JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Introgression of Leginsulin, a Cysteine-Rich Protein, and High-Protein Trait from an Asian Soybean Plant Introduction Genotype into a North American Experimental Soybean Line

Hari B. Krishnan,*[,†] Won-Seok Kim,[‡] Nathan W. Oehrle,[†] Alaa A. Alaswad,[‡,§] Ivan Baxter,[#] William J. Wiebold,[‡] and Randall L. Nelson[⊥]

[†]Plant Genetics Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Columbia, Missouri 65211, United States

[‡]Plant Science Division, University of Missouri, Columbia, Missouri 65211, United States

[§]King Abdul Aziz University, Jeddah, Saudi Arabia

[#]Plant Genetics Research Unit, Donald Danforth Plant Sciences Center, Agricultural Research Service, U.S. Department of Agriculture, St. Louis, Missouri 63132, United States

[⊥]Soybean/Maize Germplasm, Pathology, and Genetics Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Department of Crop Sciences, University of Illinois, Urbana, Illinois 61801, United States

Supporting Information

ABSTRACT: Soybean is an important protein source for both humans and animals. However, soybean proteins are relatively poor in the sulfur-containing amino acids, cysteine and methionine. Improving the content of endogenous proteins rich in sulfur-containing amino acids could enhance the nutritive value of soybean meal. Leginsulin, a cysteine-rich peptide, predominantly accumulates in Asian soybean accessions but not in most North American cultivars. By screening diverse soybean accessions from the USDA Soybean Germplasm Collection, we were able to identify one plant introduction, PI 427138, as a high-protein line with relatively high amounts of both elemental sulfur and leginsulin. We introgressed these desirable traits from PI 427138 into an experimental line with the aim of improving the overall protein content and quality of seed proteins. Biochemical characterization of inbred progenies from the cross of LD00-3309 with PI 427138 grown at six locations revealed stable ingression of high protein, high elemental sulfur, and high leginsulin accumulation. Comparison of soybean seed proteins resolved by high-resolution 2-D gel electrophoresis in combination with Delta2D image analysis software revealed preferential accumulation of a few glycinin subunits contributed to the increased protein content in the introgressed lines. Amino acid analysis revealed that even though the leginsulin introgressed lines had higher protein, leginsulin, and elemental sulfur, the overall concentration of sulfur-containing amino acids was not significantly altered when compared with the parental lines. The experimental soybean lines developed during this study (Leg-3, Leg-7, and Leg-8) lack A5, A4, and B3 glycinin subunits and could be utilized in breeding programs to develop high-quality tofu cultivars.

KEYWORDS: *leginsulin, soybean, sulfur, cysteine, methionine*

INTRODUCTION

Soybean is widely used as the major ingredient for protein in animal feeds. Soybeans are valued for their high amounts of protein with a relatively high proportion of essential amino acids. Commercial soybeans contain about 37−40% protein by dry weight. Their relatively low cost combined with their excellent nutritive value has enabled soybeans to attain elite stature as the world's dominant protein feed ingredient. However, the nutritive value of soybeans can be further enhanced if the concentration of sulfur-containing amino acids, methionine and cysteine, is elevated. The methionine and cysteine content of soybean is about 2−3%, short of the desired levels for the optimum growth of monogastric animals such as poultry and swine. Consequently, poultry and swine industries incur additional cost from augmenting feeds with synthetic methionine.[1]

Concerted attempts have been made to improve the sulfur-containing amino acid content of soybeans. One common approach has been to overexpress methionine-rich proteins in soybean seeds.[2] Townsend and Thomas were able to successfully express the 2S albumin from Brazil nut in transgenic soybeans, resulting in an overall increase in methionine content by 15−40%.[3] However, this approach was not economically successful because it was subsequently demonstrated that the Brazil nut 2S albumin is a major food allergen.[4] Methionine-rich proteins from maize have also have been expressed in transgenic soybeans.[5,6] However, these attempts have resulted in only a marginal increase in the overall sulfur amino acid content of soybean seeds.

An alternative approach involves the identification of soybean cultivars that accumulate proteins rich in sulfur-containing

Received: October 28, 2014
Revised: March 10, 2015
Accepted: March 10, 2015
Published: March 10, 2015

DOI: 10.1021/jf505202z
J. Agric. Food Chem. 2015, 63, 2862−2869

amino acids. In addition to the most abundant 7S β-conglycinin and the 11S glycinin, soybeans also accumulate a few other proteins that contain higher concentrations of methionine and cysteine. Earlier studies have demonstrated low-level accumulation of a methionine-rich protein in soybean seeds.[7,8] This 2S albumin protein contains 8.6% methionine and has been proposed as a candidate protein for overexpression in transgenic soybean seeds to elevate the methionine content.[9] Other potential proteins for overexpression in soybean include the Bowman−Birk inhibitor (BBI) and the BBI-related family of isoinhibitors. These protease inhibitors are rich in cysteine and are mainly responsible for the overall sulfur-containing amino acids of soybean seed.[10−12]

Leginsulin, a peptide consisting of 37 amino acids, belongs to the cysteine-knot family[13] and is homologous to pea (*Pisum sativum* L.) albumin (PA1b).[14] Soybean leginsulin is synthesized as a precursor polypeptide that is subjected to post-translational processing to give rise to 6 and 4 kDa peptides.[15−17] The 4 kDa peptide that is rich in cysteine residues is coined "leginsulin". We have previously shown that this peptide is preferentially soluble in 50% isopropanol and that the accumulation of this cysteine-rich peptide is drastically higher in Asian than in North American soybean accessions.[18] Furthermore, it was demonstrated that the soybean genome contains two leginsulin genes, which are differentially expressed in the embryonic axis and the cotyledons.[18] Because North American soybean cultivars accumulate low levels of leginsulin, we hypothesized that overexpression of this cysteine-rich peptide in North American soybean cultivars would improve their overall cysteine content. In this study, we crossed PI 427138, an Asian soybean accession, with an experimental North American line, resulting in lines that exhibited significantly higher protein content with improved protein quality. However, the amount of sulfur-containing amino acids was not increased.

MATERIALS AND METHODS

Reagents. Acrylamide, bis-acrylamide, ammonium persulfate, *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED), and Coomassie Brilliant Blue G-250 and R-250 were obtained from BioRad Laboratories, Inc. (Hercules, CA, USA). Protease inhibitor cocktail (Plant ProteaseArrest) was obtained from G-Biosciences (St. Louis, MO, USA). IPG strips were obtained from GE Healthcare (Piscataway, NJ, USA). Urea, thiourea, and β-mercaptoethanol were obtained from Fisher Scientific (Pittsburgh, PA, USA). Dithiothreitol (DTT), iodoacetamide, EDTA, and 2-hydroxyethyl disulfide (2-HED) were obtained from Sigma-Aldrich (St. Louis, MO, USA).

Plant Material. During the course of identification of soybean lines exhibiting high BBI activity, we serendipitously identified PI 427138 as a soybean line accumulating high amounts of leginsulin. In 2007, LD00-3309 was crossed with PI 427138 in Urbana, IL, USA. LD00-3309 is an early-maturity group IV cultivar developed at the University of Illinois and was selected because it was one of the highest yielding, publicly developed cultivars available at the time.[19] In 2008, two F1 plants were grown, and the seeds from each plant were harvested separately. In 2009, approximately 500 seeds were planted from each F1 plant. The F2 populations confirmed that each putative F1 plant was the product of cross-pollination with segregation for flower and pubescence color as expected, and 285 plants were harvested and threshed individually. In 2010, seeds from each harvested F2 plant were planted in rows 1.2 m long. Between 1 and 5 single plants were harvested from 66 rows selected for agronomic appearance for a total of 172 plants. In 2011, 27 of the F4 plant rows had a homogeneous appearance and acceptable agronomic characteristics and were bulk harvested.

Seeds from 40 F4 lines with acceptable agronomic characteristics were selected for characterization. Fifty percent isopropanol soluble proteins were isolated from these 40 F4 plants and examined for the accumulation of leginsulin. Examination of Coomassie-stained gels revealed that 15 plants accumulated high amounts of leginsulin (Supplemental Figure 1). We also measured the sulfur content of the soybean seeds from 40 plants by inductively coupled plasma−mass spectrometry. The sulfur content of these plants ranged from 3000 to 4000 ppm (Supplemental Figure 2). Seeds from the 40 plants were advanced another generation, and seeds from lines having a homogeneous appearance and acceptable agronomic characteristics were bulk harvested. From these F5 plants, we selected two lines (Leg-3 and Leg-7) that accumulated high amounts of leginsulin and high sulfur content for further characterization. Another line (Leg-8) that had high sulfur content but failed to accumulate leginsulin was also selected for further analysis.

Small plots of the three experimental lines plus the two parents were planted in two replications at Bellflower, Ivesdale, and Urbana, IL, USA, and Bradford, MO, USA, and on two soil types at Portageville, MO, USA. Seeds were harvested from all plots for analysis.

Nitrogen Content. Fifteen seeds per line from each field plot were ground to a fine powder and used for total nitrogen/protein analysis. Nitrogen content was measured using a LECO truSpec model FP-428 nitrogen analyzer (St. Joseph, MI, USA). A 200 mg sample of seed powder was combusted at 900 °C in a chamber in the presence of oxygen, leading to the release of carbon dioxide, water, and nitrogen. The gases were then passed over a potassium hydroxide aqueous solution, eliminating carbon dioxide and water. A column containing a thermal conductivity detector was then used to separate the nitrogen from any residual carbon dioxide and water, and the remaining nitrogen content was measured. The instrument was calibrated using ultrapure EDTA of known nitrogen content. The protein content of the seeds was inferred from its nitrogen content using a protein correction factor of 6.25. Three separate biological replicates were evaluated per line, and results were compared using standard deviation of the mean.[20]

Oil Content. The oil content of 10−15 seeds per line was determined using an Oxford Instruments America MQC Oilseeds Analyzer with a 26 mm probe (Concord, MA, USA). A 15-point calibration curve was built using data obtained from both pure soybean oil and soybean seeds of known oil contents (11−21%) as determined by the University of Missouri Experiment Station Chemical Laboratories (AOCS Official Method Ca 5b-71). Three separate biological replicates were evaluated per line, and results were compared using standard deviation of the mean. JMP version 9 software (SAS Institute Inc., Cary, NC, USA) was used to perform one-way analysis of variance (ANOVA) tests.[20]

Sulfur Content. Inductively coupled plasma−mass spectrometry was performed as previously described.[21] Individually weighed dry seeds were digested overnight with 2.5 mL of concentrated HNO_3 at room temperature before the samples were heated to 105 °C over 2 h and then cooled to room temperature over 2 h. Following this step, the samples were diluted with ultrapure water and transferred to 96-well autosampler plates. A PerkinElmer Elan DRC-e inductively coupled plasma mass spectrometer (Waltham, MA, USA) with an Apex Desolvation Nebulizer, a FAST sampling valve, and an Elemental Scientific SC4 DX autosampler (Omaha, NE, USA) was used for the analysis of sulfur. A liquid reference material composed of pooled samples of soybean digests was run every ninth sample to correct for ICP-MS run-to-run variation and within-run drift. All samples were normalized to the recorded dry weights.

Amino Acid Analysis. Amino acid analysis was performed at the Donald Danforth Plant Science Center Proteomics and Mass Spectrometry Facility. For the determination of the total amino acid content, the seed powder was first subjected to hydrolysis with 6 N HCl. For the quantification of methionine and cysteine, duplicate samples were first subjected to an initial oxidation step using performic acid prior to acid hydrolysis. Amino acids were quantified using the manufacturer's instructions for sample preparation in the Waters AccQ-Tag Ultra Kit and quantified on a Waters Acquity UPLC system

DOI: 10.1021/jf505202z
J. Agric. Food Chem. 2015, 63, 2862−2869

(Milford, MA, USA). Samples were run in quadruplicate and subjected to appropriate statistical analysis. Cysteine and methionine quantities are reported as cysteic acid and methionine sulfone.

1-D Electrophoresis. For 1-D electrophoretic analysis, dried mature seeds were ground into a fine powder, and 10 mg was placed into a tube for extraction using 1 mL of sodium dodecyl sulfate (SDS) sample buffer (60 mM Tris-HCl, pH 6.8, 2% SDS (w/v), 10% glycerol (v/v), and 5% 2-mercaptoethanol (v/v)) followed by boiling for 5 min. The clarified supernatant was resolved with 15% gels run using a Hoeffer SE 250 mini-Vertical electrophoresis apparatus (GE Healthcare). Separation was achieved with a constant 20 mA/gel and a typical run time of 1.2 h. Gels were removed from the cassette and placed immediately in Coomassie Blue R-250 solution.

Western Blot Analysis. Western blot analysis was performed as described earlier.[18] Briefly, total seed proteins or 50% isopropanol extracted proteins were first resolved on a 15% SDS-PAGE gels. Subsequently, the proteins were transferred to a 0.45 μm nitrocellulose membrane. After blocking with TBS (10 mM Tris-HCl, pH 7.5, 500 mM NaCl) containing 5% nonfat dry milk, the nitrocellulose membranes were incubated overnight with either lunasin or synthesized BBi peptide (N-CVDITDFCYEPCKPSEDDK) antibodies that had been diluted 1:5000 in TBST (TBS with 3% nonfat dry milk containing 0.2% Tween 20). Following three washings with TBST, the membranes were incubated with goat anti-rabbit IgG–horseradish peroxidase conjugate. Proteins reacting specifically with the antibodies were detected using the Pierce SuperSignal West Pico kit's instruction (Rockford, IL, USA).

2-D Electrophoresis. Sample preparation for 2-D electrophoretic analysis was carried out as previously described.[22] For isoelectric focusing, 300 μg of protein sample was loaded per strip using overnight in-gel rehydration. Linear gradient, 13 cm IPG strips (GE Healthcare) were brought to a rehydration volume of 250 μL with 7 M urea, 2 M thiourea, 1% CHAPS, 2% C7BzO, 5% glycerol, and 2.2% 2-HED. Strips were then passively rehydrated with the entire rehydration solution containing protein sample at 23 °C for 15 h prior to focusing. Prior to the second dimension, IEF strips were equilibrated with 5% SDS in a urea-based solution (0.05 M Tris-Cl, pH 8.8, 6 M urea, 30% glycerol, and 0.1% bromophenol blue) containing 2% DTT for 20 min and again but with 2.5% iodoacetamide for 20 min. IEF strips were placed carefully onto a Hoeffer SE600 (GE Healthcare) 16%T vertical second-dimension and secured into place with a warm 1% agarose SDS-PAGE running buffer solution (0.2% SDS). Gels were run at an initial 10 mA/gel for 1 h followed by 5 W/gel for the remainder of the run (elimination of dye front; approximately 4 h). After electrophoresis, the gels were immediately removed and fixed in 5:4:1 methanol/water/acetic acid for 30 min, followed by two brief rinses in water, and stained in a Coomassie G-250 solution overnight.

Image Acquisition and Analysis. 2-D Coomassie-stained gels were destained with multiple changes of ultrapure H_2O to remove background. Individual gels of PI 427138 and LD00-3309 were scanned separately using an Epson V700 Perfection scanner controlled through Adobe Photoshop. Images were analyzed using Decodon Delta2D v3.6 (Greifswald, Germany) to provide normalized percent spot volume data using a technique known as differential gel imaging. Essentially, each gel image was imported into Delta2D individually and overlaid to provide a combined fusion image in gray scale, whereas the individual images of each were assigned either a blue (PI 427138) or an orange (LD00-3309) color. Differences in color between the two samples indicate a higher or lower amount of that particular protein in that sample. Delta2D parameters were set to maximize spot detection (using *global* image warping and *exact* spot matching) and minimize background detection.

RESULTS

PI 427138 Accumulates High Amounts of Leginsulin. BBI and the BBI-related family of isoinhibitors are rich in cysteine and are mainly responsible for the overall sulfur-containing amino acids of soybean seed.[10−12] During the course of screening a diverse array of soybean accessions from the USDA Soybean Germplasm Collection for the identification of soybean lines with high and low contents of Bowman–Birk protease inhibitor, we serendipitously observed that soybean PI 427138 accumulated high amounts of a 4 kDa protein. We have previously shown that Asian soybean accessions accumulate higher amounts of a 4 kDa protein than the genetic base of U.S. cultivars.[18] On the basis of Western blot analysis and mass spectrometry, this protein was identified as leginsulin.[18] Even though this protein accumulates in higher amounts in Asian soybean accessions, its contribution to overall protein content of the seed is only minimal. To confirm the identity of the 4 kDa protein that is enriched in PI 427138, we performed Western blot analysis using antibodies raised against the leginsulin peptide. The antibodies reacted strongly against the 4 kDa peptide, indicating that this protein is leginsulin (Figure 1).

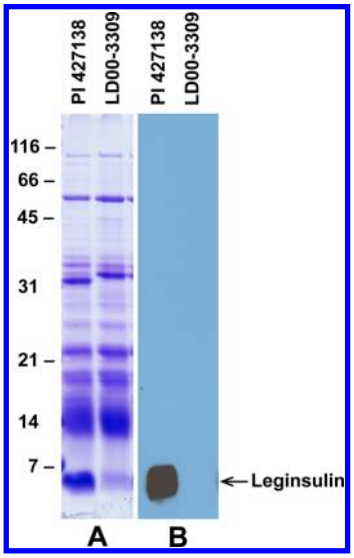

Figure 1. Accumulation of leginsulin in soybean seeds. Isopropanol (50%) soluble proteins were extracted from PI 427138 (lane 1) and LD00-3309 (lane 2), a high-yielding North American soybean cultivar, and analyzed by SDS-PAGE. Separated proteins were detected by staining the gel with Coomassie Blue (panel A) or electrophoretically transferred to a nitrocellulose membrane and reacted with leginsulin specific antibodies that were diluted 1:5000 in TBS containing 3% dry milk powder. Immunoreactive proteins were identified using anti-rabbit IgG–horseradish peroxidase conjugate followed by chemiluminescent detection (panel B). The arrow points to the 4 kDa immunoreactive peptide preferentially accumulating in PI 427138. The 4 kDa protein band in the stained gel is a mixture of the proteins, whereas the immunoreactive band shown in panel B is a single protein. The positions and sizes of the molecular weight markers in kilodaltons are shown on the left side of the figure. The same amount of protein used in panel A was used for the Western blot analysis (panel B). Also, note the size polymorphism of a prominent polypeptide (32–33 kDa) between PI 427138 and LD00-3309; however, the basis of this polymorphism is unknown (panel A).

Information on the country of origin, phenotypic data, and seed composition of PI 427138 was obtained from the USDA Germplasm Resources Information Network (GRIN) database (http://www.ars-grin.gov/cgi-bin/npgs/acc/search.pl?accid=PI+427138+). PI 427138 originated from South Korea, belongs to maturity group 0, and accumulates 48% protein and 14% oil. Earlier we reported the biochemical characterization of several high-protein soybean accessions.[23] By high-resolution two-

DOI: 10.1021/jf505202z
J. Agric. Food Chem. 2015, 63, 2862−2869

dimensional gel electrophoretic analysis, it was demonstrated that the 11S globulins (glycinin) were mainly responsible for the increased protein content of these high-protein soybean accessions. Because PI 427138 is also a high-protein accession, we wanted to examine its protein composition. For this purpose we performed high-resolution two-dimensional gel electrophoretic analysis of seed proteins of PI 427138 and compared it with LD00-3309, a high-yielding, publicly developed cultivar that contains 38% protein.[19] To monitor proteome differences between PI 427138 and LD00-3309, total seed proteins resolved by 2-DE were analyzed using Delta2D image analysis software (Figure 2). Delta2D parameters were set to maximize

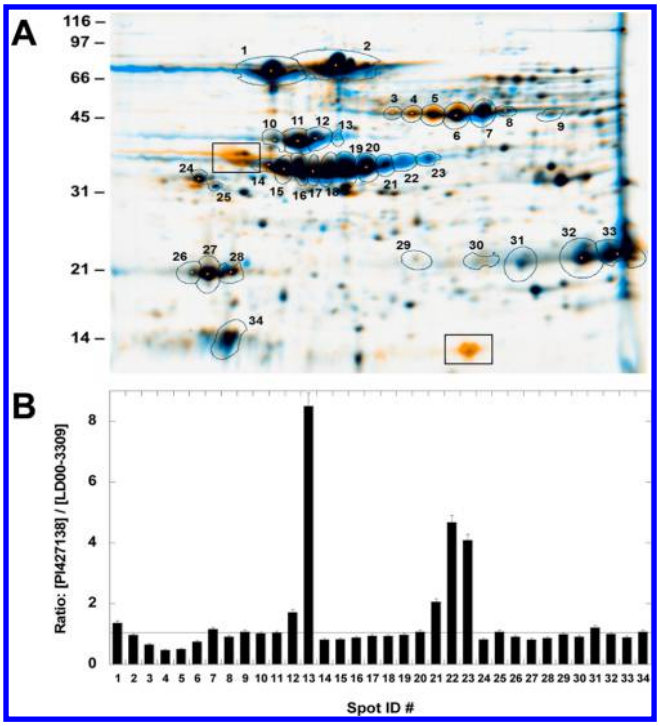

Figure 2. 2-DE comparison of total seed proteins extracted from PI 427138 and LD00-3309. Equal volumes of identically extracted seed proteins were separated using 2-DE. The Coomassie-stained gel images were analyzed using Delta2D software. Panel A shows the fusion of the two independently derived images. Gel images of PI 427138 and LD00-3309 were imported separately into Delta2D and overlaid to provide a combined fusion image. The blue color denotes those protein spots that were found to be higher in percent spot volume in PI 427138 compared with LD00-3309, whereas the orange color denotes those proteins that were found to be lower in percent spot volume. IEF separation was from pH 4 to 7, and molecular weight markers are designated in kDa. Panel B shows the ratios of percent volume [PI 427138]/[LD00-3309] of 34 spots generated from percent spot volume data collected using Delta2D image analysis software. The identities of the 34 proteins spots are shown in Supplemental Table 1. Leginsulin is not detected in these gels due to their low abundance in the total seed protein fraction.

spot detection (using global image warping and exact spot matching). A total of 34 protein spots identified previously by matrix-assisted laser desorption time of flight (MALDI-TOF) mass spectrometry as the major seed storage proteins of soybean were compared (Supplemental Table 1).[23] A comparison of the spot percent volume ratio differences clearly indicated that spots 12 (glycinin A3B4), 13 (glycinin A3B4), 21 (glycinin A1aBx precursor), 22 (proglycinin A1ab1b), and 23 (proglycinin A1ab1b) accumulated at higher amounts in PI 427138 (Figure 2). The preferential accumulation of the 11S globulins appears to the major contributing factor for the higher protein content of PI 427138.

Characterization Soybean Lines Derived from PI 427138 × LD00-3309 Crosses. The influence of environment on the seed composition traits of soybean LD00-3309, PI 427138, Leg-3, Leg-7, and Leg-8 were evaluated by growing these soybean lines at six locations, three in Illinois and three in Missouri. Total seed protein and oil content were determined with a LECO truSpec model FP-428 nitrogen analyzer and near-infrared (NIR) spectroscopy, respectively. The protein content of LD00-3309 seeds obtained from the six locations ranged from 36 to 39% (Figure 3A), whereas their oil content ranged from 19 to 21% (Figure 3B). In contrast, the protein content of PI 427138 seeds ranged from 45 to 51% (Figure 3A) and their oil content ranged from 11 to 13% (Figure 3B). The protein content of Leg-3, Leg-7, and Leg-8 ranged from 42 to 47%, which was significantly higher than that of LD00-3309, indicating the high-protein trait from PI 427138 had been successfully introgressed into these three lines (Figure 3A). The seed oil content of these soybean lines, which ranged from 14 to 17%, was intermediate when compared with the values obtained for the parents (Figure 3B). In general, soybeans grown in Columbia, MO, USA, had the highest protein content, whereas plants grown in Portageville, MO, USA, had the highest seed oil content.

The protein composition of leginsulin introgressed lines (Leg-3 and Leg-7) was also examined by 2-D gel electrophoresis (Figure 4). An examination of the Coomassie Blue stained gel revealed that PI 427138 accumulated higher amounts of 11S globulins when compared to LD00-3309. PI 427138 also failed to accumulate the Gy4 subunits. Interestingly, no substantial difference in the β-conglycinin content was observed between these two lines (Figure 4). The leginsulin introgressed lines (Leg-3 and Leg-7) also exhibited higher 11S glycinin content and, like PI 427138, lacked the Gy4 subunits (Figure 4).

Leginsulin Introgressed Lines Accumulate Lower Amounts of Bowman–Birk Protease Inhibitor. We also investigated the effect of growing location on the accumulation of leginsulin in soybean seeds. Isopropanol (50%) soluble proteins isolated from seeds grown at six locations were analyzed by SDS-PAGE (Figure 5A). A prominent 4 kDa protein was detected readily in leginsulin introgressed lines from all six locations. Western blot analysis using leginsulin antibodies confirmed the identity of this 4 kDa protein as leginsulin (Figure 5B). Previous studies have shown that the introduction of sulfur-rich proteins in soybean is often accomplished at the expense of native sulfur-rich proteins.[24] Bowman–Birk protease inhibitors are rich in cysteine and are the major contributors to the overall sulfur amino acid content of the seed. To investigate the effect of introgression of leginsulin on the accumulation of BBI, we performed Western blot analysis using antibodies raised against BBI. This analysis revealed a reduction in the accumulation of BBI in the leginsulin introgressed lines (Figure 5C). LD00-3309, which does not accumulate leginsulin, contained relatively high amounts of BBI (Figure 5C). These results suggest that accumulation of leginsulin in Leg-3 and Leg-7 occurs at the expense of BBI.

Leginsulin Introgressed Lines Contain Relatively Higher Amounts of Elemental Sulfur but the Concen-

DOI: 10.1021/jf505202z
J. Agric. Food Chem. 2015, 63, 2862–2869

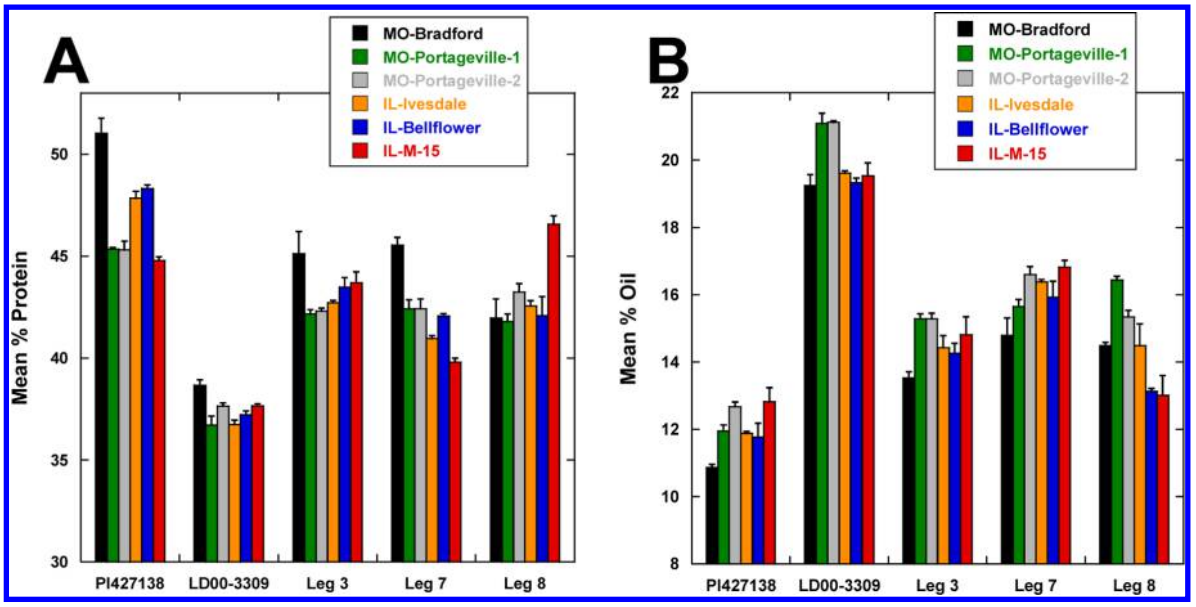

Figure 3. Mean protein and oil concentrations of soybean seeds grown in six different locations. Soybeans LD00-3309, PI 427138, Leg-3, Leg-7, and Leg-8 were grown at six locations (three in Illinois and three in Missouri). Seed protein and oil contents were determined with a LECO truSpec model FP-428 nitrogen analyzer and near-infrared (NIR) spectroscopy, respectively. Three separate biological replicates were evaluated per line. Error bars indicate the standard deviation of the mean.

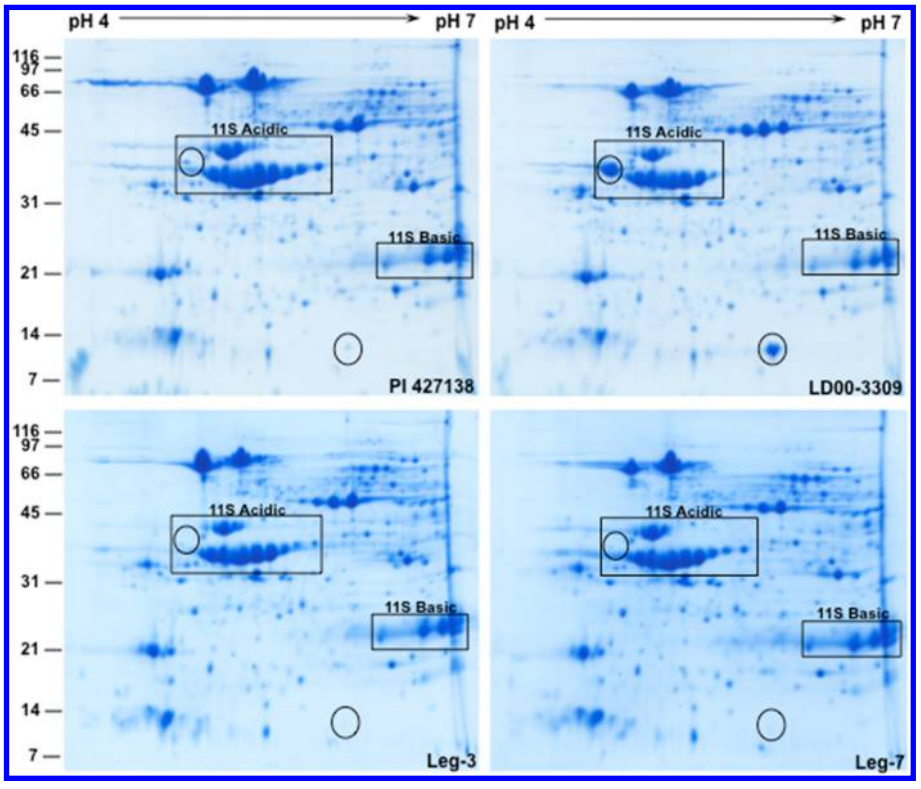

Figure 4. Comparison of the storage proteins of leginsulin introgressed soybean lines by 2-D gel electrophoresis. Total seed proteins isolated from the parents (PI 427138 and LD00-3309) and leginsulin introgressed lines (Leg-3 and Leg-7) were first separated by isoelectric focusing using a pH gradient from 4 to 7 and then by SDS-PAGE on a 13% gel. The rectangles in each panel represent the acidic and basic subunits of glycinin. The circles represent the Gy4 subunits, which are missing in PI 427138, Leg-3, and Leg-7. The gels were stained with Colloidal Coomassie Blue G-250. The positions and sizes of the molecular weight markers in kilodaltons are shown on the left side of the figure. Leginsulin is not detected in these gels due to their low abundance in the total seed protein fraction.

tration of Sulfur-Containing Amino Acids Remains Unchanged. The sulfur content of soybean seeds grown at six different locations was determined by inductively coupled plasma− mass spectrometry (Figure 6). Regardless of the location, LD00-3309 had the lowest amount of sulfur, which ranged from 3000 to 3600 ppm, whereas PI 427138 had the highest amount in most locations, ranging from 4000 to 4780 ppm (Figure 6). The sulfur content averaged from the six

DOI: 10.1021/jf505202z
J. Agric. Food Chem. 2015, 63, 2862−2869

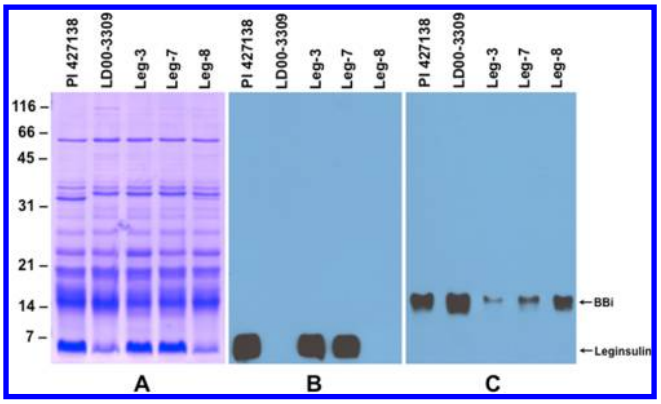

Figure 5. Leginsulin and Bowman–Birk protease inhibitor accumulation in soybean seeds. Fifty percent isopropanol soluble proteins isolated from the parents (PI 427138 and LD00-3309) and the resulting offspring (Leg-3, Leg-7, and Leg-8) were separated by SDS-PAGE on a 15% gel. Resolved proteins were detected by staining the gel with Coomassie Blue (panel A) or electrophoretically transferred to a nitrocellulose membrane and probed with leginsulin (panel B) or Bowman–Birk protease inhibitor (panel C) specific antibodies that were diluted 1:5000 in TBS containing 3% dry milk powder. Immunoreactive proteins were identified using anti-rabbit IgG–horseradish peroxidase conjugate followed by chemiluminescent detection.

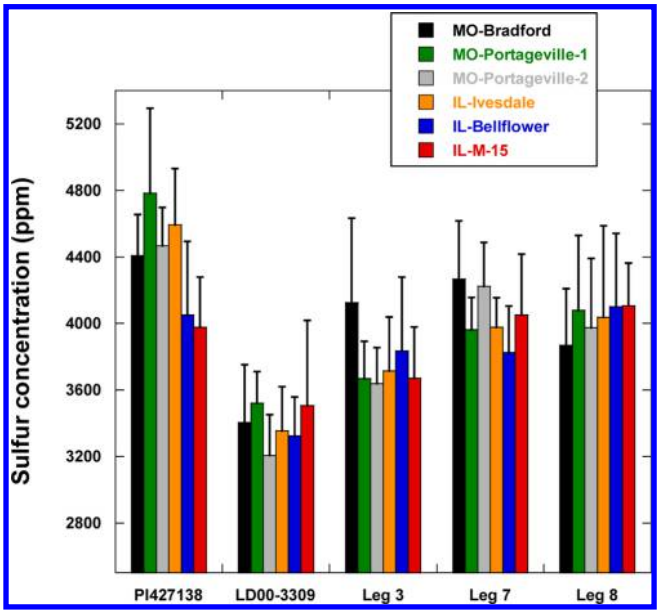

Figure 6. Elemental sulfur content of soybean seeds grown at six different locations. Soybeans LD00-3309, PI 427138, Leg-3, Leg-7, and Leg-8 were grown at six locations (three in Illinois and three in Missouri). Sulfur was determined by inductively coupled plasma–mass spectrometry. The error bars indicate the standard deviation of the mean.

locations for Leg-3, Leg-7, and Leg-8 were 3774, 4057, and 4027 ppm, respectively. These values are significantly higher than values obtained for LD00-3309. These results indicate that the high-sulfur trait from PI 427138 has been partially transferred to Leg-3, Leg-7, and Leg-8. It is also evident that growing location had only a minimal effect on sulfur accumulation in the seeds with a similar trend observed at all six locations (Figure 6).

The two soybean lines used as parents in this study, LD00-3309 and PI 427138, differ significantly in their sulfur contents. LD00-3309 averaged from the six locations had 3380 ± 264 ppm of sulfur, whereas PI 427138 had 4499 ± 113 ppm of sulfur. We were interested to examine if the differences seen in the sulfur content are also reflected in the concentration of sulfur-containing amino acids, cysteine and methionine. For this purpose we determined the amino acid content of soybean seeds of LD00-3309, PI 42713, Leg-3, Leg-7, and Leg-8. Interestingly the overall concentrations of cysteine and methionine were equivalent in all of these soybean lines (Table 1).

DISCUSSION

In this study we have successfully introgressed leginsulin from PI 427138 into a North American experimental line. Previously we have shown that leginsulin is encoded by two closely related genes (Gm13 g26330 (leginsulin 1) and Gm13 g26340 (leginsulin 2)). These two genes are located on chromosome 13 and are 5.9 kB apart.[18] Accumulation of leginsulin is drastically lower in most North American soybean cultivars than in Asian accessions.[18] Northern blot analysis has revealed that leginsulin accumulates only in the cotyledons but not in the embryonic axis in the North American soybean cultivars. In contrast, leginsulin mRNA was abundantly expressed in both the cotyledons and the embryonic axis of Asian soybean accessions. Thus, it appears the expression of leginsulin in both cotyledons and the embryonic axis may be responsible for the higher accumulation of leginsulin in Asian soybean accessions.[18]

Leginsulin, a peptide composed of 37 amino acids, is rich in cysteine residues. Our objective was to improve the overall sulfur amino acid content into adapted experimental lines by introgression of leginsulin. However, this goal was elusive because amino acid analysis revealed that the overall sulfur amino acid content (cysteine + methionine) remained unchanged (Table 1). Accumulation of Bowman–Birk protease inhibitor, another cysteine-rich protein, is reduced in leginsulin introgressed soybean lines Leg-3 and Leg-8, indicating that there is a limitation in the availability of cysteine in developing soybean seeds. Introgression of sulfur-rich proteins results in siphoning the limited cysteine pool from native sulfur-rich proteins such as Bowman–Birk protease inhibitor. Previous studies have shown the expression of Brazil nut 2S albumin, a methionine-rich protein, in transgenic soybean resulted in the reduction of Kunitz trypsin inhibitor.[24] Thus, it appears that introduction of sulfur-rich proteins into soybean alone will not be sufficient for improving the sulfur amino acid content. Metabolic engineering of sulfur assimilatory pathway will be critical for improving the availability of cysteine and methionine in developing soybean seeds.[25] Recent studies have demonstrated that such an approach can be used successfully to elevate the sulfur amino acid content of soybean seeds.[26,27]

Numerous high-protein plant introductions exist in the USDA Germplasm Collection. These plant introductions are currently being exploited as a source of high-protein alleles. The USDA ARS Soybean Genetics and Genomics Database (SoyBase) lists several potential protein QTLs located in different genomic regions. In addition, recent studies have identified new QTLs that could be used by breeders to develop high-protein lines.[28–30] We have transferred the high-protein trait of PI 427138 into Leg-3, Leg-7, and Leg-8, resulting in a substantial increase in the overall protein content. The increase

DOI: 10.1021/jf505202z
J. Agric. Food Chem. 2015, 63, 2862–2869

Table 1. Total Hydrolyzed Amino Acid Composition of Parental Seeds and Leginsulin Introgressed Crosses[a]

	mol %				
amino acid	PI 427138	LD00-3309	Leg-3	Leg-7	Leg-8
CyA	1.52 ± 0.21	1.49 ± 0.07	1.51 ± 0.01	1.45 ± 0.01	1.56 ± 0.05
His	2.16 ± 0.01	2.26 ± 0.01	2.15 ± 0.00	2.14 ± 0.02	2.24 ± 0.06
Ser	4.28 ± 0.30	5.15 ± 0.74	4.19 ± 0.17	5.27 ± 1.21	4.17 ± 0.35
Arg	6.19 ± 0.44	5.59 ± 0.01	5.87 ± 0.04	6.13 ± 0.35	5.91 ± 0.59
Gly	7.67 ± 0.06	7.56 ± 0.02	7.58 ± 0.12	7.37 ± 0.04	7.82 ± 0.18
Asp	11.63 ± 0.14	11.86 ± 0.09	11.92 ± 0.27	11.80 ± 0.11	11.85 ± 0.23
MetS	2.00 ± 0.11	1.98 ± 0.04	2.00 ± 0.04	1.94 ± 0.05	2.15 ± 0.01
Glu	16.81 ± 0.13	15.99 ± 0.16	16.55 ± 0.11	16.64 ± 0.51	16.17 ± 0.48
Thr	3.68 ± 0.12	4.07 ± 0.23	3.67 ± 0.05	3.96 ± 0.40	3.80 ± 0.25
Ala	6.23 ± 0.01	6.21 ± 0.04	6.29 ± 0.12	6.09 ± 0.01	6.38 ± 0.08
Pro	5.90 ± 0.07	5.78 ± 0.05	5.91 ± 0.01	5.91 ± 0.09	5.79 ± 0.06
Lys	5.97 ± 0.04	6.04 ± 0.05	6.16 ± 0.08	5.99 ± 0.05	6.18 ± 0.02
Tyr	2.46 ± 0.03	2.62 ± 0.00	2.18 ± 0.46	2.38 ± 0.06	2.30 ± 0.14
Val	6.24 ± 0.04	6.22 ± 0.16	6.29 ± 0.06	5.88 ± 0.45	6.27 ± 0.06
Ile	5.15 ± 0.06	5.10 ± 0.21	5.31 ± 0.01	4.99 ± 0.42	5.27 ± 0.11
Leu	7.58 ± 0.04	7.77 ± 0.11	7.82 ± 0.04	7.58 ± 0.20	7.69 ± 0.09

[a]Dry seed powder from triplicate samples was analyzed by HPLC to measure the total hydrolyzed amino acids from the parents (PI 427138 and LD00-3309) and the resulting offspring (Leg-3, Leg-7, and Leg-8). Cysteine is reported as cysteic acid (CyA), and methionine is reported as methionine sulfone (MetS). Results are the mean of three samples (n = 3), reported along with standard error (SE) of the mean.

in the protein content is, however, accompanied by a reduction in the oil content. It is well established that there is a negative correlation between seed protein and oil content.[31−35] Typically an increase of 2% protein results in a 1% decrease in oil content.[36] It has been proposed that this inverse correlation is associated with either a pair of tightly linked QTLs for protein and oil or by one pleiotropic QTL, the two alleles of which have inverse effects on both oil and protein contents.[28] Recent studies have identified new oil QTLs that have no major impact on seed protein content.[28−30] These oil-related QTLs could be exploited to increase the oil content of Leg-3, Leg-7, and Leg-8 without lowering the protein content.

Our study has identified PI 427138 as a *Gy*4 mutant. 2-D gel electrophoresis analysis of the total seed proteins from PI 427138 reveals the absence of the A5, A4, and B3 glycinin subunits, the product of the Gy4 gene. Leg-3, Leg-7, and Leg-8 generated in this study were also found to lack A5, A4, and B3 glycinin subunits (Figure 4). Tofu, a food product made of soy milk, is widely consumed in Asian countries. It is well-known that the tofu firmness, texture, and volume can be influenced by the absence of specific glycinin polypeptides.[37−39] For this reason *Gy*4 null lines such as 'Enrei' and 'Raiden' are commonly utilized in breeding programs to develop high-quality tofu cultivars. Thus, the soybean experimental lines Leg-3, Leg-7, and Leg-8 could be valuable breeding material for the production of high-quality tofu. It is not known if there is any correlation between the concentration of sulfur-containing amino acids and tofu quality. The 7S β-conglycinin is relatively low in sulfur amino acids, and its absence has been shown to alter tofu quality.[37] Additional studies are required to establish a correlation or lack thereof between the sulfur amino acid content of soybean seed proteins and tofu quality.

■ ASSOCIATED CONTENT

Supporting Information

Supplemental Table 1 and Supplemental Figures 1 and 2. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*(H.B.K.) Mail: Plant Genetics Research Unit, USDA-ARS, 108 Curtis Hall, University of Missouri, Columbia, MO 65211, USA. E-mail: Hari.Krishnan@ars.usda.gov.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. M. S. Pathan for growing soybeans at University of Missouri Fisher Delta Research Center, Portageville, MO, USA. Product names are necessary to report factually on available data; however, the University of Missouri and the USDA neither guarantee nor warrant the standard of product, and the use of the name by the University of Missouri and the USDA implies no approval of the product to the exclusion of others that may be suitable.

■ REFERENCES

(1) Imsande, J. Selection of soybean mutants with increased concentrations of seed methionine and cysteine. *Crop Sci.* **2001**, *41*, 510−515.

(2) Krishnan, H. B. Engineering soybean for enhanced sulfur amino acid content. *Crop Sci.* **2005**, *45*, 454−461.

(3) Townsend, J. A.; Thomas, L. A. Factors which influence the *Agrobacterium*-mediated transformation of soybean. *J. Cell Biochem. Suppl.* **1994**, *18A*, 78.

(4) Nordlee, J. A.; Taylor, S. L.; Townsend, J. A.; Thomas, L. A.; Bush, R. K. Identification of a Brazil-nut allergen in transgenic soybeans. *N. Engl. J. Med.* **1996**, *334*, 688−692.

(5) Dinkins, R. D.; Reddy, R.; Curtis, A.; Meurer, C. A.; Yan, B.; Trick, H.; Thhibaud-Nissen, F.; Finer, J. J.; Parrott, W. A.; Collins, G. B. Increased sulfur amino acids in soybean plants overexpressing the maize 15 kD zein protein. *In Vitro Cell Dev. Biol. Plant* **2001**, *37*, 742−747.

(6) Kim, W. S.; Krishnan, H. B. Expression of an 11 kDa methionine rich delta-zein in transgenic soybean results in the formation of two types of novel protein bodies in transitional cells situated between the vascular tissue and storage parenchyma cells. *Plant Biotechnol. J.* **2004**, *2*, 199−210.

DOI: 10.1021/jf505202z
J. Agric. Food Chem. 2015, 63, 2862−2869

(7) Kho, C. J.; de Lumen, B. O. Identification and isolation of methionine-cysteine rich proteins in soybean seeds. *Plant Foods Hum. Nutr.* **1988**, *38*, 287−296.
(8) George, A. A.; de Lumen, B. O. A novel methionine-rich protein in soybean seed: identification, amino acid composition, and N-terminal sequence. *J. Agric. Food Chem.* **1991**, *39*, 224−227.
(9) de Lumen, B. O.; Galvez, A. F.; Revilleza, M. J.; Krenz, D. C. Molecular strategies to improve the nutritional quality of legume proteins. *Adv. Exp. Med. Biol.* **1999**, *464*, 117−126.
(10) Birk, Y. The Bowman-Birk inhibitor. *Int. J. Pept. Protein Res.* **1985**, *25*, 113−131.
(11) Clarke, E. J.; Wiseman, J. Developments in plant breeding for improved nutritional quality of soya beans II. Anti-nutritional factors. *J. Agric. Sci.* **2000**, *134*, 125−136.
(12) Livingstone, D.; Beilinson, V.; Kalyaeva, M.; Schmidt, M. A.; Herman, E. M.; Nielsen, N. C. Reduction of protease inhibitor activity by expression of a mutant Bowman-Birk gene in soybean seed. *Plant Mol. Biol.* **2007**, *64*, 397−408.
(13) Nishizawa, N. K.; Mor, S.; Watanabe, Y.; Hirano, H. Ultrastuructural localization of the basic 7S globulin in soybean (*Glycine max*) cotyledons. *Plant Cell Physiol.* **1994**, *35*, 1079−1085.
(14) Rahioui, I.; Laugier, C.; Balmand, S.; Da Silva, P.; Rahbe, Y.; Gressent, F. Toxicity, binding and internalization of the pea-A1b entomotoxin in Sf9 cells. *Biochimie* **2007**, *89*, 1539−1543.
(15) Watanabe, Y.; Hirano, H. Nucleotide sequence of the basic 7S globulin gene from soybean. *Plant Physiol.* **1994**, *105*, 1019−1020.
(16) Hanada, K.; Nishiuchi, Y.; Hirano, H. Amino acid residues on the surface of soybean 4-kDa peptide involved in the interaction with its binding protein. *Eur. J. Biochem.* **2003**, *270*, 2583−2592.
(17) Hirano, H. 4-kDa peptide. In *Handbook of Biologically Active Peptides*; Kastin, A. J., Ed.; Academic Press: Burlington, MA, USA, 2006; pp 1−4.
(18) Kim, W.-S.; Jang, S.; Krishnan, H. B. Accumulation of leginsulin, a hormone-like bioactive peptide, is drastically higher in Asian than in North American soybean accessions. *Crop Sci.* **2012**, *52*, 262−271.
(19) Diers, B. W.; Cary, T. R.; Thomas, D. J.; Nickell, C. D. Registration of 'LD00-3309' soybean. *Crop Sci.* **2006**, *46*, 1384.
(20) Samuels, M. L.; Witmer, J. A. *Statistics for the Life Sciences*, 2nd ed.; Prentice-Hall: London, UK, 1999.
(21) Krishnan, H. B.; Jang, S.; Baxter, I.; Wiebold, W. J. Growing location has a pronounced effect on the accumulation of cancer chemopreventive agent Bowman-Birk Inhibitor in soybean seeds. *Crop Sci.* **2012**, *52*, 1788−1794.
(22) Krishnan, H. B.; Oehrle, N. W.; Natarajan, S. S. A rapid and simple procedure for the depletion of abundant storage proteins from legume seeds to advance proteome analysis: a case study using *Glycine max*. *Proteomics* **2009**, *9*, 3174−3188.
(23) Krishnan, H. B.; Nelson, R. L. Proteomic analysis of high protein soybean (*Glycine max*) accessions demonstrates the contribution of novel glycinin subunits. *J. Agric. Food Chem.* **2011**, *59*, 2432−2439.
(24) Streit, L. G.; Beach, L.; Register, J. C.; Jung, R., III; Fehr, W. R. Association of the Brazil nut protein gene and Kunitz trypsin inhibitor alleles with soybean protease inhibitor activity and agronomic traits. *Crop Sci.* **2001**, *41*, 1757−1760.
(25) Jez, J. M.; Krishnan, H. B. Sulfur assimilation and cysteine biosynthesis in soybean seeds: towards engineering sulfur amino acid content. In *Modification of Seed Composition to Promote Health and Nutrition*; Krishnan, H. B., Ed.; ASA-CSSA-SSSA Publishing: Madison, WI, USA, 2009; pp 249−261.
(26) Kim, W. S.; Chronis, D.; Juergens, M.; Schroeder, A. C.; Hyun, S. W.; Jez, J. M.; Krishnan, H. B. Transgenic soybean plants overexpressing *O*-acetylserine sulfhydrylase accumulate enhanced levels of cysteine and Bowman-Birk protease inhibitor in seeds. *Planta* **2012**, *235*, 13−23.
(27) Song, S.; Hou, W.; Godo, I.; Wu, C.; Yu, Y.; Matityahu, I.; Hacham, Y.; Sun, S.; Han, T.; Amir, R. Soybean seeds expressing feedback-insensitive cystathionine-y-synthase exhibit a higher content of methionine. *J. Exp. Bot.* **2013**, *64*, 1917−1926.
(28) Sun, Y.-K. *Identification of Soybean Seed Oil QTLs with Little or No Impact on Seed Protein*. M.S. thesis, Student Research in Agronomy and Horticulture, University of Nebraska, Dec 2011.
(29) Eskandari, M.; Cober, E. R.; Rajcan, I. Genetic control of soybean seed oil: II. QTL and genes that increase oil concentration without decreasing protein or with increased seed yield. *Theor. Appl. Genet.* **2013**, *126*, 1677−1687.
(30) Hwang, E.-Y.; Song, Q.; Jia, G.; Specht, J. E.; Hyten, D. L.; Costa, J.; Cregan, P. B. A genome-wide association study of seed protein and oil content in soybean. *BMC Genomics* **2014**, *15*, 1.
(31) Thorne, J. C.; Fehr, W. R. Incorporation of high-protein, exotic germplasm into soybean populations by 2- and 3-way crosses. *Crop Sci.* **1970**, *10*, 652−655.
(32) Brim, C. A.; Burton, J. W. Recurrent selection in soybeans. II. Selection for increased percent protein in seeds. *Crop Sci.* **1979**, *19*, 494−498.
(33) Hartwig, E. E.; Kilen, T. C. Yield and composition of soybean seed from parents with different protein, similar yield. *Crop Sci.* **1991**, *31*, 290−292.
(34) Helms, T. C.; Orf, J. H. Protein, oil, and yield of soybean lines selected for increased protein. *Crop Sci.* **1998**, *38*, 707−711.
(35) Cober, E. R.; Voldeng, H. D. Developing high-protein, high-yield soybean populations and lines. *Crop Sci.* **2000**, *40*, 39−42.
(36) Schwender, J.; Ohlrogge, J. B.; Shachar Hill, Y. A flux model of glycolysis and the oxidative pentosephosphate pathway in developing *Brassica napus* embryos. *J. Biol. Chem.* **2003**, *278*, 29442−29453.
(37) Poysa, V.; Woodrow, L.; Yu, K. Effect of soy protein subunit composition on tofu quality. *Food Res. Int.* **2006**, *39*, 309−317.
(38) Kim, W. S.; Ho, H. J.; Nelson, R. L.; Krishnan, H. B. Identification of several gy4 nulls from the USDA soybean germplasm collection provides new genetic resources for the development of high-quality tofu cultivars. *J. Agric. Food Chem.* **2008**, *56*, 11320−11326.
(39) Kim, W. S.; Gillman, J. D.; Krishnan, H. B. Identification of a plant introduction soybean line with genetic lesions affecting two distinct glycinin subunits and evaluation of impacts on protein content and composition. *Mol. Breed.* **2013**, *32*, 291−298.

DOI: 10.1021/jf505202z
J. Agric. Food Chem. 2015, 63, 2862−2869