Mutations in SACPD-C Result in a Range of Elevated Stearic Acid Concentration in Soybean Seed

Militza Carrero-Colón¹, Nathan Abshire², Daniel Sweeney², Erik Gaskin², Karen Hudson¹*

¹Crop Production and Pest Control Research Unit, Agricultural Research Service (ARS), United States Department of Agriculture (USDA), West Lafayette, Indiana, United States of America, ²Department of Agronomy, Purdue University, West Lafayette, Indiana, United States of America

Abstract

Soybean oil has a wide variety of uses, and stearic acid, which is a relatively minor component of soybean oil is increasingly desired for both industrial and food applications. New soybean mutants containing high levels of the saturated fatty acid stearate in seeds were recently identified from a chemically mutagenized population. Six mutants ranged in stearate content from 6–14% stearic acid, which is 1.5 to 3 times the levels contained in wild-type seed of the Williams 82 cultivar. Candidate gene sequencing revealed that all of these lines carried amino acid substitutions in the gene encoding the delta-9-stearoyl-acyl-carrier protein desaturase enzyme (SACPD-C) required for the conversion of stearic acid to oleic acid. Five of these missense mutations were in highly conserved residues clustered around the predicted di-iron center of the SACPD-C enzyme. Co-segregation analysis demonstrated a positive association of the elevated stearate trait with the SACPD-C mutation for three populations. These missense mutations may provide additional alleles that may be used in the development of new soybean cultivars with increased levels of stearic acid.


Editor: Girdhar K. Pandey, University of Delhi South Campus, India

Received November 7, 2013; Accepted April 24, 2014; Published May 20, 2014

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

Funding: Funding for this work was provided through the USDA-ARS Current Research Information System 3602-21000-006-00D. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: hudson@ars.usda.gov

Introduction

Stearic acid is one of the component fatty acids in soybean oil, comprising 2–4% of the total oil fraction. Stearic acid has a neutral effect on blood serum LDL cholesterol concentration and is therefore a desirable constituent of oils for food use [1]. Stearic acid confers a high melting temperature and oxidative stability to oils destined for end use in baking fats. Previously, to increase the proportion of stearic acid in soybean oil, the oil was subjected to hydrogenation. However, genetic manipulation of stearic acid level is more efficient and reduces the trans-fats that may be introduced by the hydrogenation process [2].

Three soybean genes have been characterized with homology to delta-9-stearoyl-acyl carrier protein desaturases (SACPDs) which are required for the conversion of stearic acid to oleic acid [3]. These genes are delimited SACPD-A, SACPD-B, and SACPD-C. SACPD-C encodes the seed-specific isoform of this enzyme, where SACPD-A and SACPD-B transcripts accumulate in all soybean tissues [1,4].

Soybeans with mutations in the SACPD-C and SACPD-B genes have been described. FAM94-41 is a spontaneously occurring change in the SACPD-C gene and results in plants with levels of stearic acid in the seed of ~9% [5]. Deletion of the SACPD-C gene in the A6 germplasm line results in up to 28% stearic acid in the seed, but the size of this deletion is uncharacterized [4,6]. Additional SACPD-C mutants have been described with a range of 10-16% stearic acid in the seeds [4,7]. SACPD-B mutants have recently been reported to contain ~10% stearic acid [8]. No mutations have been described for the SACPD-A gene. Some high stearate mutants have previously been associated with poor germination and low seed yield [9,10] however recently it was demonstrated that missense mutations in SACPD-C are not associated with poor agronomic characteristics [11]. Additional sources of germplasm carrying novel mutations in the SACPD-C gene, or novel loci which influence seed stearic acid levels are needed to circumvent this issue to enable the production of soybeans with elevated levels of stearic acid to meet the demands of the food-oil market.

Materials and Methods

Plants and growth conditions and fatty acid analysis

For screening, plants were grown in the field in West Lafayette, Indiana, as described in reference [12]. Field location GPS coordinates are latitude 40.468 degrees north, longitude minus 86.991 degrees west. Soybeans described in this study are non-transgenic, therefore no specific permits were required for growth. Fatty acid composition analysis was performed as previously described [13].

Sequencing and Genotyping

Three segments of the SACPD-C (Glyma14g27990) coding region were amplified and sequenced using the primers in Table S1. DNA sample preparation for sequencing was performed using the CTAB method [14] and sample preparation for genotyping was as previously described [13]. dCAPS genotyping [15] was performed using standard protocols with the assays developed specifically for the SACPD-C mutants provided in Table S1. To evaluate the position of substitutions, mutations were overlaid on

* E-mail: hudson@ars.usda.gov
the protein structure PDB ID 1AFR using the program Cn3D v. 4.3 [16]. Mutant SACPD-C sequences are deposited in GenBank with accession numbers KJ522450-KJ522455.

### Results and Discussion

Mutant plants with high levels of stearic acid in seeds were identified in an ongoing screen for soybean seed with altered fatty acid composition (reference [12], and unpublished data) and six lines were chosen for further characterization. These mutants were obtained from an NMU-mutagenized population in the Williams-82 genetic background [17]. Levels of stearic acid in the seed of the mutant lines ranged from 6–13% (Table 1), with the highest levels in line 18190. Line #18948 corresponds to line #14, line 18190 corresponds to #16, and line 18610 corresponds to #17 in reference [12] while the isolation of lines 15073, 14197, and 21084 has not been previously described. Line 14197 and 18948 were isolated as heterozygous M3, as revealed by segregating types of M4 seed, while the remaining lines showed a consistent level of stearic acid across multiple M4 individuals and were presumed to be homozygous isolates (Table S2 and [12]).

### Table 1. Elevated stearic acid levels in SACPD-C mutants.

<table>
<thead>
<tr>
<th></th>
<th>Palmitic Acid (16:0)</th>
<th>Stearic Acid (18:0)</th>
<th>Oleic Acid (18:1)</th>
<th>Linoleic Acid (18:2)</th>
<th>Linolenic Acid (18:3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>W82</td>
<td>10.0±0.2</td>
<td>4.0±0.1</td>
<td>20.6±0.5</td>
<td>57.0±0.5</td>
<td>8.3±0.2</td>
</tr>
<tr>
<td>15073 SACPD-C_{G224E}</td>
<td>8.1±0.4</td>
<td>12.1±1.2</td>
<td>15.9±0.6</td>
<td>55.0±1.3</td>
<td>8.8±0.6</td>
</tr>
<tr>
<td>14197 SACPD-C_{Y211C}</td>
<td>8.8±0.4</td>
<td>12.1±3.5</td>
<td>20.0±2.9</td>
<td>51.6±0.7</td>
<td>7.6±0.6</td>
</tr>
<tr>
<td>18190 SACPD-C_{A239T}</td>
<td>8.5±0.6</td>
<td>13.5±0.6</td>
<td>18.4±0.2</td>
<td>53.8±0.4</td>
<td>9.3±0.3</td>
</tr>
<tr>
<td>18948 SACPD-C_{H223R}</td>
<td>10.1±0.5</td>
<td>9.5±1.3</td>
<td>17.4±0.6</td>
<td>54.9±1.1</td>
<td>8.1±0.4</td>
</tr>
<tr>
<td>18610 SACPD-C_{A239T}</td>
<td>9.4±0.6</td>
<td>6.1±1.2</td>
<td>22.2±1.2</td>
<td>54.1±2.0</td>
<td>8.1±0.9</td>
</tr>
<tr>
<td>21084 SACPD-C_{R228H}</td>
<td>8.4±0.3</td>
<td>8.7±0.8</td>
<td>27.8±1.8</td>
<td>48.6±1.4</td>
<td>6.5±0.8</td>
</tr>
</tbody>
</table>

Fatty acid levels were averaged for n homozygous M4 or M5 lines for each mutation, and averages and standard deviations are shown. p-value (in italics) was calculated from a two-tailed, type 1 t-test for the average fatty acid level in each mutant relative to the Williams-82 (W82) wild type control. Single asterisk indicates p-values that are significant at the p<0.05 level.

doi:10.1371/journal.pone.0097891.t001

Sequencing of the SACPD-C gene from these mutant lines revealed that each carries a distinct and independent missense mutation in the coding region of SACPD-C (Figure 1 and Table 1). The sequence of the SACPD-C gene in the Williams-82 accession that serves as the germplasm background for this mutant is consistent with the previously described soybean SACPD-C sequence (Genbank #EF113911) [4] but not the Williams-82 sequence within the Phytozome database (version 9.1, www.phytozome.net) which carries a frameshift mutation early in the first exon. The exons of the SACPD-A and SACPD-B genes (Glyma07g32850 and Glyma02g15600, respectively) were also sequenced in these lines and were found to be identical to the Williams-82 sequence (not shown). As line 21084 and 18610 M3 and M4 individuals produced seed with a marginal increase in stearic acid content from ranging from 5–8%, to verify that these lines were homozygous the SACPD-C gene was sequenced from M4 plants (Table S2). All of these M4 individuals were homozygous for their respective mutations. This suggests that the mutations in these lines are less damaging to enzyme structure or function and the resultant plants display a marginal increase in stearic acid.

**Figure 1. Mutations in the SACPD-C enzyme.** Multiple sequence alignment for soybean SACPD-C with close sequence homologs showing the position of mutations in conserved locations. Other proteins shown are SACPD-A and SACPD-B from soybean, FAB2 from Arabidopsis (At3g02610/DES2), and the castor bean SACPD (30020.m000203.rco). Mutations described in this work are located at the following positions: 1. SACPD-C_{Y211C} 2. SACPD-C_{A239T} 3. SACPD-C_{G224E} 4. SACPD-C_{G224E} 5. SACPD-C_{A239T} 6. SACPD-C_{R228H}. doi:10.1371/journal.pone.0097891.g001
Visualization of the location of the affected amino acids in the determined crystal structure for the SACPD enzyme from *Ricinus communis* reveals that five of the six mutations are located in the enzyme channel domain which is composed of several alpha helixes, the function of this channel is to position the hydrocarbon chain of the fatty acid so that it can be oxidized (Figure S1) [18]. SACPD-CY211C, SACPD-CA218E, SACPD-CG224E, SACPD-CH223R are located on alpha helix six, and SACPD-CA239T is located on alpha helix seven. The amino acid substitutions introduce either changes in side chain charge or size into this region of the enzyme. These mutations may reduce the ability of a fatty acid chain to be correctly positioned within the channel domain of the enzyme. Proximity of mutations to the di-iron core and increases in polarity also correlate with greater reduction of enzymatic function [10]. SACPD-CR329I is located on a loop residue after helix ten, and this mutation results in a less dramatic increase in stearic acid level. We speculate that this residue may be involved in correct folding of the active site or possibly in interaction with other proteins.

To demonstrate that changes in the SACPD-C gene cause the elevated stearic acid phenotype in these mutant lines, cosegregation analysis was performed in segregating populations. In each case the single nucleotide polymorphism introduced by the mutation was used to develop a codominant dCAPs marker. Lines 18948 (SACPD-CY211C) and 14197 (SACPD-CY221C) were isolated as heterozygous mutants for SACPD-C mutations. M_{4} plants segregating for these SNPs were grown in the field during the 2011 growing season, and M_{4} seed was genotyped for the SNP and phenotyped for fatty acid content. Figure 2A shows association of the SACPD-CY211C mutation with elevated levels of stearic acid in the seed. Figure 2B shows cosegregation of SACPD-CY223R with elevated levels of stearic acid in the seed. Line 18190 (SACPD-CA218E), which has the highest levels of stearate, was crossed to the cultivar Prize, allowed to self-pollinate, and F_{2} plants were grown in the field (during the 2012 growing season) and genotyped. F_{3} seed from F_{2} individuals was analyzed for stearic acid content. Figure 2C shows the cosegregation of SACPD-CY211C with elevated levels of stearate. In the SACPD-CA218E x Prize segregating population, a number of homozygous mutant individuals were observed which contained stearic acid levels greater (>20% stearate) than those observed in the homozygous mutant isolate in the Williams-82 background (approximately 13%). This may be due to the segregation of a second, modifying locus distinct from that of the mutated SACPD-C gene in the SACPD-CA218E line. This genetic modifier may be in the outcross parent (Prize) genetic background, or it may be a second-site mutation in the heavily mutagenized background of the SACPD-CA218E parent line. It is estimated that the mutation frequency in the parent of SACPD-CA218E parent is on the order of 100 genic mutations per individual [17]. The seed stearate level in the Prize parent grown in parallel was 4%, similar to Williams-82. As the other SACPD genes would be good candidates to act additively with SACPD-CA218E, the SACPD-A, -B, (and -C) genes were sequenced in the Prize parent and found to encode predicted proteins with an amino acid sequence identical to the published reference sequence [4].

There is evidence that other, unknown loci in addition to SACPD-C impact stearic acid levels in soybean seed [8]. The strong mutations in SACPD-C described here are comparable to the defects observed in RG7 and RG8 which are also chemically induced point mutants in SACPD-C [7] and may therefore serve as a basis for germplasm to enhance stearic acid levels. Weaker alleles such as SACPD-CR329I and SACPD-CA239T may be useful when combined with other genes to increase stearic acid while maintaining high oleic acid levels. In addition, identification of the factors that enable the >20% stearic acid levels observed in association with SACPD-CA218E may suggest to further increase stearic acid levels for a soybean oil with enhanced functionality.

**Supporting Information**

Figure S1 Position of mutations in SACPD-C crystal structure. (TIFF)

Table S1 Primer sequences. (PDF)
Table S2  \(M_3/M_4\) phenotypes and genotypes.

Disclaimer: Product names are necessary to report factually on available data. However the USDA neither guarantees nor warrants the standard of the product, and use of the names implies no approval of the product to the exclusion of others that may also be suitable.

Acknowledgments

The authors wish to thank Dr. K Bilyeu, USDA-ARS, Columbia, MO and the Functional Foods Research Unit (USDA-ARS, Peoria, IL) for performing fatty acid analysis.

Author Contributions

Conceived and designed the experiments: KH. Performed the experiments: MC NA DS EG. Analyzed the data: MC KH DS NA. Contributed reagents/materials/analysis tools: KH. Wrote the paper: KH NA.

References