Theor Appl Genet (1986) 72, 123–128

Bean arcelin

1. Inheritance of a novel seed protein of *Phaseolus vulgaris* L. and its effect on seed composition

J. Romero Andreas, B. S. Yandell and F. A. Bliss

Department of Horticulture, University of Wisconsin, Madison, WI 53706, USA

Received September 17, 1985; Accepted October 17, 1985
Communicated by D. von Wettstein

Summary. SDS-PAGE of seed proteins from the seeds of a nondomesticated bean of Mexican origin (*Phaseolus vulgaris* L., PI 325690) revealed the presence of a novel 38 kd protein which appeared to be neither an altered phaseolin nor a lectin fraction. The protein was named arcelin, after Arcelia, the town in the state of Guerrero near which PI 325690 had been collected. The pure line, UW 325, was derived by self fertilization of the plant from a single arcelin-containing seed of PI 325690. Despite a low percentage seed phaseolin (14.6%), seed phenotype, seed germination, plant growth, pollen fertility, and percentage seed protein of UW 325 were normal. Analyses of F_2 and F_3 seeds from a single F_1 plant of the cross 'Sanilac' × PI 325690-3 revealed that arcelin expression was inherited as a single gene and that presence was dominant to absence of arcelin. The mean percentage phaseolin in the seeds of homozygous dominant *Arc/Arc* F_3 families (14.0%) was significantly lower than that of the homozygous recessive *arc/arc* seeds (44.7%). The distribution of percentage phaseolin values for seeds within segregating families was bimodal and nonoverlapping. Without exception, seeds containing arcelin (Arc+phenotype) contained a lower percentage phaseolin than seeds lacking arcelin (Arc-phenotype). Although arcelin presence was associated with low percentage phaseolin, the *Arc/Arc* and *Arc/arc* genotypes were similar for seed weight and percentage total seed protein.

Key words: *Phaseolus vulgaris* L. – Arcelin – Phaseolin – Seed proteins – Genetic variation

Introduction

Seeds of most cultivars of common bean (*Phaseolus vulgaris* L.) contain similar constituent proteins based on solubility properties: phaseolin (36–46%, by weight), globulin-2 (or G2/albumin) (5–12%), albumin (12 to 16%), prolamine (2–4%) and an alkali-soluble fraction (20–30%) (McLeester et al. 1973; Sun and Hall 1975; Ma and Bliss 1978). Lectins, defined by their ability to agglutinate blood cells, were detected in both the globulin (G2) and albumin fractions of most cultivars (Osborn et al. 1985; Brown et al. 1981). When present, they make up 6–11% of the total seed protein (Osborn et al. 1984).

The manipulation of structural and regulator genes to reduce, enhance or change the constituent fractions offers a means of producing nutritionally improved seed protein while maintaining seed yield and total protein production. The "lectinless" character (Brucher 1968), is controlled by a single recessive gene which produces not only less lectin, but also a concomitant increase in phaseolin (Osborn and Bliss 1985). Increased phaseolin levels were associated also with the 'T' phaseolin type of 'BBL 240' in 2 of 3 high-phaseolin inbred backcross progeny lines (Sullivan and Bliss 1983a). A similar association between increased phaseolin and the T-type phaseolin was observed by Hartana (1983) who used the same parents, 'Sanilac' and 'BBL 240', but employed more backcrosses to 'Sanilac' to produce near-isogenic lines. The absence of detectable phaseolin in seeds of the *Phaseolus coccineus* L. subsp. *coccineus* cv 'Mexican Red Runner' was shown to be controlled by a single recessive gene in crosses with *P. vulgaris* (Gepts and Bliss 1984).

In the present study, segregating populations of seeds were analyzed electrophoretically to determine the inheritance of expression of a novel protein band and the associated effects on other traits. The protein was named arcelin after Arcelia, the town in the state of Guerrero, Mexico near where PI 325690 had been collected (Gentry 1969).

Materials and methods

Plants and seeds

Electrophoretic screening of accessions of nondomesticated *P. vulgaris* for variability in phaseolin patterns revealed several seeds of PI 325690 which displayed a novel protein band. The plant resulting from one such seed (PI 325690-3) was allowed to self-fertilize and henceforth the progenies of this seed – homozygous for the novel protein band – were referred to as UW 325.

The original plant, PI 325690-3, was used as the pollen parent in a cross to the cultivar 'Sanilac'. A small portion of the cotyledon from the resulting F_1 seed was removed and analysed electrophoretically, and the remainder of the seed planted.

The F_2 seeds produced by self-pollination of the F_1 plant were harvested and a small chip of cotyledon was removed from each of 88 F_2 seeds for electrophoretic determination of protein phenotype. The shoot-root axes from 40 of these seeds were removed and planted, and the remainder of each seed split longitudinally to separate the two cotyledons. The seed coat which adhered to each cotyledon was retained in each analysis. F_3 seeds were divided in a similar manner. One cotyledon was used immediately for rocket immunoelectrophoresis to estimate phaseolin quantity, while the other was stored for subsequent estimation of percentage total protein by micro-Kjeldahl analysis.

Thirty-seven F_2 plants were grown from the 40 shoot-root axes planted. From among the F_3 families having adequate seed numbers, 20 were chosen at random for analyses of individual F_3 seeds for arcelin expression, seed weight, percentage total protein and percentage phaseolin. Sixteen seeds from each of the 20 F_3 families were weighed and the arcelin phenotypes determined by electrophoresis of a chip of a cotyledon.

Protein fractionation

Proteins were extracted from seeds of UW 325 at room temperature using a solution containing 0.5 M NaCl and 0.1% sodium azide at a ratio of 20 ml of solution per gram of flour. After 20 h of continuous stirring, the solution was centrifuged and the clear supernatant saved. This procedure was repeated twice using a 5 h extraction time for a total of three extractions in 30 h. The supernatants were combined and the total volume measured. The solution was then added to 5 volumes of cold (4C) distilled water to precipitate the phaseolin. This dilution resulted in a 0.08 M solution of NaCl at pH 6.41. The solution was then acidified with HCl to a pH of 4,62, and rechilled to 4 °C in a cold room. Previous experiments had shown that no more visually detectable precipitate formed below pH 4.62. The chilled solution was centrifuged at 12 k for 30 min at 4 °C in a Beckman J21 centrifuge. The supernatant was saved for dialysis, and the pellet lyophylized.

The supernatant was dialyzed against distilled water for 48 h in a 4 °C cold room with five changes of water. Dialysis of the nonphaseolin fraction of cultivated beans normally precipitates the remaining globulins referred to collectively as G2, a fraction which is normally a small component of the total protein. Dialysis of the nonphaseolin fraction of UW 325 produced a dense, flocculent precipitate unlike the faint, finely divided precipitate characteristic of G2. The dialysate, which contained arcelin and G2 proteins, was centrifuged and the pellet and remaining supernatant lyophylized.

Sodium-dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Protein extraction from a single cotyledonary chip and SDS-PAGE of the crude extract were done as described by Romero Andreas (1984).

Rocket immunoelectrophoresis for estimation of phaseolin

Individual F_2 and F_3 cotyledons were crushed between two pieces of paper, and the flour added to a preweighed centrifuge tube. Acid salt buffer (0.5 M NaCl, 0.5 M glycine, 0.025% NaN_3 , pH 2.4) was added to flour in a ratio of 50 : 1 (µl buffer to mg flour) according to the procedure of Sun et al. (1978) for the quantitative extraction of phaseolin from bean flour. Phaseolin extracts were then immunoelectrophoresed using the procedure of Laurell (1965, 1967) which was modified as described previously (Sun et al. 1978; Mutschler and Bliss 1981).

Tandem crossed immunoelectrophoresis

Tandem crossed immunoelectrophoresis was done according to Weeke (1973). The use of this method to identify immunologically identical phaseolins has been discussed previously (Mutschler 1979).

Protein determination

Percentage total protein, calculated as percentage N×6.25, was determined by the micro-Kjeldahl procedure (Association of Official Agricultural Chemists 1960).

Statistical analyses

Observed ratios for arcelin expression in segregating F_3 families were tested against an expected 3 : 1 ratio using the chi-square goodness-of-fit test. The calculations were adjusted using Yates' correction for tests having one degree of freedom. A chi-square test for heterogeneity of the data was performed also.

Analyses of variance (ANOVA) for seed weight, percentage total protein and phaseolin as a percentage of protein in seeds of F_2 and F_3 populations were done using the General Linear Model (GLM) for an unbalanced design (SAS User's Guide; Statistics, 1982 edition). Arcelin phenotype was treated as a class variable with two alternatives Arc+ and Arc–. Prior to the GLM analysis, percentage protein and percentage phaseolin values were transformed using the arcsin of the square root of 'p', where 'p' was the percentage value, in order to stabilize the variance. The values for protein and phaseolin which appear in the tables were converted back to percentages for clarity of presentation. The converted means are reported with asymmetrical confidence limits rather than standard errors because the standard errors calculated using transformed values cannot be meaningfully converted (Sokal and Rohlf 1981).

F_3 families were divided into two groups: those segregating for arcelin and those homozygous for arcelin expression. Each group was analyzed as a whole (interfamily analysis), followed by separate analysis for each family within a group (intrafamily analysis).

Results

Identification of the variant PI 325690-3

Six seeds from PI 325690 contained an intensely staining protein band (arcelin) with an apparent molecular

weight slightly heavier than the 31,000 dalton band characteristic of G2/albumin or lectin polypeptides. This intensely staining protein has not been reported previously in seeds of either domesticated or other wild beans.

Electrophoresis of 40 seeds resulting from the self-pollination of one of the variant seeds, PI 325690-3, revealed identical seed protein profiles, with all seeds containing arcelin. The resulting seeds were considered to constitute a pure line, named UW 325 to distinguish it from the original accession (PI 325690).

Fractionation of seed proteins from UW 325

The results of the physical extraction showed that the seed protein of UW 325 was composed of 12% phaseolin, 73% "G2" (shown to include arcelin) and 15% acid albumins. The phaseolin fraction showed no contamination with other proteins (Fig. 1), while the "G2" fraction contained G2/albumin (lectin), a small amount of phaseolin and the protein producing the intensely staining band (arcelin).

The immunological identity of phaseolin from UW 325 and that from 'Tendergreen' was demonstrated by tandem immunoelectrophoresis (not shown). Therefore, rocket immunoelectrophoresis using 'Tendergreen' phaseolin as a standard was a valid method for quantifying phaseolin from UW 325. Rocket immunoelectrophoresis and the physical extraction of protein from seeds of UW 325 yielded similar phaseolin values of 14.6% and 12.2%, respectively.

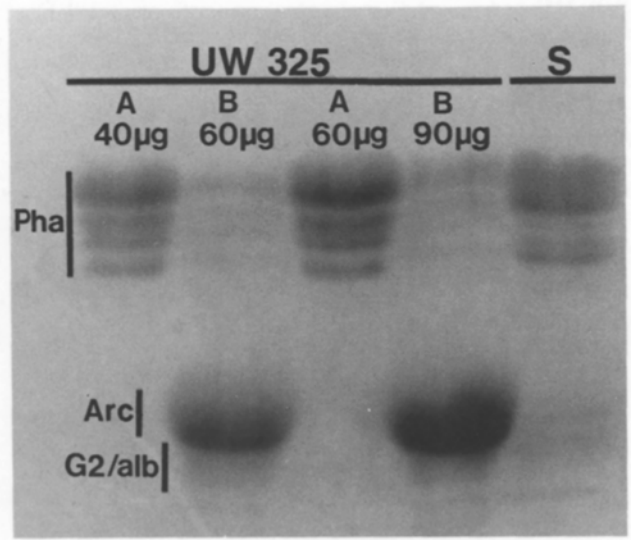

Fig. 1. One-dimensional electrophoretic patterns of phaseolin and 'G2' fraction polypeptides from UW 325 and a 'total' seed protein extract from 'Sanilac' (S). **A** 40 µg and 60 µg of protein from the phaseolin (*Pha*) fraction; **B** 60 µg and 90 µg of protein from the '*G2*' fraction, containing arcelin (*Arc*), *G2*/albumins (*G2/alb*) or lectin

Inheritance of arcelin protein expression

Using one-dimensional SDS-PAGE, each seed was classified phenotypically as either Arc+ or Arc–. The observed F_2 ratio of 67 Arc+ seeds to 21 Arc– seeds fit a 3 : 1 segregation ratio (Table 1), suggesting single gene inheritance, with presence of arcelin being dominant (Fig. 2a). The segregation ratios for each of 10 segregating F_3 families and the combined ratios did not deviate significantly from an expected 3 : 1 ratio (Table 1). The homozygous dominant genotype resulting in the presence of arcelin (Arc+ phenotype) was designated *Arc/Arc,* while the homozygous recessive which showed no arcelin (Arc–) was designated *arc/arc.*

Effects associated with arcelin expression

A population of 40 F_2 seeds from the F_1 of 'Sanilac' × PI 325690-3 was examined to determine whether differences in percentage phaseolin, percentage total protein and seed weight were associated with expression of arcelin protein. Comparisons between Arc+ and Arc– F_2 seeds showed that those seeds containing arcelin had significantly less percentage phaseolin (15.9 g versus 41.3 g/100 g protein) than seeds without arcelin (Table 2). The values for percentage phaseolin (g/100 g protein) of the F_2 seeds showed a bi-modal distribution with no overlap between classes. Low percentage phaseolin was associated invariably with the presence of arcelin. However, no significant differences were found between Arc+ and Arc– phenotypes for either percentage total protein or mean seed weight.

F_3 seeds in 20 families (3 *Arc/Arc,* 7 *arc/arc,* 10 segregating) were analyzed for the same traits. Allowing for the variability among families, there were differences between Arc+ and Arc– seeds, taken from their respective homozygous families, in mean seed weight ($P = 0.049$). As in the F_2 seeds, differences in percentage phaseolin were much larger, whether com-

Table 1. Segregation for expression of arcelin protein in the F_2 and F_3 generations resulting from the cross 'Sanilac' × PI325690-3

Parent or generation	Observed no. of seeds		Chi square	*P*
	Arc+	Arc –		
Sanilac ×	24			
PI 325690-3 ×		24		
F_1	1			
F_2	67	21	0.06	> 0.50
F_3 combined[a]	161	50	0.23	> 0.50

[a] Heterogeneity Chi square = 6.815; df = 9; $P > 0.50$

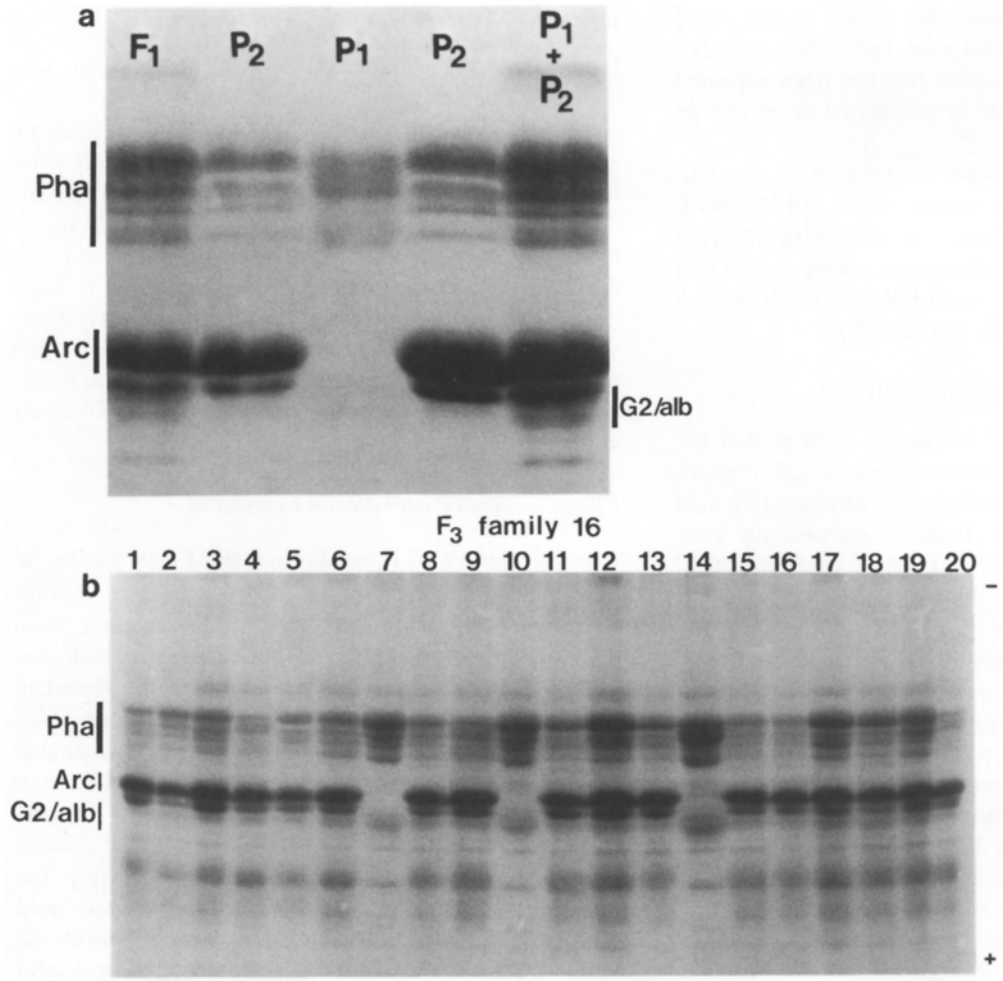

Fig. 2 a, b. Electrophoretic patterns of seed proteins from 'Sanilac', 'UW 325', and F_1 and F_3 seed from the cross 'Sanilac × UW325'; Pha-phaseolin; Arc-arcelin; G2/alb G2/albumin (lectin). **a** P1 'Sanilac'; P2 'UW 325'; F1 'Sanilac × UW 325'; *P1 + P2*–mixture of equal amounts of Sanilac and UW 325 seed protein; **b** *Arc+* and *Arc–* seed phenotypes in F_3 family 16. The seed protein profiles in *lanes 7, 10,* and *14* are *Arc–*, the remaining profiles are *Arc+*

Table 2. Comparison on mean seed weight, percentage total protein and percentage phaseolin between Arc+ and Arc – seeds in one F_2 population, seven *arc/arc* and three *Arc/Arc* F_3 families, and ten segregating F_3 families, resulting from the cross 'Sanilac' × PI 325690-3

Parent or generation	Total seeds	Arc (+/–)	Seed weight (mg)	Total protein (g/100 g flour)	Phaseolin (g/100 g protein)
F_2	12	–	85.5 ± 8.8[a]	21.5 (20.9, 22.1)[b]	41.3 (40.7, 42.0)
	28	+	81.9 ± 5.7	21.4 (20.7, 22.0)	15.9 (15.7, 16.0)
F_3 Homoz.[c]	112	–	119.1 ± 2.1	25.0 (24.8, 25.2)	44.7 (44.2, 45.3)
	48	+	111.3 ± 3.2	25.7 (25.4, 26.0)	14.0 (13.4, 14.6)
F_3 Segreg.[d]	50	–	103.6 ± 2.3	25.0 (24.7, 25.2)	45.4 (44.5, 46.3)
	161	+	99.9 ± 1.3	24.4 (24.2, 24.5)	19.1 (18.7, 19.5)
'Sanilac'[e]		–	223	26.4	46.7
UW 325		+	42	25.5	14.6

[a] Standard error of the mean
[b] Asymmetrical 95% confidence interval
[c] 112 *arc/arc* seeds vs. 48 *Arc/Arc* seeds
[d] 50 Arc – seeds vs. 161 Arc+ seeds
[e] Parental values are from analysis of bulk seed lots rather than from single seeds

Table 3. Range of values for percentage phaseolin (g phaseolin/100 g protein) in Arc+ and Arc− seeds sampled from either homozygous or heterozygous F_3 families of the cross 'Sanilac' × PI 325690-3

Family no.	No. of seeds		Range of values (g phaseolin/100 g protein)			
	Arc+	Arc−	Arc+		Arc−	
			Lowest	Highest	Lowest	Highest
7	16	0	3.9	16.0		
26	16	0	9.5	20.0		
35	16	0	12.1	22.8		
30	0	16			28.7	51.2
14	0	16			32.2	58.3
31	0	16			33.2	62.9
40	0	16			36.2	57.6
24	0	16			37.3	54.2
11	0	16			38.2	58.3
36	0	16			41.2	59.2
25	13	7	4.7	27.6	38.2	50.3
20	16	4	8.4	23.7	37.7	57.1
19	15	5	9.1	34.8	49.1	58.7
23	17	3	10.1	26.5	28.8	36.5
22	17	3	11.6	26.1	31.4	48.1
10	21	10	11.7	29.0	39.3	50.5
16	17	3	11.9	30.9	40.3	47.7
3	14	6	14.3	28.4	40.4	54.3
4	17	3	14.3	30.0	43.0	52.6
33	14	6	14.4	33.7	43.4	52.9

parisons were made between Arc+ and Arc− phenotypes in homozygous families or in segregating families (Table 2). Regression of percentage protein on percentage phaseolin showed a slight relationship ($P = 0.038$), which was unaffected by presence or absence of arcelin.

Separate analyses were performed on each of the F_3 families to investigate the relationships among percentage phaseolin, percentage protein and mean seed weight as mediated by arcelin expression. Of the 10 homozygous families, three showed positive correlations between phaseolin and seed weight; one a negative correlation. Three families showed a positive association between phaseolin and percentage protein, and three a slight negative association. Analyses of the 10 segregating families revealed no apparent relationship between arcelin and either seed weight or percentage total protein, though some differences overall were seen.

Comparisons between the 10 homozygous families (either *Arc/Arc* or *arc/arc*) showed that the distributions of percentage phaseolin values did not overlap between the two phenotypic groups. Rankings of each seed within each segregating F_3 family for percentage phaseolin showed that seeds containing arcelin had the lowest percentage phaseolin within that family without exception (Table 3).

Discussion

A novel seed protein was discovered in several seeds of PI 325690, an accession of nondomesticated common bean collected near Arcelia, Guerrero, Mexico. Differential solubility studies showed that the protein co-isolated with the G2 fraction. The amount extracted, as well as the appearance of the protein on one-dimensional SDS-PAGE suggested that the protein occurred in the seeds in large quantities. Other studies suggest that arcelin is neither altered phaseolin nor lectin, but a distinct protein unlike other bean seed storage proteins in amino acid composition, isoelectric point, and N-terminal amino acid sequence (Blake and Bliss 1983). Since the seeds of UW 325 contained 25.5% protein, a value typical for common beans, other proteins apparently comprised a major part of the seed protein.

Seed weight, percentage total protein and percentage phaseolin are usually maternally influenced, quantitative traits (Kelly and Bliss 1975; Mutschler and Bliss 1981; Sullivan and Bliss 1983b). The distribution of percentage phaseolin values in the segregating populations was consistent with the expected effect of a single dominant gene on a metric trait. Low percentage phaseolin was associated with arcelin presence regardless of seed weight, percentage total protein, genetic background, and gene dosage. The presence or absence of arcelin produced no consistent effects on either seed size or percentage total protein. The bimodal distribution of percentage phaseolin values in segregating families demonstrated that arcelin presence and low percentage phaseolin were highly heritable and that maternal influence was much less than is usually observed.

It is possible that the structural gene for arcelin is tightly linked to a trans-acting regulator gene for phaseolin, and the association between arcelin and phaseolin is not a cause and effect relationship. If linkage is the case, then the regulator gene must be dominant in order to explain the observed inheritance of the low percentage phaseolin trait. A dominant regulator gene might produce a gene product which somehow interferes with the initiation of phaseolin synthesis or the accumulation of phaseolin in a highly specific manner. Alternatively, the association between arcelin expression and low percentage phaseolin could be a competition effect, due to preferential synthesis of arcelin, earlier initiation of synthesis, or preferential packaging.

The dramatic change in the seed protein composition of Arc+ lines had no apparent effect on seed germination, plant growth or reproductive fitness. Arc+ seeds were visually indistinguishable from Arc− seeds.

Acknowledgements. The authors gratefully acknowledge the helpful comments of Drs. O. E. Nelson and W. H. Gabelman (University of Wisconsin-Madison) and Dr. T. C. Osborn (ARCO Plant Cell Research Institute) on this manuscript, and

the assistance of Ken Kmiecik in conducting this research. Funds for these investigations were provided by the Graduate School and the College of Agricultural and Life Sciences, University of Wisconsin-Madison and the ARS of the USDA under grant no. 81-CRCR-1-0604 of the Competitive Grants Office.

References

Association of Official Agricultural Chemists (1960) Official methods of analysis, 9th edn. The Association, Washington DC, p 643

Blake T, Bliss FA (1983) Purification and characterization of a novel major bean seed protein. Agron Abstr, p 55

Brown JWS, Osborn TC, Bliss FA, Hall TC (1981) Genetic variation in the subunits of globulin-2 and albumin seed proteins of French bean. Theor Appl Genet 60:245–250

Brucher O (1968) Absence of phytohaemagglutinin in wild and cultivated beans from South America. Proc Tropical Region Am Soc Hortic Sci 12:68–85

Gentry HS (1969) Origin of the common bean, *Phaseolus vulgaris*. Econ Bot 23:55–69

Gepts P, Bliss FA (1984) Enhanced available methionine concentration associated with higher phaseolin levels in common bean seeds. Theor Appl Genet 69:47–53

Hartana A (1983) Genetic variability in seed protein levels associated with two phaseolin protein types in common bean (*Phaseolus vulgaris* L.). MSc Thesis, University of Wisconsin, Madison

Kelly JD, Bliss FA (1975) Heritability estimates of percentage seed protein and available methionine and correlations with yield in dry beans. Crop Sci 15:753–757

Laurell CB (1965) Antigen antibody crossed electrophoresis. Anal Biochem 10:358–361

Laurell CB (1967) Quantitative immunoelectrophoresis in antibody containing agarose gel. In: Peters H (ed) Proteins in biological fluids, vol 14. Elsevier, Amsterdam, pp 499–502

Ma Y, Bliss FA (1978) Seed proteins of common bean. Crop Sci 17:431–437

McLeester RC, Hall TC, Sun SM, Bliss FA (1973) Comparison of globulin proteins from *Phaseolus vulgaris* with those of *Vicia faba*. Phytochemistry 2:85–93

Mutschler MA (1979) The genetic control of globulin-1 seed protein and its relationship to total protein content and quality in dry bean (*Phaseolus vulgaris* L.); and male sterility in the dry bean (*Phaseolus vulgaris* L.). PhD Thesis, University of Wisconsin, Madison

Mutschler MA, Bliss FA (1981) Inheritance of bean seed globulin content and its relationship to protein content and quality. Crop Sci 21:289–294

Osborn TC, Bliss FA (1985) Effects of genetically removing lectin seed protein on horticultural and seed characteristics of common bean. J Am Soc Hortic Sci 110:484–488

Osborn TC, Brown JWS, Bliss FA (1985) Bean lectin. 5. Quantitative genetic variation in seed lectins of *Phaseolus vulgaris* and its relationship to qualitative lectin variation. Theor Appl Genet 70:22–31

Osborn TC, Mannen J-F, Brown JWS, Bliss FA (1984) Bean lectin. 4. Genetic variation in the non-denatured structure of lectins from different *Phaseolus vulgaris* L. cultivars. Theor Appl Genet 67:547–552

Romero-Andreas J (1984) Genetic variability in the seed phaseolin of nondomesticated bean (*Phaseolus vulgaris* var. 'aborigineus') and the inheritance and physiological effects of arcelin, a novel seed protein. PhD Thesis, University of Wisconsin, Madison

Sokal RR, Rohlf FJ (1981) Assumptions of analysis of variance. In: Biometry. Freeman, San Francisco, pp 418–419

Sullivan JG, Bliss FA (1983a) Expression of enhanced seed protein content in inbred backcross lines of common bean. J Am Soc Hortic Sci 108:787–791

Sullivan JG, Bliss FA (1983b) Genetic control of quantitative variation in phaseolin seed protein of common bean. J Am Soc Hortic Sci 108:782–787

Sun SM, Hall TC (1975) Solubility characteristics of globulins from *Phaseolus* seeds in regard to their isolation and characterization. J Agric Food Chem 23:184–189

Sun SM, Mutschler MA, Bliss FA, Hall TC (1978) Protein synthesis and accumulation in bean cotyledons during growth. Plant Physiol 61:918–923

Weeke, B (1973) Rocket Immunoelectrophoresis. Scand J Immunol (Suppl 21) 2:37–46