Published March, 1991

Yield and Composition of Soybean Seed from Parents with Different Protein, Similar Yield

E. E. Hartwig* and T. C. Kilen

ABSTRACT

Soybean [*Glycine max* (L.) Merr.] seed yields often show a negative association with seed protein. The objective of this study was to determine if this association could be minimized by crossing a high and a normal protein line of equal seed yield. An F_2 population of 1000 plants was grown. Plants were harvested individually and seed evaluated for oil content using the nuclear magnetic resonance technique. Two bulk populations were developed: one included the 8% having the highest oil and the other the 8% having lowest oil. With the high negative correlation between seed protein and oil, the low oil population was expected to provide lines having high seed protein concentration. The two populations were advanced in bulk through the F_6 generation, after which 200 plants were harvested individually from each population and 200 F_7 lines grown from each. The F_7 lines were harvested individually and seed analyzed for seed protein and oil. The 18 lines having the highest protein and the 18 lines having the highest oil were selected for further evaluation in replicated trials. Lines were evaluated in five environments for seed yield, protein, and oil. The overall mean seed yield of the high protein lines averaged 6% less than the overall mean of the lines selected for high oil; however, when seed yield of the two highest protein lines at each environment were compared with the two highest oil lines, the high protein lines had a 1% advantage in seed yield, an 18% advantage in seed protein, and a 20% reduction in oil. The results demonstrate the potential for developing soybean lines high in seed protein and equal in seed yield to lines having high oil content.

SOYBEAN is grown primarily for its seed protein and oil. Seed oil concentration was emphasized in several of the earlier cultivar releases in the USA. These include 'Lincoln' (14), 'Roanoke' (3), and 'Dorman' (4), released in 1943, 1946, and 1952, respectively. Seed protein of soybean cultivars currently grown in the USA is 405 g kg^{-1} and oil is 210 g kg^{-1} on a dry-matter basis.

Soybean is unique as a producer of high-quality vegetable protein, but has many competitors as producers of vegetable oil. During the 5-yr period 1965 to 1969 the ratio of price per unit of oil to protein was 10:9. This ratio had shifted to 10:11.6 for the 5-yr period 1985 to 1989. In determining price per unit of protein we have attributed the entire value of soybean meal to the protein. Soybean has been priced on the basis of average content of oil and protein. Situations have occurred where soybean produced in specific regions has been discounted in the market because of low oil content. No situation has developed where premiums have been paid for higher protein soybean. Technology has been developed that makes rapid determination of soybean seed composition possible. Based on present trends, soybean cultivars having higher seed protein would have greater value per unit than typical cultivars now in production. A high negative correlation (−0.80 to −0.88) exists between seed protein and oil.

The seed protein of soybean lines within the southern U.S. germplasm collection ranges from 360 to 500 g kg^{-1} and oil ranges from 150 to 240 g kg^{-1} (11), based on dry matter. Several studies (13,1) have reported a negative correlation between seed yield and seed protein. Nonadapted germplasm lines were the contributors of genes for higher protein. Many of the germplasm lines are primitive types having poor agronomic characteristics and prone to shatter at maturity. Seed yield for germplasm lines of Maturity Groups V and VI evaluated at Stoneville, MS (5) averaged 65% of that for adapted cultivars. Hartwig (5) concluded that when a nonadapted germplasm line was used as a parent, regardless of the character to be transferred, two backcrosses to an adapted type were needed for obtaining a moderate number of lines with the desired character and yield approaching the adapted parent.

Productive cultivars or lines have been developed that have a moderate increase in seed protein. 'Tracy' (6), released in 1973, has produced slightly greater seed yields than 'Davis' (2) [3-yr average 1973 Uniform Soybean Tests, Southern States (11)] and 435 g kg^{-1} protein vs. 395 g kg^{-1} for Davis. Tracy was released primarily for its productivity and high level of resistance to phytophthora rot caused by *Phytophthora megasperma* Drechs. f. sp. *glycinea* Kuan & Erwin. Hartwig and Hinson (10) developed backcross-2 lines having 'Bragg' (12) as the recurrent parent, which equaled Bragg in seed yield and had a seed protein concentration of 435 g kg^{-1} in comparison with 404 g kg^{-1} for Bragg.

This study was conducted to determine relative seed yield of soybean lines differing in seed protein concentration when they were developed from parents having nearly similar seed yields.

MATERIALS AND METHODS

Parents for the study were 'Bedford' (9) and D77-5769. These lines were considered to be similar in seed yield, but distinctly different in seed composition (Table 1). Bedford, which was selected for resistance to soybean cyst nematode,

E.E. Hartwig and T.C. Kilen, USDA-ARS, Soybean Prod. Res. Unit, P.O. Box 196, Stoneville, MS 38776. Joint contribution from USDA-ARS and Mississippi Agric. and For. Exp. Stn. Received 4 May 1990. *Corresponding author.

Published in Crop Sci. 31:290–292 (1991).

Table 1. Seed yield, seed composition, and maturity for Bedford and D77-5769 soybean.

Location	Strain	Seed yield	Seed Protein	Seed Oil	Maturity from Bedford
		kg ha^{-1}	g kg^{-1}		d
Stoneville 1978	Bedford	2775	397	189	
	D77-5769	2648	473	162	+2
	LSD (0.05)	538			
Regional Preliminary VI 1979†	Bedford	2406	392	192	
	D77-5769	2420	463	166	+1
	LSD (0.05)	235	13	6	

† Preliminary Group VI nursery was grown at Holland, VA; Plymouth, NC; Portageville, MO; Keiser, AR; Stoneville, MS (two soil types); Belle Mina AL; and Jay, FL.

Heterodera glycines Ichinohe, Races 3 and 4, has an average seed composition representative of cultivars currently in production. Line D77-5769 (Fig. 1) was developed from a multiple crossing program to develop a high protein line with the desired agronomic characteristics, including nonshattering at maturity and resistance to the diseases bacterial pustule caused by *Xanthomonas campestris* pv. *glycines* (Nakano) Dye, and phytophthora rot.

In the initial crosses, germplasm lines having a high concentration of seed protein were crossed with adapted lines. A pedigree breeding program was utilized to develop advanced F_5 lines that were evaluated in replicated plantings for seed yield and seed composition. Selection in the F_3 and F_4 generations was for plant type, resistance to diseases, and seed holding. Lines selected as parents for second-cycle crosses had 450 g kg^{-1} seed protein or higher, moderate seed holding, resistance to bacterial pustule, field resistance to phytophthora rot, and the highest seed yield possible. Selection procedures were similar in each later cycle.

One of the parents in the development of D77-5769 was D61-3505 (protein 451 g kg^{-1}, oil 181 g kg^{-1}), a selection from D49-2491^6 × PI174862. The inheritance of seed protein concentration is not fully understood. In our studies we found the seed protein concentration of seed from F_1 plants to be at the parental midpoint. Line D61-3505 was developed using a backcrossing program in which we assumed three loci, and thus the F_1 plant would have eight gametic combinations for seed protein concentration. After BC1, we used 24 BC F_1 plants as male parents. Seed composition was determined from each plant. The F_1 plants for the next cycle were grown from the male most closely approximating the original F_1 plant. The segregation of genes for seed composition, along with segregation among genes contributing to seed yield, indicates the complexity of developing highly productive germplasm appreciably higher in seed protein concentration.

One thousand F_2 plants were harvested individually from the cross Bedford × D77-5769. Seed from individual plants was analyzed for oil concentration using a nondestructive nuclear magnetic resonance (NMR) procedure. Selection for low oil was considered an adequate procedure for high seed protein selection, based on the high negative correlation between seed protein and oil. Seed from the 8% of the plants with the highest oil content were bulked to make up one composite and seed from the 8% with the lowest oil content made up a second composite. The two composite populations were advanced in bulk in the F_3, F_4, F_5, and F_6 generations. Two hundred F_7 lines were grown from each bulk population, harvested individually, and analyzed for seed protein and oil. The 18 lines having highest protein concentration and the 18 lines having the highest oil concentration were selected for evaluation in replicated trials. Each strain was grown in four-row plots 6 m in length, with 5 m of the two center rows harvested. Rows were spaced 90 cm apart. Data were obtained for seed yield, protein, and oil. Concentration of seed protein and oil was determined using near infrared reflectance (NIR) equipment. Data were analyzed by analyses of variance. The soybean lines were evaluated in five environments, labeled A to E, respectively: Stoneville, MS, clay soil, 1987; Stoneville, loam, 1988; Stoneville, clay, 1988; Raymond, MS, 1988; and Plymouth, NC, 1988.

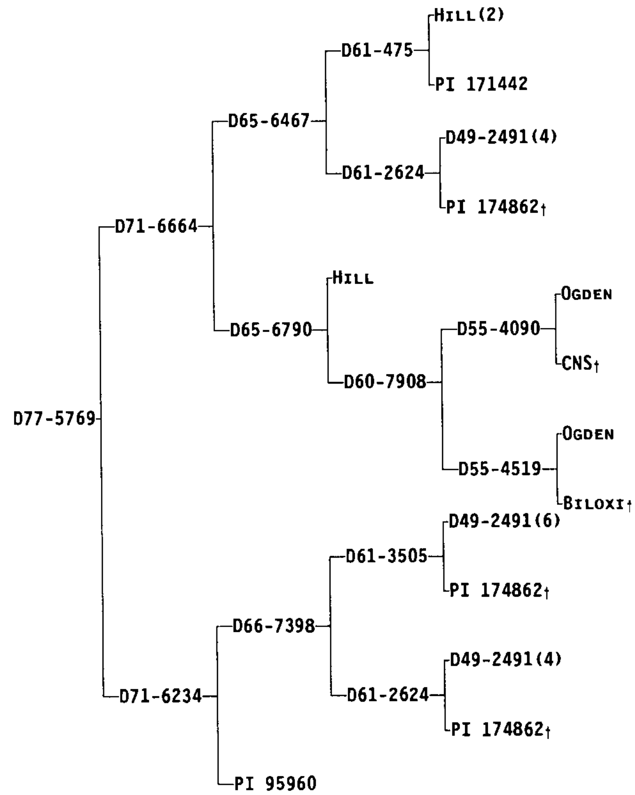

Fig. 1. Pedigree of D77-5769 showing step-wise incorporation of resistance to pest problems and high protein to develop a productive line. † Sources of genes for hybrid seed protein concentration.

RESULTS AND DISCUSSION

The mean seed yield and seed composition for the 18 high protein lines and the 18 high oil lines grown in five environments are reported in Table 2. The high protein lines had a mean seed yield 94% of that for the high oil lines. This seed yield advantage was consistent across environments. The difference was significant ($P \leq 0.05$) in four environments, but was not in Environment D. Environment C, Stoneville clay, 1988, had the lowest protein concentration. This planting was made 10 June. All other plantings were made in mid-May. The high oil lines also had the lowest protein content in Environment C. As a mean of the five environments, the protein lines had a 12% (49 g kg^{-1}) advantage in protein and a 14% (29 g kg^{-1}) reduction in oil. Thus, the ratio of protein gain to oil reduction was 1.7:1. The correlation coefficient for seed yield and protein concentration in Environment B was −0.21.

Table 3 reports the seed yield and seed composition for the two lines having the highest protein concentration in each environment and for the two lines hav-

Table 2. Mean seed yield and seed composition for 18 high protein lines and 18 high oil lines selected from the cross Bedford × D77-5769 evaluated at five environments.

	High protein lines			High oil lines		
		Seed			Seed	
Environment	Seed yield	Protein	Oil	Seed yield	Protein	Oil
	kg ha^{-1}	—— g kg^{-1} ——		kg ha^{-1}	—— g kg^{-1} ——	
A†	2413‡	460	172	2554	408	202
B	2695	450	174	2835	404	204
C	2580	436	181	2762	383	216
D	2204	474	186	2325	424	217
E	2614	462	172	2883	419	192
Mean	2501	457	177	2672	408	206
percent, in relationship to high oil lines	94	112	86	100	100	100

† A = Stoneville, MS, clay, 1987; B = Stoneville, MS, loam, 1988; C = Stoneville, MS, clay, 1988; D = Raymond, MS, 1988; E = Plymouth, NC, 1988.

‡ Mean seed yields of high protein and high oil lines differ significantly at the 0.05 level at Environments A, B, C, and E but not at D based on a LSD. Mean protein and oil levels difference exceed the 0.01 level of confidence in each environment based on an F-test.

Table 3. Seed yield and composition for the two highest protein lines and two highest oil lines from the cross Bedford × D77-5769 when grown at five environments.

	High protein lines			High oil lines		
		Seed			Seed	
Environment	Seed yield	Protein	Oil	Seed yield	Protein	Oil
	kg ha^{-1}	—— g kg^{-1} ——		kg ha^{-1}	—— g kg^{-1} ——	
A†	2514[2]‡	472	160	2513[7]	390	209
	2648[4]	480	162	2508[11]	403	206
B	2782[2]	463	169	2782[8]	396	211
	2715[4]	463	172	2785[10]	400	210
C	2769[1]	450	178	2736[8]	374	231
	2836[4]	454	173	2762[10]	360	229
D	2058[1]	484	179	1970[8]	421	231
	2224[6]	484	172	2204[10]	424	222
E	2776[5]	474	167	2998[8]	408	203
	2850[3]	471	168	2768[9]	404	200
Mean	2617	469	172	2598	398	215
Percent	101	118	80	100	100	100

† A = Stoneville, MS, clay, 1987; B = Stoneville, MS, loam, 1988; C= Stoneville, MS, clay, 1988; D = Raymond, MS, 1988; E = Plymouth, NC, 1988.

‡ Soybean lines: [1]D86-25, [2]D86-34, [3]D86-57, [4]D86-102, [5]D86-182, [6]D86-195, [7]D86-221, [8]D86-260, [9]D86-285, [10]D86-324, [11]D86-37.

LSD within environments and for the means, seed yield differences between high protein and high oil levels are NS; protein and oil differences are significant at 0.01 level of confidence, based on LSD.

ing the highest oil concentration in each environment. Mean seed yield for the two groups across the five environments was nearly similar. The high protein lines had a 1% advantage in seed yield, an 18% (71 g kg^{-1}) advantage in seed protein and a 20% (43 g kg^{-1}) reduction in oil in comparison with the high oil lines.

Line D86-102 ranked highest in protein concentration in three environments and had the highest overall mean. Across the five environments its mean seed yield was 2553 kg ha^{-1}, protein 466 g kg^{-1}, and oil 174 g kg^{-1}. Line D86-260 ranked highest in oil concentration in four of the environments and had the higher overall mean. For the five environments, its seed yield was 2494 kg ha^{-1}, seed protein 401 g kg^{-1} and oil 213 g kg^{-1}. The high protein line, D86-102, had a 2.4% advantage in seed yield, a 16% (65 g kg^{-1}) advantage in seed protein and an 18% (39 g kg^{-1}) reduction in oil in comparison with D86-206 (differences in seed yield NS, $P < 0.05$; differences in protein and oil concentration significant at $P = 0.01$ level).

The performance of D86-102 and D86-260 demonstrates that when using parents of nearly similar seed yielding capacity, but distinctly different seed composition, it is possible to develop lines similar in seed yield, but distinctly higher in seed protein and lower in oil.

Line D86-102 has a 16% greater concentration of seed protein than D86-260. This increase in seed protein concentration is substantially greater than the 8% increase in seed protein concentration that the cultivar Tracy or its subline 'Tracy-M' (7) (tolerant to the herbicide metribuzin [4-amino-6-(1,1-dimethylethyl)-3-(methylthio)-1,2,4-triazin-5(4*H*)-one]) have compared with Davis. We conclude that should a marketing strategy develop favoring soybean seed having a high concentration of seed protein, it would be possible to supply this market with highly productive cultivars.

ACKNOWLEDGMENTS

We appreciate the assistance of J.W. Burton and Joe E. Askew in managing the plantings at Plymouth, NC, and Raymond, MS, respectively.

REFERENCES

1. Burton, J.W. 1985. Breeding soybeans for improved protein quantity and quality. p. 361–367. *In* R. Shibles (ed.) 3rd Soybean Res. Conf., Ames, IA. 12–17 Aug. 1984. Westview Press, Boulder, CO.
2. Caviness, C.E., and H.J. Walters. 1966. Registration of Davis soybeans. Crop Sci. 6:502.
3. Hartwig, E.E. 1946. Roanoke, a new soybean for North Carolina. Information sheet. Feb.
4. Hartwig, E.E. 1952. Dorman, a new soybean variety for the Mid-South. Miss. Farm Res. 15:1–7.
5. Hartwig, E.E. 1972. Utilization of soybean germplasm strains in a soybean improvement program. Crop. Sci. 12:856–859.
6. Hartwig, E.E. 1974. Registration of Tracy soybeans. Crop Sci. 14:777.
7. Hartwig, E.E., W.L. Barrentine, and C.J. Edwards, Jr. 1980. Registration of Tracy-M soybeans. Crop Sci. 20:825.
8. Hartwig, E.E., and C.J. Edwards, Jr. 1986. Evaluation of soybean germplasm maturity groups V–X. USDA-ARS report.
9. Hartwig, E.E., and J.M. Epps. 1978. Registration of Bedford soybeans. Crop Sci. 18:915.
10. Hartwig, E.E., and K. Hinson. 1972. Association between chemical composition of seed and seed yield. Crop Sci. 12:829–830.
11. Hartwig, E.E., and K.W. Jamison. 1973. The Uniform Soybean Tests, Southern States. USDA-ARS, Delta Branch Exp. Stn., Stoneville, MS.
12. Hinson, K., and E.E. Hartwig. 1964. Registration of Bragg and Hardee soybeans. Crop Sci. 4:664.
13. Shannon, G., Jr., J.R. Wilcox, and A.L. Probst. 1972. Estimated gains from selection for protein and yield in the F_4 generation of six soybean populations. Crop Sci. 12:824–826.
14. Weber, C.R. 1944. Lincoln, a new variety high in yield and oil content. Soybean Dig. 4:6–7.