2009

Agronomic, seed composition and molecular characteristics of soybean lines with novel genes for modified glycinin and beta-conglycinin content

Jonathan Jenkinson
Iowa State University

Agronomic, seed composition, and molecular characteristics of soybean lines with novel genes for modified glycinin and beta-conglycinin content

by

Jonathan Earl Jenkinson

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Plant Breeding

Program of Study Committee:
Walter Fehr, Major Professor
Jack Dekkers
Michael Lee
Thomas Lübberstedt
Duška Stojšin
Lester Wilson

Iowa State University

Ames, Iowa

2009

TABLE OF CONTENTS

ABSTRACT

Soybean [*Glycine max* (L.) Merr.] genotypes with increased beta-conglycinin (BC) and reduced glycinin (Gly) may be beneficial to food product quality and to human health. A soybean line, B2G2, developed in Japan, has the recessive alleles *gy1*, *gy2*, *gy3*, *gy4*, and *gy5* that control Gly content. Its seed protein has about 50 g kg^{-1} BC and 0 g kg^{-1} Gly. Monsanto Co. has used B2G2 to develop lines with different combinations of recessive and wild-type alleles at the five Gy loci. The first objective of this research was to determine the influence of different combinations of the *gy* and *Gy* alleles on the content of Gly and BC and their impact on agronomic traits. The protein content, oil content, protein composition, yield, and maturity of 19 lines in each of eight genotypic classes representing different *gy* allele combinations were compared in five populations. The BC content was greatest for lines with *gy1*,*2*, *gy4*, and *gy5* and for lines with the five *gy* alleles. Mean grain yield was significantly different among the eight genotypic classes; however, the differences were not consistent among populations. It should be possible to develop lines with *gy* alleles that have increased BC, reduced Gly, and similar protein content, oil content, yield, and maturity compared to lines with *Gy* alleles.

The second objective of this research was to determine the effect of locations and planting dates on protein composition in lines with different combinations of *gy* and *Gy* alleles. Planting dates did not significantly affect protein composition, but the means of locations were significantly different for all of the protein components, except for the A124 subunit of Gly. Grain produced at different locations would be expected to differ in BC

content, but planting dates would not be expected to significantly affect that protein component.

The third objective of this research was to determine if DNA markers located on the same linkage group as the genes controlling the α subunit of BC were associated with a reduction in the α subunit and an increase in the α′ subunit content. The protein composition of seven genotypes with cytosine (C) nucleotides at two closely linked marker loci (CC-CC) and seven genotypes with thymine (T) and adenine (A) nucleotides at the same marker loci (TT-AA) were compared. The α subunit content of the CC-CC genotypes was significantly lower by 3.1 g kg^{-1}, and all CC-CC genotypes had an α : α′ ratio of < 1.0. In two populations developed by crossing a CC-CC parent with a TT-AA parent, the marker genotype was associated with the α : α′ ratio. There were 7 % of the TT-AA plants and 33 % of the CC-CC plants that had an α : α′ ratio of < 1.0. These two markers should be useful in selecting for the desired α : α′ ratio, but it may be possible to develop markers that are more predictive.

CHAPTER 1

INTRODUCTION

Soybeans are an important source of high quality and inexpensive plant protein. Approximately 15 % of world soybean production is processed to provide food ingredients such as protein isolates and concentrates or used directly in food products (Zarkadas et al., 2007). The utilization of soybean protein in food products may be increased by improvements in the functional or nutritional properties of soybean protein.

The two major storage proteins of soybean, glycinin (Gly) and beta-conglycinin (BC), differ in functional and nutritional properties. Higher levels of BC are associated with better emulsification properties, improved low-temperature gelling (Rickert et al., 2004) and potential human health benefits (Kohno et al., 2006). Genes controlling the production of some of the subunits of Gly and BC have been identified (Beilinson et al., 2002; Yoshino et al., 2002). Soybean breeding lines with altered Gly and BC content and different complements of Gly and BC subunits have been developed (Ishikawa et al., 2006; Poysa et al., 2006; Yagasaki et al., 1997). One of these breeding lines, B2G2, has recessive alleles for all of the glycinin genes (*gy1* to *gy5*) and contains 0 g kg^{-1} of the seed protein as Gly and approximately 50 g kg^{-1} as BC (Yagasaki et al., 1996). This line has been used in breeding programs in North America. These programs are using molecular markers developed for the *gy* alleles to incorporate the alleles and the resulting protein composition change into elite North American cultivars (Poysa et al., 2006; Wu et al., 2007; Yu et al., 2005). Cultivars derived from these efforts should offer unique protein compositions to the soybean protein industry. The affect of the *gy* alleles on agronomic traits such as grain yield and on seed

traits such as protein and oil content is not known. One objective of this study was to measure the effect of different combinations of *gy* alleles on protein composition and agronomic traits.

Successful commercial production of cultivars with altered protein components will require consistent achievement of the desired protein composition over a range of environmental conditions. The second objective of this research was to assess the stability of protein composition in different growing environments.

Increasing the quantity of BC will have some functional benefits in various food products. A further enhancement would be to improve the quality of BC. Manzoni et al. (2003) identified the α′ subunit of BC as being largely responsible for some of the health benefits associated with soybean protein. BC has three subunits: α, α′, and β. The CG-1 gene (Genbank Accession No. AB030838) controls the production of the α′ subunit, while the CG-2 and CG-3 genes (Genbank Accession No. AB051865) control the production of the α subunit (Ishikawa et al., 2006). The specific genes that control the β subunit production have not been identified (Ishikawa et al., 2006). Most cultivars contain more α than α′, however, North American cultivars with α′ content greater than the α content have recently been identified (Bringe et al., 2007). The third objective of this study was to determine if existing molecular markers were associated with the altered α : α′ ratio.

LITERATURE REVIEW

Structure and importance of soybean protein

The two major storage proteins of soybean are glycinin (Gly) and beta-conglycinin (BC). Gly accounts for approximately 35-40 % and BC 25-30 % of total seed proteins. Glycinin is composed of five subunits: A1aB2, A2B1a, A1bB1b, A5A4B3, and A3B4, each of which has an acidic and basic polypeptide linked by a disulfide bond. The BC fraction is composed of three subunits: α, α′, and β. There is a strong negative correlation between Gly and BC content, ranging from -0.83 to -0.92 (Ogawa et al., 1989; Mizuno et al., 1994; Fehr et al., 2003).

Glycinin and BC have different functional and nutritional properties. Differences in Gly and BC composition have been shown to affect food properties such as gelation, solubility and emulsification (Cai and Chang, 1999; Guo et al., 2002; Kuipers et al., 2006; Rickert et al., 2004). Higher levels of BC and reduced levels of Gly result in improved low-temperature gelling, increased protein solubility, and better emulsification capacity and emulsion stability (Rickert et al., 2004). These properties would be beneficial to food products such as beverages with pH between 6 and 7 and meat products that are typically cooked to temperatures close to the optimum gelling temperature of BC. It also has been shown that BC has potential benefits to human health. In human and animal studies, diets enriched in BC have reduced blood serum levels of cholesterol, triglycerides and insulin (Aoyama et al., 2001; Moriyama et al., 2004; Kohno et al., 2006) and inhibited arterial plaque formation (Adams et al., 2004). Additional studies have shown that the α′ subunit of BC may be largely responsible for the health benefits associated with BC. Lovati et al.

(1998) demonstrated that BC was much more effective than Gly in upregulating low-density lipoprotein (LDL) uptake and metabolism in human liver cells. It was further shown that protein made from soybeans lacking the α′ subunit had no effect on LDL receptor activity (Manzoni et al., 1998), and that purified α′ subunit was more effective than whole BC in the uptake and degradation of LDL in both human liver cells and in rats (Manzoni et al., 2003; Duranti et al., 2004).

Production of soybean cultivars with increased BC levels would offer functional benefits in food products containing soybean protein and potentially be beneficial to human health. Increasing the level of the α′ subunit may enhance the potential health benefit.

Inheritance of reduced glycinin content

Five major genes (*Gy1*, *Gy2*, *Gy3*, *Gy4*, and *Gy5*) control the production of the A1aB2, A2B1a, A1bB1b, A5A4B3, and A3B4 glycinin subunits (Nielsen et al., 1989). All of these genes have been mapped to specific linkage groups in the soybean genome (Cho et al., 1989; Nielsen et al., 1989; Diers et al., 1994; Beilinson et al., 2002). A study using a mapping population derived from two cultivars, Forrest and Raiden, with variant alleles for the five genes indicated that *Gy1* and *Gy2* were tightly linked and inherited together while the remaining three genes were inherited independently (Cho et al., 1989). Yagasaki et al. (1996) studied the protein SDS-PAGE patterns and concluded that *Gy1*, *Gy2* and *Gy3* co-segregated. However, the SDS-PAGE method did not separate the proteins produced by the three genes and it was impossible to discriminate segregation of *Gy1*, *Gy2* and *Gy3*. Using DNA probes, Beilinson et al.(2002) confirmed that *Gy1* and *Gy2* were located on the same linkage group with only 3 kb between them, while *Gy3*, *Gy4*, and *Gy5* were located on different linkage groups from each other and from *Gy1* and *Gy2*.

A soybean line, B2G2, lacking all the glycinin subunits was produced by crossing a gamma-irradiated *Glycine max* line with the *gy1* to *gy4* alleles to a *Glycine soja* line with the gy5 allele (Yagasaki et al., 1996). In addition, eight lines were developed from the cross that contained various combinations of mutant *gy* alleles. A line carrying all wild-type *Gy* alleles (*Gy1 – Gy5*) contained 41.2 % of total protein as Gly and 37.6 % as BC. A line carrying all mutant Gly alleles (*gy1 - gy5*) contained 2.8 % of total protein as Gly and 71.1 % as BC.

Agronomic traits of reduced glycinin lines

The soybean line B2G2 with the five *gy* alleles has been used in crosses to develop breeding lines with various combinations of g*y1*,*2*, *gy3*, *gy4*, and g*y5* alleles (Poysa et al., 2006). There are several publications on the affect of reduced Gly content on food quality (Nagano et al., 1996; Cai and Chang, 1999; Bian et al., 2003). There are no publications concerning how the *gy* alleles and a reduction in Gly content influence grain yield in soybeans.

Stability of seed protein composition in lines with reduced glycinin

There are few studies that have investigated the stability of Gly and BC content in multiple genotypes over different locations and planting dates. Murphy and Resurreccion (1984) measured total Gly and BC content in ten genotypes in a single location each of two years. They concluded that both years and locations affect protein composition to a greater extent than do genetic differences. Conclusions cannot be made from their data on the effect of location within a year or the relative importance of year and location because the genotypes were not all grown at the same location each year. Fehr et al. (2003) measured total Gly and BC content as well as the subunits of each in 14 cultivars over three years and eight locations each year. Significant differences were found among cultivars for total Gly

and total BC and all of the subunits with the exception of the A3B4 subunit of Gly. In their study, years and locations were not significant factors for most of the protein subunits and, most importantly, the cultivar x location interaction was not significant for any protein composition trait. The cultivar x year interaction was only significant for the A3B4 subunit of Gly. It is important to note that all of the cultivars in their study were within the normal range for Gly and BC content. Poysa et al. (2006) created a set of lines containing different combinations of *gy* alleles. These lines were grown at one location in 2002 and two locations in 2003. Various seed traits and food processing traits were evaluated with respect to genotype, location and year effects but data on protein composition at each location and year were not reported. For food processing traits such as tofu hardness and firmness, the differences among genotypes had a greater effect than environment differences, and year effects were more important than location effects. Krishnan et al. (2005) reported that nitrogen and sulfur fertility affect the accumulation of Gly and BC in field and greenhouse conditions. They indicated that differences in growing conditions or agronomic practices should be expected to affect protein composition.

Genes controlling beta-conglycinin composition

A multi-gene family controls the expression of the three β-conglycinin subunits. The *CG-1* (GenBank accession AB030838) gene controls expression of the α′ subunit and is not linked to other BC genes, while *CG-2* and *CG-3* (GenBank accession AB051865) are inverse repeats of each other separated by about 3 kb and together control the expression of the α subunit (Ishikawa et al., 2006). The specific genes that control β subunit production have not been identified. In normal soybean cultivars in both Asia and North America, the content of the α subunit is about 120 g kg^{-1}, the α′ subunit about 90 g kg^{-1}, and the β subunit

about 60 g kg^{-1} of total seed protein. While these values may vary with cultivars and environments, the ratio of the three subunits remains very consistent at $\alpha : \alpha' : \beta = 2.2 : 1.5 : 1.0$ (Maruyama et al., 1999). North American cultivars with a $\alpha : \alpha'$ ratio of < 1.0 have recently been identified (Bringe et al., 2007). By combining this trait with the increased BC trait, it should be possible to develop lines with greatly increased α' content. It would be desirable to be able to select for this trait using molecular markers.

REFERENCES

Adams, M.R., D.L. Golden, A.A. Franke, S.M. Potter, H.S. Smith, and M.S. Anthony. 2004. Dietary soy β-conglycinin (7S globulin) inhibits atherosclerosis in mice. J Nutr. 134:511-516.

Aoyama, T., M. Kohono, T. Saito, K. Fukui, K. Takamatsu, T. Yamamoto, Y. Hashimoto, M. Hirotsuka, and M. Kito. 2001. Reduction by phytate-reduced soybean β-conglycinin of plasma triglyceride level of young and adult rats. Biosci. Biotechnol. Biochem. 65: 1071-1075.

Beilinson, V., Z. Chen, R.C. Shoemaker, R.L. Fischer, R.B. Goldberg, and N.C. Nielsen. 2002. Genomic organization of glycinin genes in soybean. Theor. Appl. Gen. 104:1132-1140.

Bian, Y., D.J. Myers, K. Dias, M.A. Lihono, S. Wu, P.A. Murphy. 2003. Functional properties of soy protein fractions produced using a pilot-scale process. J. Am. Oil Chem. Soc. 80:545-549.

Bringe, N., J. Jenkinson, and K. Wu. 2007. Increased alpha-prime beta-conglycinin soybeans. U.S. Patent application 60/971336. Date filed: 11 Sept 2007.

Cai, T., and K.C. Chang. 1999. Processing effect on soybean storage proteins and their relationship with tofu quality. J. Agric. Food Chem. 47:720 – 727.

Cho, T.J., C.S. Davies, and N.C. Nielsen. 1989. Inheritance and organization of glycinin genes in soybean. Plant Cell 1:329-337.

Diers, B.W., V. Beilinson, N.C. Nielsen, and R.C. Shoemaker. 1994. Genetic mapping of the *Gy4* and *Gy5* glycinin genes in soybean and analysis of a variant of *Gy4*. Theor. Appl. Genet. 89:297-304.

Duranti, M., M.R. Lovati, V. Dani, A. Barbiroli, A. Scarafoni, S. Castiglioni, C. Ponzone, and P. Morazzoni. 2004. The α′ subunit from soybean 7S globulin lowers plasma lipids and upregulates liver β-VLDL receptors in rats fed a hypercholesterolemic diet. J. Nutr. 134:1334-1339.

Fehr, W.R., J.A. Hoeck, S.L. Johnson, P.A. Murphy, J.D. Nott, G.I. Padilla, and G.A. Welke. 2003. Genotype and environment influence on protein components of soybean. Crop Sci. 43:511-514.

Guo, S.T., C. Tsukamoto, K. Takahashi, K. Yagasaki, Q.X. Nan, and T. Ono. 2002. Incorporation of soymilk lipid into soy protein coagulum by the addition of calcium chloride. J. Food Sci. 67: 3215-3219.

Ishikawa, G., Y. Takada, and T. Nakamura. 2006. A PCR-based method to test for the presence or absence of β–conglycinin α′- and α-subunits in soybean. Mol.

Breeding. 17:365-374.

Kohno, M., M. Hirotsuka, M. Kito, and Y. Matsuzawa. 2006. Decreases in the serum triacylglycerol and visceral fat mediated by dietary soybean β-conglycinin. J. Atheroscler. Thromb. 13:247-255.

Krishnan, H.B., J.O. Bennett, W.-S. Kim, A.H. Krishnan, and T.P. Mawhinney. 2005. Nitrogen lowers the sulfur amino acid content of soybean (*Glycine max* [L.] Merr.) by regulating the accumulation of Bowman-Birk protease inhibitor. J. Agric. Food Chem. 53:6347 – 6354.

Kuipers, B.J.H., G.A. van Koningsveld, A.C. Alting, F. Driehuis, A.G.J. Voragen, and H. Gruppen. 2006. Opposite contributions of glycinin- and β-conglycinin- derived peptides to the aggregation behavior of soy protein isolate hydrolysates. Food Biophys. 1:178-188.

Lovati, M.R., C. Manzoni, E. Gianazza, and C.R. Sirtori. 1998. Soybean protein products as regulators of liver low-density lipoprotein receptors. I. Identification of active β-conglycinin subunits. J. Agric. Food Chem. 46:2474-2480.

Manzoni, C., M.R. Lovati, E. Gianazza, Y. Morita, and C.R. Sirtori. 1998. Soybean protein products as regulators of liver low-density lipoprotein receptors. II. α-α′ rich commercial soy concentrate and α′ deficient mutant differently affect low-density lipoprotein receptor activation. . J. Agric. Food Chem. 46:2481-2484.

Manzoni, C., M. Duranti, I. Eberini, H. Scharnag, W. Marz, S. Castiglioni, and M. Lovati. 2003. Sucbellular localization of soybean 7S globulin in HepG2 cells and LDL receptor up-regulation by its constituent subunit. J. Nutr. 133:2149-2155.

Maruyama, N., R. Sato, Y. Wada, Y. Matsumura, H. Goto, E. Okuda, S. Nakagawa, and

S. Utsumi. 1999. Structure-physicochemical function relationships of soybean β-conglycinin constituent subunits. J. Agric. Food Chem., 47:5278 – 5284.

Mizuno, Y., K. Takahashi, H. Namai, and K. Kitamura. 1994. Changes of seed components with the absence of α- and α′- subunits of 7S globulin in soybean. Japan. J. Breed. 44: 201-204.

Moriyama, T., K. Kishimoto, K. Nagai, R. Urade, T. Ogawa, S. Utsumi, N. Maruyama, and M. Maebuchi. 2004. Soybean β-conglycinin diet suppresses serum triglyceride levels and normal and genetically obese mice by induction of β-oxidation, down-regulation of fatty acid synthase, and inhibition of triglyceride absorption. Biosci. Biotechnol. Biochem. 68: 352-359.

Murphy, P.A., and A.P. Resurreccion. 1984. Varietal and environmental differences in soybean glycinin and β-conglycinin content. J. Agric. Food Chem. 32:911-915.

Nagano, T., Y. Fukuda T. Akasaka. 1996. Dynamic viscoelastic study on the gelation of 7S globulins from soybeans. J. Agric. Food Chem. 44: 3484-3488.

Nielsen, N.C., C.D. Dickinson, T.J. Cho, V.H. Thanh, B.J. Scallon, R.L. Fischer, T.L. Sims, and R.B. Goldberg. 1989. Characterization of the glycinin gene family. Plant Cell. 1:313-328.

Ogawa, T., E. Tayama, K. Kitamura, and N. Kaizuma. 1989. Genetic improvement of seed storage proteins using three variant alleles of 7S globulin in soybean (*Glycine max* L.). Japan. J. Breed. 39:137-147.

Poysa, V., L. Woodrow, and K. Yu. 2006. Effect of soy protein composition on tofu quality. Food Res. Int. 39:309-317.

Rickert, D.A., L.A. Johnson, and P.A. Murphy. 2004. Functional properties of improved

glycinin and β-conglycinin fractions. J. Food Sci. 69:303-311.

Wu, K., T. Horejsi, J. Byrum, N. Bringe, J. Yang, D. Pei, and R. Reiter. Agronomically elite soybeans with high beta-conglycinin content. U.S. Patent application 11/517186. Date issued: 11 Sept 2007.

Yagasaki, K., N. Kaizuma, and K. Kitamura. 1996. Inheritance of glycinin subunits and characterization of glycinin molecules lacking the subunits in soybean (*Glycine max* (L.) Merr.) Breeding Science. 46:11-15.

Yagasaki, K., T. Takagi, M. Sakai, and K. Kitamura.1997. Biochemical characterization of soybean protein consisting of different subunits of glycinin. J. Agric. Food Chem. 45:656 – 660.

Yoshino, M., A. Kanazawa, K. Tsutsumi, I. Nakamura, K. Takahashi, and Y. Shimamoto. 2002. Structural variation around the gene encoding the α subunit of β-conglycinin and correlation with the expression of the α subunit. Breeding Science. 52:285-292.

Yu, K., V. Poysa, M. Haffner, B. Zhang, and L. Woodrow. 2005. Absence of the A4 peptide in the G4 glycinin subunit of soybean cultivar Enrei is caused by a point mutation in the *Gy4* gene. Gen. and Mol. Biol. 28:440-443.

Zarkadas, C.G., C. Gagnon, V. Poysa, S. Khanizadeh, E.R. Cober, V. Chang, and S. Gleddie. 2007. Protein quality and identification of the storage protein subunits of tofu and null soybean genotypes, using amino acid analysis, one- and two-dimensional gel electrophoresis, and tandem mass spectrometry. Food Res. Int. 40: 111-128.

CHAPTER 2

AGRONOMIC AND SEED CHARACTERISTICS OF SOYBEAN LINES WITH ALLELES FOR MODIFIED GLYCININ CONTENT

From a paper submitted to *Crop Science*

ABSTRACT

Soybean [*Glycine max* (L.) Merr.] genotypes have been developed with increased beta-conglycinin (BC) and reduced glycinin (Gly) to improve the quality and health benefits of food products containing soybean protein. The changes in protein composition are due to the five recessive alleles, *gy1*,*gy2*, *gy3*, *gy4*, and *gy5*, that modify Gly content. The objective of this study was to evaluate the influence of the *gy* alleles on protein composition and agronomic traits. Nineteen lines in each of eight genotypic classes involving different combinations of *gy* alleles from five populations were evaluated in 2008. The BC content of lines was maximized with the *gy1*,*2*, *gy4*, and *gy5* alleles, but the *gy3* allele also was required to obtain 0 g kg^{-1} of Gly. The effect of the recessive alleles on the increase in BC content was greatest for *gy1*,*2*, intermediate for *gy5*, and least for *gy3*. There were significant differences in mean grain yield among the eight genotypic classes; however, the differences were not consistent among the populations. It should be possible to develop soybean lines with increased BC and decreased Gly content that have yield, maturity, and protein and oil contents similar to cultivars with normal protein composition.

INTRODUCTION

The two major storage proteins of soybean are glycinin (Gly) and β-conglycinin (BC). Gly accounts for ~ 40 % and BC ~ 25 % of total seed proteins (Nielsen et al., 1989). Glycinin is composed of five subunits: A1aB2, A2B1a, A1bB1b, A5A4B3, and A3B4, each of which has an acidic (A) and basic (B) polypeptide linked by a disulfide bond. The production of these subunits is controlled by five major genes *Gy1* to *Gy5* (Nielsen et al., 1989). *Gy1* and *Gy2* will be referred to hereinafter as *Gy1,2* because they are separated by only about 3 kb of DNA and are inherited together, while the remaining three genes are inherited independently (Cho et al., 1989; Beilinson et al., 2002). The BC fraction is composed of three subunits: α, α′, and β. There is a strong negative correlation between Gly and BC content, but neither Gly or BC has a significant correlation with total seed protein (Ogawa et al., 1989; Mizuno et al., 1994; Fehr et al., 2003).

A soybean line, B2G2, containing recessive alleles at the five *Gy* loci and lacking all the Gly protein subunits was developed in Japan (Yagasaki et al., 1996). They combined gamma-radiation-induced mutations in *Gy1,2* and *Gy3* with a recessive allele in *Gy4* commonly found in Japanese cultivars and a recessive allele in *Gy5* obtained from a *Glycine soja* line. They observed that the seed protein of B2G2 had about 50 g kg^{-1} BC and 0 g kg^{-1} Gly.

Differences in Gly and BC content have been shown to affect food properties such as gelation, solubility, and emulsification (Cai and Chang, 1999; Guo et al., 2002 ; Kuipers et al., 2006; Rickert et al., 2004). Higher levels of BC and reduced levels of Gly result in improved low-temperature gelling, increased protein solubility, better emulsification capacity, and emulsion stability (Rickert et al., 2004). It also has been reported that BC has

potential benefits to human health. In human and animal studies, diets enriched in BC have reduced serum levels of cholesterol, triglycerides and insulin (Aoyama et al., 2001; Moriyama et al., 2004; Kohno et al., 2006) and inhibited arterial plaque formation (Adams et al., 2004). Lovati et al. (1998) demonstrated that BC was much more effective than Gly at up-regulating low-density lipoprotein uptake and metabolism in human liver cells. Hydrolysates of soybeans reduced in Gly and increased in BC also have been found to be effective in controlling fat accumulation in adipocytes (Martinez-Villaluenga et al., 2008) and inhibiting leukemia cell growth in vitro (Wang et al., 2008). In contrast, Gly hydrolysates increased fat accumulation in adipocytes (Martinez-Villaluenga et al., 2008).

No studies have reported on the impact of the individual *gy1*,*2*, *gy3, gy4*, and *gy5* alleles on agronomic traits. Successful commercial production of soybean protein with elevated BC and reduced Gly would require the development of cultivars whose seed yield, maturity, protein content, and oil content was similar to cultivars with conventional Gly and BC contents. The objectives of this study were to determine the influence of the *gy* alleles on the content of Gly and BC and their impact on agronomic traits.

MATERIALS AND METHODS

Five segregating populations were developed by crossing five conventional cultivars with only wild-type *Gy* alleles to a line, MV0118, with the five recessive *gy* alleles. All of the parents were developed by the Monsanto Co.. The common cultivars had relative maturities that ranged from 2.4 to 3.6 and the line MV0118 had a relative maturity of 2.5. The pedigree of MV0118 was MV32//B2G2/MV19. MV32 and MV19 were two cultivars developed by Monsanto Co. that had only wild-type *Gy* alleles. The soybean line B2G2 with the five *gy* alleles was developed in Japan (Yagasaki et al., 1996).

The parents were crossed at the Monsanto Caribe Research Farm near Isabela, PR, in March 2006. The F1 seeds from the five crosses were planted in rows 3.3 m long with 9 seeds m^{-1} during November 2006 at the Monsanto Research Farm in Kihei, HI. Leaf tissue was collected from each F1 plant and DNA from the tissue was analyzed using a direct molecular marker based on a single nucleotide polymorphism (SNP) in the *Gy4* gene (Wu et al., 2007). Plants that were heterozygous for the marker were considered to be hybrids. The F2 seeds from hybrid plants within each population were bulked together. All of the F2 seeds from each population were planted in rows 3.3 m long rows with 18 seeds m^{-1} during February 2007 in Kihei, HI. Leaf tissue was collected from each F2 plant in the five populations. The DNA from each sample was analyzed using SNP markers for *Gy1, 2, 3, 4* and *5* (Wu et al., 2007). The primer sequence used to select *gy1,2* was based on the deletion of the Gy1,2 wild-type sequence and the markers for *gy3* and *gy5* were SNPs tightly linked to the gene of interest. Based on the marker results, F2 plants were assigned to one of eight genotypic classes and 30 F3 seeds from each individual were planted in a progeny row 3.3 m long with 10 seeds m^{-1} in Kihei, HI during July 2007. One F3 plant was harvested from each F2:3 line and threshed individually. Five F4 seeds from each F3 plant were used for sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis to measure protein composition and the remaining F4 seeds from each plant were planted in a row 4 m long with a range from 11 to 15 seeds m^{-1} at Massai Agricultural Services in Rancagua, Chile during November 2007. The 19 F3:4 lines in each genotypic class from each of the populations that were uniform for maturity were harvested individually in bulk.

For the field tests in 2008, the 19 F3:5 lines of the eight genotypic classes for each population were grown as five separate experiments. The experiments were grown as

randomized complete-block design with two replications planted at Bloomington, IL, on 22 May; Oxford, IN, on 22 May; Findlay, OH, on 24 May; and Cambridge, IA, on 15 June. The soil types are a Tama silt loam (fine-silty, mixed, superactive, mesic Typic Argiudoll) at Bloomington, a Darroch silty loam (fine-loamy, mixed, superactive, mesic Aquic Argiudoll) at Oxford, a Hoytville clay (fine, illitic, mesic Mollic Epiaqualf) at Findlay and a Coland clay loam (fine-loamy, superactive, mesic Cumulic Endoaquoll) at Cambridge. The plots at each location were two rows 4.5 m long with 76 cm between rows. The seeding rate was 24 seeds m^{-1}. Each plot was evaluated for maturity, grain yield and grain moisture content. Maturity was recorded as number of days after 31 August when 95 % of the pods on the main stem had reached their mature color. The seed from each plot was harvested in bulk with a self-propelled plot combine (ALMACO, Nevada, IA) and the weight and moisture content were recorded. Grain yield for each plot was adjusted to 130 g kg^{-1} of moisture. A 250-g seed sample from each plot at Bloomington, Oxford, and Findlay was analyzed by near infra-red transmittance with an Infratec 1220 grain analyzer (Foss, Eden Prairie, MN) to determine protein and oil content.

For SDS-PAGE analysis of each line in each genotypic class, five F4 seeds from each F3 plant harvested were ground with a Thomas-Wiley model 4 mill (Thomas-Wiley, Swedesboro, NJ) using a 2 mm screen. The ground sample was reground in the same mill using a 1 mm screen. The SDS-PAGE analysis was performed using the method described by Martinez-Villaluenga et al. (2008). This method dissociates each of the acidic polypeptides (A1a, A1b, A2, A3, A4, and A5) from the basic polypeptides (B1a, B1b, B2, B3, and B4), and produces separate bands for each of the three subunits of BC and the A3 and basic

subunits of Gly, while the A1a, A1b, A2 and A4 subunits of Gly migrate together and form one band that was labeled A124 in this study.

The protein composition data were analyzed with PROC GLM of SAS version 9.1 (SAS Institute, 2003). Genotypic class was considered a fixed effect and the individual lines within each class were considered replications of each class. Linear regression analysis of yield on maturity was performed using PROC REG of SAS version 9.1 (SAS Institute, 2003). Each experiment at each location, except for pop5, had a significant linear regression coefficient; therefore, the yield of each plot in pop1 to pop4 was adjusted by the linear regression coefficient. The adjusted yield, maturity, protein, and oil data were analyzed as a randomized complete-block design using the mixed model procedure of SAS version 9.1(SAS Institute, 2003). Locations and genotypes were considered fixed effects and replications were considered random effects. Differences between means were determined using Tukey's honestly significant difference test (Tukey, 1949).

RESULTS AND DISCUSSION

There were significant differences among the eight genotypic classes in each population for all the traits that were evaluated. The range among classes for seed protein content was 11 g kg^{-1} and for oil content was 12 g kg^{-1} (Table 1). None of the classes with one or more *gy* alleles were consistently different from class 1 with only wild-type alleles across populations for the two traits. This indicated that in a cultivar development program, modification of Gly and BC content with the *gy* alleles would not be expected to have an unfavorable impact on protein or oil content.

Based on a comparison of classes 4 and 8, it was necessary to have the *gy1,2* alleles present to achieve the greatest reduction in Gly and two of its subunits and the greatest

increase in total BC and its subunits (Tables 2 and 3). Averaged across populations, Class 4 with the wild-type allele *Gy1,2* had 221 g kg^{-1} total Gly, 114 g kg^{-1} A124, and 107 g kg^{-1} of the basic subunit while class 8 with the *gy1,2* allele had 0 g kg^{-1} for the three traits. The 0 g kg^{-1} of A3 in class 4 indicated that the *gy1,2* allele was not essential for minimizing that subunit. The decrease in Gly associated with *gy1,2* was accompanied by a significant increase in total BC and its subunits based on comparisons of classes that differed at the *Gy1,2* locus. For example, the total BC content of class 5 ranged from 25 to 39 % higher than class 1, and the BC content of class 8 ranged from 35 to 51 % higher than class 4 (Table 3). The relative proportions of the three subunits of BC were not changed due to any of the *gy* alleles.

The *gy3* allele was necessary to minimize total Gly and the A124 and basic subunits based on a comparison of classes 6 and 8 (Table 2). There was some content of the three traits present in class 6 with only the wild-type allele *Gy3*, but none of the three traits occurred in class 8 with the *gy3* allele. The *gy3* allele was not necessary to minimize the content of the A3 subunit based on a comparison of the two classes. Although the *gy3* allele had a role in minimizing Gly and two of its subunits, it was not necessary for maximizing total BC and its subunits based on a comparison of class 6 with the *Gy3* allele and class 8 with the *gy3* allele (Table 3). The difference between the classes 6 and 8 for total BC and α' was only significant for pop1 in which class 8 was greater than class 6. No significant differences between the two classes were observed in any of the populations for the α and β subunits. Jenkinson and Fehr (2010a) obtained the same result in a comparison of one of more lines with the genotypes of classes 6 and 8 across four locations and three planting dates. The two studies indicated that it would be necessary to incorporate the *gy3* allele in a

cultivar to minimize total Gly and two of its subunits, but not to maximize total BC and its subunits.

The impact of the *gy4* allele on Gly and BC was not evaluated to the same extent as the other *gy* alleles because it was necessary to limit the number of genotypic classes in the study (Table 4). The recessive allele was present in all of the classes, except 1 and 5 (Table 4). The *gy4* allele has been routinely used by the Monsanto Co. in its cultivar development program to achieve the elevated BC and reduced Gly associated with genotypic classes 6, 7, and 8.

The *gy5* allele had a significant role in reducing Gly and its A3 and basic subunits based on comparisons of classes 7 and 8 in the five populations (Table 2). There was some content of the three traits in class 7 with the *Gy5* allele, but none present in class 8 with the *gy5* allele. The *gy5* allele was necessary for maximizing total BC and its α and α′ subunits in the populations, although in pop1 and pop2 the total BC content of class 7 was not significantly different from that one of class 6 or class 8 (Table 3). The value of the BC increase associated with *gy5* has to be balanced with the breeding costs of selecting for the allele.

There were significant differences in the A3 subunit content among the four classes with the *Gy5* allele. In all populations, class 7 had the greatest A3 content while Class 1 had the lowest A3 content. Classes 3 and 5 had A3 contents intermediate to classes 1 and 7. Martinez-Villaluenga et al.(2008) also found that soybean lines with reduced total Gly and A124 content had higher A3 content than lines with normal levels Gly and A124.

In summary, the *gy1,2* allele had the greatest affect on the BC increase and Gly decrease, the *gy5* allele was intermediate, and the *gy3* allele had the smallest affect. There

was a significant negative correlation of -0.89 ($P < 0.05$) between total Gly and total BC. This was similar to the correlation between the two traits of -0.84 ($P < 0.05$) reported by Jenkinson and Fehr (2010a) and -0.92 ($P < 0.05$) reported by Fehr et al. (2003).

For all of the populations, there were significant differences in mean grain yield among classes and among lines within classes (Table 6). The yields of lines in classes 6, 7, and 8 were of particular interest because they had the highest BC, which would be of interest to the food industry. Averaged across populations, the mean yields of the three classes were not consistently different from class 1 that had the wild-type *Gy* alleles of conventional cultivars. In pop1 and pop5, the yields of classes 6, 7, and 8 were not significantly different from each other or class 1. The mean yield of class 6 was only significantly less than class 1 in pop2, while class 7 was significantly less in pop2 and pop4 and class 8 was significantly less in pop2 and pop3. Averaged across populations, the mean yields were 2861 g kg^{-1} for class 1, 2845 g kg^{-1} for class 6, 2813 g $kg^{-1}$for class 7, and 2740 g kg^{-1} for class 8. Expressed as a percentage of class 1, the mean yield of class 6 was 0.6 % less, class 7 was 1.7 % less, and class 8 was 4.2 % less.

Another assessment of the differences in seed yield among classes 1, 6, 7, and 8 was made by determining the number of lines that were significantly better than the mean in each population based on the least significant differences at the 0.05 probability level (Table 4). Averaged across populations, class 1 had 0.4 more selected lines than class 6, 2.0 more than class 7, and 4.2 more than class 8.

There were significant differences in mean maturity among classes and among lines within classes for all the populations (Table 4). The differences among classes ranged from 2.0 d in pop5 to 4.2 d in pop2. In all the populations, the ranges for maturity overlapped for

all classes. It should be possible to develop cultivars with different combinations of *gy* alleles that have similar maturity to cultivars with wild-type Gy alleles.

This study indicated that soybean cultivars with increased levels of BC can be developed using *gy* alleles that control Gly production. Each *gy* allele affected total BC and its subunits differently, which makes it possible to develop cultivars with a range of BC and Gly for the food industry. It should be possible to develop cultivars with the genotype of class 6 that had the highest BC without sacrificing seed yield. It would be slightly more difficult to select high yielding cultivars with the genotype of class 7 and more difficult for cultivars with the five *gy* alleles of class 8. Classes 6 and 7 would be less difficult to manage in a breeding program because they require one less *gy* allele than class 8. The choice among the three classes would depend on the amount of total Gly and the subunits desired by the food industry. Class 8 would be necessary if the highest BC and lowest Gly was required. Classes 6 and 7 could be useful if some Gly and its subunits were acceptable.

ACKNOWLEDGEMENTS

The author thanks personnel at the Monsanto soybean breeding stations in Bloomington, IL; Findlay, OH; Huxley, IA; and Oxford, IN, for their assistance with the field testing; Nathan Durnell for conducting the SDS-PAGE analysis; personnel at the Monsanto facility in Ankeny, IA, for conducting the protein and oil analysis; and Jim Behm and Kyle Smith for assistance with the development of the tested lines.

Table 1. Mean seed protein and oil content of lines for eight genotypic classes from five soybean populations grown in three environments in 2008.

Trait	Glycinin alleles	Class	Population 1	2	3	4	5
			----------g kg^{-1} seed†----------				
Protein	*Gy*1,2 *Gy*3 *Gy*4 *Gy*5	1	416a§	415abc	412b	408abc	419ab
	*Gy*1,2 *Gy*3 *gy*4 *gy*5	2	410cd	418a	410b	402d	417ab
	*Gy*1,2 *gy*3 *gy*4 *Gy*5	3	414abc	418a	419a	410a	415b
	*Gy*1,2 *gy*3 *gy*4 *gy*5	4	413abcd	416ab	414b	409ab	418ab
	*gy*1,2 *Gy*3 *Gy*4 *Gy*5	5	416ab	409c	414b	406bcd	410c
	*gy*1,2 *Gy*3 *gy*4 *gy*5	6	409d	412bc	413b	405bcd	419ab
	*gy*1,2 *gy*3 *gy*4 *Gy*5	7	413abcd	415abc	413b	406abc	421a
	*gy*1,2 *gy*3 *gy*4 *gy*5	8	411bcd	413abc	411b	404cd	416ab
Oil	*Gy*1,2 *Gy*3 *Gy*4 *Gy*5	1	206ab	209a	206a	206abc	200a
	*Gy*1,2 *Gy*3 *gy*4 *gy*5	2	207a	201b	204ab	209a	197b
	*Gy*1,2 *gy*3 *gy*4 *Gy*5	3	205ab	201b	199d	204cde	200a
	*Gy*1,2 *gy*3 *gy*4 *gy*5	4	203bc	201b	200d	202de	196b
	*gy*1,2 *Gy*3 *Gy*4 *Gy*5	5	206ab	202b	203bc	207ab	202a
	*gy*1,2 *Gy*3 *gy*4 *gy*5	6	206a	202b	201cd	205bcd	194bc
	*gy*1,2 *gy*3 *gy*4 *Gy*5	7	201c	197c	199d	203de	192c
	*gy*1,2 *gy*3 *gy*4 *gy*5	8	202bc	198bc	200d	201e	194bc

† Protein and oil content on a moisture basis of 0 g kg^{-1}.

§ Means within a column and trait followed by the same letter were not significantly different at the 0.05 probability level based on Tukey's honestly significant difference (Tukey, 1949).

Table 2. Glycinin composition of seed from F3 plants for eight genotypic classes from five soybean populations grown in Hawaii in 2008.

Trait	Glycinin alleles	Class	Population 1		Population 2		Population 3		Population 4		Population 5	
			g kg^{-1}†	%‡	g kg^{-1}	%	g kg^{-1}	%	g kg^{-1}	%	g kg^{-1}	%
Gly§	*Gy1,2 Gy3 Gy4 Gy5*	1	295a¶		339a		317a		315a		331a	
	Gy1,2 Gy3 gy4 gy5	2	210b		245bc		251b		213c		237b	
	Gy1,2 gy3 gy4 Gy5	3	296a		267b		270b		259b		246b	
	Gy1,2 gy3 gy4 gy5	4	219b		242c		214c		216c		212c	
	gy1,2 Gy3 Gy4 Gy5	5	175c		166d		157d		161d		193c	
	gy1,2 Gy3 gy4 gy5	6	29e		32f		40f		41f		32e	
	gy1,2 gy3 gy4 Gy5	7	83d		82e		87e		78e		82d	
	gy1,2 gy3 gy4 gy5	8	0f		0g		0g		0g		0f	
A3	*Gy1,2 Gy3 Gy4 Gy5*	1	28c	9.7c	23d	6.8d	26c	8.4d	26c	8.5d	24d	7.3d
	Gy1,2 Gy3 gy4 gy5	2	0d	0.0d	0e	0.0e	0d	0.0e	0d	0.0e	0e	0.0e
	Gy1,2 gy3 gy4 Gy5	3	28c	9.6c	29c	11.4c	34b	13.4c	32b	13.2c	32c	13.3c
	Gy1,2 gy3 gy4 gy5	4	0d	0.0d	0e	0.0e	0d	0.0e	0d	0.0e	0e	0.0e
	gy1,2 Gy3 Gy4 Gy5	5	36b	21.4b	34b	20.8b	33b	22.0b	34b	21.3b	39b	20.3b
	gy1,2 Gy3 gy4 gy5	6	0d	0.0d	0e	0.0e	0d	0.0e	0d	0.0e	0e	0.0e
	gy1,2 gy3 gy4 Gy5	7	42a	50.6a	41a	49.8a	41a	49.1a	38a	51.5a	43a	52.9a
	gy1,2 gy3 gy4 gy5	8	0d	0.0d	0e	0.0e	0d	0.0e	0d	0.0e	0e	0.0e
A124	*Gy1,2 Gy3 Gy4 Gy5*	1	114ab	38.3c	147a	43.0c	136a	42.9b	130a	41.2b	141a	42.5b
	Gy1,2 Gy3 gy4 gy5	2	109b	51.8a	129b	52.9a	132a	52.8a	112b	52.2a	126b	53.5a
	Gy1,2 gy3 gy4 Gy5	3	123a	41.3b	110c	40.7c	104b	37.7c	102b	38.5b	94d	37.9c
	Gy1,2 gy3 gy4 gy5	4	115ab	52.6a	122b	50.3b	110b	51.5a	112b	51.9a	112c	52.6a
	gy1,2 Gy3 Gy4 Gy5	5	44c	24.2d	45d	27.0e	41c	26.2d	40c	25.4c	53e	27.2d
	gy1,2 Gy3 gy4 gy5	6	11d	38.7c	12e	37.1d	19d	44.9b	17d	39.6b	13f	38.7c
	gy1,2 gy3 gy4 Gy5	7	0e	0.0e	0e	0.0f	0e	0.0e	0e	0.0d	0g	0.0e

Table 2. Continued

Trait	Glycinin alleles	Class	Population 1		Population 2		Population 3		Population 4		Population 5	
			g kg^{-1}†	%‡	g kg^{-1}	%	g kg^{-1}	%	g kg^{-1}	%	g kg^{-1}	%
	gy1,2 gy3 gy4 gy5	8	0e	0.0e	0e	0.0f	0e	0.0e	0e	0.0d	0g	0.0e
Basic	*Gy1,2 Gy3 Gy4 Gy5*	1	153a	52.0b	170a	50.2bc	154a	48.7bc	158a	50.4bcd	166a	50.2bc
	Gy1,2 Gy3 gy4 gy5	2	101b	48.2c	116b	47.1c	119c	47.2c	101c	47.8d	110bc	46.5d
	Gy1,2 gy3 gy4 Gy5	3	145a	49.1c	128b	47.9c	132b	48.9bc	126b	48.4cd	120b	48.8cd
	Gy1,2 gy3 gy4 gy5	4	104b	47.4c	120b	49.7bc	104d	48.5bc	104c	48.1cd	101c	47.4d
	gy1,2 Gy3 Gy4 Gy5	5	96b	54.5b	88c	52.2b	83e	51.7ab	87d	53.3b	101c	52.5b
	gy1,2 Gy3 gy4 gy5	6	18d	60.8a	20e	62.9a	20g	55.1a	23f	60.4a	20e	61.3a
	gy1,2 gy3 gy4 Gy5	7	42c	49.4c	42d	50.2bc	46f	50.9bc	40e	51.3bc	39d	47.1d
	gy1,2 gy3 gy4 gy5	8	0e	0.0d	0f	0.0d	0h	0.0d	0g	0e	0f	0.0e

† Contents of protein components expressed as g kg^{-1} of seed protein.

‡ Contents of subunits expressed as a percentage of total glycinin.

§ Total glycinin

¶Means within a column and trait followed by the same letter were not significantly different at the 0.05 probability level based on Tukey's honestly significant difference (Tukey, 1949).

Table 3. Beta-conglycinin composition of seed from F3 plants for eight genotypic classes from five soybean populations grown in Hawaii in 2008.

Trait	Glycinin alleles	Class	Population 1		Population 2		Population 3		Population 4		Population 5	
			g kg^{-1}†	%‡	g kg^{-1}	%	g kg^{-1}	%	g kg^{-1}	%	g kg^{-1}	%
BC§	*Gy1,2 Gy3 Gy4 Gy5*	1	247e¶		232e		238e		239e		226e	
	Gy1,2 Gy3 gy4 gy5	2	277d		265d		267d		294d		283d	
	Gy1,2 gy3 gy4 Gy5	3	254e		254d		278d		272d		278d	
	Gy1,2 gy3 gy4 gy5	4	277d		261d		291d		282d		299cd	
	gy1,2 Gy3 Gy4 Gy5	5	309c		311c		324c		324c		315c	
	gy1,2 Gy3 gy4 gy5	6	408a		381ab		401a		400a		417a	
	gy1,2 gy3 gy4 Gy5	7	360b		363b		351b		354b		374b	
	gy1,2 gy3 gy4 gy5	8	374b		393a		404a		393a		405a	
α	*Gy1,2 Gy3 Gy4 Gy5*	1	113e	45.9ab	113e	48.8a	113f	47.7a	112e	46.8a	109e	48.0a
	Gy1,2 Gy3 gy4 gy5	2	128d	46.3ab	127d	47.8ab	126e	47.3ab	135d	46.1a	132d	47.0ab
	Gy1,2 gy3 gy4 Gy5	3	120de	47.1a	116e	45.9cd	129de	46.5ab	125d	46.0a	128d	46.5abc
	Gy1,2 gy3 gy4 gy5	4	124de	44.6b	123de	47.2abc	139d	47.8a	129d	45.5a	138cd	46.2abc
	gy1,2 Gy3 Gy4 Gy5	5	144c	46.8a	146c	46.9bcd	154c	47.7a	150c	46.3a	147c	46.8ab
	gy1,2 Gy3 gy4 gy5	6	186a	45.5ab	172b	45.1d	183a	45.6b	187a	46.6a	186a	44.6c
	gy1,2 gy3 gy4 Gy5	7	163b	45.3ab	165b	45.4cd	169b	48.3a	165b	46.8a	170b	45.2c
	gy1,2 gy3 gy4 gy5	8	171b	45.6ab	183a	46.4bcd	190a	47.0ab	177ab	45.0a	185a	45.8c
α′	*Gy1,2 Gy3 Gy4 Gy5*	1	81f	33.0cd	73e	31.4b	78e	33.1c	77e	32.4c	75e	33.0c
	Gy1,2 Gy3 gy4 gy5	2	95e	34.3abc	90d	34.1a	94d	35.2a	99d	33.7bc	96d	33.9bc
	Gy1,2 gy3 gy4 Gy5	3	83f	32.6d	86d	34.0a	94d	33.7bc	92d	34.0b	95d	34.5abc
	Gy1,2 gy3 gy4 gy5	4	96e	34.7ab	90d	34.5a	102cd	35.1a	100d	35.8a	104cd	34.7ab
	gy1,2 Gy3 Gy4 Gy5	5	106d	34.3abc	108c	34.8a	109c	33.7bc	110c	34.1b	108c	34.4abc
	gy1,2 Gy3 gy4 gy5	6	141a	34.4abc	132a	34.9a	137a	34.0abc	139a	34.7ab	141a	34.0bc

Table 3. Continued

Trait	Glycinin alleles	Class	Population 1		Population 2		Population 3		Population 4		Population 5	
			g kg^{-1}†	%‡	g kg^{-1}	%	g kg^{-1}	%	g kg^{-1}	%	g kg^{-1}	%
	gy1,2 gy3 gy4 Gy5	7	122c	34.0bcd	122b	33.6a	122b	34.6abc	122b	34.3b	130b	34.7ab
	gy1,2 gy3 gy4 gy5	8	133b	35.6a	137a	35.0a	142a	35.1a	136a	34.6ab	144a	35.7a
β	*Gy1,2 Gy3 Gy4 Gy5*	1	53c	21.1a	46c	19.7a	46c	19.3ab	50e	20.8a	43d	19.0a
	Gy1,2 Gy3 gy4 gy5	2	54c	19.4a	48bc	18.0a	47c	17.5b	60cd	20.3a	55c	19.2a
	Gy1,2 gy3 gy4 Gy5	3	52c	20.2a	52bc	20.1a	55bc	19.8ab	55de	20.0a	54c	19.1a
	Gy1,2 gy3 gy4 gy5	4	58c	20.7a	48bc	18.3a	50c	17.1b	52de	18.7a	57c	19.1a
	gy1,2 Gy3 Gy4 Gy5	5	59c	18.9a	58b	18.2a	61b	18.6ab	64c	19.6a	59c	18.8a
	gy1,2 Gy3 gy4 gy5	6	81a	20.0a	77a	20.0a	82a	20.4a	75ab	18.6a	89a	21.4a
	gy1,2 gy3 gy4 Gy5	7	75ab	20.7a	76a	21.0a	60b	17.1b	67bc	18.9a	75b	20.0a
	gy1,2 gy3 gy4 gy5	8	71b	18.9a	73a	18.6a	72a	17.9ab	80a	20.5a	76b	18.6a

† Contents of protein components expressed as g kg^{-1} of seed protein.

‡ Contents of subunits expressed as a percentage of total beta-conglycinin.

§ Total beta-conglycinin

¶Means within a column and trait followed by the same letter were not significantly different at the 0.05 probability level based on Tukey's honestly significant difference (Tukey, 1949).

Table 4. Mean and range for yield and maturity of lines for eight genotypic classes from five soybean populations grown in four locations in 2008.

			Adjusted Yield†		Maturity	
Population	Glycinin alleles	Class	Mean	Range	Mean	Range
			----------kg ha^{-1}--------		---------days‡-------	
1	*Gy1,2 Gy3 Gy4 Gy5*	1	2901cd§	2426 - 3417**	19.9d	13.9 - 24.6**
	Gy1,2 Gy3 gy4 gy5	2	3081ab	2687 - 3381**	21.2c	15.5 - 26.3**
	Gy1,2 gy3 gy4 Gy5	3	2911cd	2608 - 3355**	22.9b	16.8 - 27.3**
	Gy1,2 gy3 gy4 gy5	4	2772d	2557 - 3133**	23.2b	15.8 - 27.0**
	gy1,2 Gy3 Gy4 Gy5	5	3169a	2720 - 3492**	21.6c	16.1 - 26.0**
	gy1,2 Gy3 gy4 gy5	6	2935bc	1939 - 3368**	20.1d	14.0 - 24.6**
	gy1,2 gy3 gy4 Gy5	7	2994bc	2733 - 3274**	24.0a	16.4 - 24.6**
	gy1,2 gy3 gy4 gy5	8	2913cd	2579 - 3157**	22.5b	16.0 - 26.5**
2	*Gy1,2 Gy3 Gy4 Gy5*	1	2891a	2398 - 3270**	22.8e	18.4 - 30.9**
	Gy1,2 Gy3 gy4 gy5	2	2767b	2341 - 3147**	23.8d	17.6 - 32.6**
	Gy1,2 gy3 gy4 Gy5	3	2835ab	2281 - 3224**	25.5bc	17.0 - 34.5**
	Gy1,2 gy3 gy4 gy5	4	2790b	2379 - 3211**	25.8bc	21.1 - 30.8**
	gy1,2 Gy3 Gy4 Gy5	5	2885a	2496 - 3229**	26.3ab	18.9 - 31.5**
	gy1,2 Gy3 gy4 gy5	6	2777b	2451 - 3040**	23.4de	18.6 - 30.0**
	gy1,2 gy3 gy4 Gy5	7	2769b	2354 - 3199**	27.0a	21.4 - 34.5**
	gy1,2 gy3 gy4 gy5	8	2764b	2427 - 3014**	25.0c	17.8 - 31.8**
3	*Gy1,2 Gy3 Gy4 Gy5*	1	2784ab	2159 - 3158**	25.1ab	16.4 - 29.1**
	Gy1,2 Gy3 gy4 gy5	2	2729ab	2321 - 3128**	24.7bc	20.0 - 30.1**
	Gy1,2 gy3 gy4 Gy5	3	2521c	1932 - 2949**	24.2cd	16.0 - 31.5**
	Gy1,2 gy3 gy4 gy5	4	2631bc	2044 - 3052**	23.7de	17.6 - 30.6**
	gy1,2 Gy3 Gy4 Gy5	5	2806a	1916 - 3344**	22.8f	16.5 - 28.0**
	gy1,2 Gy3 gy4 gy5	6	2750ab	1896 - 3072**	23.3ed	17.0 - 28.9**
	gy1,2 gy3 gy4 Gy5	7	2735ab	2252 - 3216**	25.5a	16.4 - 31.9**
	gy1,2 gy3 gy4 gy5	8	2505c	1888 - 2874**	25.6a	16.6 - 31.6**
4	*Gy1,2 Gy3 Gy4 Gy5*	1	2836abc	2283 - 3101**	25.6b	21.0 - 31.5**
	Gy1,2 Gy3 gy4 gy5	2	2893a	2335 - 3263**	27a	23.8 - 31.3**
	Gy1,2 gy3 gy4 Gy5	3	2721bcde	2393 - 3136**	25.7b	16.8 - 29.5**
	Gy1,2 gy3 gy4 gy5	4	2637e	1948 - 3206**	25.4b	17.3 - 29.9**
	gy1,2 Gy3 Gy4 Gy5	5	2795abcd	2170 - 3354**	23.9e	16.4 - 29.4**
	gy1,2 Gy3 gy4 gy5	6	2859ab	2255 - 3419**	25.1bc	17.6 - 29.0**
	gy1,2 gy3 gy4 Gy5	7	2687de	1880 - 3166**	24.2de	17.1 - 29.6**
	gy1,2 gy3 gy4 gy5	8	2714cde	1775 - 3232**	24.7cd	17.1 - 29.5**
5	*Gy1,2 Gy3 Gy4 Gy5*	1	2891b	2365 - 3141**	26.3b	22.5 - 32.3**
	Gy1,2 Gy3 gy4 gy5	2	2823bc	2311 - 3095**	27ab	23.4 - 30.0**
	Gy1,2 gy3 gy4 Gy5	3	2878b	2342 - 3284**	27.1ab	22.5 - 32.0**
	Gy1,2 gy3 gy4 gy5	4	2717c	2116 - 3081**	26.3b	22.8 - 29.0**
	gy1,2 Gy3 Gy4 Gy5	5	3027a	2553 - 3615**	26.9ab	24.8 - 29.4**

Table 4. Continued

Population	Glycinin alleles	Class	Adjusted Yield† Mean	Adjusted Yield† Range	Maturity Mean	Maturity Range
			----------kg ha^{-1}--------		----------days‡--------	
	gy1,2 Gy3 gy4 gy5	6	2903b	2438 - 3552**	26.7ab	24.0 - 28.8**
	gy1,2 gy3 gy4 Gy5	7	2878b	2552 - 3208**	26.3b	22.8 - 29.1**
	gy1,2 gy3 gy4 gy5	8	2806bc	2346 - 3114**	28.3a	24.6 - 30.3**

*, ** Significant difference among lines within a class at the 0.05 and 0.01 probability level, respectively.

† Grain yield on a moisture basis of 130 g kg^{-1} after adjustment for maturity based on linear regression analysis.

‡ Days after 31 August.

§ Means within a column followed by the same letter were not significantly different at the 0.05 probability level based on Tukey's honestly significant difference (Tukey, 1949).

Table 5. Number of lines with yields significantly better than the population mean for eight genotypic classes from five soybean populations grown in four locations in 2008.

		Population					
Glycinin alleles	Class	1	2	3	4	5	$\bar{x}$
Gy1,2 Gy3 Gy4 Gy5	1	4†	12	9	7	6	7.6
Gy1,2 Gy3 gy4 gy5	2	9	6	8	10	5	7.6
Gy1,2 gy3 gy4 Gy5	3	3	8	2	5	8	5.2
Gy1,2 gy3 gy4 gy5	4	1	5	6	3	2	3.4
gy1,2 Gy3 Gy4 Gy5	5	12	10	10	9	12	10.6
gy1,2 Gy3 gy4 gy5	6	8	4	7	8	9	7.2
gy1,2 gy3 gy4 Gy5	7	4	4	8	5	7	5.6
gy1,2 gy3 gy4 gy5	8	2	6	1	6	2	3.4

† Significant differences from the mean were determined with Tukey's honestly significant difference test (Tukey, 1949) at the 0.05 probability level.

REFERENCES

Adams, M.R., D.L. Golden, A.A. Franke, S.M. Potter, H.S. Smith, and M.S. Anthony. 2004. Dietary soy β-conglycinin (7S globulin) inhibits atherosclerosis in mice. J. Nutr. 34:511-516.

Aoyama, T., M. Kohono, T. Saito, K. Fukui, K. Takamatsu, T. Yamamoto, Y. Hashimoto, M. Hirotsuka, and M. Kito. 2001. Reduction by phytate-reduced soybean β-conglycinin of plasma triglyceride level of young and adult rats. Biosci. Biotechnol. Biochem. 65: 1071-1075.

Beilinson, V., Z. Chen, R.C. Shoemaker, R.L. Fischer, R.B. Goldberg, and N.C. Nielsen. 2002. Genomic organization of glycinin genes in soybean. Theor. Appl. Gen. 104:1132-1140.

Cai, T., and K.C. Chang. 1999. Processing effect on soybean storage proteins and their relationship with tofu quality. J. Agric. Food Chem. 47:720 – 727.

Cho, T.J., C.S. Davies, and N.C. Nielsen. 1989. Inheritance and organization of glycinin genes in soybean. Plant Cell 1:329-337.

Fehr, W.R., J.A. Hoeck, S.L. Johnson, P.A. Murphy, J.D. Nott, G.I. Padilla, and G.A. Welke. 2003. Genotype and environment influence on protein components of soybean. Crop Sci. 43:511-514.

Guo, S.T., C. Tsukamoto, K. Takahashi, K. Yagasaki, Q.X. Nan, and T. Ono. 2002. Incorporation of soymilk lipid into soy protein coagulum by the addition of calcium chloride. J. Food Sci. 67: 3215-3219.

Jenkinson, J.E., and W.R. Fehr. 2010a. Environmental stability of beta-conglycinin and glycinin in soybean. Crop Sci.50: in review.

Kohno, M., M. Hirotsuka, M. Kito, and Y. Matsuzawa. 2006. Decreases in the serum triacylglycerol and visceral fat mediated by dietary soybean β-conglycinin. J. Atheroscler. Thromb. 13:247-255.

Kuipers, B.J.H., G.A. van Koningsveld, A.C. Alting, F. Driehuis, A.G.J. Voragen, and H. Gruppen. 2006. Opposite contributions of glycinin- and β-conglycinin- derived peptides to the aggregation behavior of soy protein isolate hydrolysates. Food Biophys. 1:178-188.

Lovati, M.R., C. Manzoni, E. Gianazza, and C.R. Sirtori. 1998. Soybean protein products as regulators of liver low-density lipoprotein receptors. I. Identification of active β-conglycinin subunits. J. Agric. Food Chem. 46:2474-2480.

Martinez-Villaluenga, C., N.A. Bringe, M.A. Berhow, and E. Gonzalez de Meija. 2008. β-Conglycinin embeds active peptides that inhibit lipid accumulation in 3T3-L1 adipocytes in vitro. J. Agric. Food Chem. 56: 10533-10543.

Mizuno, Y., K. Takahashi, H. Namai, and K. Kitamura. 1994. Changes of seed components with the absence of α- and α′- subunits of 7S globulin in soybean. Japan. J. Breed. 44: 201-204.

Moriyama, T., K. Kishimoto, K. Nagai, R. Urade, T. Ogawa, S. Utsumi, N. Maruyama, and M. Maebuchi. 2004. Soybean β-conglycinin diet suppresses serum triglyceride levels and normal and genetically obese mice by induction of β-oxidation, downregulation of fatty acid synthase, and inhibition of triglyceride absorption. Biosci. Biotechnol. Biochem. 68: 352-359.

Nielsen, N.C., C.D. Dickinson, T.J. Cho, V.H. Thanh, B.J. Scallon, R.L. Fischer, T.L. Sims, and R.B. Goldberg. 1989. Characterization of the glycinin gene family. Plant Cell. 1:313-328.

Ogawa, T., E. Tayama, K. Kitamura, and N. Kaizuma. 1989. Genetic improvement of seed storage proteins using three variant alleles of 7S globulin in soybean (*Glycine max* L.). Japan. J. Breed. 39:137-147.

Rickert, D.A., L.A. Johnson, and P.A. Murphy. 2004. Functional properties of improved glycinin and β-conglycinin fractions. J. Food Sci. 69:303-311.

SAS Institute. 2002-2003. The SAS system for windows: Release 9.1. SAS Inst., Cary, NC.

Tukey, J.W. 1949. Comparing individual means in the analysis of variance. Biometrics. 5:99-114.

Wang, W., N.A. Bringe, M.A. Berhow, and E. Gonzalez de Meija. 2008. β-conglycinins among sources of different soybean varieties that inhibit leukemia cells in vitro. J. Agric. Food Chem. 56:4012-4020.

Wu, K., T. Horejsi, J. Byrum, N. Bringe, J. Yang, D. Pei, and R. Reiter. Agronomically elite soybeans with high beta-conglycinin content. U.S. Patent application 11/517186. Date issued: 11 Sept 2007.

Yagasaki, K., N. Kaizuma, and K. Kitamura. 1996. Inheritance of glycinin subunits and characterization of glycinin molecules lacking the subunits in soybean (*Glycine max* (L.) Merr.) Breeding Sci. 46:11-15.

CHAPTER 3

ENVIRONMENTAL STABILITY OF BETA-CONGLYCININ AND GLYCININ CONTENT IN SOYBEAN

From a paper submitted to *Crop Science*

ABSTRACT

Soybean [*Glycine max* (L.) Merr.] genotypes with increased beta-conglycinin (BC) and reduced glycinin (Gly) may have benefits in food production and to human health. Stability of BC and Gly content over different environmental conditions will be an important factor in the production of these genotypes. The objective of this study was to evaluate the effect of locations and planting dates on the protein composition of soybean genotypes with *gy* alleles for reduced Gly and increased BC contents. Six soybean genotypes with different *gy* allele combinations and one common cultivar with only wild-type *Gy* alleles were grown at four locations with three planting dates at each location in 2008. There were no significant differences among planting dates for the mean contents of BC, Gly, or their subunits. The means of locations were significantly different for all of the protein components, except for the A124 subunit of Gly. The genotype with four *gy* alleles produced 0 g kg^{-1} of Gly across locations and planting dates, but only three of the *gy* alleles were required to maximize BC content. Grain produced at different locations would be expected to differ in BC content, but planting dates would not be expected to significantly affect that protein component.

INTRODUCTION

Beta-conglycinin (BC) has been shown to improve some functional properties of soybean protein (Rickert et al., 2004) and to have potential benefits to human health such as reduction in cholesterol (Aoyama et al., 2001), control of fat accumulation (Martinez-Villaluenga et al., 2008) and in vitro inhibition of leukemia cell growth (Wang et al., 2008). The seed protein of common North American soybean cultivars typically contains ~ 25 g kg^{-1} BC and ~ 40 g kg^{-1} glycinin (Gly) (Nielsen et al., 1989). Beta-conglycinin is made up of three subunits: α, α′, and β. Glycinin has five subunits: A1aB2, A2B1a, A1bB1b, A5A4B3, and A3B4, each of which has an acidic (A) and basic (B) polypeptide linked by a disulfide bond. The production of these subunits is controlled by the major genes *Gy1* to *Gy5*, respectively (Nielsen et al., 1989). *Gy1* and *Gy2* are separated by approximately 3 kb of DNA and are inherited as a single locus, while the remaining three genes are inherited independently (Cho et al., 1989; Beilinson et al., 2002).

A soybean line, B2G2, containing recessive alleles at the four Gy loci was developed in Japan (Yagasaki et al., 1996). They combined gamma-radiation-induced mutations in *Gy1,2* and *Gy3* with a recessive allele of *Gy4* commonly found in Japanese cultivars and a recessive allele of *Gy5* obtained from a *Glycine soja* line. They observed that the seed protein of B2G2 had ~ 50 g kg^{-1} BC and 0 g kg^{-1} Gly.

The *gy* alleles obtained from B2G2 have been incorporated into lines to produce genotypes with increased BC content that are adapted to North American environments. Successful production of cultivars with increased BC contents will require an understanding of the influence of environment on the protein component. The objective of this study was to

determine the environmental stability of BC, Gly, and their subunits in soybean lines containing different combinations of *Gy* and *gy* alleles.

MATERIALS AND METHODS

The seven genotypes used in the study were developed by the Monsanto Co.. Genotype 1 with all wild-type *Gy* alleles was the cultivar A3525 with a maturity of 3.5. Genotype 2 had *gy* alleles at the five loci and a maturity of 2.5. Genotypes 3-1 and 3-2 with a maturity of 3.0 had the glycinin alleles *gy1,2*, *Gy3*, *gy4*, and *gy 5*, while genotypes 4-1, 4-2, and 4-3 had the alleles *gy1,2*, *Gy3*, *gy4*, and *Gy5* and a maturity of 3.4, 3.0, and 2.8, respectively. The source of the *gy* alleles was the soybean line B2G2. The pedigree of genotype 2 was MV32//B2G2/MV19. MV32 and MV19 were cultivars developed by the Monsanto Co. that had the wild-type *Gy* alleles. The other five genotypes used in the study were developed by crossing genotype 2 to five different cultivars with wild-type *Gy* alleles and selecting for the presence of the *gy* alleles with molecular markers (Wu et al., 2007). The genotypes were planted at the following four locations in 2008 with planting dates in parentheses: Cambridge, IA (5 May, 22 May, 18 June), Huxley, IA (5 May, 22 May, 18 June), Bloomington, IL (14 June, 24 June, 10 July), and Galena, MD (7 May, 28 May, 18 June). The soil types are a Spillville loam (fine-loamy, mixed, superactive mesic Cumulic Hapludoll) at Cambridge, an Ipava silt loam (fine, smectitic, mesic Aquic Argiudoll) at Bloomington, a Nicollet loam (fine-loamy, mixed, superactive, mesic Aquic Hapludoll) at Huxley, and a Matapeake silt loam (fine-silty, mixed, semiactive, mesic Typic Hapludult) at Galena. The experiment was grown as a randomized complete-block design with a split-plot arrangement of planting dates as main plots and genotypes as the subplots. There were two replications at each location. Each plot was a single row 1 m long with 70 cm between plots

and a seeding rate of 25 seeds m^{-1}. The seed from each plot was harvested in bulk and a 250 g sample was analyzed by near infra-red transmittance with an Infratec 1220 grain analyzer (Foss, Eden Prairie, MN) to determine total protein and oil content. A 15-seed sample from each plot was used to measure protein composition by sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE). The seeds were ground with a Thomas-Wiley model 4 mill (Thomas-Wiley, Swedesboro, NJ) using a 2 mm screen. The ground sample was reground in the same mill using a 1 mm screen.

The SDS-PAGE analysis was performed using the method described by Martinez-Villaluenga et al. (2008). This method dissociates each of the acidic polypeptides (A1a, A1b, A2, A3, A4, and A5) from the basic polypeptides (B1a, B1b, B2, B3, and B4), and produces separate bands for each of the three subunits of BC and the A3 and basic subunits of Gly, while the A1a, A1b, A2 and A4 subunits of Gly migrate together and form one band that was labeled A124 in this study.

Analyses of variance were conducted using PROC MIXED of SAS version 9.1 (SAS Institute, 2003) for each protein component. Locations, planting dates, and genotypes were considered fixed effects, and replications were considered a random effect. Differences between means were determined using Tukey's honestly significant difference test (Tukey, 1949).

RESULTS AND DISCUSSION

There were significant differences among the means of locations for the protein and oil content of the seven genotypes (Tables 1 and 2). The range among locations was 34 g kg^{-1} for protein and 21 g kg^{-1} for oil. The differences among planting dates for the two traits were not significant. The range among planting dates was 12 g kg^{-1} for protein and 9 g kg^{-1}

for oil. Although there were significant differences among genotypes for protein and oil content, none of the genotypes with recessive *gy* alleles represented by more than one line were consistently different than the wild-type for either trait. The significant genotype x location and genotype x planting date interactions for protein and oil were associated with changes in the ranking of genotypes.

The mean content of total BC and its subunits were significantly different among the four locations, but not among the three planting dates (Table 1). The range among locations was 42 g kg^{-1} for total BC, 11 g kg^{-1} for α, 13 g kg^{-1} for α′, and 22 g kg^{-1} for β, and the range among planting dates was 6 g kg^{-1} for total BC, 3 g kg^{-1} for α, 5 g kg^{-1} for α′, and 2 g kg^{-1} for β (Table 3). Research by the food industry will be needed to determine if the observed differences among locations would be important for various applications.

The six genotypes with *gy1,2* and *gy4* alleles were significantly different in total BC and its subunits than genotype 1 with only *Gy* alleles (Tables 1 and 3). Genotypes 2, 3-1, and 3-2 had significantly more total BC than genotypes 4-1 to 4-2 and at least 1.8 fold more BC than genotype 1 when averaged across locations and planting dates. The lack of significant differences between genotype 2 with only *gy* alleles and genotypes 3-1 and 3-2 with the Gy3 allele indicated that *gy3* did not significantly alter BC and its subunits when *gy1,2*, *gy4,* and gy5 were present. The same result was observed by Jenkinson and Fehr (2010b) in four of five populations in which lines with and without *gy3* were compared. Genotypes 4-1, 4-2, and 4-3 with the wild-type alleles *Gy3* and *Gy5* were all significantly lower in BC and its subunits than genotype 2, which indicated that the mutant allele *gy5* was important for maximizing total BC and the subunits. Jenkinson and Fehr (2010b) also found that *gy5* had an important role in maximizing the contents of the four seed components.

The absence of significant interactions of genotypes with locations and planting dates for BC, α, and α′ would facilitate selection for the traits in a breeding program (Table 1). Data obtained for lines grown at one location were reliable for identifying the lines with the highest levels of the three traits (Table 3). At all the locations and planting dates, genotypes 4-1, 4-2, and 4-3 had lower contents of the three traits than genotypes 2, 3-1, and 3-2 that had the highest mean contents across locations. Selection for the β subunit would be more difficult because the range in values among lines with different recessive alleles was small and the genotype x location and genotype x planting date interactions were significant.

The increase in the contents of total BC and its subunits for genotypes 2 to 4-3 were achieved through a reduction in Gly content (Table 4). The significant negative correlation between BC and Gly of -0.84 was similar to the negative correlations of -0.92 reported by Fehr et al. (2003) and of -0.89 reported by Jenkinson and Fehr (2010b). The range in mean Gly contents among genotypes was from zero for genotype 2 with only *gy* alleles to 370 g kg^{-1} for genotype 1 with only wild-type alleles. Both the *gy3* and *gy5* alleles were important for reducing Gly. Genotypes 3-1 and 3-2 with the *Gy3* allele had significantly greater Gly than genotype 2. The addition of *Gy5* in genotypes 4-1 to 4-3 resulted in about a two-fold increase in total Gly compared with genotypes 3-1 and 3-2, primarily through an increase in the A3 subunit. Although the *gy3* allele did not seem necessary to maximize BC or to reduce the A3 subunit of Gly, it was necessary to minimize total Gly through its impact on the content of the A124 and basic subunits (Table 4). The presence of the *Gy3* allele in genotypes 3-1 and 3-2 resulted in a significant increase in A124 and the basic subunits compared with genotype 2. Jenkinson and Fehr (2010b) observed a similar impact of *gy3* on Gly and its subunits in their evaluation of lines from five segregating populations. It is not

known at this time if the amount of Gly, A124, and the basic subunit caused by the *gy3* allele is important enough in food applications to warrant incorporating the allele into soybean cultivars developed for their elevated total BC.

There were significant differences among locations for the contents of total Gly and the A3 and basic subunits; however, planting dates were not significantly different for total Gly or the three subunits (Table 1). The range among locations was 16 g kg^{-1} for total Gly, 5 g kg^{-1} for A124, 6 g kg^{-1} for A3, and 7 g kg^{-1} for the basic subunit, and the range among planting dates was 16 g kg^{-1} for total Gly, 6 g kg^{-1} for A124, 1 g kg^{-1} for A3, and 9 g kg^{-1} for the basic subunit (Table 4). There was no production of Gly or the subunits by genotype 2 or of the A3 subunit in genotypes 3-1 and 3-2 for any of the locations or planting dates. This indicated that expression of the *gy* alleles in those genotypes was not influenced by the environment. The significant genotype x location interactions for Gly and its subunits were associated with changes in the magnitude of the differences between genotype 1 and the genotypes with *gy* alleles and with changes in rank for genotypes 4-1 to 4-3 among locations.

In summary, the *gy* alleles for altered Gly content were effective for changing the content of BC across locations and planting dates. For commercial grain production, planting dates would not be expected to have a significant impact on the content of total BC and it subunits for cultivars with *gy* alleles. Grain produced in different locations would be expected to differ in BC content. The importance of the differences among locations will require evaluation by the food industry of grain with contents of BC and Gly similar to those observed among locations in this study. The effect of different years on protein composition in lines with different combinations of *gy* and wild-type alleles will need to be evaluated.

ACKNOWLEDGEMENTS

The author thanks personnel at the Monsanto soybean breeding stations in Bloomington, IL, Galena, MD, and Huxley, IA, for their assistance with the field testing, Nathan Durnell for conducting the SDS-PAGE analysis, and personnel at the Monsanto facility in Ankeny, IA, for conducting the protein and oil analysis.

Table 1. Mean squares from the combined analysis of variance for seed composition of seven genotypes grown at four locations on three planting dates in 2008.

				Beta-conglycinin				Glycinin			
Sources of variation	df	Protein	Oil	Total	α	α′	β	Total	A124	A3	Basic
Location (L)	3	9336**	3860**	14093*	907*	1316*	3860*	1953*	189	446*	1084**
Rep /L (R/L)	4	55	18	950	119	85	339	130	125	38	28
Planting date (P)	2	2183	1207	554	153	407	28	3542	496	31	833
L × P	6	756	229	694	236	82	449	864	185	38	211
P × R/L	8	1365**	667**	779*	205	321**	153*	1062**	251*	37	162
Genotype (G)	6	1634**	964**	132929**	21338**	17582**	11492**	339172**	67007**	22576**	73133**
G × L	18	156**	44**	588	115	82	141*	598*	127**	93**	270**
G × P	12	71*	32**	281	50	42	149*	563	49	27	183
G × L × P	36	33	21	540	119	56	75	341	52	22	115
Error	72	35	12	342	110	69	66	290	52	22	110

*, ** Significant at the 0.05 and 0.01 probability levels, respectively.

Table 2. Mean protein and oil content of seven soybean genotypes at three planting dates and four locations in 2008.

Trait	Glycinin alleles	Genotype	Planting date			Location†				$\bar{x}$
			1	2	3	1	2	3	4	
			g kg^{-1}							
Protein	*Gy12, Gy3, Gy4, Gy5*	1	417ab‡	407ab	421ab	406ab	433a	407a	414a	415a
	gy12, gy3, gy4, gy5	2	410ab	409ab	414ab	406ab	435a	384a	418a	411ab
	gy12, Gy3, gy4, gy5	3-1	412ab	412a	422ab	408ab	430ab	405a	419a	415a
	gy12, Gy3, gy4, gy5	3-2	398b	396b	406b	394b	415c	383a	407a	400b
	gy12, Gy3, gy4, Gy5	4-1	421a	414a	432a	414a	438a	409a	428a	422a
	gy12, Gy3, gy4, Gy5	4-2	416ab	410a	422ab	413a	437a	397a	418a	416a
	gy12, Gy3, gy4, Gy5	4-3	396b	395b	412ab	394b	419bc	379a	412a	401b
$\bar{x}$			410§	406	418	405¶	429	395	417	411
Oil	*Gy12, Gy3, Gy4, Gy5*	1	195abc	201ab	190a	197a	183ab	195a	206a	195ab
	gy12, gy3, gy4, gy5	2	184c	185d	181a	181b	168c	195a	190b	183d
	gy12, Gy3, gy4, gy5	3-1	192abc	192c	185a	192a	178ab	195a	196ab	190bc
	gy12, Gy3, gy4, gy5	3-2	203a	201a	193a	198a	186a	204a	207a	199a
	gy12, Gy3, gy4, Gy5	4-1	189bc	193bc	183a	190ab	175bc	194a	196ab	189cd
	gy12, Gy3, gy4, Gy5	4-2	200ab	203a	194a	197a	185a	207a	208a	199a
	gy12, Gy3, gy4, Gy5	4-3	203a	204a	191a	199a	186a	208a	205ab	199a
$\bar{x}$			195§	197	188	193¶	180	200	201	193

† Location 1 = Cambridge, IA; 2 = Huxley, IA; 3 = Bloomington, IL; 4 = Galena, MD.

‡ Means within a column followed by the same letter were not significantly different at the 0.05 probability level based on Tukey's honestly significant difference (Tukey, 1949).

§ Differences among planting dates were not significantly different at the 0.05 probability level for protein and oil content.

¶ Minimum significant difference among locations at the 0.05 probability level for protein content was 9 g kg^{-1} and for oil content was 2 g kg^{-1}.

Table 3. Mean content of total beta-conglycinin and its subunits of seven soybean genotypes at three planting dates and four locations in 2008.

Trait	Glycinin alleles	Genotype	Planting date			Location†				
			1	2	3	1	2	3	4	$\bar{x}$
			----------g kg^{-1}----------							
BC‡	*Gy*1,2, *Gy*3, *Gy*4, *Gy*5	1	260c§	248c	248c	256c	265d	246c	241c	252e
	*gy*1,2, *gy*3, *gy*4, *gy*5	2	465a	453a	461a	470a	486a	426a	457a	459a
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-1	453ab	453a	459a	471a	472ab	420a	457a	455ab
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-2	464a	456a	469a	464a	491a	447a	451a	463a
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-1	428ab	434ab	432ab	447ab	459ab	417a	402b	431c
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-2	436ab	435ab	431ab	438ab	450bc	417a	431ab	434bc
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-3	415b	399b	396a	409b	423c	380b	401b	403d
$\bar{x}$			417¶	411	414	422#	435	393	406	414
α	*Gy*1,2, *Gy*3, *Gy*4, *Gy*5	1	122d	118c	117d	122d	119e	120d	116e	119c
	*gy*1,2, *gy*3, *gy*4, *gy*5	2	203a	202a	204a	205a	210ab	197ab	202a	203a
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-1	203ab	204a	203ab	205a	207abc	200a	200ab	203a
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-2	204a	203a	203ab	200ab	213a	203a	196ab	203a
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-1	184c	183b	181c	191abc	189cd	180bc	169d	182b
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-2	187bc	189ab	185bc	182bc	192bcd	188abc	185bc	187b
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-3	184c	177b	172c	176c	187d	176c	172cd	178b
$\bar{x}$			184¶	182	181	183#	188	180	177	182
α′	*Gy*1,2, *Gy*3, *Gy*4, *Gy*5	1	92c	90c	93c	90c	97c	91c	87c	91c
	*gy*1,2, *gy*3, *gy*4, *gy*5	2	168a	161a	169a	170a	177a	154ab	162a	166a
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-1	173a	165a	169a	174a	177a	157ab	167a	169a

Table 3. Continued

Trait	Glycinin alleles	Genotype	Planting date			Location†				
			1	2	3	1	2	3	4	$\bar{x}$
			g kg^{-1}							
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-2	169a	163a	171a	169a	175a	164a	163a	168a
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-1	139b	139b	145b	140b	150b	140b	135b	141b
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-2	143b	140b	147b	147b	149b	137b	141b	143b
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-3	145b	143b	142b	145b	145b	141b	141b	143b
$\bar{x}$			147¶	143	148	148#	153	140	142	146
β	*Gy*1,2, *Gy*3, *Gy*4, *Gy*5	1	46d	41d	38c	44d	48e	36d	38b	41c
	*gy*1,2, *gy*3, *gy*4, *gy*5	2	93abc	90bc	88ab	95bc	99bcd	75c	93a	90b
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-1	77c	84c	88ab	92bc	88d	63c	89a	83b
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-2	92abc	91bc	96ab	95bc	104bc	80bc	92a	93b
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-1	104ab	113a	106ab	115a	121a	98a	98a	108a
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-2	107a	107ab	99ab	108ab	110ab	93ab	106a	104a
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-3	87bc	80c	82b	88c	91cd	64c	88a	83b
$\bar{x}$			87¶	86	85	91#	94	72	86	86

† Location 1 = Cambridge, IA; 2 = Huxley, IA; 3 = Bloomington, IL; 4 = Galena, MD.

‡ Total beta-conglycinin

§ Means within a column followed by the same letter were not significantly different at the 0.05 probability level based on Tukey's honestly significant difference (Tukey, 1949).

¶ Differences among planting dates were not significantly different at the 0.05 probability level for BC, α, α′, and β content.

Minimum significant difference among locations at the 0.05 probability level were 36 g kg^{-1} for BC, 13 g kg^{-1} for α, 11 g kg^{-1} for α′, and 19 g kg^{-1} for β.

Table 4. Mean content of glycinin and its subunits of seven soybean genotypes at three planting dates and four locations in 2008.

Trait	Glycinin alleles	Genotype	Planting Date			Location†				
			1	2	3	1	2	3	4	$\bar{x}$
			g kg^{-1}							
Gly‡	*Gy*1,2, *Gy*3, *Gy*4, *Gy*5	1	367a§	352a	392a	353a	378a	380a	372a	370a
	*gy*1,2, *gy*3, *gy*4, *gy*5	2	0d	0d	0d	0d	0d	0d	0d	0d
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-1	74c	52c	69c	53c	66c	82c	60c	65c
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-2	67c	74c	69c	66c	73c	80c	61c	70c
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-1	153b	146b	165b	137b	159b	147b	175b	154b
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-2	164b	157b	175b	162b	178b	151b	169b	165b
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-3	164b	158b	180b	153b	181b	161b	175b	167b
$\bar{x}$			141¶	134	150	132 #	148	143	144	142
A124	*Gy*1,2, *Gy*3, *Gy*4, *Gy*5	1	161a	153a	166a	153a	163a	153a	171a	160a
	*gy*1,2, *gy*3, *gy*4, *gy*5	2	0c	0d	0c	0d	0d	0c	0c	0c
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-1	31b	23c	30b	27bc	30bc	32b	24b	28b
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-2	32b	32b	33b	35b	38b	30b	27b	32b
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-1	26b	23c	28b	22c	26c	24b	32b	26b
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-2	25b	23c	30b	25bc	33bc	24b	23b	26b
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-3	28b	22c	30b	25bc	30bc	27b	25b	27b
$\bar{x}$			43¶	39	45	41 #	46	41	43	43
A3	*Gy*1,2, *Gy*3, *Gy*4, *Gy*5	1	37b	34b	36b	35b	38b	34b	38b	36b
	*gy*1,2, *gy*3, *gy*4, *gy*5	2	0c	0c	0c	0c	0c	0c	0c	0c
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-1	0c	0c	0c	0c	0c	0c	0c	0c
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-2	0c	0c	0c	0c	0c	0c	0c	0c
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-1	56a	57a	64a	55a	63a	53a	65a	59a
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-2	64a	62a	62a	62a	62a	54a	73a	63a
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-3	61a	63a	63a	54a	66a	56a	72a	62a

Table 4. Continued

Trait	Glycinin alleles	Genotype	Planting Date			Location†				
			1	2	3	1	2	3	4	$\bar{x}$
			g kg^{-1}							
$\bar{x}$			31¶	31	32	29 #	33	28	35	31
Basic	*Gy*1,2, *Gy*3, *Gy*4, *Gy*5	1	169a	165a	190a	166a	177a	193a	163a	175a
	*gy*1,2, *gy*3, *gy*4, *gy*5	2	0d	0d	0d	0e	0d	0e	0d	0d
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-1	39c	29c	35c	26d	33c	48cd	32c	35c
	*gy*1,2, *Gy*3, *gy*4, *gy*5	3-2	35c	39c	36c	31d	35c	46d	34c	36c
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-1	71b	66b	73b	61c	71b	70bc	79b	70b
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-2	74b	73b	83b	76b	84b	73b	74b	77b
	*gy*1,2, *Gy*3, *gy*4, *Gy*5	4-3	75b	73b	87b	74b	84b	77b	78b	78b
$\bar{x}$			66¶	63	72	62 #	69	72	66	67

† Location 1 = Cambridge, IA; 2 = Huxley, IA; 3 = Bloomington, IL; 4 = Galena, MD.

‡ Total glycinin.

§ Means within a column followed by the same letter were not significantly different at the 0.05 probability level based on Tukey's honestly significant difference (Tukey, 1949).

¶ Differences among planting dates were not significantly different at the 0.05 probability level for Gly, A124, A3 and basic content.

Minimum significant difference among locations at the 0.05 probability level were 14 g kg^{-1} for Gly, 14 g kg^{-1} for A124 , 6 g kg^{-1} for A3 and 2 g kg^{-1} for basic.

REFERENCES

Aoyama, T., M. Kohono, T. Saito, K. Fukui, K. Takamatsu, T. Yamamoto, Y. Hashimoto, M. Hirotsuka, and M. Kito. 2001. Reduction by phytate-reduced soybean β-conglycinin of plasma triglyceride level of young and adult rats. Biosci. Biotechnol. Biochem. 65: 1071-1075.

Beilinson, V., Z. Chen, R.C. Shoemaker, R.L. Fischer, R.B. Goldberg, and N.C. Nielsen. 2002. Genomic organization of glycinin genes in soybean.104:1132-1140.

Cho, T.J., C.S. Davies, and N.C. Nielsen. 1989. Inheritance and organization of glycinin genes in soybean. Plant Cell 1:329-337.

Fehr, W.R., J.A. Hoeck, S.L. Johnson, P.A. Murphy, J.D. Nott, G.I. Padilla, and G.A. Welke. 2003. Genotype and environment influence on protein components of soybean. Crop Sci. 43:511-514.

Jenkinson, J.E., and W.R. Fehr. 2010b. Agronomic and seed characteristics of soybean lines with alleles for modified glycinin content. Crop Sci.50: in review.

Martinez-Villaluenga, C., N.A. Bringe, M.A. Berhow, and E. Gonzalez de Meija. 2008. β-Conglycinin embeds active peptides that inhibit lipid accumulation in 3T3-L1 adipocytes in vitro. J. Agric. Food Chem. 56: 10533-10543.

Nielsen, N.C., C.D. Dickinson, T.J. Cho, V.H. Thanh, B.J. Scallon, R.L. Fischer, T.L. Sims, and R.B. Goldberg. 1989. Characterization of the glycinin gene family. Plant Cell. 1:313-328.

Rickert, D.A., L.A. Johnson, and P.A. Murphy. 2004. Functional properties of improved glycinin and β-conglycinin fractions. J. Food Sci. 69:303-311.

SAS Institute. 2002-2003. The SAS system for windows: Release 9.1. SAS Inst., Cary, NC.

Tukey, J.W. 1949. Comparing individual means in the analysis of variance. Biometrics. 5:99-114.

Wang, W., N.A. Bringe, M.A. Berhow, and E. Gonzalez de Meija. 2008. β-conglycinins among sources of different soybean varieties that inhibit leukemia cells in vitro. J. Agric. Food Chem. 56:4012-4020.

Wu, K., T. Horejsi, J. Byrum, N. Bringe, J. Yang, D. Pei, and R. Reiter. Agronomically elite soybeans with high beta-conglycinin content. U.S. Patent application 11/517186. Date issued 11-Sept 2007.

Yagasaki, K. N. Kaizuma, and K. Kitamura. 1996. Inheritance of glycinin subunits and characterization of glycinin molecules lacking the subunits in soybean (*Glycine max* (L.) Merr.) Breeding Sci. 46:11-15.

CHAPTER 4

DNA MARKERS ASSOCIATED WITH ALTERED PROPORTIONS OF BETA-CONGLYCININ SUBUNITS

ABSTRACT

The beta-conglycinin (BC) portion of soybean [*Glycine max* (L.) Merr.] seed protein typically contains more of the α subunit than the α′ subunit, but soybean genotypes with α′ as the predominant subunit have been identified. Significant human health benefits, such as cholesterol reduction, may be associated with the α′ subunit. The objective of this study was to determine if DNA markers located on the same linkage group as the genes controlling α subunit expression were associated with a change in the α : α′ ratio from >1.0 to <1.0.

The CG-2 and CG-3 genes control the production of the α subunit and are separated by only about 3 kb of DNA. Two soybean genotypes with a α : α′ ratio of < 1.0 had CC nucleotides at each of two marker loci on the same linkage group as CG-2 and CG-3. These two genotypes, five other genotypes with the same marker alleles (CC-CC), and seven genotypes with TT nucleotides at one marker locus and AA nucleotides at the second marker locus (TT-AA) were grown at one location in 2007. The α subunit content of the CC-CC genotypes was significantly lower than the TT-AA genotypes by 3.1 g kg^{-1}. The α : α′ ratio of the CC-CC gentoypes was always < 1.0 and of the TT-AA group was always > 1.0. Two segregating populations were formed by crossing a CC-CC parent with two TT-AA parents. In these populations, the two markers were associated with the α : α′ ratio. Selection for the CC-CC marker class would increase the frequency of individuals with an α : α′ ratio of < 1.0.

INTRODUCTION

Glycinin (Gly) and beta-conglycinin (BC) are the two major storage proteins of soybean. Glycinin accounts for ~ 40% and BC ~ 25% of total seed proteins (Nielsen et al., 1989). The subunits of Gly and the genes controlling Gly subunit production have been characterized (Cho et al., 1989; Diers et al., 1994; Beilinson et al., 2002). Beta-conglycinin has three subunits: α, α′, and β. The CG-1 gene (Genbank Accession No. AB030838) controls the production of the α′ subunit, while the CG-2 and CG-3 genes (Genbank Accession No. AB051865) control the production of the α subunit (Ishikawa et al., 2006). The specific genes that control β subunit production have not been identified. The sequences of CG-2 and CG-3 are found on soybean linkage group I (www.phytozome.org) and are separated by only about 3 kb of DNA. CG-1 has been mapped to soybean linkage group F (Chen and Shoemaker, 1998).

There is a strong negative correlation between Gly and BC. Fehr et al (2003) reported a correlation of -0.92 in cultivars with normal levels of both proteins. In genotypes with the recessive *gy1*, *gy2*, *gy3*, *gy4*, and *gy5* alleles associated with reduced Gly the correlation ranged from -0.85 to -0.89 (Jenkinson and Fehr, 2010a; Jenkinson and Fehr, 2010b). Genotypes with reduced Gly and increased BC contents are being developed because BC has been shown to improve some functional properties of soybean protein such as low-temperature gelation, protein solubility and emulsification stability (Rickert et al., 2004). BC also has been associated with benefits to human health such as reduced levels of cholesterol, triglycerides and insulin, (Aoyama et al., 2001; Moriyama et al., 2004; Kohno et al., 2006), reduced arterial plaque formation (Adams et al., 2004), and inhibition of leukemia cell

growth in vitro (Wang et al., 2008). Additional studies have demonstrated that the α′ subunit may be largely responsible for the health benefits associated with BC. Lovati et al. (1998) demonstrated that BC was more effective than Gly at upregulating low-density lipoprotein (LDL) uptake and metabolism in human liver cells. It was further shown that protein made from soybeans lacking α′ had no effect on LDL receptor activity (Manzoni et al., 1998), and that purified α′ was more effective than whole BC in in the uptake and degradation of LDL in both human liver cells and in rats (Manzoni et al., 2003; Duranti et al., 2004).

Genotypes with normal and altered levels of Gly and BC usually have the three BC subunits present in the same proportions: the α subunit is ~ 45 % of total BC, α′ is ~ 35 % and β is ~ 20 % (Maruyama et al., 1999). Recently, North American cultivars have been identified that have an α′ content greater than α content, which results in an α : α′ ratio of <1.0. (Bringe et al., 2007). This change in subunit proportions could be due to a genetic change causing either a reduction in α subunit content or an increase in α′ subunit content. The increased α′ content, when combined with the increased total BC trait, may be useful for developing soybean lines with improved human health properties. Two SNP markers, #11 and #15, were identified by Monsanto Co. that are located on LG I in the region of the CG-2 and CG-3 sequences. If the change in subunit proportions is caused by a variant allele of CG-2 and CG-3 then these markers could be used to select for the desired α : α′ ratio. The objective of this study was to determine if these markers were associated with the change in the α : α′ ratio from >1.0 to <1.0.

MATERIALS AND METHODS

Two soybean genotypes developed by Monsanto Co., MV0060 and MV0064, were previously found to have an $\alpha : \alpha'$ ratio of <1.0 compared to a ratio of >1.0 for other conventional cultivars (Table 1). The CG-2 and CG-3 gene sequences associated with production of the α subunit were found on soybean linkage group (LG) I using the BLAST tool of www.phytozome.org. CG-2 and CG-3 and separated by approximately 3 kb of DNA and may be inherited as a single locus (Yoshino et al., 2002). Marker #11 is for a C / T base change on LG I and marker #15 is for a C / A base change on LG I. MV0060 and MV0064 have the C nucleotide for both markers. Many other genotypes of similar maturity have a T nucleotide at marker #11 and an A nucleotide at marker #15.

MV0060, MV0064 and five other genotypes with the CC-CC marker and seven genotypes with the TT-AA marker were planted in one plot each in Bloomington, IL, in 2007. All the genotypes were developed by Monsanto Co. and ranged in relative maturity from 2.8 to 4.1. Each plot was a single row 1 m long with 70 cm between plots and a seeding rate of 25 seeds m^{-1}. Leaf tissue was sampled from ten plants in each row and the DNA from each sample was analyzed with marker #11 to confirm each parent was homozygous for the marker. The seed from each plot was harvested in bulk. A 15-seed sample from each plot was used to measure protein composition by sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE). The seeds were ground with a Thomas-Wiley model 4 mill (Thomas-Wiley, Swedesboro, NJ) using a 2 mm screen. The ground sample was reground in the same mill using a 1 mm screen.

The SDS-PAGE analysis was performed on two replicates of each ground sample using the method described by Martinez-Villaluenga et al. (2008). This method dissociates

each of the acidic polypeptides (A1a, A1b, A2, A3, A4, and A5) from the basic polypeptides (B1a, B1b, B2, B3, and B4), and produces separate bands for each of the three subunits of BC and the A3 and basic subunits of Gly, while the A1a, A1b, A2 and A4 subunits of Gly migrate together and form one band that was labeled A124 in this study.

Analyses of variance were conducted using PROC MIXED of SAS version 9.1 (SAS Institute, 2003) for each protein component. Marker type and genotypes were considered fixed effects, and replications were considered a random effect. Differences between genotype means were determined using Tukey's honestly significant difference test (Tukey, 1949).

Two segregating populations were developed by crossing MV0064 with the CC-CC marker to two lines, MV0124 and MV0040, with the TT-AA marker and $\alpha : \alpha' > 1.0$ (Table 1). All of the parents were developed by the Monsanto Co.. The relative maturity of MV0064 was 3.2, of MV0124 was 2.4 and of MV0040 was 3.5. The parents were crossed at the Monsanto Research Farm in Kihei, HI, in April 2007. The F1 seeds from the two crosses were planted in rows 3.3 m long with 9 seeds m^{-1} during August 2006 at the Monsanto Research Farm in Kihei, HI. Leaf tissue was collected from each F1 plant and DNA from the tissue was analyzed using marker #11. Plants that were heterozygous for the marker were considered to be hybrids. The F2 seeds from hybrid plants within each population were bulked together. Approximately 350 F2 seeds from each population were planted in six 3.3 m long rows with 18 seeds m^{-1} during December 2007 in Kihei, HI. Leaf tissue was collected from each F2 plant in the two populations. The DNA from each tissue sample was analyzed with SNP markers #11 and #15 using a SNPlex assay that can distinguish the two homozygous genotypes and the heterozygous genotype (Tobler et al., 2005). A single pod

was harvested from each plant that had the CC-CC or the TT-AA marker. Within each population, the seed for each marker was bulked. There were no plants for which a recombination between the two marker loci was observed. Ten rows per marker type per population and one row of each of the three parents were planted during June 2008 at the Monsanto Breeding Station at Huxley, IA. Each row was 1 m long with 70 cm between plots and a seeding rate of 25 seeds m^{-1}. Leaf tissue was collected from 187 F3 plants in the MV0124/MV0064 population, and from 230 F3 plants in the MV0040/MV0064 population. The DNA from each tissue sample was analyzed using SNP markers #11 and #15. From the MV0124/MV0064 population, 180 F3 plants and from the MV0040/MV0064 population 220 F3 plants were harvested individually. Seed of each parent was harvested in bulk. Five F4 seeds from each plant and for each parent were used for SDS-PAGE analysis. One replicate per plant and two replicates per parent of SDS-PAGE was performed as described previously.

A chi-square test of independence was used to test for association between the two markers and the altered $\alpha : \alpha'$ trait. The null hypothesis was that the proportion of plants with each level of $\alpha : \alpha'$ was the same for each marker class. The expected value for a marker class / $\alpha : \alpha'$ category, for example TT-AA plants with an $\alpha : \alpha'$ of < 1.0, was calculated in each population by multiplying the proportion of plants in the TT-AA marker class by the proportion of plants with an $\alpha : \alpha'$ of < 1.0 and then multiplying by the total number of plants in that population. This calculation was repeated for each of the remaining three marker class / $\alpha : \alpha'$ categories. A chi-square value was calculated for each of the four categories in each population by squaring the difference between the observed and expected values for each category and dividing by the expected value. The total chi-square test statistic for each

population was calculated by summing the four chi-square values. The critical chi-square value for α=0.01 and 1 degree of freedom was 6.63.

RESULTS AND DISCUSSION

The mean α subunit content of the CC-CC class was significantly lower than the mean of the TT-AA class and five of the CC-CC genotypes were significantly lower in α content than all of the TT-AA genotypes (Table 1). The α subunit content expressed as a percentage of total BC also was significantly lower for the CC-CC class and five of the CC-CC genotypes were lower than all the TT-AA genotypes for this trait.

The mean α' subunit content and α' percentage of total BC of the CC-CC class were significantly higher than the TT-AA class, but none of the individual genotypes were different from each other for α' content and only one CC-CC genotype, MV0061, had a higher α' percentage of total BC than all the TT-AA genotypes. All of the CC-CC genotypes had an α:α' ratio of < 1.0, while all of the TT-AA genotypes were > 1.0, and five of the CC-CC genotypes were significantly lower than all the TT-AA genotypes for this trait.

The mean β subunit content and β percentage of total BC were significantly higher for the CC-CC class compared to the TT-AA class; however, none of the CC-CC lines were significantly different from all of the TT-AA lines. For example, MV0064 and MV0065 with the highest and lowest β subunit content and percentage were both in the CC-CC class.

The contents of total BC, total Gly and the A3 and basic subunits of Gly were not different between classes or genotypes (Table 1 and 2). The content of the A124 subunit of glycinin was higher for the CC-CC class, but six of the TT-AA genotypes were not significantly different in A124 content from all of the CC-CC genotypes (Table 2.)

Jenkinson and Fehr (2010b) found that the relative proportions of the three BC subunits were not changed due to the recessive *gy1*, *gy2*, *gy3*, *gy4*, and *gy5* alleles for reduced glycinin content. In that study and in Jenkinson and Fehr (2010a), all the genotypes had greater α subunit content compared with the α′ subunit, and the β subunit content was the lowest of the three subunits. In the present study, the α′ subunit of the CC-CC class, expressed in g kg^{-1} of seed protein and as percentage of total BC, was greater than the α subunit of the CC-CC class (Table 1). The CC-CC allele may be associated with an α:α′ ratio of < 1.0 because all of the CC-CC genotypes in this study had an α:α′ ratio of < 1.0 and all of the TT-AA genotypes were > 1.0 for this trait.

The two TT-AA parents used to create the populations in this study both had an α:α′ ratio of > 1.0, and the CC-CC parent was < 1.0 (Table 3). In both populations the chi-square independence test indicated an association between the marker class and the α:α′ ratio ($p < 0.01$). Less than seven percent of the TT-AA plants in both populations had an α:α′ ratio of < 1.0. Averaged across both populations greater than 82 % of the plants with an α:α′ ratio of < 1.0 had the CC-CC marker genotype. These results indicated that selecting for the CC-CC allele increases the probability of obtaining plants with the desired α:α′ ratio.

In summary, the markers used in this study were associated with one or more alleles affecting the relative proportions of the α and α′ subunits of BC. Further studies will be needed to determine the mode of inheritance of the reduced α:α′ trait, to develop markers more closely linked to the alleles which control the trait, and to determine if the trait can be combined with alleles which affect glycinin production to further increase α′ subunit content.

ACKNOWLEDGEMENTS

The author thanks personnel at the Monsanto soybean breeding stations in Bloomington,IL, and Huxley, IA for assistance with the field work, Nathan Durnell for conducting the SDS-PAGE analysis, and personnel at the Monsanto facility in Ankeny, IA for conducting the marker analysis.

Table 1. Mean content of beta-conglycinin and its subunits of 14 soybean genotypes grown at one location in 2007.

Marker class	Genotype	BC†	α		α′		α : α′	β	
		-----g kg^{-1}‡-----		%§	g kg^{-1}	%		g kg^{-1}	%
TT-AA	MV0040	28.0a¶	11.0bc	39.7ab	9.3a	33.3cd	1.2ab	6.1bc	21.8ef
	MV0053	25.4a	10.9bc	42.7ab	8.6a	33.7cd	1.3a	6.0bc	23.7cdef
	MV0054	27.5a	11.3ab	41.0ab	9.8a	35.5bcd	1.2abc	6.5bc	23.5cdef
	MV0055	26.4a	11.8ab	44.8a	9.3a	35.1bcd	1.3a	5.4bc	20.2f
	MV0057	24.6a	10.2bc	41.5ab	8.5a	34.6cd	1.2abc	5.9bc	24.0cdef
	MV0058	27.9a	11.5ab	41.1ab	9.2a	32.8cd	1.3a	7.3abc	26.2bcdef
	MV0059	27.9a	12.8a	45.8a	9.0a	32.3d	1.4a	6.2bc	22.1def
	$\bar{x}$	26.8	11.3**	42.4**	9.1**	33.9**	1.3**	6.2*	23.0*
CC-CC	MV0060	28.1a	9.5cd	33.9bcd	10.4a	37.1bcd	0.9bc	8.2ab	29.1abcde
	MV0061	24.8a	7.1e	28.7d	11.5a	46.2a	0.7c	6.2bc	25.1bcdef
	MV0062	26.1a	8.1de	30.9cd	10.2a	38.9bc	0.8c	7.9abc	30.3abcd
	MV0063	24.8a	7.3e	29.4d	9.3a	37.7bcd	0.8c	8.2ab	32.9ab
	MV0064	26.2a	10.4bc	39.6abc	10.9a	41.3ab	0.9bc	5.0c	19.0f
	MV0065	27.7a	7.5e	27.2d	10.5a	38.0bcd	0.7c	9.7a	34.9a
	MV0066	24.6a	7.2e	29.4d	9.7a	39.1bc	0.8c	7.8abc	31.6abc
	$\bar{x}$	26.0	8.2	31.3	10.3	39.7	0.8	7.6	29.0

*, ** Significant difference between marker class means at the 0.05 and 0.01 probability level, respectively.

† Total beta-conglycinin.

‡ Contents of protein subunits expressed as g kg^{-1} of seed protein.

§ Contents of protein subunits as a percentage of total beta-conglycinin.

¶ Means within a column followed by the same letter were not significantly different at the 0.05 probability level based on Tukey's honestly significant difference (Tukey, 1949).

Table 2. Mean content of glycinin and its subunits of 14 soybean genotypes grown at one location in 2007.

Marker class	Genotype	Gly†	A124		A3		Basic	
		------g kg^{-1}‡-----		%§	g kg^{-1}	%	g kg^{-1}	%
TT-AA	MV0040	33.6a¶	14.4bc	46.6a	3.2a	10.4a	16.0a	51.7a
	MV0053	35.4a	15.8abc	48.0a	3.5a	10.5a	16.1a	48.4a
	MV0054	35.2a	16.0abc	48.0a	3.7a	11.1a	15.5a	46.4a
	MV0055	35.7a	15.3bc	46.8a	3.2a	9.8a	17.2a	51.8a
	MV0057	36.2a	15.8abc	47.2a	3.6a	10.8a	16.8a	49.8a
	MV0058	34.4a	15.7bc	49.1a	3.4a	10.6a	15.3a	48.2a
	MV0059	31.5a	14.1c	48.6a	3.1a	10.5a	14.4a	49.8a
	$\bar{x}$	34.5	15.3*	47.7	3.4	10.5	15.9	49.4
CC-CC	MV0060	32.6a	14.3bc	47.8a	3.0a	9.8a	15.4a	51.2a
	MV0061	36.7a	17.1abc	49.4a	3.6a	10.4a	16.0a	46.4a
	MV0062	39.6a	17.3abc	48.0a	4.1a	11.2a	18.3a	50.8a
	MV0063	36.2a	15.9abc	47.7a	4.0a	11.8a	16.4a	49.3a
	MV0064	42.6a	19.2abc	45.3a	3.2a	7.6a	20.3a	47.2a
	MV0065	39.6a	17.6abc	47.7a	3.6a	9.7a	18.4a	49.8a
	MV0066	38.1a	17.2abc	49.2a	3.6a	10.3a	17.4a	49.9a
	$\bar{x}$	37.9	16.9	47.8	3.5	10.1	17.4	49.2

* Significant difference between marker class means at the 0.05 probability level.

† Total glycinin.

‡ Contents of protein subunits expressed as g kg^{-1} of seed protein.

§ Contents of protein subunits expressed as a percentage of total glycinin

¶ Means within a column followed by the same letter were not significantly different at the 0.05 probability level based on Tukey's honestly significant difference (Tukey, 1949).

Table 3. Chi-squared (χ^2) test for association between marker alleles on linkage group I and $\alpha : \alpha'$ in F3 plants and the parent genotypes for two soybean populations at one location in 2008.

			$\alpha : \alpha'$					
			Observed			Expected		
Genotype	Marker class	Ratio	<1.0	>1.0	Total	<1.0	>1.0	χ^2
			----------no.----------					
MV0124	TT-AA	1.6a†						
MV0064	CC-CC	0.8c						
MV0124/MV0064	TT-AA		6	86	92	17.9	74.1	20.1**
	CC-CC		29	59	88	17.1	70.9	
	Total		35	145	180			
MV0040	TT-AA	1.2b						
MV0064	CC-CC	0.8c						
MV0040/MV0064	TT-AA		5	107	112	14.3	97.7	14.2**
	CC-CC		23	84	107	13.7	93.3	
	Total		28	191	219			

** Significant association between marker class and $\alpha : \alpha'$ at the 0.01 probability level.
† Means within a column followed by the same letter were not significantly different at the 0.05 probability level based on Tukey's honestly significant difference (Tukey, 1949).

REFERENCES

Adams, M.R., D.L. Golden, A.A. Franke, S.M. Potter, H.S. Smith, and M.S. Anthony. 2004. Dietary soy β-conglycinin (7S globulin) inhibits atherosclerosis in mice. J. Nutr. 34:511-516.

Aoyama, T., M. Kohono, T. Saito, K. Fukui, K. Takamatsu, T. Yamamoto, Y. Hashimoto, M. Hirotsuka, and M. Kito. 2001. Reduction by phytate-reduced soybean β-conglycinin of plasma triglyceride level of young and adult rats. Biosci. Biotechnol. Biochem. 65: 1071-1075.

Beilinson, V., Z. Chen, R.C. Shoemaker, R.L. Fischer, R.B. Goldberg, and N.C. Nielsen. 2002. Genomic organization of glycinin genes in soybean. Theor. Appl. Gen. 104:1132-1140.

Bringe, N., J. Jenkinson, and K. Wu. 2007. Increased alpha-prime beta-conglycinin soybeans. U.S. Patent application 60/971336. Date filed: 11-Sept 2007.

Cho, T.J., C.S. Davies, and N.C. Nielsen. 1989. Inheritance and organization of glycinin genes in soybean. Plant Cell 1:329-337.

Chen, Z., and R.C. Shoemaker. 1998. Four genes affecting seed traits in soybeans map to linkage group F. J. Hered. 89: 211-215.

Diers, B.W., V. Beilinson, N.C. Nielsen, and R.C. Shoemaker. 1994. Genetic mapping of the *Gy4* and *Gy5* glycinin genes in soybean and analysis of a variant of *Gy4*. Theor. Appl. Genet. 89:297-304.

Duranti, M., M.R. Lovati, V. Dani, A. Barbiroli, A. Scarafoni, S. Castiglioni, C. Ponzone,

and P. Morazzoni. 2004. The α′ subunit from soybean 7S globulin lowers plasma lipids and upregulates liver β-VLDL receptors in rats fed a hypercholesterolemic diet. J Nutr. 134:1334-1339.

Fehr, W.R., J.A. Hoeck, S.L. Johnson, P.A. Murphy, J.D. Nott, G.I. Padilla, and G.A. Welke. 2003. Genotype and environment influence on protein components of soybean. Crop Sci. 43:511-514.

Ishikawa, G., Y. Takada, and T. Nakamura. 2006. A PCR-based method to test for the presence or absence of β–conglycinin α′- and α-subunits in soybean. Mol Breeding. 17:365-374.

Jenkinson, J.E., and W.R. Fehr. 2010a. Environmental stability of beta-conglycinin and glycinin in soybean. Crop Sci.50: in review.

Jenkinson, J.E., and W.R. Fehr. 2010b. Agronomic and seed characteristics of soybean lines with alleles for modified glycinin content. Crop Sci.50: in review.

Kohno, M., M. Hirotsuka, M. Kito, and Y. Matsuzawa. 2006. Decreases in the serum triacylglycerol and visceral fat mediated by dietary soybean β-conglycinin. J Atheroscler. Thromb. 13:247-255.

Lovati, M.R., C. Manzoni, E. Gianazza, and C.R. Sirtori. 1998. Soybean protein products as regulators of liver low-density lipoprotein receptors. I. Identification of active β-conglycinin subunits. J. Agric. Food Chem. 46:2474-2480.

Manzoni, C., M.R. Lovati, E. Gianazza, Y. Morita, and C.R. Sirtori. 1998. Soybean protein products as regulators of liver low-density lipoprotein receptors. II. α-α′ rich commercial soy concentrate and α′ deficient mutant differently affect

low-density lipoprotein receptor activation. . J. Agric. Food Chem. 46:2481-2484.

Manzoni, C., M. Duranti, I. Eberini, H. Scharnag, W. Marz, S. Castiglioni, and M. Lovati. 2003. Sucbellular localization of soybean 7S globulin in HepG2 cells and LDL receptor up-regulation by its constituent subunit. J. Nutr. 133:2149-2155.

Martinez-Villaluenga, C., N.A. Bringe, M.A. Berhow, and E. Gonzalez de Meija. 2008. β-Conglycinin embeds active peptides that inhibit lipid accumulation in 3T3-L1 adipocytes in vitro. J. Agric. Food Chem. 56: 10533-10543.

Maruyama, N., R. Sato, Y. Wada, Y. Matsumura, H. Goto, E. Okuda, S. Nakagawa, and S. Utsumi. 1999. Structure-physicochemical function relationships of soybean β-conglycinin constituent subunits. J. Agric. Food Chem., 47:5278 – 5284.

Moriyama, T., K. Kishimoto, K. Nagai, R. Urade, T. Ogawa, S. Utsumi, N. Maruyama, and M. Maebuchi. 2004. Soybean β-conglycinin diet suppresses serum triglyceride levels and normal and genetically obese mice by induction of β-oxidation, downregulation of fatty acid synthase, and inhibition of triglyceride absorption. Biosci. Biotechnol. Biochem. 68: 352-359.

Nielsen, N.C., C.D. Dickinson, T.J. Cho, V.H. Thanh, B.J. Scallon, R.L. Fischer, T.L. Sims, and R.B. Goldberg. 1989. Characterization of the glycinin gene family. Plant Cell. 1:313-328.

Rickert, D.A., L.A. Johnson, and P.A. Murphy. 2004. Functional properties of improved glycinin and β-conglycinin fractions. J. Food Sci. 69:303-311.

SAS Institute. 2002-2003. The SAS system for windows: Release 9.1. SAS Inst., Cary, NC.

Tobler, A.R., S. Short, M.R. Andersen, T.M. Paner, J.C. Briggs, S.M. Lambert, P.P. Wu, Y Wang, A.Y. Spoonde, R.T. Koehler, N. Peyret, C. Chen, A.J. Broomer, D.A. Ridzon, H.Zhou, B.S. Hoo, K.C. Hayashibara, L.N. Leong, C.N. Ma, B.B. Rosenblum, J.P. Day, J.S. Zigle, F.M. De La Vega, M.D. Rhodes, K.M. Hennessy, and H.M. Wenz. 2005. The SNPlex genotyping system: A flexible and scalable platform for SNP genotyping. J. Biom. Tech. 16: 398-406.

Tukey, J.W. 1949. Comparing individual means in the analysis of variance. Biometrics. 5:99-114.

Wang, W., N.A. Bringe, M.A. Berhow, and E. Gonzalez de Meija. 2008. β-conglycinins among sources of different soybean varieties that inhibit leukemia cells in vitro. J. Agric. Food Chem. 56:4012-4020.

CHAPTER 5

GENERAL CONCLUSIONS

Soybean lines with the *gy1*,*2*, *gy4*, and *gy5* alleles had total beta-conglycinin (BC) contents of approximately 400 g kg^{-1} of seed protein and glycinin (Gly) contents of approximately 35 g kg^{-1} of seed protein. The *gy3* allele was not required to mazimize BC content, but was required to achieve 0 g kg^{-1} Gly.

The differences in seed yield and maturity among lines with various combinations of *gy* alleles were not consistent across the five populations. It may be easier to develop lines with the *gy1*,*2*, *gy4*, and *gy5* alleles, compared to lines that also have the *gy3* allele, with seed yield similar to cultivars with normal protein composition. On average across the five populations lines with *gy1*,*2*, *gy4*, and *gy5* yielded 0.6 % less and had the same number of lines with yield significantly greater than the population mean compared to lines with all Gy alleles, Lines with *gy1*,*2*, *gy3*, *gy4*, and *gy5* yielded 4.2 % less and had four fewer lines with yield significantly greater than the population mean compared to lines with all Gy alleles.

Location had a significant effect on all of the protein components, except for the A124 subunit of Gly. There were no significant differences among planting dates for the mean contents of BC, Gly and their subunits. Therefore, grain produced in different locations should be expected to differ in BC content, but grain produced from different plantings at the same location may not differ in BC content.

Two molecular markers on the same linkage group as the CG-2 and CG-3 genes that control production of the α subunit of BC were evaluated for association with a change in the α:α′ ratio. Genotypes with the CC-CC marker alleles for these two markers all had an α:α′

ratio of < 1.0 and all of the TT-AA genotypes had a ratio of > 1.0 . In two populations formed by crossing a CC-CC parent with two TT-AA parents, there was a significant association between the marker allele and the $\alpha : \alpha'$ ratio trait. Less than 7% of the TT-AA plants in both populations had an $\alpha : \alpha'$ ratio of < 1.0, while 33 % of the CC-CC plants had an $\alpha : \alpha'$ ratio of < 1.0. It should be possible to select for the desired $\alpha:\alpha'$ ratio using these two markers.

APPENDIX A

Agronomic and seed characteristics of soybean lines with alleles for modified glycinin content: Analyses of variance and regression coefficients

Table A1. Analysis of variance for eight protein components in genotypes of eight classes from five populations grown in one environment in 2007.

			Mean Squares																	
Population	Source	df	BC†		α		α′		β		Gly		A3		A124		Basic	
1	Class	7	686.6	**	136.6	**	95.9	**	24.7	**	2486.1	**	64.2	**	597.8	**	614.3	**
	Error‡	144	9.6		2.4		1.4		2.3		9.3		0.1		2.7		2.6	
	CV (%)		9.9		10.9		10.9		23.9		18.7		17.9		25.7		19.6	
2	Class	7	772.7	**	140.0	**	107.0	**	34.9	**	2838.0	**	59.1	**	744.7	**	671.3	**
	Error	144	10.5		2.3		1.1		2.6		13.6		0.2		3.4		4.0	
	CV (%)		10.5		10.7		10.2		27.0		21.4		31.0		26.3		23.3	
3	Class	7	729.0	**	151.1	**	94.1	**	29.7	**	2521.6	**	63.9	**	657.5	**	581.8	**
	Error	144	12.9		3.1		1.8		2.3		12.9		0.3		3.3		4.2	
	CV (%)		11.3		11.7		12.3		25.7		21.5		32.5		27.0		25.0	
4	Class	7	647.6	**	135.3	**	87.7	**	21.9	**	2342.5	**	59.9	**	578.2	**	558.9	**
	Error	144	13.1		3.5		1.7		2.0		15.1		0.3		4.1		4.1	
	CV (%)		11.3		12.6		11.7		22.4		24.2		32.4		31.5		25.4	
5	Class	7	863.3	**	150.8	**	114.7	**	43.6	**	2550.3	**	70.3	**	646.4	**	610.8	**
	Error	144	15.5		3.8		2.0		2.8		11.3		0.1		3.3		3.1	
	CV (%)		12.1		12.9		12.6		26.4		20.1		21.8		26.9		21.3	

** Significant at the 0.01 probability level.

† g kg^{-1} of seed protein for each protein component.

‡ The error variation is from differences between individual lines within each marker class.

Table A2. Analysis of variance for protein and oil content in genotypes of eight classes from five populations grown in three environments in 2008.

			Mean Squares			
Population	Source	df	Protein†		Oil†	
1	Class	7	4.3	**	2.3	**
	Location	2	181.3	**	48.6	**
	Error	445	0.9		0.4	
	CV %		2.3		3.1	
2	Class	7	5.5	**	6.9	**
	Location	2	89.8	**	46.7	**
	Error	445	1.9		0.6	
	CV %		3.3		3.8	
3	Class	7	3.8	**	4.2	**
	Location	2	46.3	**	22.9	**
	Error	445	1.0		0.4	
	CV %		2.4		3.2	
4	Class	7	3.9	**	3.9	**
	Location	2	15.1	**	29.4	**
	Error	445	0.9		0.4	
	CV %		2.3		3.1	
5	Class	7	5.9	**	7.3	**
	Location	2	78.1	**	27.8	**
	Error	445	0.8		0.4	
	CV %		2.2		3.2	

** Significant at the 0.01 probability level.

† g kg^{-1} of seed

Table A3. Expected mean squares for a randomized complete block design over locations

Location (L) †	$\sigma^2 + grc\Phi^2_L + gc\sigma^2_{R/L}$
Rep /L ‡	$\sigma^2 + gc\sigma^2_{R/L}$
Class (C)	$\sigma^2 + grl\Phi^2_C$
Genotype / C	$\sigma^2 + rl\Phi^2_{G/C}$
C × L	$\sigma^2 + gr\Phi^2_{CL}$
G × L	$\sigma^2 + cr\Phi^2_{GL}$
Error	σ^2

† Locations, classes, and genotypes were fixed effects.

‡ Replications were a random effect.

Table A4. Analysis of variance for five populations across four environments in 2008.

			Mean Squares			
Population	Source	df	Adjusted Yield†		Maturity‡	
1	Location (L)	3	1747819.1	**	15578.9	**
	Rep/L (R/L)	4	1097545.9	**	685.5	**
	Class (CL)	7	2288539.7	**	324.1	**
	Geno/CL (G/C)	126	470138.7	**	63.7	**
	CL*L	21	362225.4	**	9.7	**
	G*L	54	120486.4		4.8	
	Error	982	105500.2		4.4	
	CV %		11.0		9.6	
2	L	3	11715168.4	**	19486.2	**
	R/L	4	3079751.9	**	49.6	**
	CL	7	438037.3	**	337.0	**
	G/CL	126	346221.4	**	124.0	**
	CL*L	21	127542.0		9.3	**
	G*L	54	158745.7		5.7	*
	Error	982	118847.4		4.0	
	CV %		12.3		8.0	
3	L	3	25386968.0	**	22498.5	**
	R/L	4	4530799.1	**	22.4	**
	CL	7	2073876.5	**	162.7	**
	G/CL	126	654736.8	**	105.9	**
	CL*L	21	341439.0	**	5.2	
	G*L	54	155051.9	*	3.7	
	Error	982	107240.4		3.6412	
	CV %		12.2		7.8	
4	L	3	23405720.2	**	19622.8	
	R/L	4	3491031.3	**	16.0	
	CL	7	1266172.7	**	146.3	
	G/CL	126	660928.4	**	80.1	
	CL*L	21	291419.8	*	7.2	
	G*L	54	210499.7		7.1	
	Error	982	177309.3		6.2584	
	CV %		15.2		9.9	

Table A4. Continued

			Mean Squares			
Population	Source	df	Adjusted Yield†		Maturity‡	
5	L	3	20315713.4	**	15856.9	**
	R/L	4	1644331.9	**	1.4	
	CL	7	704513.2	**	63.6	**
	G/CL	126	431081.5	**	26.9	**
	CL*L	21	357729.9	**	41.2	**
	G*L	54	150819.7		33.2	**
	Error	982	163916.8		11.5	
	CV %		14.1		12.6	

*, ** Significant at the 0.05 and 0.01 probability level, respectively.

† Grain yield on a moisture basis of 130 g kg^{-1} after adjustment for maturity based on linear regression analysis.

‡ Days after 31 August.

Table A5. Coefficients from linear regression of yield on maturity for five populations grown at four locations in 2008.

	Population				
Location	1	2	3	4	5
1	69**	28**	52**	40**	-2
2	30**	34**	29**	51**	14
3	39**	-20**	27**	21**	8
4	89**	46**	50**	53**	10

** Significant at the 0.01 probability level.

APPENDIX B

Environmental stability of beta-conglycinin and glycinin content in soybean: Expected mean squares and analyses of variance

Table B1. Expected mean squares for a split-plot design over locations with planting dates as the main plots and genotypes as the subplots.

Location (L) †	$\sigma^2 + grp\Phi^2_L + gp\sigma^2_{R/L}$
Rep /L (R/L) ‡	$\sigma^2 + gp\sigma^2_{R/L}$
Planting date (P)	$\sigma^2 + lrg\Phi^2_P + g\sigma^2_{R/LP}$
L × P	$\sigma^2 + gr\Phi^2_{LP} + g\sigma^2_{R/LP}$
P × R/L	$\sigma^2 + g\sigma^2_{R/LP}$
Genotype (G)	$\sigma^2 + lrp\Phi^2_G$
G × L	$\sigma^2 + rp\Phi^2_{GL}$
G × P	$\sigma^2 + lr\Phi^2_{GP}$
G × L × P	$\sigma^2 + r\Phi^2_{PGL}$
Error	σ^2

† Locations planting dates and genotypes were fixed effects.

‡ Replications was a random effect.

Table B2. Mean squares from the combined analysis of variance for seed composition of seven genotypes grown at four locations on three planting dates in 2008.

				Beta-conglycinin				Glycinin			
Sources of variation	df	Protein	Oil	Total	α	α′	β	Total	A124	A3	Basic
Location (L)	3	9336**	3860**	14093*	907*	1316*	3860*	1953*	189	446*	1084**
Rep /L (R/L)	4	55	18	950	119	85	339	130	125	38	28
Planting date (P)	2	2183	1207	554	153	407	28	3542	496	31	833
L × P	6	756	229	694	236	82	449	864	185	38	211
P × R/L	8	1365**	667**	779*	205	321**	153*	1062**	251*	37	162
Genotype (G)	6	1634**	964**	132929**	21338**	17582**	11492**	339172**	67007**	22576**	73133**
G × L	18	156**	44**	588	115	82	141*	598*	127**	93**	270**
G × P	12	71*	32**	281	50	42	149*	563	49	27	183
G × L × P	36	33	21	540	119	56	75	341	52	22	115
Error	72	35	12	342	110	69	66	290	52	22	110
CV%		3.2	4.7	5.0	6.0	6.3	11.5	14.9	20.4	15.6	17.9

*, ** Significant at the 0.05 and 0.01 probability levels, respectively.

APPENDIX C

DNA Markers associated with altered proportions of beta-conglycinin subunits: Expected mean squares and analyses of variance

Table C1. Expected mean squares and analysis of variance for seven genotypes in each of two marker classes grown at one location in 2007.

Source	Expected Mean Squares	df	Mean Squares													
			BC†	α†	α′†	α : α′	β†	Gly†	A3†	A124†	Basic †	α‡	α′‡	β‡		
Marker class (M)	$\sigma^2 + gr\Phi^2_M$	1	4.5	71.3 **	10.9 **	1.6 **	13.0 **	78.2	0.2	18.7 *	16.8	856.9 **	240.1 **	245.4 **		
Genotype/M	$\sigma^2 + mr\Phi^2_{G/M}$	12	4.0	2.4 **	0.7	0.2 **	2.7 **	17.7	0.2	2.9	3.7	22.7 **	11.5 **	32.7 **		
Replication	$\sigma^2 + mg\sigma^2_r$	1	2.6	0.3	0.2	0.0	0.8	25.3	0.0	0.0	3.5	0.2	13.7 *	4.6		
Error	σ^2	13	3.1	0.2	0.6	0.0	0.6	19.0	0.2	0.7	3.1	4.7	2.6	4.1		
CV%			6.7	4.2	7.7	6.3	11.2	12.8	11.9	5.2	10.6	5.9	4.3	7.8		

*, ** Significant at the 0.05 and 0.01 probability level, respectively.

† g kg^{-1} of seed protein

‡ % of total BC

ACKNOWLEDGEMENTS

To my wife Alison, and our sons, Evan and Connor, for their love, support and sacrifice. To my parents, Bob and Janet, for making sure I was ready for a new challenge.

I thank Walter Fehr for his advice, patience and commitment to education. I thank Alejandro Hernandez, Bob Buehler, Brett Bussler, Duška Stojšin and Mike Hawbaker for their guidance and mentoring. I thank Monsanto's Soybean Breeding group, under the leadership of Bob Buehler and Ted Crosbie, for affording me this incredible opportunity, and supporting me throughout my graduate studies. I thank Curtis Scherder, a fellow graduate student, and Neal Bringe, a fellow soybean protein enthusiast, for their help and advice. I would also like to thank the numerous people in Soybean Breeding and in Technology at Monsanto, in particular Jim Behm and Kyle Smith, who helped with the collection of the data used in this research.