

Article

Development and characterization of a soybean experimental line lacking the β' subunit of β -conglycinin and G1, G2, and G4 glycinin

Bo Song, Nathan W. Oehrle, Shanshan Liu, and Hari B. Krishnan

J. Agric. Food Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jafc.7b05011 • Publication Date (Web): 11 Dec 2017

Downloaded from <http://pubs.acs.org> on December 26, 2017

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.



ACS Publications

Development and characterization of a soybean experimental line lacking the α' subunit of β -conglycinin and G1, G2, and G4 glycinin

Bo Song^{1,2}, Nathan W. Oehrle², Shanshan Liu^{1,*}, and Hari B. Krishnan^{2,3,*}

¹Key Laboratory of Soybean Biology at the Chinese Ministry of Education, Northeast Agricultural University, Harbin, China.

²Plant Genetics Research Unit, Agricultural Research Service, U.S. Department of Agriculture, University of Missouri, Columbia, MO, USA.

³Plant Science Division, University of Missouri, Columbia, MO, USA

Corresponding authors

*Plant Genetics Research Unit, Agricultural Research Service, U.S. Department of Agriculture, University of Missouri, Columbia, MO 65211. E-mail: Hari.Krishnan@ARS.USDA.GOV

*Key Laboratory of Soybean Biology at the Chinese Ministry of Education, Northeast Agricultural University, Harbin, China. E-mail: ars336699@aliyun.com

ABSTRACT: A soybean experimental line (BSH-3) devoid of a subset of seed storage proteins was developed by crossing a mutant donor line ‘HS99B’ with a Chinese cultivar ‘Dongnong47’ (DN47). One dimensional and high-resolution 2-D gel electrophoresis revealed the absence of G1 (A₁B₂), G2 (A₂B_{1a}), and G4 (A₅A₄B₃) glycinin and the α' subunit of β -conglycinin in BSH-3 seeds. Despite the lack of these abundant seed proteins, BSH-3 seeds still accumulated 38% protein. BSH-3 seeds also accumulated high levels of free amino acids as compared to DN47 seeds, particularly arginine, and the amount of several essential amino acids were significantly elevated in BSH-3 seeds. Elevated accumulation of α and β -subunit of β -conglycinin, G5 glycinin, Kunitz trypsin inhibitor, and Bowman-Birk protease inhibitor indicates seed proteome rebalancing in BSH-3 seeds. Immunoblot analysis using sera from soybean allergic patients demonstrated the complete lack of a major allergen (α' subunit of β -conglycinin) in BSH-3 seeds. However, elevated levels of other allergens were found in BSH-3 seeds due to proteome rebalancing. Transmission electron microscopy observation of mature seeds of BSH-3 revealed striking differences in the appearance of the protein storage vacuoles when compared to DN47.

KEY WORDS: Allergens, glycinins, β -conglycinin, proteome rebalancing, protein storage vacuole, soybean

INTRODUCTION

The two major commercially important components of soybean seeds are the protein and oil. North American soybean cultivars contain approximately 34-37% protein and 19% oil. Soybeans are an exceptional protein source since they contain a well-balanced amino acid profile with the exception of sulfur-containing amino acids.^{1,2} On account of the high protein quality and content, soybeans are extensively used in animal feed and more recently in aquaculture. The abundant seed storage proteins of soybeans are classified into 7S and 11S globulins based on their sedimentation coefficients.^{3,4} The 7S globulins are referred to as β -conglycinin and are composed of three subunits, α' (72 kDa), α (70 kDa) and β (52 kDa).⁵ The β -conglycinin are glycoproteins and are encoded by multigene families. The 11S globulins are named as glycinin and are encoded by at least five genes (*gy1*, *gy2*, *gy3*, *gy4*, and *gy5*).⁶ Recently, a few additional glycine genes have been recognized some of which are either not expressed or expressed at very low levels.^{7,8} The products of five glycinin genes are classified into Group I and Group II based on their sequence homology. Additionally, Group II glycinins have been subdivided into Group IIa and Group IIb. Glycinins are synthesized as larger precursor proteins which are subsequently processed into an acidic and basic subunit that are held together by sulfide bonds. The members of the three groups of glycinins, though they exhibit high amino acid sequence homology, reveal differences at the carboxyl terminal end of the acidic subunit.⁸

Glycinin and β -conglycinin account for about 70% of the total seed protein content of soybean.⁹ These two group of proteins are mainly responsible for the nutritive value of soybean seed proteins and impact the quality of food products derived from soybean. For example, the quantity and quality of tofu, which is made from soymilk, is influenced by the 7S and 11S protein subunit composition.¹⁰ Naturally occurring mutants or radiation induced mutations that

affect the accumulation of soybean seed storage proteins have been well documented in the literature.¹¹⁻¹⁶ Japanese researchers have developed a soybean line that lacks all the glycinin and β -conglycinin subunits by integrating different mutations through crossbreeding.¹⁵ Similarly, soybean lines lacking β -conglycinin^{17,18} or both glycinin and β -conglycinin were developed by RNA interference.^{19,20} Even though these soybean lines lacked the major seed storage proteins, they were able to grow and reproduce normally indicating that these proteins are dispensable.

Some of the soybean seed storage protein mutants have been biochemically characterized.^{15,20} Soybean mutant lines that lacked both the glycinin and β -conglycinin were reported to contain equivalent nitrogen content when compared to that of the wild-type cultivars.^{15,20} The absence of major seed storage proteins was compensated by the accumulation of free amino acids, especially arginine.¹⁵ The absence of the abundant seed proteins resulted in preferential increases in the accumulation of lipoxygenase, sucrose binding protein, agglutinin and the basic 7S globulin.¹⁵ Analyses utilizing systems biology techniques (proteomics, metabolomics, and transcriptomics) indicated that rebalancing of protein composition in mutant line is accomplished with small alterations to the seed transcriptome and metabolome.²⁰ Here, we report the development and characterization of a soybean experimental line (BSH-3) that is devoid of several abundant seed storage proteins.

MATERIALS AND METHODS

Plant materials. Chinese cultivar ‘Dongnong47’ (DN47), a high-oil elite soybean cultivar, was selected as a recurrent parent. A soybean line (‘RiB’, also referred as ‘HS99B’) that lacks α' and α subunits of β -conglycinin and some components of group I and IIa glycinin^{14,15} was used

as an introgression line. For breeding purpose DN47 was used as the female parent and ‘HS99B’ as the male parent. A scheme outlining the steps involved in the development of a soybean experimental line (BSH-3) is depicted in Supplemental Figure 1. DN47 and BSH3 were grown in a field at the soybean experimental plot at the research farm of the Key Laboratory of Soybean Biology and Breeding/Genetics of the Chinese Agriculture Ministry. Blocks comprising four rows of plants for each line were 3 m in length and were positioned 70 cm apart, with a plant distance of 14 cm in a randomized complete block design (RCBD) with three replications.

Seed protein, oil and fatty acid profile. Dry seeds of DN47 and BSH3 were harvested at maturity and stored at room temperature. Seeds from ten plants were pooled and ground to a fine powder with a coffee grinder. These seed samples were utilized for protein and amino acid analyses. Total nitrogen/protein of soybean seed was measured using the LECO model FP-428 nitrogen analyzer (LECO Corporation, Michigan, USA). The protein content of seeds was determined by multiplying the nitrogen content value with a conversion factor of 6.25. Three biological replicates were evaluated per line. The oil content was quantified by near-infrared reflectance (NIR) spectroscopy (Tecator AB, Hoganas, Sweden). The fatty acid profiles of soybean were determined by gas chromatograph.²¹ Briefly, crushed seeds were extracted overnight with 5 mL of chloroform: hexane: methanol (8:5:2, v/v/v). Fatty acids from 100 μ L aliquots of the extract were methylated with 75 μ L of methanolic sodium methoxide:petroleum ether:ethyl ether (1:5:2, v/v/v). Fatty acids were separated utilizing Agilent Series 6890 capillary gas chromatograph (Palo Alto, CA, USA) that was fitted with an AT-Silar capillary column (Alltech Associates, Deerfield, IL, USA). Standard fatty acid mixtures were used for determining relative amounts of each fatty acid.

Protein isolation and immunoblot analysis. To isolate total seed proteins, 10 mg of dry seed powder was extracted with 1ml of sample extraction buffer (2% SDS, 60mM Tris-HCl, pH 6.8, 10% glycerol and 5% β -mecaptoethanol), followed by boiling at 100° C for 5 min. The samples were clarified by centrifugation at 15800xg for 5 min. The resulting clear supernatant was designated as the total seed protein fraction and used for SDS-PAGE analysis. Sera from five patients with documented allergy to soybean were obtained from Plasma Lab International (Everett, WA). ImmunoCAP *in vitro* quantitative assay showed 7.1-24.5 kU/L soybean-specific IgE in the sera of these patients. Sera from these five soybean-allergic patients were pooled and used for immunoblot analyses at 1:1000 dilution. Electrophoresis and immunoblot analysis was performed as described earlier.¹⁸

Trypsin and chymotrypsin inhibitor activity assays. Protease inhibitor assays were performed following established protocol.²² Briefly, dry seed powder (50 mg) was extracted with 1 mL of 50 mM Tris-HCl, pH 8.0 by vigorously mixing on a vortexer for 10 min at room temperature. Following centrifugation at 15800xg for 5 min the clear supernatant was used to measure the trypsin or chymotrypsin inhibitory activity. Trypsin inhibitor activity was measured in 2 ml Eppendorf tubes. Trypsin (20 μ g) was added to the assay mixture (Tris-HCl, 20 mM CaCl_2 , pH 8.2) and incubated for 15 min at 37° C. To this assay mixture 1 mM BAPNA was added and incubated for additional 10 min. The reaction was terminated by the addition of 30% acetic acid and the absorbance at 410 nm was recorded. One trypsin inhibitor unit (TIU) was defined as the amount of inhibitor that reduces the absorbance of the non-inhibited reaction by 0.01. Chymotrypsin inhibitor activity was measured by incubating known amounts to seed protein extracts with chymotrypsin (80 μ g) in the assay buffer (Tris-HCl, 20 mM CaCl_2 , pH 7.8) at 37° C for 15 min. N-glutaryl-L-phenylalanine-4-nitroanilide (GLUPHEPA, 1 mM) was added

and the assay mixture was left for additional 45 min at 37° C. Reactions were terminated by the addition of 30% acetic acid and absorbance at 410 nm was recorded. One chymotrypsin inhibitor unit (CIU) was defined as the amount of inhibitor that decreases the absorbance of the non-inhibited reaction by 0.01.

Two-dimensional gel electrophoresis. Two-dimensional gel electrophoresis of soybean seed proteins was performed as described earlier.²³ Coomassie stained gels were analyzed for proteome differences using Delta2D image analysis software.

Amino acid analysis. Total amino acid content was determined by hydrolysis of soybean seed powder with 6N HCl for 22 h in sealed evacuated tubes at a constant boiling temperature of 110° C. An amino acid analyzer (Hitachi L-8000, Hitachi, Tokyo, Japan) was used to determine the amino acid composition of the hydrolysates. For the quantification of methionine and cysteine, duplicate samples were first subjected to an initial oxidation step using performic acid prior to acid hydrolysis. Samples were run in triplicates and subjected to appropriate statistical analysis. Free amino acid content was measured as described earlier.²²

Transmission electron microscopy. Dry seeds of DN47 and BSH-3 seeds were germinated on 1% water plates for 12 h in a 30° C incubator. The seeds were dissected into 2-4 mm cubes with a razor and immediately fixed in 2.5% glutaraldehyde buffered at pH 7.2 with 50 mM sodium phosphate for 4 h at room temperature. Following the primary fixation, the seed samples were rinsed four times in sodium phosphate buffer. Half of the seed samples were transferred to new vials and post-fixed for 1 h with 1% aqueous osmium tetroxide. The seed tissue was extensively washed in distilled water and dehydrated in a graded acetone series and infiltrated with Spurr's resin as described earlier.²⁴ Thin sections of the seed tissue were cut, collected on

cooper grids and stained with 0.5% uranyl acetate and 0.4% lead citrate. The stained sections were examined at 80 kV under JEOL 1200 EX (Tokyo, Japan) transmission electron microscope.

RESULTS

Development of soybean experimental line (BSH-3). β -conglycinin, the 7S globulins of soybean, is composed of three subunits of 72, 70 and 52 kDa, respectively. All three subunits of β -conglycinin have been identified as allergens²⁵ and the elimination of these subunits have been shown to improve nutritional quality and functionality of soybean proteins.^{26,27} We are interested in developing soybean cultivars lacking one or all subunits of β -conglycinin. Earlier we have described the development of advanced backcrossing breeding lines that are devoid of allergenic α -subunit of β -conglycinin.^{22,28,29} In addition to β -conglycinin, some of the 11S glycinins have also been shown to be food allergens.³⁰ Thus, the development of soybean lines lacking the allergenic subunits of both the 7S β -conglycinin and the 11S glycinin will be desirable. Till now such hypoallergenic soybean lines have not been developed in a Chinese soybean genetic background. In this study, we have used HS99B, a soybean line lacking several subunits of both 7S β -conglycinin and the 11S glycinin as male parent, and DN47, a high-oil elite Chinese cultivar, as a female parent in backcross breeding program as outlined in Supplemental Figure 1. To Transfer the desirable trait (lack of the allergenic subunits of 7S β -conglycinin and the 11S glycinin subunits) into Chinese soybean genetic background, successive backcross was carried out from 2008 to 2010. BC_2F_2 generation was developed and selfed to achieve homozygosity (BC_2F_6 progeny) for the deficiency of the allergenic subunits of 7S β -conglycinin and the 11S glycinin subunits. The introgression of the desirable trait was monitored throughout the breeding scheme by SDS-PAGE. Soybean lines with the described protein composition were further

selected based on their agronomic performance. Finally, an advanced breeding line (BSH-3) with desirable agronomic characteristics was selected for further characterization.

BSH-3 fail to accumulate several abundant seed storage proteins. The protein composition of DN47 and BSH-3 was examined by 1-D gel electrophoresis. A comparison of the total seed protein profiles demonstrates the complete absence of the α' subunit of β -conglycinin and a drastic reduction in the accumulation of the acidic and basic subunits of glycinin (Figure 1). The absence of these abundant seed proteins in BSH-3 seeds resulted in enhanced accumulation of α and β -subunits of β -conglycinin and A3 glycinin subunit. We also fractionated the seed proteins by calcium precipitation²³ into seed storage and non-storage protein fractions. An examination of the seed storage protein fraction of DN47 and BSH-3 seeds confirmed similar changes that were observed with the total seed protein fraction (Figure 1). Interestingly, a comparison of non-storage protein fraction revealed no drastic changes in their profiles, though a slight increase in few proteins were also observed (Figure 1).

Additional comparison of DN47 and BSH-3 seeds was also performed by high-resolution 2-D gel electrophoresis. The seed proteins of DN47 were resolved into numerous discrete protein spots. Earlier studies have identified 2-D resolved protein spots by mass spectrometry.^{31,32} The most abundant protein spots of DN47 represented the different subunits of glycinin and β -conglycinin. In contrast, 2-D gel analysis of BSH-3 seeds revealed that several abundant protein spots were missing. BSH-3 failed to accumulate the α' subunit of β -conglycinin, glycinin G1, G2 and G4. This observation was confirmed by overlaying the DN47 and BSH-3 2-D protein profiles using Delta2D software (Figure 2). The absence of these proteins was compensated by an increased accumulation of the α - and β -subunit of β -conglycinin and glycinin G5 (Figure 2).

Protein, oil and fatty acid profile of DN-47 and BSH-3 seeds. To investigate if loss of abundant seed proteins resulted in any alterations of seed components we examined the total protein and oil content of these seeds. Total nitrogen/protein content of DN47 and BSH-3 seeds were measured using LECO nitrogen analyzer. BSH-3 seeds, in spite of lacking several abundant seed proteins, had a higher protein content than DN47 (Supplemental Table 1). Determination of oil content with near-infrared reflectance spectroscopy revealed that DN47 had 3% higher oil content than BSH-3 (Supplemental Table 1). Analysis of the five major fatty acids of soybean seeds by gas chromatography showed differences in their profiles (Supplemental Table 1). A slight decrease in palmitic acid (16:0), stearic acid (18:0), and oleic acid (18:1) was observed in BSH-3 seeds when compared to that of DN47. In contrast, the amount of linoleic acid and linolenic acid was significantly higher in BSH-3 seeds (Supplemental Table 1).

BSH-3 seeds accumulate high levels of free amino acids. Previous studies have shown that soybean mutants lacking the seed storage proteins accumulate high levels of free amino acids.^{15,20} To examine if similar situation also occurred in BSH-3 we determined the total and free amino acid content (Table 1). An examination of the total amino acid composition revealed a slight increase in several of the essential amino acids in BSH-3 when compared to that of DN-47 (Table 1). The sulfur-containing amino acid (Met + Cys) content of BSH-3 was 59% higher than that of DN47. A comparison of the free amino acid content revealed significant increases in the concentration of several amino acids in BSH-3 seeds especially arginine and glutamic acid (Table 1). Interestingly, the concentration of free methionine and free cysteine levels were lower in BSH-3 seeds (Table 1).

Loss of abundant seed proteins results in altered levels of allergens in BSH-3 seeds.

Our earlier study has demonstrated that all the three subunits of β -conglycinin are potential allergens.²⁵ The experimental line developed in this study (BSH-3) lacks some of the known allergens such as G1 and G2 glycinin and the α' subunit of β -conglycinin. In order to examine if the loss of these proteins in BSH-3 has altered the profile of other soybean allergens, we performed immunoblot analysis. Ig-E antibodies from the sera from pooled soybean-sensitive patients reacted against the α' subunit of β -conglycinin, which was completely absent in BSH-3 seeds (Figure 3). However, the Ig-E antibodies also cross-reacted with few other proteins that were not detected in DN-47 seeds. Since there is no strict correlation between IgE binding activity and allergenicity, it is not known if the cross-reactive proteins have any clinical relevance. Serum from pooled soybean-sensitive pigs reacted against all three subunits of β -conglycinin present in DN-47 seeds (Figure 3). The reaction against the α and β -conglycinin was much stronger in BSH-3 seed extracts. A recent study has shown glycinin G5 is also a potential soybean allergen.³³ Earlier, we have purified this protein and raised polyclonal antibodies.³⁴ We utilized this antibody to examine any differences in the accumulation of this allergenic protein in DN47 and BSH-3 seeds. Immunoblot analysis clearly demonstrates that BSH-3 seeds accumulated higher amounts of this potential allergen when compared to that of DN47 seeds (Figure 3).

BSH-3 seeds accumulate elevated levels of KTi and BBi. Previously we have shown 50% isopropanol can be used to isolate trypsin and chymotrypsin inhibitors from soybean seeds.²² 50% isopropanol extracted proteins were used in western blot analysis to compare the KTi and BBi levels in DN47 and BSH-3 seeds. Both KTi and BBi levels were markedly elevated in BSH-3 seeds when compared to that of DN47 (Figure 4). Similarly, trypsin and chymotrypsin

inhibitor activity measurement also clearly indicated that BSH-3 seeds exhibited higher trypsin and chymotrypsin inhibitor activities (Figure 4).

Transmission electron microscopy of BSH-3 seeds reveals changes in the appearance of protein storage vacuoles. Previous studies have shown that suppression of 11S β -conglycinin proteins promotes the formation of endoplasmic reticulum derived protein bodies in soybean seeds.¹⁷ To examine if the loss of several abundant seed proteins in BSH-3 led to the formation of protein bodies we carried out electron microscopy observation of thin sections of DN47 and BSH-3 seeds. Even though small protein bodies were sporadically observed they were not routinely observed in BSH-3 seeds. Earlier we have shown that the morphology and structural integrity of protein storage vacuoles (PSV) is affected by post-fixation of soybean seeds with osmium tetroxide.²⁴ Electron microscopy observation of osmicated thin sections of soybean seeds revealed prominent oil bodies and large protein storage vacuoles, the storage compartment for the native glycinin and β -conglycinin (Figure 5). The protein storage vacuoles in DN47 seeds were completely filled and had a uniform granular appearance. They also accumulated phytate crystals, which appear as small holes within the PSV (Figure 5). Examination of thin-sections of osmicated BSH-3 seeds also revealed similar structural features as observed with DN47. In contrast, a comparison of DN47 and BSH-3 non-osmicated tissues revealed a drastic appearance of PSV. Like DN47, the BSH seeds contained numerous PSV containing phytate crystals. However, the appearance of PSV in BSH-3 were strikingly different from that of DN47. Narrow bands of light staining areas alternating with dark staining areas were prominently seen in these PSV (Figure 5). In contrast, the PSV in DN47 had uniform granular appearance without alternating dark and light staining regions (Figure 5).

DISCUSSION

In this study, we have developed an advanced soybean breeding line, BSH-3 devoid of several abundant seed proteins. High-resolution 2-D gel electrophoresis revealed the absence of G1, G2, and G4 glycinin and the α' subunit of β -conglycinin in BSH-3 seeds. Interestingly, in spite of the absence of several abundant seed proteins, the overall protein content of BSH-3 is higher than the parental line, DN47. Analysis of BSH-3 seed protein composition reveals that the loss of abundant seed proteins has resulted in proteome rebalancing. The loss of these abundant proteins resulted in enhanced accumulation of α and β subunits of β -conglycinin and G5 glycinin subunit. Proteome rebalancing due to loss of abundant seed storage proteins has been previously reported.^{15,20} Japanese workers have developed a soybean line lacking both the glycinin and β -conglycinin, yet the nitrogen content of these seeds was found to be similar to that of the wild-type cultivars.¹⁵ The absence of the abundant seed proteins, however, resulted in preferential increase in the accumulation of lipoxxygenase, sucrose binding protein, agglutinin and the basic 7S globulin.¹⁵ Similarly, a transgenic soybean line lacking both glycinin and beta-conglycinin (SP-) was developed by RNA interference.²⁰ Just as observed by the Japanese researchers,¹⁵ these transgenic soybean line also revealed an increase in the accumulation of a few proteins. Thus, in cases where the abundant seed proteins are removed by either RNAi or integrating mutations it leads to proteomic rebalancing by promoting the accumulation of less abundant seed proteins to maintain overall nitrogen content of the seeds.^{15,20}

The absence of major seed storage proteins is compensated by the accumulation of free amino acids.^{15,20} We also found that there was a significant increase in the accumulation of free amino acids, especially arginine in the seeds of BSH-3 seeds. A marginal increase in several of the total amino acid was found in BSH-3 compared to DN47. Statistically significant increases

in both of the essential and non-essential amino acid concentration was found in BSH-3 seeds compared to DN47 seeds. The sulfur amino acids (Met and Cys) content was also elevated seeds due to proteome rebalancing in BSH-3 seeds. The effect of removing the abundant seed proteins on the overall sulfur amino acid content of soybean seeds is debatable. A comparison of the amino acid composition of soybean seed reveals that glycinins are relatively rich in sulfur amino acids than β -conglycinins.¹ Thus, it is to be expected that elimination of β -conglycinin could led to increased levels of sulfur-containing amino acids in soybean seeds.^{35,36} However, some studies have shown that elimination of β -conglycinin did not significantly elevate total sulfur amino acid composition of soybean seeds.^{15,18,20} In this study, we have observed a significant increase the sulfur amino acid content in BSH-3 seeds. This may be due an increase in the accumulation of Bowman-Birk protease inhibitor, a cysteine-rich protein.

Soybean seed storage proteins, glycinin and β -conglycinin, which are synthesized in the endoplasmic reticulum (ER) are transported to protein storage vacuoles (PSV). In cereals, the abundant seed proteins (prolamin) are directly deposited within the lumen of rough ER resulting in the formation protein bodies (PB). Protein bodies are normally found in soybean. However, the appearance of ER-derived protein bodies has been reported in transgenic soybeans where the accumulation of α' and α subunits of β -conglycinin were suppressed.¹⁷ The role of different glycinin and β -conglycinin subunits in the formation of PB have also been investigated.^{36,37} Electron microscopic observation revealed a close relationship between the formation of PB and glycinin subunits. ER-derived PB were not seen in soybean mutants lacking glycinin group 1 subunits, while the frequency of ER-derived PB was greatly increased in mutants accumulating higher amounts of glycinin group I subunits.^{36,37} Previous ultrastructural studies of soybean seed storage protein mutants have not observed any alterations in the morphology of PSV. In the

current study, we observed a striking difference in the appearance of PSV in BSH-3 seeds. Interestingly, these differences were evident only when post-fixation with osmium tetroxide was omitted. Earlier we have demonstrated the profound effect of osmium tetroxide on the appearance and structural integrity of PSV in soybeans.²⁴ Osmium tetroxide primarily interact with lipids and forms cross-links with proteins facilitating their structural integrity. Since BSH-3 seeds fail to accumulate several abundant proteins, the absence of post-fixation with osmium tetroxide may have resulted in weekend cross-linking of proteins in PSV leading to their altered morphology.

Soybean seed proteins have been identified as significant food allergens. Approximately 2% adults and 5-8% of infants in USA and Europe are allergic to soybeans.³⁹ About 33 soybean proteins ranging in molecular weight from 7 to 71 kDa have been reported to bind IgE antibodies from soybean allergic patients. Gly m Bd 30 K, Gly m Bd 28 K, Gly m Bd 60 K, and Gly1 and Gly2 glycinin proteins have been identified and studied in some detail.^{30,40-42} Soybean meal is the predominant protein source in animal feed. Research has shown soybean meal can elicit immunological response in monogastric animals and may affect the optimal growth of poultry and livestock.^{43,44} β -conglycinin, the 7S storage protein of soybean, has been identified as the most allergenic protein in the soybean meal. On account of the importance of soybean in animal feed and human diet, attempts have been made to develop hypoallergenic soybeans. Two approaches have been undertaken to develop hypoallergenic soybeans. The first approach involves identification of mutants from soybean germplasm collections. Currently mutants lacking the individual or all the three subunits of β -conglycinin has been identified and exploited in breeding programs to develop hypoallergenic soybean cultivars. Biotechnology approaches have also been employed to eliminate the suppression of allergens in soybean. Gene-silencing

techniques have been employed to eliminate the p34/Gly m Bd 30 K⁴⁵ and Gly m 5.¹⁸ Interestingly, the elimination of abundant seed proteins results in proteome rebalancing. In the current study, we have observed that the elimination of major allergens (α' subunit of β -conglycinin and subset of glycinin) in BSH-3 seeds has resulted in an increase in the concentration of other allergens. In situations where the abundant seed proteins are eliminated it is important to monitor the levels of all potential allergens. Thus, development of hypoallergenic soybeans will be a major challenge on account of significant intrinsic allergenicity of several of the soybean proteins.

ABBREVIATIONS

ER, endoplasmic reticulum; BBi, Bowman Birk protease inhibitor; chymotrypsin inhibitor; GLUPHEPA, N-glutaryl-L-phenylalanine-4-nitroanilide; KTi, Kuntiz trypsin inhibitor; NIR, near-infrared reflectance; PB, protein bodies; PSV, protein storage vacuole.

ACKNOWLEDGEMENTS

This research was supported by China Research Service : Ministry of Science and Technology of China (2016YFD0100504 and 2017RAQXJ104) and a State Scholarship Fund provided to Bo Song by the China Scholarship Council, and funds from the USDA-Agricultural Research Service. 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.

Supporting Information Available: [Schematic outline of development of soybean experimental line (BSH-3) derived from DN47 x HS99B crosses (Figure S1), Protein, oil and fatty acid content of DN47 and BSH-3 seeds (Table S1).

REFERENCES

- (1) Krishnan, H. B. Engineering soybean for enhanced sulfur amino acid content *Crop Sci.* **2005**, *45*, 454–461.
- (2) 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.
- (3) Nielsen, N. C. Soybean Seed Composition. In *Soybean: Genetics, Molecular Biology and Biotechnology*; Verma, D. P. S., Shoemaker, R. C., Eds.; CABI: Wallingford, U.K., 1996.
- (4) Krishnan, H. B. Biochemistry and molecular biology of soybean seed storage proteins. *J. New Seeds* **2000**, *2*, 1–25.
- (5) Thanh, V. H.; Shibasaki, K. β -Conglycinin from soybean proteins. Isolation and immunological and physicochemical properties of the monomeric forms. *Biochim. Biophys. Acta* **1977**, *490*, 370–384.
- (6) Nielsen, N. C.; Dickinson, C. D.; Cho, T. J.; Thanh, B. H.; Scallan, B. J.; Fischer, R. L.; Sims, T. L.; Drews, G. N.; Goldberg, R. B. Characterization of the glycinin gene family. *Plant Cell* **1989**, *1*, 313–328.
- (7) Beilinson, V.; Chen, Z.; Shoemaker, R.C.; Fischer, R.L.; Goldberg, R.B.; Nielsen, N.C. Genomic organization of *glycinin* genes in soybean. *Theor. Appl. Genet.* **2002**, *104*, 1132–1140.

- 377 (8) Li, C.; Zhang, Y.M. Molecular evolution of glycinin and β -conglycinin gene families in
378 soybean (*Glycine max* L. Merr.). *Heredity* **2011**, *106*, 633–641.
- 379 (9) Hill, J. E.; Breidenbach, R.W. Proteins of soybean seeds. I. Isolation and characterization of
380 the major components. *Plant Physiol.* **1974**, *53*, 742–746.
- 381 (10) Poysa, V.; Woodrow, L.; Yu, K. Effect of soy protein subunit composition on tofu
382 quality *Food Res. Int.* **2006**, *39*, 309–317.
- 383 (11) Kitamura, K.; Kaizuma, N. Mutant strains with low level of subunits of 7S globulin in
384 soybean (*Glycine max* Merr.) seed. *Jpn. J. Breed.* **1981**, *31*, 353–359.
- 385 (12) Ladin, B.F.; Doyle, J.J.; Beachy, R.N. Molecular characterization of deletion mutation
386 affecting the α' subunits of β -conglycinin. *J. Mol. Appl. Genet.* **1984**, *2*, 372–389.
- 387 (13) Takahashi, K.; Banba, H.; Kikuchi, A.; Ito, M.; Nakamura, S. An induced mutant line
388 lacking the α -subunit of β -conglycinin in soybean (*Glycine max* (L.) Merrill). *Jpn. J. Breed.*
389 **1994**, *44*, 65–66.
- 390 (14) Yagasaki, K.; Kaizuma, N.; Kitamura, K. Inheritance of glycinin subunits and
391 characterization of glycinin molecules lacking the subunits in soybean (*Glycine max* (L.)
392 Merr.). *Breeding Sci.* **1996**, *46*, 11–15.
- 393 (15) Takahashi, M.; Uematsu, Y.; Kashiwaba, K.; Yagasaki, K.; Hajika, M.; Matsunaga, R.;
394 Komatsu, K.; Ishimoto, M. Accumulation of high levels of free amino acids in soybean
395 seeds through integration of mutations conferring seed protein deficiency. *Planta* **2003**, *217*,
396 577–586.

- 397 (16) Kim, W.-S.; Gillman, J.; Krishnan, H., Identification of a plant introduction soybean line
398 with genetic lesions affecting two distinct glycinin subunits and evaluation of impacts on
399 protein content and composition. *Mol. Breeding* **2013**, *32*, 291-298.
- 400 (17) Kinney, A.J.; Jung, R.; Herman, E.M. Cosuppression of the α subunits of β -conglycinin
401 in transgenic soybean seeds induces the formation of endoplasmic reticulum-derived protein
402 bodies. *Plant Cell* **2001**, *13*, 1165–1178.
- 403 (18) Kim, W-S.; Jez, J.M.; Krishnan, H.B. Effects of proteome rebalancing and sulfur
404 nutrition on the accumulation of methionine rich δ -zein in transgenic soybeans. *Front. Plant*
405 *Sci.* **2014**, *5*:633. doi: 10.3389/fpls.2014.00633
- 406 (19) Schmidt, M.A.; Herman, E.M. Proteome rebalancing in soybean seeds can be exploited to
407 enhance foreign protein accumulation. *Plant Biotech. J.* **2008**, *6*, 832–842.
- 408 (20) Schmidt, M.A.; Barbazuk, W.B.; Sandford, M.; May, G.; Song, Z.; Zhou, W.; Nikolau,
409 B.J.; Herman, E.M. Silencing of soybean seed storage proteins results in a rebalanced
410 protein composition preserving seed protein content without major collateral changes in the
411 metabolome and transcriptome. *Plant Physiol.* **2011**, *156*, 330–345.
- 412 (21) Bilyeu, K.; Palavalli, L.; Sleper, D.; Beuselinck, P. Mutations in soybean microsomal
413 omega-3 fatty acid desaturase genes reduce linolenic acid concentration in soybean seeds.
414 *Crop Sci.* **2005**, *45*, 1830–1836.
- 415 (22) Song, B.; Oehrle, N.W.; Liu, S.; Krishnan, H.B. Characterization of seed storage proteins
416 of several perennial soybean species. *J. Agric. Food Chem.* **2016**, *64*, 8499-8508.

- 417 (23) Krishnan, H.B.; Oehrle, N.W.; Natarajan, S.S. A rapid and simple procedure for the
418 depletion of abundant storage proteins from legume seeds to advance proteome analysis: A
419 case study using *Glycine max*. *Proteomics* **2009**, *9*, 3174-3188.
- 420 (24) Krishnan, H.B. Preparative procedures markedly influence the appearance and structural
421 integrity of protein storage vacuoles in soybean seeds. *J. Agric. Food Chem.* **2008**, *56*,
422 2907-2912.
- 423 (25) Krishnan, H.B.; Kim, W.S.; Jang, S.; Kerley, M. All three subunits of soybean β -
424 conglycinin are potential food allergens. *J. Agric. Food Chem.* **2009**, *57*, 983-943.
- 425 (26) Nakamura, T.; Utsumi, T.; Mori, T. Cultivar differences in gelling characteristics of
426 soybean Glycinin. *J. Agric. Food Chem.* **1984**, *32*, 647-651.
- 427 (27) Ogawa, A.; Samoto, M.; Takahashi, K. Soybean allergens and hypoallergenic soybean
428 products. *J. Nutr. Sci. Vitaminol.* **2000**, *46*, 271-279.
- 429 (28) Song, B.; Shen, L.W.; Wei, X.S.; Guo, B.W.; Tuo, Y.; Tian, F.D.; Li, W.B.; Liu, S.S.
430 Marker-assisted backcrossing of a null allele of the α -subunit of soybean (*Glycine max*) β -
431 conglycinin with a Chinese soybean cultivar. *Plant Breeding* **2014**, *133*, 638-648.
- 432 (29) Song, B.; An, L.; Han, Y.; Gao, H.; Ren, H.; Zhao, X.; Wei, X.; Krishnan, H.B.; Liu, S.
433 Transcriptome profile of near-isogenic soybean lines for β -conglycinin α -subunit deficiency
434 during seed maturation. *PLoS One* **2016**, *11*(8):e0159723.
- 435 (30) Xiang, P.; Beardslee, T. A.; Zeece, M. G.; Markwell, J.; Sarath, G. Identification and
436 analysis of a conserved immunoglobulin E binding epitope in soybean G1a and G2a and
437 peanut Ara h 3 glycinins. *Arch. Biochem. Biophys.* **2002**, *408*, 51-57.

- 438 (31) Mooney, B.P.; Krishnan, H.B.; Thelen, J.J. High-throughput peptide mass fingerprinting
439 of soybean seed proteins: automated workflow and utility of UniGene expressed sequence
440 tag databases for protein identification. *Phytochemistry* **2004**, *65*, 1733-1744.
- 441 (32) Hajduch, M.; Ganapathy, A.; Stein, J.W.; Thelen, J.J. A Systematic proteomic study of
442 seed filling in soybean. Establishment of high-resolution two-dimensional reference maps,
443 expression profiles, and an interactive proteome database. *Plant Physiol.* **2005**, *137*, 1397-
444 1419.
- 445 (33) Gagnon, C.; Poysa, V.; Cober, E.R.; Gleddie, S. Soybean allergens affecting North
446 American patients identified by 2D gels and mass spectrometry. *Food Anal. Methods* **2010**,
447 *3*, 363-374.
- 448 (34) Gillman, J.D.; Kim, W-S.; Song, B.; Oehrle, N.W.; Tawari, N.R.; Liu, S.; Krishnan, H.B.
449 Whole genome resequencing identifies the molecular genetic cause for the absence of a Gy5
450 glycinin protein in soybean PI 603408. *G3, Genes/Genomes/Genetics* **2017**, *5*, 2345-2352.
- 451 (35) Ogawa, T.; Tayama, E.; Kitamura, K.; Kaizuma, N. Genetic improvement of seed storage
452 protein using three variant alleles of 7S globulin subunits in soyabean. (*Glycine max* L.).
453 *Jpn. J. Breed.* **1989**, *39*, 137-147.
- 454 (36) Panthee, D.R.; Kwanyuen, P.; Sams, C.E.; West, D.R.; Saxton, A.M.; Pantalone, V.R.
455 Quantitative trait loci for β -conglycinin (7S) and glycinin (11S) fractions of soybean storage
456 protein. *J. Am. Oil Chem. Soc.* **2004**, *81*, 1005–1012.

- 457 (37) Herman, E.M.; Schmidt, M.A. ER to vacuole trafficking (ERvt) provides an alternate
458 pathway for endomembrane progression for transfer to the vacuole. *Plant Physiol.* **2004**,
459 *136*, 3440–3446.
- 460 (38) Mori, T.; Maruyama, N.; Nishizawa, K.; Higasa, T.; Yagasaki, K.; Ishimoto, M.; Utsumi,
461 S. The composition of newly synthesized proteins in the endoplasmic reticulum
462 determines the transport pathways of soybean seed storage proteins. *Plant J.* **2004**, *40*, 238–
463 249.
- 464 (39) Herman, E.M.; Burks, A.W. The impact of plant biotechnology on food allergy. *Curr.*
465 *Opinion Biotechnology* **2011**, *22*, 224–30.
- 466 (40) Ogawa, T.; Tsuji, H.; Bando, N.; Kitamura, K.; Zhu, Y.L.; Hirano, H.; Nishikawa, K. Ide
467 ntification of the soybean allergenic protein, Gly m Bd 30K, with the soybean seed 34-kDa
468 oil-body-associated protein *Biosci., Biotechnol., Biochem.* **1993**, *57*, 1030–1033.
- 469 (41) Shibasaki, M.; Suzuki, S.; Tajima, S.; Nemoto, H.; Kuroume, T. Allergenicity of major
470 component proteins of soybean. *Int. Arch. Allergy Appl. Immunol.* **1980**, *61*, 441–448.
- 471 (42) Ogawa, T.; Bando, N.; Tsuji, T.; Nishikawa, K.; Kitamura, K. α -Subunit of β -
472 conglycinin, an allergenic protein recognized by IgE antibodies of soybean-sensitive
473 patients with atopic dermatitis *Biosci., Biotechnol., Biochem.* **1995**, *59*, 831–833.
- 474 (43) Zheng, S.; Qin, G.; Tian, H.; Sun, Z. 2014. Role of soybean β -conglycinin subunits as
475 potential dietary allergens in piglets. *Vet. J.* **2014**, *199*, 434–438.
- 476 (44) Wu, J.J.; Cao, C.M.; Ren, D.D.; Zhang, Y.; Kou, Y.N.; Ma, L.Y.; Feng, S.B.; Li, Y.;
477 Wang, X.C. (2016) Effects of soybean antigen proteins on intestinal permeability, 5-

hydroxytryptamine levels and secretory IgA distribution in the intestine of weaned piglets.

Ital. J. Anim. Sci. **2016**, *15*, 174-180.

(45) Herman, E.M.; Helm, R.M.; Jung, R.; Kinney, A.J. Genetically modification removes and immunodominant allergen form soybean. *Plant Physiol.* **2003**, *132*, 36–43.

FIGURE CAPTIONS

Figure 1. SDS-PAGE analysis of of DN47 and BSH-3 seed proteins. Total seed proteins (Panel A), calcium precipitated storage proteins (Panel B) and residual proteins after calcium precipitation (Panel C) were fractionated by SDS-PAGE on a 12% gel and visvuaized with Coomassie Blue. Proteins in BSH-3 whose abundance is either increased (shown with a green arrow) or decreased (shown with a red arrow) are shown (Panel B). Note unlike the calcium precipitated seed storage proteins, the protein profile of residual proteins after calcium precipitation are very similar (Panel C). M, protein markers, lane 1, DN-47; and lane 2, BSH-3.

Figure 2. Overlay of 2D gels of DN-47 and BSH-3 seed proteins using Delta2D software. Seed proteins of DN-47 and BSH-3 were individually separated by isoelectric focusing on IPG strips and then by SDS-PAGE on 13.5% gels. Images of Coomassie Blue stained gels were imported into Delta2D, assigned two different colors (green = DN-47; red = BSH-3) and overlaid to provide a combined fusion image. The red color denotes those protein spots that were found to be higher in BSH-3. Yellow demonstrates similar protein quantities in each, green color demonstrates absence of that particular protein species in BSH-3 seeds. Spots 1, 2 and 3

represent the 7S β -conglycinin subunits and spots 4-8 represents the different glycinin subunits. Spots 9 and 10 represents the KTi and BBi, respectively.

Figure 3. Detection of allergens in DN47 and BSH-3 seeds. Total seed proteins were separated on a 10% SDS-PAGE gels. Resolved proteins were stained with Coomassie Blue (Panel A) or transferred to nitrocellulose membrane and incubated with pooled sera from soybean-sensitive patients (panel B) or sera from pooled pig sera (panel C) or glycinin A3-specific antibodies. Immunoreactive proteins were detected by chemiluminescence using anti-human IgE horseradish peroxidase conjugate, anti-pig IgG horseradish peroxidase conjugate and anti-rabbit IgG horseradish peroxidase conjugate, respectively. M, protein markers, lane 1, DN-47; and lane 2, BSH-3.

Figure 4. Kunitz (KTi) and Bowman Birk (BBi) protease inhibitors in DN47 and BSH-3 seeds. 50% isopropanol-soluble proteins were resolved on a 15% SDS-PAGE gels. Resolved proteins were stained with Coomassie Blue (Panel A) or transferred to nitrocellulose membrane and incubated with BBi peptide antibodies (panel B) or KTi antibodies (panel C). Trypsin (KTi) and chymotrypsin inhibitor (BBi) activity was measured using N-benzoyl-L-arginine ethyl ester and N-glutaryl-L-phenylalanine-4-nitroanilide as substrates, respectively (panel D). Inhibitor units were defined as the amount of inhibitor that decreases the absorbance of the non-inhibited reaction by 0.1. Significantly different means are indicated by ** ($p \leq 0.05$).

Figure 5. Transmission electron microscopy observation of DN47 and BSH-3 seeds. PSV, protein storage vacuole; OB, oil bodies.

Table 1. Protein-bound and free amino acid content of DN47 and BSH-3 seeds.

	A.A.		F.A.A	
	DN47	BSH-3	DN47	BSH-3
Essential amino acid				
Thr	1.40±0.08	1.48±0.03	0.0117±0.0015	0.0170±0.0017*
Val	1.51±0.06	1.70±0.02**	0.0147±0.0035	0.0147±0.0006
Met	0.38±0.08	0.67±0.04**	0.0127±0.0035	0.0057±0.0015*
Ile	1.56±0.10	1.54±0.03	0.0037±0.0006	0.0067±0.0006*
Leu	2.74±0.03	3.00±0.01**	0.0050±0.0000	0.0070±0.0000*
Phe	1.82±0.04	1.88±0.02	0.0327±0.0116	0.0280±0.0010
Lys	2.29±0.03	2.44±0.02**	0.0160±0.0000	0.0200±0.0010*
T.E.A.A	11.70±0.15	12.69±0.06**	0.0963±0.0189	0.0990±0.0036
Non-essential amino acid				
Asp	3.93±0.07	4.13±0.06*	0.0553±0.0015	0.0627±0.0035
Ser	1.81±0.06	2.13±0.04**	0.0107±0.0006	0.0103±0.0006
Glu	5.98±0.15	6.55±0.07**	0.0540±0.0035	0.1003±0.0021*
Gly	1.46±0.02	1.59±0.01**	0.0093±0.0006	0.0123±0.0012*
Ala	1.48±0.02	1.53±0.03*	0.0133±0.0032	0.0183±0.0015*
Cys	0.49±0.04	0.72±0.03**	0.0337±0.0055	0.0300±0.0017
Tyr	1.10±0.09	1.06±0.01	0.0107±0.0097	0.0157±0.0040
His	0.91±0.03	0.98±0.01*	0.0040±0.0010	0.0163±0.0025*
Arg	2.44±0.11	2.99±0.06**	0.0647±0.0081	0.5660±0.0671*
Pro	1.77±0.07	1.88±0.02	0.0110±0.0020	0.0180±0.0026*
T.A.A.	33.06±0.26	36.26±0.37**	0.3617±0.0169	0.9507±0.0787*
T.S.A.A.	0.87±0.11	1.39±0.07**	0.0463±0.0090	0.0357±0.0006
Total N	38.37±0.73	40.68±0.14**		

Dry seed powder from triplicate samples were analyzed by HPLC to measure the total and free amino acid from DN47 and BSH-3. Significantly different means are indicated by * ($p\leq0.05$) or ** ($p\leq0.01$).

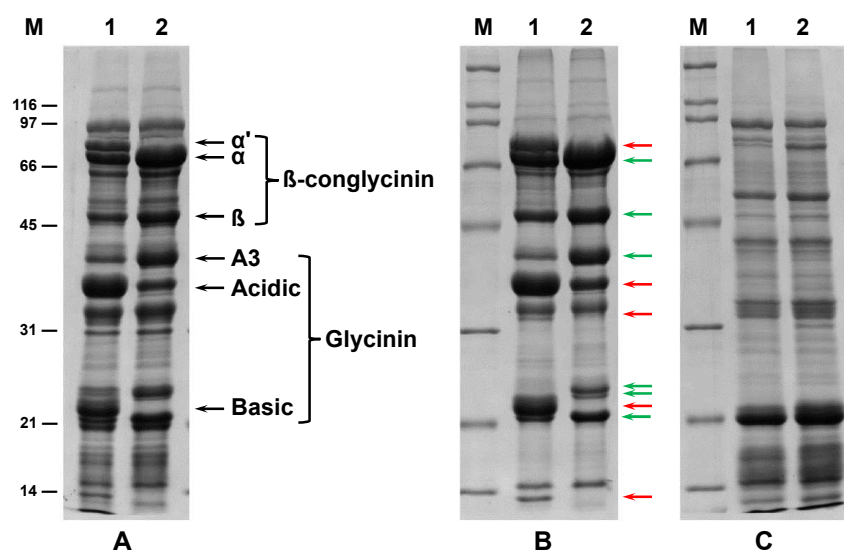
Figure 1

Figure 2

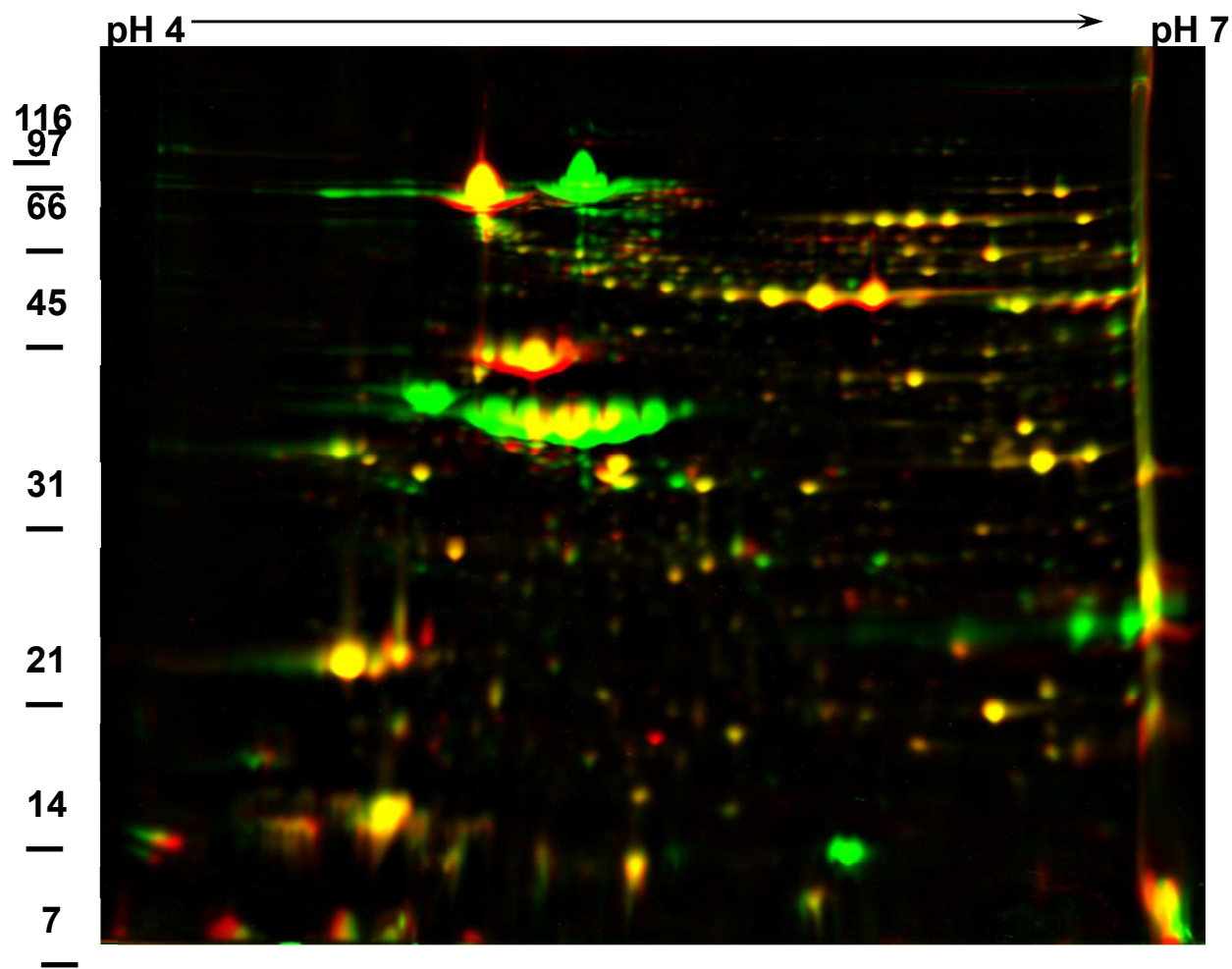


Figure 3

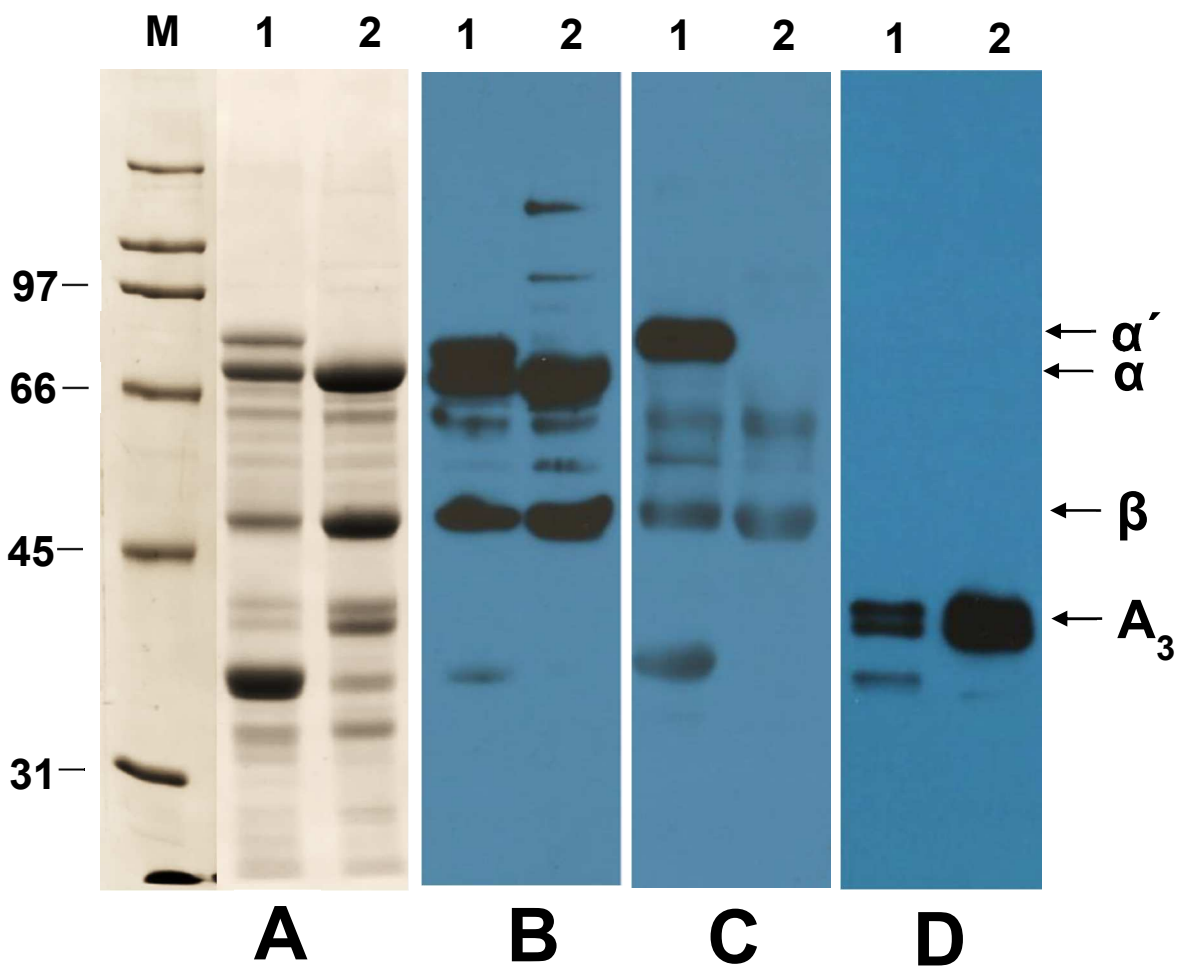


Figure 4

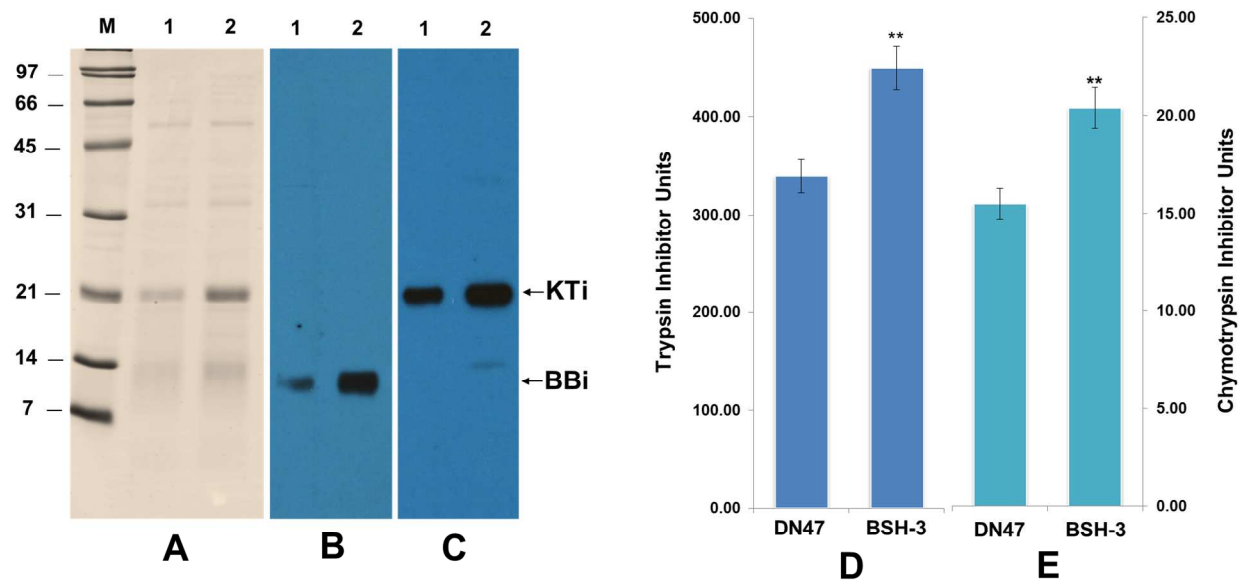
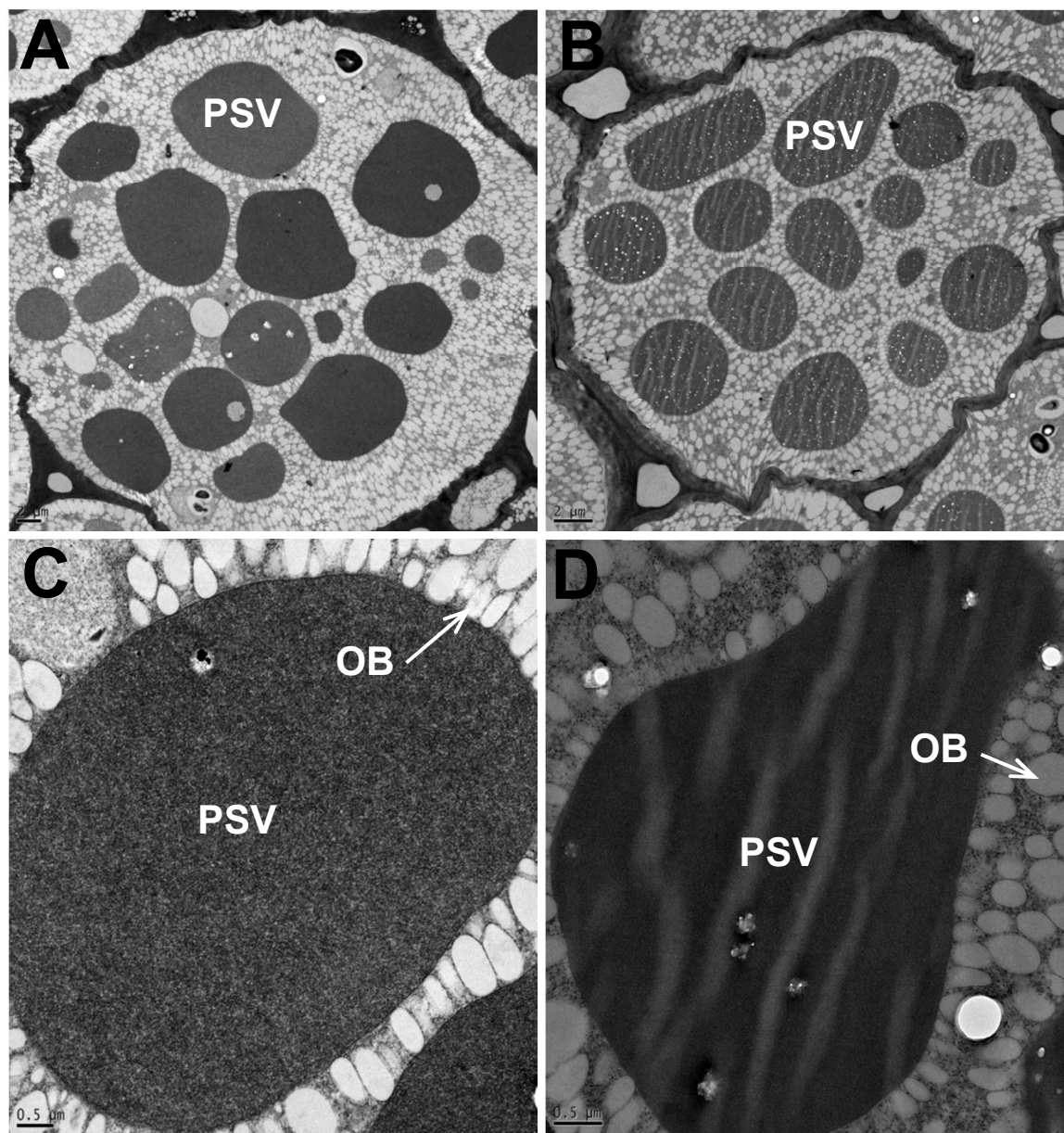


Figure 5



Graphic for the table of contents.

