Published March, 1997

Relationships between Seed Yield and Seed Protein in Determinate and Indeterminate Soybean Populations

James R. Wilcox* and Zhang Guodong

ABSTRACT

The development of soybean [*Glycine max* (L.) Merr.] cultivars with high seed protein has been hampered by the inverse relationship between seed yield and seed protein concentration. Determinate and indeterminate near-isolines of soybean are usually similar in seed protein but differ in growth type. This study was conducted to determine if relationships between seed yield and seed protein differed in populations of these two plant types. Determinate and indeterminate progenies were identified in two crosses between high-protein indeterminate breeding lines and determinate strains with average seed protein. The F_4-derived progenies were evaluated for 2 yr in three-replicate yield tests at West Lafayette, IN. Seed protein and oil concentration were determined on seed from individual entries in each replication. Mean seed protein concentrations were similar for the two plant types in each cross. Regression of seed protein on seed yield was not significant among determinate progenies in either cross. In contrast, negative slopes of regression lines were significant when seed protein was regressed on seed yield for indeterminate progenies in each cross. The three highest yielding determinate entries from one cross were similar in yield to the determinate cultivar Charleston and varied from 449 to 478 g kg^{-1} seed protein compared with Charleston with 420 g kg^{-1} protein. In the second cross, the three highest yielding determinate progenies were similar in yield to Charleston; two of these had higher seed protein than Charleston. The three highest yielding indeterminate lines in both crosses were similar in yield to the indeterminate check cultivars Edison and Flyer and varied in seed protein from 418 to 435 g kg^{-1} compared with 414 g kg^{-1} protein for Edison and 418 g kg^{-1} for Flyer. The data demonstrate that in these two crosses, determinate progenies were a better source of selections that combined high seed yield with high seed protein than were indeterminate progenies.

Soybean is grown for the oil in the seed and for the high protein meal after the oil is extracted. Composition for these major seed constituents has, with few exceptions, remained unchanged at ≈400 g kg^{-1} protein and 210 g kg^{-1} oil, on a moisture-free basis, in commercially grown cultivars during the 70 yr soybean has been produced in the USA. There is considerable genetic variability for seed protein concentration among germplasm accessions of soybean. The major constraint to the development of soybean cultivars with increased seed protein has been the inverse relationship between seed protein and seed yield in both determinate and indeterminate soybean (Burton, 1985; Hartwig and Hinson, 1972; Sebern and Lambert, 1984; Wehrmann et al., 1987).

Determinate and indeterminate soybean adapted to production in the northern USA vary in growth characteristics (Bernard, 1972). Determinate plants terminate stem growth abruptly at the onset of flowering. This results in a short plant with few nodes and thick stems compared with indeterminate plants. Cooper (1985) has developed several high yielding determinate cultivars, adapted to production in the northern USA, that are very resistant to lodging because of their short stature. In contrast, indeterminate plants continue stem growth, node, and leaf production for several weeks after the onset of flowering. The stem and upper leaves become progressively smaller and number of pods per node decreases near the top of indeterminate plants. This results in tall plants that produce pods during a period of several weeks.

Seed protein concentration is usually very similar in determinate and indeterminate near-isolines of soybean. Bernard (1972) reported that seed protein and oil concentrations were virtually identical in determinate and indeterminate isolines of 'Clark' and 'Harosoy.' In contrast, Hicks et al. (1969), evaluating these same isolines at various plant populations during 2 yr, reported the determinate and indeterminate isolines were similar in seed oil concentration but determinate isolines averaged lower seed protein content than indeterminate isolines. Escalante and Wilcox (1993) reported no difference in mean seed protein concentration between 10 determinate and indeterminate F_6-derived near-isolines of soybean developed from the crosses 'Hobbit' × 'Dawson', Hobbit × 'Lakota', Dawson × CX663-37-2-2, and 'Amcor' × CX663-37-2-2. The parentage of CX663-37-2-2 is ('Williams'^{5} × 'Wayne' *r* Rpm Rpsl × a determinate selection from 'Amsoy' × PI 219.782. The range in seed protein across seven nodes of determinate plants was similar to the range across 11 nodes of indeterminate plants.

The purpose of this study was to determine if relationships between seed yield and seed protein concentration differed among populations of determinate and indeterminate soybean. The information would be useful in developing breeding strategies to develop productive soybean cultivars with high seed protein concentration.

MATERIALS AND METHODS

Determinate and indeterminate progenies varying in seed protein concentration were developed from two soybean crosses, CX1314 (CX1038-63 × HC84-553-1) and CX1412 (Charleston × CX1039-99). The indeterminate breeding line CX1038-63 was the highest protein (477 g kg^{-1}) selection from the cross 'Cutler 71' × CX797-21. CX1039-99, also indeterminate, was the highest protein (465 g kg^{-1}) selection from the cross CX797-115 × Cutler 71. The breeding lines CX797-21 and −115 are high-protein selections from the backcross Cutler 71(2) × 'Pando'. HC84-553-1, with 414 g kg^{-1} seed protein, is a determinate selection from the cross Hobbit × K74-104-76-205 and was evaluated in Uniform Test III of The Uniform Soybean Tests Northern States 1988 (Wilcox, 1988). Charleston is a high-yielding, determinate cultivar that averages 420 g kg^{-1} seed protein concentration

J.R. Wilcox, USDA-ARS, Crop Production and Pathology Res. and Dep. of Agron., Purdue Univ., West Lafayette, IN 47905-1150; and G. Zhang, Dep. of Agron., Purdue Univ., West Lafayette, IN 47901-1150. Journal Paper no. 14736 of the Purdue Univ. Agric. Res. Programs. Received 2 Aug. 1995. *Corresponding author (jwilcox@dept.agry.purdue.edu).

Published in Crop Sci. 37:361–364 (1997).

Abbreviation: APV, approximate processed values.

(Cooper et al., 1995). Progenies from these two crosses were advanced by single seed descent from the F_2 through the F_4 generation. Individual F_4 plants were harvested and seeds from these random F_4 plants were evaluated for protein and oil concentration to determine the variability among progenies for these traits. The F_4 progeny were grown in F_5 plant rows, 2 m long, spaced 0.75 m apart. Determinate and indeterminate F_5 plant rows were selected on the basis of good agronomic type and to ensure variability in seed protein.

Progenies from the two crosses were evaluated in separate tests on a Chalmers soil pine-silty, mixed, mesic Typic Haplaquoll) at the Purdue Univ. Agronomy Research Center near West Lafayette, IN. Thirty-six determinate and 40 indeterminate lines varying in seed protein from 417 to 500 g kg^{-1} were included from the cross CX1314, and 25 determinate and 20 indeterminate lines varying in seed protein from 405 to 485 g kg^{-1} were included from the cross CX1412. These lines plus parents and the check entries Edison, Flyer, and Resnik, cultivars currently in production, were evaluated in three replications of a randomized complete-block design in each of 2 yr. Randomization was restricted in each block so that determinate and indeterminate selections were grouped to avoid competitive effects of tall indeterminate on short determinate selections. Four-row plots, 5 m long with a 0.6-m spacing between rows, were planted with 450 seeds. Plots were end trimmed to 4 m prior to harvest.

Data were recorded on the center two rows of each plot for maturity (days after 31 August when 95% of pods were brown), and plant height, measured from the soil surface to the tip of the main stem, in centimeters. The center two rows of each plot were harvested and seed yield recorded in grams per plot. From each replication, a 25-g sample of seed was analyzed for protein and oil concentration, using infrared transmission, by the USDA-ARS Natl. Center for Agric. Utilization Res. at Peoria, IL. Analyses of variance were calculated in which entries, or progenies, and years were considered random effects. In the analysis of variance, mean squares for entries were partitioned to determine variability between and within determinate and indeterminate genotypes. Bartlett's test for homogeneity of variances was used to compare variances for seed protein among all $F_{4:5}$ breeding lines vs. those included in these tests. Covariance analyses were computed to determine if there were relationships between plant maturity and either seed yield or seed protein concentration in each of the populations. Regression analyses, based on mean values across replications and years, were used to evaluate relationships between specific traits for both determinate and indeterminate entries. Approximate processed values (APV) for specific genotypes were calculated based on the formula developed by W.D. Hanson (Leffel, 1990). Dollar values for oil and meal in this formula were the October 1993 to September 1994 average monthly price of \$602 t^{-1} for oil and \$183 t^{-1} for 440 g kg^{-1} soybean meal.

RESULTS AND DISCUSSION

Both F-tests and Bartlett's tests for homogeniety of variances were used to compare variances among progenies for seed protein concentration for the population of $F_{4:5}$ progeny rows vs. the $F_{4:5}$ progeny rows selected for inclusion in these tests (Table 1). The F-tests showed no significant differences between variances of the population vs. the selections for either determinate or indeterminate progenies in the two crosses. None of the chi-square values from the Bartlett's Test exceeded the critical chi-square values for homogeneity, indicating variances for the population and the selections were homogeneous.

Determinate progenies did not differ significantly from indeterminate progenies in mean days to maturity or seed protein concentration in cross CX1314 (Table 2). Determinate progenies averaged shorter in mature plant height, as would be expected, and lower in both seed yield and seed oil concentration than indeterminate progenies. Mean seed yields and seed protein concentrations of determinate and indeterminate progenies of the CX1412 cross were not significantly different. In this cross, determinate progenies averaged a significant 3 d earlier in maturity, were shorter in mature plant height, and lower in seed oil concentration than indeterminate progenies. Similar results were reported by Hicks et al. (1969). In their study, determinate and indeterminate isolines of Clark and Harosoy soybean were similar in seed yield, averaged across several plant populations during a 2-yr period.

There was significant variability among both determinate and indeterminate entries for each trait measured in both crosses. The range of variability for each trait among determinate and indeterminate progenies of the two crosses is shown in Table 2. The range in maturity was 13 and 14 d among determinate and indeterminate progenies of CX1314 and 14 and 19 d among determinate and indeterminate progenies of CX1412. Covariance

Table 1. Results of Bartlett's test for homogeneity of variances for seed protein among all determinate and indeterminate $F_{4:5}$ progenies and selected $F_{4:5}$ progenies from two soybean crosses.

Cross and population	No.	Mean	Variance	Chi-square	Critical chi-square
		g kg^{-1}			
CX1314 Determinate					
Population	56	438	484.55	0.3665	3.84
Selections	36	439	582.39		
CX1314 Indeterminate					
Population	141	437	507.37	0.0815	3.84
Selections	40	441	471.38		
CX1412 Determinate					
Population	58	445	360.65	0.9423	3.84
Selections	26	443	498.18		
CX1412 Indeterminate					
Population	108	448	401.54	0.9423	3.84
Selections	20	452	642.67		

Table 2. Means and ranges, averaged across three replications in each of two years, for characteristics of determinate and indeterminate soybean strains from two crosses.

	Determinate		Indeterminate	
Population and trait	Mean	Range	Mean	Range
CX1314				
Maturity, mo.-day	9-22 NS	9-17 to 9-30	9-23	9-18 to 9-29
Plant height, cm	70**	64 to 76	102	86 to 111
Yield, kg ha^{-1}	3010**	2579 to 3317	3308	2960 to 3528
Protein, g kg^{-1}	453 NS	420 to 486	452	417 to 500
Oil, g kg^{-1}	174**	151 to 195	186	178 to 291
CX1412				
Maturity (mo.-day)	9-23**	9-17 to 10-1	9-26	9-17 to 10-6
Plant height (cm)	73**	67 to 81	115	105 to 121
Yield (kg ha^{-1})	3408 NS	3133 to 3766	3474	3118 to 3870
Protein (g kg^{-1})	440 NS	407 to 485	436	405 to 470
Oil (g kg^{-1})	180**	161 to 199	189	174 to 208

**, NS Differences between means of determinate and indeterminate lines significant at the 0.01 probability level and nonsignificant, respectively.

analyses between the traits plant maturity and both seed yield and seed protein concentration demonstrated no significant effects of plant maturity on these two traits.

The regression of seed yield on plant height was not significant for either determinate or indeterminate progenies in these two crosses (Table 3). This demonstrated that the significant variability in seed yield among progenies within plant types was not associated with differences in plant height. This is in contrast to data of Wilcox and Sedlyama (1981) who reported that the rate of increase of seed yield per unit of plant height was three times greater among determinate than among indeterminate progenies from a single soybean cross.

The regression of seed protein concentration on seed yield was not significant in determinate progenies of either cross (Table 3). This indicates that among determinate progenies, lines could be identified combining high seed yield and high seed protein. The absence of a negative association between seed yield and seed protein among progenies that varied this widely in seed protein concentration is a unique finding. There was a significant negative association between seed yield and seed protein among the indeterminate progenies of these two crosses. Among the indeterminate progenies, seed protein concentration decreased 5 g kg^{-1} for every 100 kg ha^{-1} increase in seed yield. This inverse relationship between seed yield and seed protein is typical for progenies of most crosses between parents differing widely in seed protein and has limited progress in developing high-yielding soybean cultivars that have high protein concentration (Burton, 1985).

The inverse relationship, as measured by regression coefficient, between seed oil and protein concentration varied from -1.39 to -1.95 among determinate and indeterminate progenies of these crosses (Table 3). The strong inverse relationship between seed protein and oil is consistent with data from virtually all soybean populations (Burton, 1985). Slopes of the regression lines in both crosses were significantly greater for the indeterminate than for the determinate progenies, indicating a greater reduction in seed protein per unit increase in seed oil for the indeterminate progenies.

Table 3. Slopes (b) and standard errors (SE) for dependent vs. independent variables in determinate and indeterminate progenies from two soybean crosses.

	Determinate		Indeterminate	
Cross and traits	b	SE	b	SE
CX1314				
Yield (kg ha^{-1}) vs. plant height (cm)	7.78 NS	8.78	2.12 NS	4.78
Protein (g kg^{-1}) vs. yield (kg ha^{-1})	0.01 NS	0.02	−0.05*	0.02
Protein (g kg^{-1}) vs. oil (g kg^{-1})	−1.39**	0.13	−1.65**	0.09
CX1412				
Yield (kg ha^{-1}) vs. plant height (cm)	1.90 NS	10.24	0.74 NS	13.28
Protein (g kg^{-1}) vs. yield (kg ha^{-1})	−0.03 NS	0.02	−0.05**	0.01
Protein (g kg^{-1}) vs. oil (g kg^{-1})	−1.48**	0.17	−1.95**	0.21

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

Entries with superior yield and seed protein concentration were identified among both determinate and indeterminate progenies of each cross. In addition, progenies differed in protein production per unit area that was calculated for all progenies of each cross (Table 4). The three highest yielding determinate entries in CX1314 were similar in yield to lines with the highest seed protein and were not significantly different from the yield of the

Table 4. Characteristics of parents, check entries, and of the three determinate and indeterminate entries with the highest yield, protein concentration, or protein per unit area in each of two soybean crosses, averaged across three replications and two years.

Entry	Yield	Protein	Oil	Protein	APV†
	kg ha^{-1}	— g kg^{-1} —		kg ha^{-1}	\$ ha^{-1}
CX1314 Determinate					
Yield-156	3317	467	174	1550	864
-162	3271	478	155	1564	831
-177	3259	449	186	1464	848
Protein-148	2910	486	152	1414	742
-155	3184	486	163	1545	832
-194	3152	480	160	1511	811
Protein (kg ha^{-1})-162	3271	478	155	1564	831
-156	3317	467	174	1550	864
-155	3184	486	163	1545	832
Charleston	3574	420	197	1499	912
HC85-553-1	3156	414	210	1310	820
LSD 0.05	339	9	8	150	88
CX1314 Indeterminate					
Yield-39	3528	419	203	1476	909
-53	3525	429	196	1514	909
-11	3522	420	201	1484	908
Protein-2	3152	500	160	1577	835
-56	3374	486	163	1641	881
-91	2995	475	177	1423	793
Protein (kg ha^{-1})-56	3374	486	163	1641	881
-5	3454	472	176	1631	910
-95	3401	474	177	1615	900
Edison	3798	414	199	1571	965
CX1038-63	3128	477	164	1496	808
LSD 0.05	303	8	7	138	81
CX1412 Determinate					
Yield-123	3766	407	190	1533	929
-135	3698	451	178	1668	948
-145	3687	431	191	1590	944
Protein-148	3144	485	161	1524	817
-151	3280	482	161	1580	848
-149	3406	472	169	1613	886
Protein (kg ha^{-1})-135	3698	451	178	1668	948
-141	3583	459	161	1648	898
-149	3406	472	169	1613	886
Charleston	3721	416	196	1547	942
LSD 0.05	253	8	7	116	
CX1412 Indeterminate					
Yield-60	3902	418	199	1633	998
-21	3870	435	191	1684	996
-55	3684	430	188	1583	935
Protein-9	3131	470	178	1471	824
-45	3355	466	174	1565	872
-61	3496	460	180	1608	911
Protein (kg ha^{-1})-21	3870	435	191	1684	996
-60	3902	418	199	1633	998
-61	3496	460	180	1608	911
Edison	3600	402	204	1446	909
Flyer	3768	418	201	1576	967
Resnik	3305	414	205	1368	850
CX1039-99	3149	465	177	1470	822
LSD 0.05	264	7	6	111	68

† APV = approximate processed value based on the Oct. 1993–Sept. 1995 average monthly price of \$602 t^{-1} for oil and \$183 t^{-1} for 440 g kg^{-1} soybean meal.

determinate cultivar, Charleston. However, yields of the three high protein entries were all significantly lower than the yield of Charleston. Protein production per unit area was similar for all entries with superior yield, protein concentration, or protein per hectare. Of the three lines with high protein per hectare, two were among the highest yielding lines and one was among the three lines with high protein. The two highest yielding lines, CX1314-156 and -162 had seed protein concentrations of 467 and 478 g kg^{-1}, respectively. The cultivar Charleston had the highest APV of any of the determinate entries. The three entries with the highest yield and the highest protein per hectare had APV that did not differ significantly from Charleston.

Similar results were observed for the three indeterminate strains from cross CX1314 selected for superior yield, protein concentration, and protein production per unit area (Table 4). High yielding entries were similar in yield to the cultivar Edison, but the three strains with the highest protein were lower in seed yield than Edison. All high yielding indeterminate strains were significantly lower in protein than the high yielding determinate strains of the cross. None of the three indeterminate lines that produced the highest protein per hectare were among the highest yielding lines. CX1314-56, however, was among the three indeterminate entries with the highest seed protein and the three entries that produced the highest protein per unit area. The three highest yielding strains and two strains with the highest protein per hectare did not differ significantly in APV from the cultivar Edison.

High yielding, determinate entries of CX1412 were similar in yield to Charleston, and two of these averaged significantly higher seed protein than Charleston. The three determinate lines with the highest seed protein were lower in seed yield than Charleston. One of the lines producing the highest amount of protein per hectare was among the highest yielding lines and one was among the highest protein lines. Only one line from these two crosses, CX1412-135, produced significantly more protein per hectare than the high yielding check cultivar, Charleston. This line combined high seed yield with 451 g kg^{-1} seed protein. The three high yielding lines and high protein per hectare lines did not differ significantly in APV from Charleston.

Among the indeterminate progenies of cross CX1412, the highest yielding lines were similar in yield to the cultivar Flyer and two of the three were significantly higher in seed protein than Flyer (Table 4). The three entries with the highest seed protein were lower in seed yield than Flyer, but CX1412-61 was similar in yield to Edison and Resnik. Two of the entries producing the highest protein per hectare were among the highest yielding lines; one was among the highest protein lines. Two of the high yielding lines were among the three highest protein per hectar lines and had APV values significantly greater than Edison but not significantly different from Flyer.

The variability in APV and the LSD values for APV are determined by the variability in protein and oil among entries, years, and replications. Constants in the calculation of APV are the price of soybean oil at \$602 t^{-1} and of soybean meal at \$183 t^{-1}.

The data demonstrate that in these two crosses, determinate progenies were a better source of genotypes that combined high seed protein with high seed yield than were indeterminate progenies. The commonly reported inverse relationship between seed yield and seed protein did not exist among the determinate progenies but was present among the indeterminate progenies (Table 3). The three highest yielding determinate selections from CX1314 were similar in yield to Charleston and had seed protein concentrations of 449 to 478 g kg^{-1} (Table 4). The determinate selection CX1412-135 was similar in yield to Charleston and averaged 451 g kg^{-1} seed protein. None of the high yielding indeterminate selections in either cross averaged more than 435 g kg^{-1} seed protein. The successful development of determinate soybean strains that combine high seed yield and high seed protein concentration indicates that determinate selections may offer the best opportunities to develop high yielding cultivars with seed protein concentrations in excess of 450 g kg^{-1}.

ACKNOWLEDGMENTS

The authors acknowledge the support of the Indiana Soybean Development Council who provided part of the funding for this research.

REFERENCES

Bernard, R.L. 1972. Two genes affecting stem termination in soybeans. Crop Sci. 12:235–239.

Burton, J.W. 1985. Breeding soybeans for improved protein quantity and quality. p. 361–367. *In* R. Shibles (ed.) Proc of the World Soybean Res. Conf. III. Ames, IA. 12–17 Aug. 1984. Westview Press, Inc., Boulder, CO.

Cooper, R.L. 1985. Breeding semidwarf soybeans. p. 289–311. *In* J. Janick (ed.) Plant Breed. Rev. Vol. 3. AVI, Westport, CT.

Cooper, R.L., R.J. Martin, S.K. St. Martin, A. Calip-DuBois, R.J. Fioritto, and A.F. Schmitthenner. 1995. Registration of 'Charleston' soybean. Crop Sci. 35:593.

Escalante, E.E., and J.R. Wilcox. 1993. Variation in seed protein among nodes of determinate and indeterminate soybean near-isolines. Crop Sci. 33:1166–1168.

Hartwig, E.E., and K. Hinson. 1972. Association between chemical composition of seed and seed yield of soybeans. Crop Sci. 12: 829–830.

Hicks, D.R., J.W. Pendleton, R.L. Bernard, and T.J. Johnston. 1969. Response of soybean plant types to planting patterns. Agron. J. 61:290–293.

Leffel, R.C. 1990. Economic models and breeding strategies for soybean improvement. J. Prod. Agric. 3:582–586.

Sebern, N.A., and J.W. Lambert. 1984. Effect of stratification for percent protein in two soybean populations. Crop Sci. 24:225–228.

USDA. 1995. Oilseeds: World market & trends. Circular Ser. FOP 11-95. Foreign Agricultural Service, Washington, DC.

Wehrmann, V.K., W.R. Fehr, S.R. Cianzio, and J.F. Cavins. 1987. Transfer of high seed protein to high-yielding soybean cultivars. Crop Sci. 27:927–931.

Wilcox, J.R. 1988. The uniform soybean tests northern states 1988. USDA-ARS and Dep. of Agronomy, Purdue Univ., West Lafayette, IN.

Wilcox, J.R., and T. Sediyama. 1981. Interrelationships among height, lodging and yield in determinate and indeterminate soybeans. Euphytica 30:323–326.