Theor Appl Genet (1984) 68:253-257

THEORETICAL AND APPLIED GENETICS

© Springer-Verlag 1984

Inheritance of alleles for Cgy_1 and Gy_4 storage protein genes in soybean *

K. Kitamura **, C. S. Davies and N. C. Nielsen

United States Department of Agriculture, Agricultural Research Service, Agronomy Department, Purdue University, West Lafayette, IN 47907, USA

Received November 17, 1983
Communicated by D. von Wettstein

Summary. A cultivar lacking the glycinin subunit $A_5A_4B_3$ ('Raiden') was crossed with one lacking the α'-subunit of β-conglycinin ('Keburi'). Analysis of F_2 and F_3 progeny indicated that the missing bands of the $A_5A_4B_3$ and the α'-subunit were each controlled by a recessive allele of two independently segregating genes. Gene symbols Gy_4/gy_4 and Cgy_1/cgy_1 were proposed for the genes which confer the presence or absence of the glycinin and conglycinin subunits, respectively.

Key words: Glycinin mutant – β-Conglycinin mutant – Storage protein mutant – Protein quality – Soybean

Introduction

Cultivated soybean varieties generally contain 40 to 50% protein on a dry weight basis. Of this, about 70% is accounted for by two components in the globulin fraction which have the trivial names glycinin and β-conglycinin. Both have low sulfur amino acid contents and this contributes to the nutritional deficiencies of soybeans (Derbyshire et al. 1976). Of these two storage proteins, glycinin contains more sulfur than β-conglycinin (Millerd 1975). In addition, glycinin and β-conglycinin are each composed of multiple subunits which have a 3- to 4-fold difference in their S-amino acid composition (Staswick et al. 1981; Thanh and Shibasaki 1977). This raises the possibility that soybean nutritional quality might be improved through genetic manipulation, provided that genetic variation for these subunits can be found.

In the course of screening soybean germplasm collections, several accessions were identified which lacked specific glycinin or β-conglycinin subunits. One genetic type, represented in this study by the cultivar 'Raiden', lacks glycinin subunit $A_5A_4B_3$ (Staswick and Nielsen 1983; Kitamura et al. 1980), whereas a second type, the cultivar 'Keburi', lacks the α'-subunit of β-conglycinin (Kitamura and Kaizuma 1981). We now describe the inheritance of these subunits as determined from screening segregating progeny by SDS-polyacrylamide gel electrophoresis.

Materials and methods

Crosses between 'Raiden' and 'Keburi' were made with 'Keburi' as the pollen parent, and F_2 seeds were obtained from F_1 plants grown during the spring of 1981. Sixty-two of 120 individual F_2 seeds were bisected with a razor blade to remove 15–20 mg of the cotyledon. Seed portions containing the embryo were planted and F_3 seed was harvested from the individual plants.

In preparation for analysis, the individual F_2 and F_3 seeds were crushed in a coin envelope with a hammer. Fifteen mg of the resulting powder were added to 2 ml of a sample buffer (0.05 M TRIS-HCl (pH 8), 0.2% sodium dodecyl sulfate (SDS) and 5 M urea) and macerated in a mortar and pestle. Samples were allowed to stand for 30 min and then were centrifuged to remove debris. The seed proteins were separated by SDS slab gel electrophoresis using a 2.5 cm, 5% polyacrylamide stacking gel and 7.5 cm, 7 to 13% polyacrylamide gradient separation gel. Electrophoresis was for 30 min at 100 volts followed by 3.5 h at 125 volts. The gels were stained in 0.225% Coomassie blue in water-ethanol-acetic acid (50:43:7) and destained with water-ethanol-acetic acid (17:2:1).

* Cooperative research of USDA-ARS and the Indiana Agric. Exp. Stn., Purdue Univ., West Lafayette, IN 47907, USA. Indiana Agric. Exp. Stn. Journal Article 9675. Financial support from the American Soybean Association Research Foundation is gratefully acknowledged

** Present address: Faculty of Agriculture, Iwate University, 3-18-8 Ueda, Morioka, Japan

F_3 seed samples from each of forty-nine F_2 plants of known genotype were analyzed for protein at the USDA Northern Regional Research Center at Peoria, Illinois. All nine possible F_2 genotypes were represented.

Results

Many of the major polypeptides of both glycinin and β-conglycinin can be unambiguously identified after electrophoretic separation (Fig. 1). The electrophoretic pattern produced by 'Keburi' lacked the α'-subunit of β-conglycinin, whereas the pattern produced by 'Raiden' lacked the A_5 and B_3 subunits of glycinin. A_4, which is also lacking in 'Raiden' (Staswick et al. 1981), migrates with A_{1a}, A_{1b} and A_2 in this electrophoretic system, and therefore its absence could not be scored reliably. F_1 seeds contained both the α' and A_5+B_3 polypeptides, which indicated that presence of the subunits was dominant to absence (data not shown).

The α' subunit in the F_2 generation segregated approximately 3 present : 1 absent (Table 1), which is consistent with the hypothesis that a single recessive gene conferred the absence of this subunit. Absence of A_5 was always accompanied by a missing B_3 band and vice versa. As in the case of the α' subunit, a 3 : 1 ratio for presence and absence of the A_5 and B_3 polypeptides occurred in the F_2 (Fig. 1, Table 1), which suggested that a single recessive gene also conditioned the absence of both A_5 and B_3.

The inheritance of these patterns was further investigated by electrophoretic analysis of F_3 seeds produced by the F_2 seed embryos with cotyledons of known phenotype (Table 2). In both cases, segregation followed the expected 3 : 1, 0 : 1, and 1 : 0 ratios for heterozygous, homozygous dominant, and homozygous recessive genotypes, respectively, as anticipated with single recessive genes.

A few F_3 progeny of homozygous recessive F_2 plants showed the dominant phenotype (i.e., bands representing the α' or A_5B_3 subunits; Table 2). These are probably due to outcrossing because the plants were field-grown. Alternatively, this result could be attributed to modifer genes, although we consider that unlikely.

Joint segregation of the two electrophoretic patterns in the F_2 followed closely the expected 9 : 3 : 3 : 1 ratio for independent inheritance, and a recombination frequency of 45 ± 6.25 was estimated by maximum-likelihood (Allard 1956). These data are consistent with the hypothesis that the two genes are inherited independently.

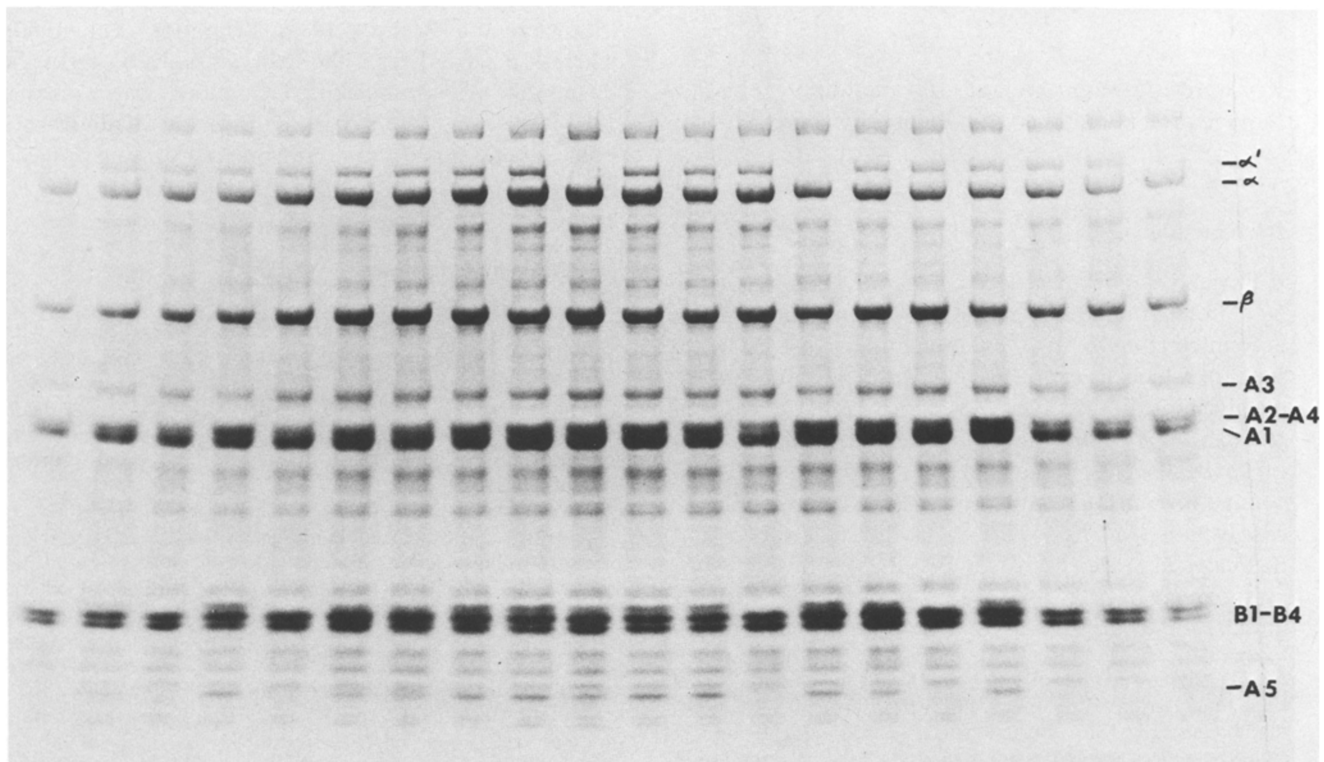

Fig. 1. Sample of SDS-PAGE analysis of seed proteins in the F_2-progeny of 'Raiden' × 'Keburi' (sample in each lane from one seed). Sample preparation and electrophoretic conditions are specified in methods. Both parental types are represented in the electrophoretic pattern; 'Raiden' (gy_4gy_4) *8th lane from right,* 'Keburi' (cgy_1cgy_1) *7th lane from right*

Table 1. Observed and expected F_2 segregation for the genes Cgy_1/cgy_1 and Gy_4/gy_4 from the cross 'Raiden' × 'Keburi'

Gene(s)	Observed	Expected (indep.)	χ^2	P
Segregation for Cgy_1	88:32	90:30	0.18	>0.50
Segregation for Gy_4	94:26	90:30	0.71	>0.25
Joint segregation	68:26:20:6	67.5:22.5:22.5:7.5	1.13	>0.75

Table 2. Observed and expected F_3 segregation for the genes Cgy_1/cgy_1 and Gy_4/gy_4 from the cross 'Raiden' × 'Keburi'

Gene	Proposed F_2 genotype[a]	No. of F_2 plants	Observed	Expected	χ^2	P
Cgy_1	$Cgy_1\ Cgy_1$	21	236:0	236:0	–	–
	$Cgy_1\ cgy_1$	27	248:78	244.5:81.5	0.200	>0.90
	$cgy_1\ cgy_1$	14	4:170	0:174	–	–
Gy_4	$Gy_4\ Gy_4$	16	209:0	209:0	–	–
	$Gy_4\ gy_4$	30	262:88	262.5:87.5	0.003	>0.99
	$gy_4\ gy_4$	16	1:180	0:181	–	–

[a] Based on the segregation ratio of 10–12 F_3 seeds from each F_2 plant

While polypeptides B_3 and A_5 can be separated by electrophoresis in polyacrylamide gels that contain sodium dodecyl sulfate, A_4 migrates to the same relative position as A_{1a}, A_{1b} and A_2 (Moreira et al. 1981). When urea and SDS are both included in the gels, however, the mobility of A_4 in the gels decreases and it becomes separated from all other acidic glycinin polypeptides (Nielsen, unpublished data). Consistent with an earlier report (Staswick and Nielsen 1983), when the proteins in seed extracts of Raiden were separated in the later type of electrophoretic system, A_5, A_4 and B_3 were all missing. Likewise, all recombinants among F_3 seeds, from crosses between 'Raiden' and 'Keburi' in which A_5 and B_3 were missing also lacked A_4 (Fig. 2). Thus, the coordinate loss of all three polypeptides is conditioned by the same recessive allele.

It was of interest to know if the lack of $A_5A_4B_3$ and α'-subunits would detectably alter seed protein content. To address this question, the protein content of seed samples from each of 49 F_2-plants with known genotype was determined. No significant differences in percentage protein between the genotypes were found ($F = 1.08$, $P > 0.05$). Samples ranged between 37.6 and 50.5% protein.

Discussion

Previous studies have shown that glycinin consists of six subunits (Moreira et al. 1981), each of which is

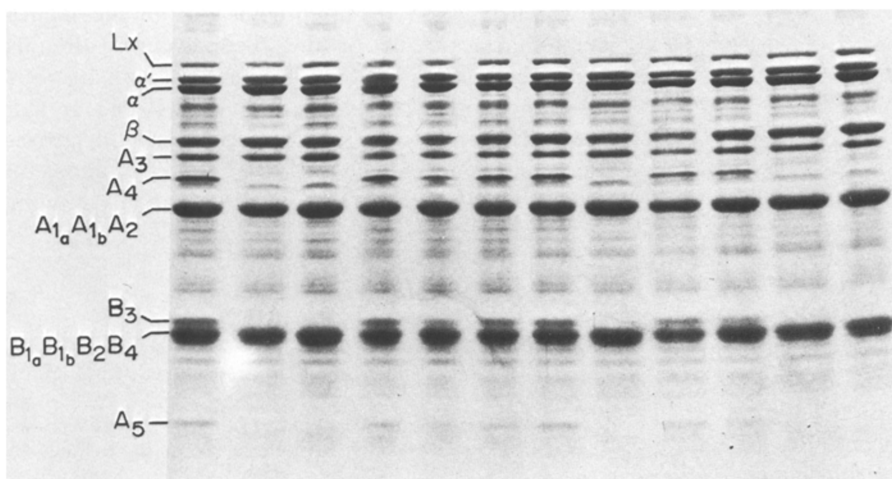

Fig. 2. Sample of SDS-PAGE analysis of seed proteins in the F_3-progeny of 'Raiden' × 'Keburi' (sample in each lane from one seed). The seeds are segregation for Gy_4gy_4 gene. The gel contains 5 M urea which allows visualization of the missing A_4 polypeptide in addition to the A_5 and B_3 polypeptides. All three polypeptides were similarly missing when seed extracts from 'Raiden' were analyzed

composed of one or more acidic polypeptide components which are attached to a basic one by both covalent and non-covalent interactions (Staswick et al. 1981). A series of homologous acidic polypeptide components (A_{1a}, A_{1b}, A_2, A_3, A_4, A_5, A_6) and a series of homologous basic ones (B_{1a}, B_{1b}, B_2, B_3, B_4) have been purified and identified on the basis of differences in their primary structures (Moreira et al. 1981). The initial gene product for each glycinin subunit is a single, high molecular weight precursor which is modified both co- and post-translationally to yield the linked components (Tumer et al. 1981).

Certain properties of A_5 and A_4 are at variance with those of the other acidic components. Polypeptide A_5 is smaller ($M_r \sim 10,000$) than all of the others ($M_r \sim 35,000$ to 42,000), but is covalently bound to B_3 by a single disulfide bond to yield a M_r 30,000 subunit instead of one of M_r 57,000 (Staswick and Nielsen 1983). It is not clear how the smaller subunit is related to the larger ones and how they can occupy equivalent positions in the glycinin complex. Polypeptide A_4, on the other hand, is not linked to a basic polypeptide component by a disulfide linkage like the other acidic ones even though it has the same approximate molecular weight($M_r \sim 37,000$). While NH_2-terminal sequence of A_4 exhibits about 25% homology with that region in the other acidic polypeptides, that degree of homology is substantially lower than the level of homology exhibited in this region (e.g., $>50\%$) by the other acidic polypeptides. Thus, the way in which A_4 contributes to the glycinin complex is also in question.

When considered together with structural data presented elsewhere (Staswick 1982; Staswick and Nielsen 1983; Staswick et al. 1983), the genetic data we have presented provide a logical answer to both of the inconsistencies outlined above. Polypeptide A_5 is homologous with the NH_2-terminal sequences of all acidic components except A_4 (Staswick et al. 1983). The A_4 polypeptide, on the other hand, is homologous with a segment of the A_2 primary structure which begins 100 residues from the NH_2-terminus (Staswick 1982). This suggests that polypeptides A_5 and A_4 together account for the acidic component of a glycinin subunit, with A_5 preceding A_4 in the precursor sequence. This hypothesis predicts that a single genetic defect would result in the coordinate loss of A_5, A_4 and the basic polypeptide associated with this subunit. B_3 has previously been shown to be linked to A_5 by a disulfide bond at a position equivalent to those occupied in other glycinin subunits (Staswick 1982). The null-allele in 'Raiden' is therefore consistent with the predicted type of genetic defect.

The two null-alleles we have described genetically in this work are the first variants proven to affect glycinin and β-conglycinin, the two storage proteins which account for about 70% of the seed protein in soybeans. We use the symbol Gy_4 to denote the gene that controls the presence of $A_5A_4B_3$ in glycinin. The numeral denotes A_4 and is used in an effort to maintain a correspondence between the biochemical and genetic nomenclatures. We chose an acidic polypeptide because the coding sequence for the acidics precedes that of the basics in the DNA precursor. Furthermore, the A_4 subunit was selected because it is larger than A_5. The symbol Cgy_1 was chosen to denote the α'-subunit from β-conglycinin. From biochemical studies (Nielsen 1984), it is clear that each of these genes belong to small families of genes in the soybean genome (Goldberg et al. 1981). The numerical subscripts in the gene symbols anticipate the identification of other members in each of these gene families. Indeed, genes which affect the presence or mobility of a large number of the subunits of both glycinin and β-conglycinin have been identified and will be presented elsewhere (Davis and Nielsen, in preparation).

Genes which cause substantial reductions in the prolamin fraction of various cereals have been reported (for a review, see Larkins 1983) and studied intensively in an effort to use the high lysine phenotypes associated with them to improve seed nutritional quality. While the gy_4gy_4 and cgy_1cgy_1 genotypes each result in the loss of proteins like the genes that affect cereal prolamins, another important characteristic clearly differentiates the effect of the two types of genes. Most of the high lysine variants in cereals appear to result from changes in regulatory genes and cause pleiotropic effects. In addition to reducing the expression of families of structural genes for prolamins, they reduce both the total protein and starch content in the seeds. In contrast, the two genes in soybean appear to affect individual gene products. As such, they may not have the detrimental effects on seed yield that the high lysine genes have in cereals.

The identification of cgy_1 and gy_4 alleles has important ramifications for future attempts to improve the nutritional quality of seeds. There are 2- to 4-fold variations in the methionine content of the various glycinin and β-conglycinin subunits. $A_5A_4B_3$ has one of the lowest methionine contents of the glycinin subunits, whereas the α'-subunit has one of the highest methionine contents of the β-conglycinin subunits. Additional null-alleles may be identified which specifically affect other subunits. Although it is encouraging that percentage protein was not appreciably decreased after elimination of the $A_5A_4B_3$ and α'-subunits, the effect of such alteration on nutritional quality remains to be evaluated.

References

Allard RW (1956) Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia 24:235–278

Derbyshire E, Wright DJ, Boulter D (1976) Legumin and vicilin, storage proteins of legume seeds. Phytochemistry 15:3–24

Goldberg RB, Hoschek G, Tam SH, Ditta GS, Breibenbach RW (1981) Abundance, diversity, and regulation of mRNA sequence sets in soybean embryogenesis. Dev Biol 83:201–217

Kitamura K, Kaizuma N (1981) Mutant strains with low level of 7S globulin in soybean (*Glycine max* Merr.) seed. Jpn J Breed 31:353–359

Kitamura K, Toyokawa Y, Harada K (1980) Polymorphism of glycinin in soybean seeds. Phytochemistry 19:1841–1843

Larkins B (1983) Genetic engineering of seed storage proteins. In: Kosuge T, Meredith CP, Hollaender A (eds) Genetic engineering of plants. Plenum, New York, pp 93–118

Millerd A (1975) Biochemistry of legume seed proteins. Annu Rev Plant Physiol 26:53–72

Moreira MA, Hermodson MA, Larkins BA, Nielsen NC (1981) Comparison of the primary structure of the acidic polypeptides of glycinin. J Biol Chem 210:633–642

Nielsen NC (1984) Chemistry of legume storage proteins. Philos Trans R Soc (in press)

Staswick P (1982) Characterization of the subunits of glycinin and a complete primary structure of the A_2B_1A subunit. PhD Thesis Purdue University, West Lafayette Ind

Staswick PE, Hermodson MA, Nielsen NC (1981) Identification of the acidic and basic subunit complexes of glycinin. J Biol Chem 256:8752–8755

Staswick PE, Nielsen NC (1983) Characterization of a soybean cultivar lacking certain glycinin subunits. Arch Biochem Biophys 223:1–8

Staswick PE, Broue P, Nielsen NC (1983) Glycinin composition of several perennial species related to soybean. Plant Physiol 72:1114–1118

Thanh VH, Shibasaki K (1977) β-conglycinin from soybean proteins: isolation and immunological and physiochemical properties of the monomeric forms. Biochim Biophys Acta 490:370–384

Tumer NE, Thanh VH, Nielsen NC (1981) Purification and characterization of mRNA from soybean seeds: identification of glycinin and β-conglycinin precursors. J Biol Chem 256:8756–8760