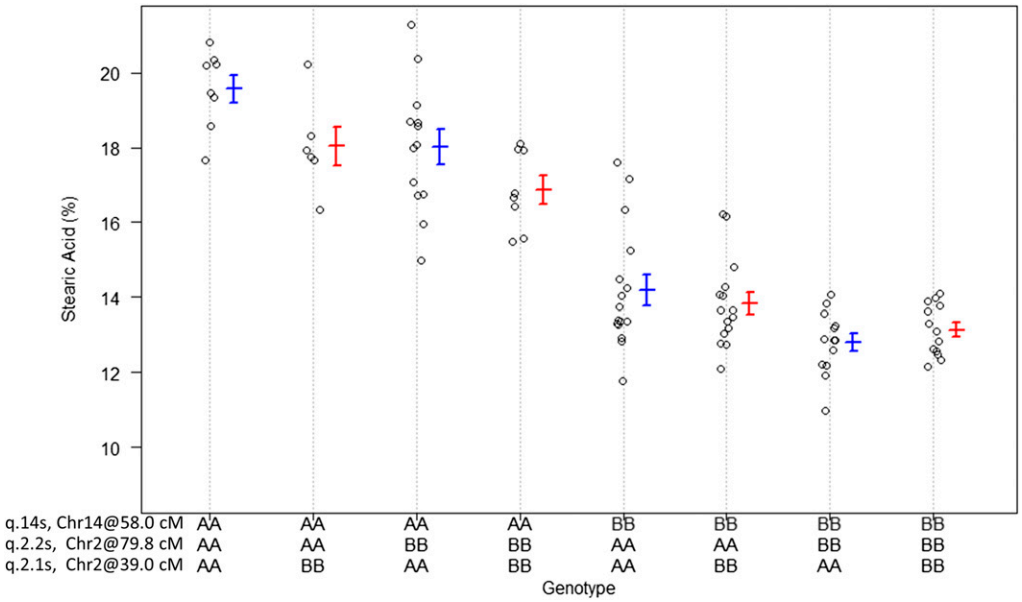

Figure 3 Interaction plot between *q.2.1s q.2.2s* and *q.14s*. Genotype "AA" indicates homozygosity for A6 alleles, allele "BB" indicates homozygosity for alleles from 194D. Middle bar indicates mean of genotypes, top and bottom bars indicate ± 1 SEM. Parental stearic acid contents averaged over all locations were: A6 = 22.4 ± 3.8%, 194D = 11.4 ± 1.6%. Chr, chromosome.

the variation, while *q.19h* accounted for 10.42% of the phenotypic variation. *Q.19h* mapped to nearly the same genomic region as the known maturity gene *E3* (Glyma19g224200) and linked gene *Dt1* (Glyma19g194300) (Watanabe *et al.* 2009; Tian *et al.* 2010). There were no published QTL near *q.10h*; however, at least one report (Wang *et al.* 2004) identified two chromosome 10 microsatellite markers associated with plant height within 10 cM of *q.10h*.

DISCUSSION

QTL mapping method selection in the soybean literature

The ability to detect QTL is entirely dependent on population structure and the level of genetic diversity within that population, as well as the number and quality of markers used in the analysis. Considering the multitude of available mapping strategies and software, as well as usage of individualized linkage maps of differing genetic lengths, it is difficult to compare historical mapping results with our study. However, several of our QTL in this study were positionally confirmed. In the public genetic database for soybean, there are very few QTL mapping studies published that analyzed seed stearic acid, in relation to the number of QTL published for other oil composition traits (www.soybase.org, accessed 05-01-2016). Four of the studies utilized the CIM method and successfully detected associations between markers and the stearic acid trait (Hyten *et al.* 2004; Panthee *et al.* 2006; Reinprecht *et al.* 2006; Wang *et al.* 2012). All other stearic acid QTL published in the database were determined by simple marker regression (Diers and Shoemaker 1992; Xie *et al.* 2011).

Difficulties in detecting QTL

A major confounding effect in our study was the extreme differences between the parent line A6, which exhibits a remarkably large deletion on chromosome 14 encompassing the entirety of the *SACPD-C* gene (as well as ~142 other genes), and the missense mutation in 194D (Gillman *et al.* 2014). The extreme phenotype values paired with the large deletion swelled the LOD score for the QTL on chromosome 14; LOD = 40.84 for stearic acid, which is extraordinarily high in the context of quantitative trait mapping. A QTL of such major effect makes it extremely difficult to detect other, relatively minor effect QTL. In addition, because of the genomic deletion on chromosome 14, any markers on chromosome 14 that showed a significant association with fatty acid traits flank the deletion. This experiment is the first to conduct QTL mapping using *SACPD-C* mutants, and as *SACPD*-C is a known factor in determining seed stearic acid levels, we can safely make the association between *q.14p*, *q.14s*, *q.14o*, *q.14e*, *q.14n*, and *q.14oi* and the genomic deletion that contains *SACPD-C*. The QTL on chromosome 14 was the same as discussed for palmitic, stearic, oleic, and linoleic acids, and is almost certainly due to the differences between the missense *SACPD-C* in 194D and the deletion in A6. *Q.14p*, *q.14s*, *q.14o*, *q.14e*, *q.14n*, and *q.14oi* correspond to *SACPD*-C (Glyma14g27990).

Ultimately, we chose to continue with the final method described, SIM with covariate analysis using BLUPs for our analysis. By calculating BLUPs, we were able to take environmental variance into account, which strengthened the association between each marker and the stearic acid phenotype. We sought to further reduce the residual phenotypic variation by conducting covariate analysis. When a locus of strong effect exists in a population, it can mask the associations between any small effect loci and confound understanding of genetic architecture behind the phenotype (Xu 2003; Mackay 2001; Broman and Sen 2009). Using this method, we successfully identified an additional 20 QTL in our study, and strengthened the association of QTL with phenotypes. In this way, we were able to dampen the effects of the chromosome 14 QTL, thereby strengthening the associations between other markers and each trait.

We noted a pronounced effect of the A6 chromosomal deletion on seed stearic acid: ~18% for all lines with the chromosome 14 deletion, + ~5% stearic acid relative to lines with the 194D missense mutation. This difference was greater than expected based on our previous studies with other *SACPD-C* genomic deletions of smaller size (Gillman *et al.* 2014). Although unproven, our results strongly suggest that the A6 genomic deletion, which encompasses *SACPD-C*, also contains another gene (or genes) whose loss also elevates stearic acid content from ~12.5 to ~18%. Identification of such gene(s) is impossible using our mapping population and, as the A6 chromosomal deletion contains at least 142 additional genes, it is a daunting task for reverse genetic approaches.

Table 6 Epistatic interactions between QTL

Trait	QTL	Chromosome	Marker 1	Marker 2	LOD	R^2	*P*-value
Stearic acid	*q.2.1s* × *q.14s*	2 × 14	c2.loc39	c14.loc58	11.14	9.97	1.07E−09
Linolenic acid	*q.8n* × *q.13n*	8 × 13	c8.loc150	c13.loc118.2	5.18	7.45	0.000235

QTL, quantitative trait loci; LOD, logarithm of the odds.

Minor effect QTL for other fatty acid traits and plant height

We detected QTL for several annotated and molecularly characterized fatty acid biosynthetic genes as well as one major plant height/maturity gene. Although the significant marker may differ slightly for each trait due to maximum LOD optimization in model selection, for the sake of discussion, the QTL on chromosome 2 for palmitic, stearic, and linolenic acids can be considered to be identical because their support intervals overlap very significantly. *Q.2p* may be near the map position of *FAD3-B* (Glyma02g39230), which participates in the desaturation of C18:2 (linoleic acid) to C18:3 (linolenic acid) (Bilyeu *et al.* 2003; Gillman and Bilyeu 2012); however, our interval for this particular QTL was quite wide so the association is tenuous. This same level of overlap was seen with the chromosome 14 QTL found for all five fatty acid species, as well as for total seed oil. We expected to see this high level of colocation because of the close genetic association among fatty acid traits due to genes that participate in the fatty acid biosynthetic pathway ($|r|$ values ranging from 0.23 to 0.81). This level of overlap is also highly indicative of the strength and reliability of our QTL results.

Q.5p is near the genomic region of *FATB1-A* (Glyma05g12300), which catalyzes the conversion of 16:0-ACP to C16:0 (palmitic acid) (Cardinal *et al.* 2007).

Q.19h is almost certainly *E3* (Glyma19g224200), which affects flowering time in soybean in combination with the other "*E*" genes to determine flowering time/plant maturity (Watanabe *et al.* 2009). In the context of dissection of genetic architecture of complex traits, these results are purely confirmatory, *i.e.*, they serve to provide confidence in our ability to detect QTL.

Environmental influence and heritability of fatty acid composition

It is a strongly accepted fact that seed oil traits are highly dependent on environmental conditions during seed fill, particularly temperature (Kumar *et al.* 2006). A6's stearic acid phenotype was much lower in the three Missouri environments we studied (23.4 ± 4.4, 20.9 ± 2.9, and 22.3 ± 5.5%) as compared with the original report of 28.1% stearic acid (Hammond and Fehr 1983; Lundeen *et al.* 1987), and this difference may be due to different environmental conditions in Northern Missouri compared to Ames, Iowa during seed fill/maturation. Fatty acid traits are relatively simple from a genetic perspective, and the target of ~20% stearic acid is almost reached with three QTL from A6 (Figure 3), although we noted relatively large environmental variation (Table 2). However, heritability values (h^2) for fatty acid traits in our study were quite strong ($0.72 < h^2 < 0.89$), and twelve out of the 176 lines examined exhibited stearic acid levels > 20% consistently across environments, which suggests that the target may be achievable, at least in certain environments.

Impact of the chromosome 14 deletion on yield, seed traits, and plant morphology

Early studies have definitively demonstrated that A6 alleles for elevated stearic acid are also associated with decreased seed yield and unacceptable agronomic performance (Lundeen *et al.* 1987), presumably due to the large genomic deletion responsible for elevated stearic acid. In our recent work, we demonstrated that loss of SACPD-C activity is inextricably correlated with nodule morphological abnormalities, early nodule senescence, and an impaired ability to fix nitrogen (Gillman *et al.* 2014; Krishnan *et al.* 2016). Although we did not measure seed yield in our present study, we noted a decrease in seed oil content in lines that inherited the A6 chromosome 14 deletion (Table 5).

We saw no correlation between plant height and seed stearic acid content in this study, and our goal was a rapid QTL analysis for seed composition rather than a true seed yield study. Nevertheless, it is expected that the large chromosome 14 deletion will result in an unavoidable yield penalty as has been previously described (Lundeen *et al.* 1987). In the continued effort to develop a high stearic acid soybean variety that meets the current breeding target (> 20% stearic acid), the genomic deletion present in A6 provided valuable genetic information, but deletions/mutations from A6 *per se* will almost certainly not be of direct use in development of germplasm or cultivars with > 20% seed stearic acid and acceptable seed yield.

Identification of minor effect QTL for stearic acid

We identified three additional minor effect seed stearic acid QTL (Table 5), which do not appear to overlap with deletions identified in our previous work (Gillman *et al.* 2014), and it is tempting to report genes within the other QTL intervals whose annotations would suggest they may be associated with oil biosynthesis. However, the confidence intervals for our QTL are still quite large and the effects are relatively small. Further research would be needed to fine map the minor effect QTL. Moreover, the vast majority of gene annotation in soybean is almost completely inferred based on protein similarity to *Arabidopsis* genes (http://soybase.org/genomeannotation/, accessed 05-01-2016), and even in Arabidopsis only about 40% of enzyme and transporter genes have strong evidence for function (Niehaus *et al.* 2015)

Conclusions

We evaluated the genetic architecture for extremely elevated stearic acid (> 20%) through GBS/QTL analysis of RILs produced by reciprocal crosses between two mutant lines. Numerous minor effect QTL were identified for soybean fatty acids, total oil, and for plant height. In contrast to *a priori* expectation, a major effect QTL for stearic acid corresponded to a large genomic deletion from line A6, which explained ~57% of the phenotypic variance for seed stearic acid. However, the loss of SACPD-C activity *per se* boosts seed stearic acid from ~3% stearic acid to ~12.5%. These results suggest that the large genomic deletion in A6 contains both *SACPD-C* and an unidentified gene; the loss of both in line A6 increases seed stearic acid content to ~18% of seed oil. A combination of the major QTL and two minor effect QTL was able to boost seed stearic acid to above 20% of seed oil and meet the market target, but this elevation is expected to occur at a significant cost to seed yield. These genetic results will help guide and inform breeding efforts to increase soybean oil oxidative stability.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to emeritus professor Walter Fehr of Iowa State University for the generous donation of line

A6, which was instrumental in this study. The authors would also like to acknowledge the superb technical assistance of US Department of Agriculture, Agricultural Research Service (USDA-ARS) technicians Jeremy Mullis and Alexandria Berghaus. The project was partially funded by a grant from the United Soybean Board (project # 1420-632-6605). Mention of a trademark, vendor, or proprietary product does not constitute a guarantee or warranty of the product by the USDA and does not imply its approval to the exclusion of other products or vendors that may also be suitable. The USDA-ARS, Midwest Area, is an equal opportunity, affirmative action employer and all agency services are available without discrimination.

LITERATURE CITED

Bachlava, E., R. E. Dewey, J. W. Burton, and A. J. Cardinal, 2009 Mapping and comparison of quantitative trait loci for oleic acid seed content in two segregating soybean populations. Crop Sci. 49(2): 433.

Bilyeu, K., J. D. Gillman, and A. R. LeRoy, 2011 Novel *FAD3* mutant allele combinations produce soybeans containing 1% linolenic acid in the seed oil. Crop Sci. 51(1): 259.

Bilyeu, K. D., L. Palavalli, D. A. Sleper, and P. R. Beuselinck, 2003 Three microsomal omega-3 fatty-acid desaturase genes contribute to soybean linolenic acid levels. Crop Sci. 43(5): 1833.

Bolon, Y. T., W. J. Haun, W. W. Xu, D. Grant, M. G. Stacey *et al.*, 2011 Phenotypic and genomic analyses of a fast neutron mutant population resource in soybean. Plant Physiol. 156(1): 240–253.

Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss *et al.*, 2007 TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23(19): 2633–2635.

Broman, K. W., and S. Sen, 2009 *A Guide to QTL Mapping with R/qtl in Statistics for Biology and Health*. Springer, New York.

Buhr, T., S. Sato, F. Ebrahim, A. Xing, Y. Zhou *et al.*, 2002 Ribozyme termination of RNA transcripts down-regulate seed fatty acid genes in transgenic soybean. Plant J. 30(2): 155–163.

Cardinal, A. J., J. W. Burton, A. M. Camacho-Roger, J. H. Yang, R. F. Wilson *et al.*, 2007 Molecular analysis of soybean lines with low palmitic acid content in the seed oil. Crop Sci. 47(1): 304–310.

Chen, Q.-s., Z.-c. Zhang, C.-y. Liu, D.-w. Xin, H.-m. Qiu *et al.*, 2007 QTL analysis of major agronomic traits in soybean. Agric. Sci. China 6(4): 399–405.

Cooper, J. L., B. J. Till, R. G. Laport, M. C. Darlow, J. M. Kleffner *et al.*, 2008 TILLING to detect induced mutations in soybean. BMC Plant Biol. 8(1): 9.

Csanadi, G., J. Vollmann, G. Stift, and T. Lelley, 2001 Seed quality QTLs identified in a molecular map of early maturing soybean. Theor. Appl. Genet. 103(6–7): 912–919.

Diers, B. W., and R. C. Shoemaker, 1992 Restriction fragment length polymorphism analysis of soybean fatty acid content. J. Am. Oil Chem. Soc. 69(12): 1242–1244.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto *et al.*, 2011 A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. PLoS One 6(5): e19379.

Eskandari, M., E. R. Cober, and I. Rajcan, 2012 Genetic control of soybean seed oil: I. QTL and genes associated with seed oil concentration in RIL populations derived from crossing moderately high-oil parents. Theor. Appl. Genet. 126(2): 483–495.

Fehr, W.R., 2007 Breeding for modified fatty acid composition in soybean. Crop Science 47(Suppl.3): S-72–S.87.

Frankel, E. N., 1991 Review. Recent advances in lipid oxidation. J. Sci. Food Agric. 54(4): 495–511.

Gillman, J. D., and K. Bilyeu, 2012 Genes and alleles for quality traits on the soybean genetic/physical map, pp. 67–96 in *Designing Soybeans for 21st Century Markets*, edited by Willson, R. F.. AOCS Press, Urbana, IL.

Gillman, J. D., M. G. Stacey, Y. Cui, H. R. Berg, and G. Stacey, 2014 Deletions of the SACPD-C locus elevate seed stearic acid levels but also result in fatty acid and morphological alterations in nitrogen fixing nodules. BMC Plant Biol. 14(1): 143.

Hammond, E. G., and W. R. Fehr, 1983 Registration of A6 germplasm line of soybean. Crop Sci. 23(1): 192.

Hunter, J. E., J. Zhang, and P. M. Kris-Etherton, 2009 Cardiovascular disease risk of dietary stearic acid compared with trans, other saturated, and unsaturated fatty acids: a systematic review. Am. J. Clin. Nutr. 91(1): 46–63.

Hyten, D. L., V. R. Pantalone, A. M. Saxton, M. E. Schmidt, and C. E. Sams, 2004 Molecular mapping and identification of soybean fatty acid modifier quantitative trait loci. J. Am. Oil Chem. Soc. 81(12): 1115–1118.

Kim, H.-K., Y.-C. Kim, S.-T. Kim, B.-G. Son, Y.-W. Choi *et al.*, 2010 Analysis of quantitative trait loci (QTLs) for seed size and fatty acid composition using recombinant inbred lines in soybean. J. Life Sci. 20(8): 1186–1192.

Krishnan, H.B., A.A. Alaswad, N.W. Oehrle, and J.D. Gillman, 2016 Deletion of the SACPD-C locus alters the symbiotic relationship between bradyrhizobium japonicum USDA110 and soybean, resulting in elicitation of plant defense response and nodulation defects. Mol. Plant Microbe Interact. DOI: 10.1094/MPMI-08-16-0173-R.

Kumar, V., A. Rani, S. Solanki, and S. M. Hussain, 2006 Influence of growing environment on the biochemical composition and physical characteristics of soybean seed. J. Food Compos. Anal. 19(2–3): 188–195.

La, T. C., S. M. Pathan, T. Vuong, J.-D. Lee, A. M. Scaboo *et al.*, 2014 Effect of high-oleic acid soybean on seed oil, protein concentration, and yield. Crop Sci. 54(5): 2054–2062.

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14): 1754–1760.

Li, H., T. Zhao, Y. Wang, D. Yu, S. Chen *et al.*, 2011 Genetic structure composed of additive QTL, epistatic QTL pairs and collective unmapped minor QTL conferring oil content and fatty acid components of soybeans. Euphytica 182(1): 117–132.

Li, Z., R. F. Wilson, W. E. Rayford, and H. R. Boerma, 2002 Molecular mapping genes conditioning reduced palmitic acid content in N87–2122–4 Soybean. Crop Sci. 42(2): 373.

Liang, H.-z., Y.-l. Yu, S.-f. Wang, Y. Lian, T.-f. Wang *et al.*, 2010 QTL mapping of isoflavone, oil and protein contents in soybean (Glycine max L. Merr.). Agric. Sci. China 9(8): 1108–1116.

Lundeen, P. O., W. R. Fehr, E. G. Hammond, and S. R. Cianzio, 1987 Association of alleles for high stearic acid with agronomic characters of soybean. Crop Sci. 27(6): 1102–1105.

Mackay, T. F. C., 2001 The genetic architecture of quantitative traits. Annu. Rev. Genet. 35(1): 303–339.

Mansur, L. M., K. G. Lark, H. Kross, and A. Oliveira, 1993 Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (Glycine max L.). TAG 86(8): 907–913.

Niehaus, T. D., A. M. K. Thamm, V. de Crécy-Lagard, and A. D. Hanson, 2015 Proteins of unknown biochemical function: a persistent problem and a roadmap to help overcome it. Plant Physiol. 169(3): 1436–1442.

Panthee, D. R., V. R. Pantalone, and A. M. Saxton, 2006 Modifier QTL for fatty acid composition in soybean oil. Euphytica 152(1): 67–73.

Pham, A.-T., J.-D. Lee, J. G. Shannon, and K. D. Bilyeu, 2010 Mutant alleles of FAD2–1A and FAD2–1B combine to produce soybeans with the high oleic acid seed oil trait. BMC Plant Biol. 10(1): 195.

Poland, J. A., P. J. Brown, M. E. Sorrells, and J.-L. Jannink, 2012 Development of high-density genetic maps for Barley and Wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One 7(2): e32253.

Qi, Z.-m., Q. Wu, X. Han, Y.-n. Sun, X.-y. Du *et al.*, 2011 Soybean oil content QTL mapping and integrating with meta-analysis method for mining genes. Euphytica 179(3): 499–514.

Rahman, S. M., Y. Takagi, K. Miyamoto, and T. Kawakita, 1995 High stearic acid soybean mutant induced by x-ray irradiation. Biosci. Biotechnol. Biochem. 59(5): 922–923.

Rahman, S. M., Y. Takagi, and T. Kinoshita, 1997 Genetic control of high stearic acid content in seed oil of two soybean mutants. Theor. Appl. Genet. 95(5–6): 772–776.

Reinprecht, Y., V. W. Poysa, K. Yu, I. Rajcan, G. R. Ablett *et al.*, 2006 Seed and agronomic QTL in low linolenic acid, lipoxygenase-free soybean (Glycine max (L.) Merrill) germplasm. Genome 49(12): 1510–1527.

Renzyaeva, T. V., 2013 On the role of fats in baked flour goods. Foods Raw Mater. 1(1): 19–25.
Ribeiro, A. P. B., R. Grimaldi, L. A. Gioielli, and L. A. G. Gonçalves, 2009 Zero trans fats from soybean oil and fully hydrogenated soybean oil: physico-chemical properties and food applications. Food Res. Int. 42(3): 401–410.
Ruddle, P., R. Whetten, A. Cardinal, R. G. Upchurch, and L. Miranda, 2012 Effect of a novel mutation in a Δ9-stearoyl-ACP-desaturase on soybean seed oil composition. Theor. Appl. Genet. 126(1): 241–249.
Ruddle, P., A. Cardinal, R. G. Upchurch, C. Arellano, and L. Miranda, 2013a Agronomic effects of mutations in two soybean Δ9–stearoyl-acyl carrier protein-desaturases. Crop Sci. 53(5): 1887.
Ruddle, P., R. Whetten, A. Cardinal, R. G. Upchurch, and L. Miranda, 2013b Effect of Δ9-stearoyl-ACP-desaturase-C mutants in a high oleic background on soybean seed oil composition. Theor. Appl. Genet. 127(2): 349–358.
Schmutz, J., S. B. Cannon, J. Schlueter, J. Ma, T. Mitros *et al.*, 2010 Genome sequence of the palaeopolyploid soybean. Nature 463: 178–183.
Shannon, L. M., B. S. Yandell, and K. Broman, 2013 Users guide for new BCsFt tools for R/qtl. Available at: http://cran.r-project.org/web/packages/qtl/vignettes/bcsft.pdf.
Shibata, M., K. Takayama, A. Ujiie, T. Yamada, J. Abe *et al.*, 2008 Genetic relationship between lipid content and linolenic acid concentration in soybean seeds. Breed. Sci. 58(4): 361–366.
Spencer, M. M., D. Landau-Ellis, E. J. Meyer, and V. R. Pantalone, 2004 Molecular markers associated with linolenic acid content in soybean. J. Am. Oil Chem. Soc. 81(6): 559–562.
Swarts, K., H. Li, J.A. Romero Navarro, D. An, M.C. Romay *et al.*, 2014 Novel methods to optimize genotypic imputation for low-coverage, next-generation sequence data in crop plants. The Plant Genome 7(3).
Tarancón, P., S. M. Fiszman, A. Salvador, and A. Tárrega, 2013 Formulating biscuits with healthier fats. Consumer profiling of textural and flavour sensations during consumption. Food Res. Int. 53(1): 134–140.
Tian, Z., X. Wang, R. Lee, Y. Li, J. E. Specht *et al.*, 2010 Artificial selection for determinate growth habit in soybean. Proc. Natl. Acad. Sci. USA 107(19): 8563–8568.
Van Camp, D., N. H. Hooker, and C.-T. J. Lin, 2012 Changes in fat contents of US snack foods in response to mandatory trans fat labelling. Public Health Nutr. 15(06): 1130–1137.
Wang, D., A. M. Procopiuk, B. W. Diers, and G. L. Graef, 2004 Identification of putative QTL that underlie yield in interspecific soybean backcross populations. Theor. Appl. Genet. 108(3): 458–467.
Wang, X., G.-L. Jiang, M. Green, R. A. Scott, D. L. Hyten *et al.*, 2012 Quantitative trait locus analysis of saturated fatty acids in a population of recombinant inbred lines of soybean. Mol. Breed. 30(2): 1163–1179.
Warner, K., and M. Gupta, 2005 Potato chip quality and frying oil stability of high oleic acid soybean oil. J. Food Sci. 70(6): s395–s400.
Watanabe, S., R. Hideshima, Z. Xia, Y. Tsubokura, S. Sato *et al.*, 2009 Map-based cloning of the gene associated with the soybean maturity locus E3. Genetics 182(4): 1251–1262.
Wilson, R. F., 2004 Seed Composition, pp. 621–677 in *Soybeans: Improvement, Production, and Uses*, edited by Boerma, H. R., and J. Specht. American Society of Agronomy, Madison, WI.
Xie, D., Y. Han, Y. Zeng, W. Chang, W. Teng *et al.*, 2011 SSR- and SNP-related QTL underlying linolenic acid and other fatty acid contents in soybean seeds across multiple environments. Mol. Breed. 30(1): 169–179.
Xu, S., 2003 Theoretical basis of the beavis effect. Genetics 165(4): 2259–2268.
Yu, S., J. Derr, T. D. Etherton, and P. M. Kris-Etherton, 1995 Plasma cholesterol-predictive equations demonstrate that stearic acid is neutral and monounsaturated fatty acids are hypocholesterolemic. Am. J. Clin. Nutr. 61(5): 1129–1139.
Zhang, P., J. W. Burton, R. G. Upchurch, E. Whittle, J. Shanklin *et al.*, 2008 Mutations in a Δ–Stearoyl-ACP-desaturase gene are associated with enhanced stearic acid levels in soybean seeds. Crop Sci. 48(6): 2305.

Communicating editor: J. Ma