

- Veldboom, L.R. 1994. Genetic analysis of inbred and topcross progeny of an elite, single-cross maize population. Ph.D. diss. (Diss. Abstr. 95-18451). Iowa State Univ., Ames.
- Veldboom, L.R., M. Lee, and W.L. Woodman. 1994. Molecular marker-facilitated studies in an elite maize population: I. Linkage analysis and determination of QTL for morphological traits. *Theor. Appl. Genet.* 88:7-16.
- Veldboom, L.R., and M. Lee. 1996a. Genetic mapping of quantitative trait loci in maize in stress and nonstress environments: I. Grain yield and yield components. *Crop Sci.* 36:1310-1319.
- Veldboom, L.R., and M. Lee. 1996b. Genetic mapping of quantitative trait loci in maize in stress and nonstress environments: II. Plant height and flowering. *Crop Sci.* 36:1320-1327.
- Zeng, Z.B. 1994. Precision mapping of quantitative trait loci. *Genetics* 136:1457-1468.
- Zuber, M.S., and L.L. Darrah. 1980. 1979 U.S. germplasm base. p. 234-249. *In* Proc. of the 30th Annu. Corn and Sorghum Ind. Res. Conf., Am. Seed Trade Assoc., Chicago. 9-11 Dec. 1980. Washington, DC.

Developing High-Protein, High-Yield Soybean Populations and Lines

E. R. Cober* and H. D Voldeng

ABSTRACT

Developing high-yielding, high-protein soybean [*Glycine max* (L.) Merr.] lines is difficult because of the inverse relationship between seed yield and seed protein content. The objective of this study was to evaluate single cross and rapid back cross breeding methods to achieve both high seed yield and high protein content. 'AC Proteus' was used as the high-protein source and 'Maple Glen' was used as the adapted high-yielding parent. Six single plants were randomly selected from each of 149 F₃ progeny rows to develop 886 single cross derived lines. Reciprocal back cross populations were made using 60 F₁ plants from the single cross and the high-yielding parent, Maple Glen. Ten single plant selections were randomly taken from each of 80 F₁-derived families to develop 800 back cross derived lines. In 1994, all lines were tested in the field. About 20% of lines were retained for testing through this breeding project from 1994 to 1996 although selection intensity differed across populations and years. In 1997, six single cross and nine back cross lines were tested at six locations in Eastern Canada. The seed yield and protein content of the single cross lines were not significantly different from the back cross derived lines. All the selected lines had higher seed protein content than did Maple Glen. Both breeding strategies produced lines with significantly higher seed protein content than Maple Glen. However, none of these lines had significantly higher seed yield than Maple Glen or a recently released high protein cultivar, AC Proteina. These populations exhibited a very low or no association between seed yield and protein ($r = -0.06$ to -0.21). Therefore, the parents may be useful sources of alleles which do not exhibit the usual pleiotrophic effect of low seed yield and high seed protein. The use of a back cross to the adapted parent following the single cross was not beneficial in the development of high yielding lines.

THE DEVELOPMENT of high-protein soybean cultivars may improve the competitiveness of livestock producers if farm-produced, full-fat soybean can substitute for soybean meal. Dairy producers may even be able to include high-protein, uncooked soybean in feed rations without a production penalty (McNiven et al., 1994).

High-yielding, high-protein soybean cultivars are difficult to develop because of the inverse relationship between seed yield and seed protein content. Burton (1987) reported inverse relationships in all but two populations in his review. Wilcox and Cavins (1995) summa-

rized a number of breeding studies and reported a moderate to strong inverse relationship between seed yield and protein content ranging from $r = -0.23$ to -0.86 .

Sources of high protein used in breeding programs are typically unadapted plant introductions. A theoretical consideration on the introgression of favorable alleles from less adapted germplasm has been presented by Dudley (1982). He concluded that when one parent contains a larger number of favorable alleles, the use of a back cross to the adapted parent would generate enhanced populations. If the parents are diverse and the desired trait in the exotic parent is controlled by few loci, additional back crosses would be beneficial.

A number of studies of single and back crossing strategies to combine high seed yield and high protein have been completed in soybean. Thorne and Fehr (1970) concluded that 3-way crosses were superior to 2-way crosses. Hartwig and Hinson (1972) concluded that a second back cross further enhanced the probability of finding a high-yielding, high-protein line. Wehrmann et al. (1987) concluded that selection for high protein between the two back cross generations increased the seed protein content, in the resulting progeny, while maintaining seed yield. Wilcox and Cavins (1995) were able to develop a BC₃ line which had high seed protein and recovered the yield of the recurrent parent. It was necessary to select for high protein between each back cross and as a result it took over 20 yr to develop this line.

The objective of this study was to evaluate single cross and rapid back cross breeding methods to achieve both high seed yield and protein content.

MATERIALS AND METHODS

To produce the single cross population (X3585), the high-protein parent AC Proteus (Voldeng et al., 1996a) was crossed to the high-yield parent Maple Glen (Voldeng et al., 1996b) in 1991. A total of 60 F₁ seeds were produced. The F₂ generation was grown in a field (Grenville loam, coarse-loamy, mixed, frigid Typic Eutrochrepts) at the Central Experimental Farm (Ottawa, ON) and 149 single plant selections were randomly taken. In 1993, 149 F₃ progeny rows were grown and usually six single plants were randomly taken from each progeny row. In 1994, 886 F_{3/4} progeny rows were grown in modified augmented design (MAD; Lin and Poushinsky, 1985) trials with 4.5-m single row plots. The back cross was made in 1992 with the 60 F₁ plants from the single cross (X3585) and Maple Glen. Reciprocal crosses were made and 76 F₁ seeds were

Eastern Cereal and Oilseed Research Centre, Agric. & Agri-Food Canada, Ottawa, Ontario, Canada, K1A 0C6. ECORC Contribution No. 991410. Received 20 Jan. 1999. *Corresponding author (coberer@em.agr.ca).

Table 1. Agronomic performance, seed composition, and pedigrees of parent and check soybean cultivars from two trials grown at the Central Experimental Farm, Ottawa, ON, in 1996.

Cultivar	Yield	Maturity	Prot	Oil	Pedigree
	kg ha ⁻¹	d	g kg ⁻¹		
AC Proteus	3341	117	521	151	Merit/PI 153293/PI 189950/3/3*Maple Arrow
Maple Arrow	3730	119	441	205	Harosoy 63/PI 438477
AC Proteina	3668	119	498	168	AC Proteus/3/Weber//Maple Presto/Williams
Maple Glen	3610	121	457	201	PI 438477//Portage/Amsoy/3/Premier
LSD (0.05)	215	3	8	4	

produced for X3658 (X3585/Maple Glen) and 32 F₁ seeds were produced for X3659 (Maple Glen/X3585). BC₁-F₁ plants were grown in 1992 and 108 F₁-derived families were advanced through the F₂ and F₃ generations using single pod descent maintaining a minimum family size of 70 plants. In 1993, 108 F_{1,4} families were planted in single rows in the field. A number of families were segregating for dwarf plants and were discarded. A total of 51 families for X3658 and 29 families for X3659 were retained. Ten single plants were randomly taken from each F₁-derived family. In 1994, a total of 800 F_{4,5} progeny rows were grown in MAD trials with single row plots. The 1686 single and back cross derived lines were grown in MAD trials with five subplots per main plot, i.e., each main plot check was bordered by two experimental plots on each side. Eighty subcheck plots were grown.

Following the single row MAD trials grown in 1994, selection was made for high protein on a line basis in the single cross population and on an F₁-derived family basis in the back cross populations. In 1995, single- and back cross derived lines were grown in MAD trials, with four row 4.5-m plots, at two locations on the Central Experimental Farm with the same main plot size and proportion of subcheck plots as in 1994. High-protein lines were selected on a line basis and tested in 1996 in simple rectangular lattice trials grown at two locations on the Central Experimental Farm (Table 1). Following selection for high protein and high yield, six single- and nine back cross lines were tested in 1997 as part of simple lattice trials grown at Ottawa and Inkerman in eastern Ontario, at Listowel and Guelph in southwestern Ontario, and at St-Bruno and Acadie in Quebec. Maple Glen and AC Proteina, a recently released high-protein cultivar, were included in these tests. AC Proteina replaced AC Proteus as a check because of higher productivity. The set of lines tested in 1997 was analyzed as a randomized complete block by AGROBASE 98 (Agronomix Software Inc., Winnipeg, MB) with location, replicate, and line considered random effects. Seed protein and oil content were determined by near infrared transmittance. Protein yield was calculated from seed yield and seed protein content. Maturity was determined when 95% of pods had reached their mature pod color.

To investigate the relationship between seed yield and seed protein and oil, correlation coefficients were calculated for

the unselected lines grown in 1994 and the selected lines grown in 1996. The correlation coefficient was also calculated for seed oil and seed protein using the 1996 data. Estimates of genetic variance and broad sense heritability were made for the two populations with variance components computed from the multiple location trials in 1996 and 1997.

RESULTS AND DISCUSSION

Approximately 20% of the lines were advanced to the next level of testing each generation (Table 2). In 1996, 42 single cross and 42 back cross derived lines were tested in two lattice trials. Generally seed protein content was 20 to 30 g kg⁻¹ higher than usual for the parent and check cultivars (Table 1); however, the selected lines (Fig. 1) contained more seed protein than Maple Glen, the oilseed parent. The single cross lines yielded on average 3451 kg ha⁻¹ with seed protein content of 497 g kg⁻¹ which was not significantly different from the back cross mean yield of 3211 kg ha⁻¹ and seed protein content of 507 g kg⁻¹. In 1997, six single cross and nine back cross derived lines, tested at six locations, had seed protein significantly greater than Maple Glen, the adapted parent (Table 3). No significant differences between single and back cross derived lines were observed for seed yield or seed protein yield. There were both single cross and back cross derived lines that yielded similarly to Maple Glen while having

Table 2. The number of lines and families retained in testing for high protein and high yield in both single and back cross soybean populations.

Year	Single cross		Back cross			
	Lines	F ₂ -derived families	X3658†		X3659†	
			Lines	F ₁ -derived families	Lines	F ₁ -derived families
1994	886	149	510	51	290	29
1995	122	68	190	19	70	7
1996	42	28	19	7	23	5
1997	6	5	0	0	9	3

† X3658 = X3585/Maple Glen; X3659 = Maple Glen/X3585; X3585 = AC Proteus/Maple Glen.

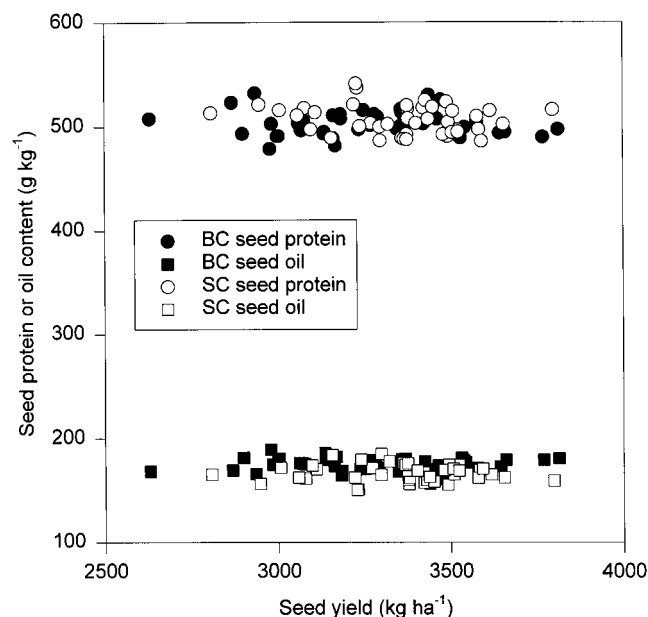


Fig. 1. Seed yield vs. seed protein and oil content for 42 single cross (SC) and 42 back cross (BC) derived soybean lines tested in two trials on the Central Experimental Farm, Ottawa in 1996.

Table 3. Agronomic performance of single cross (SC) and back cross (BC) soybean tested at six locations in Eastern Canada in 1997.

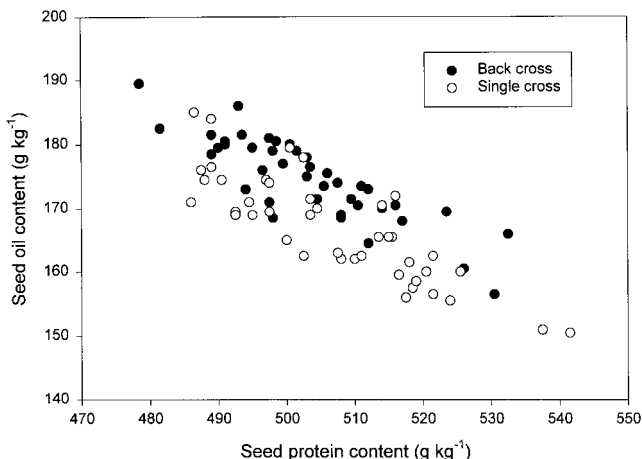
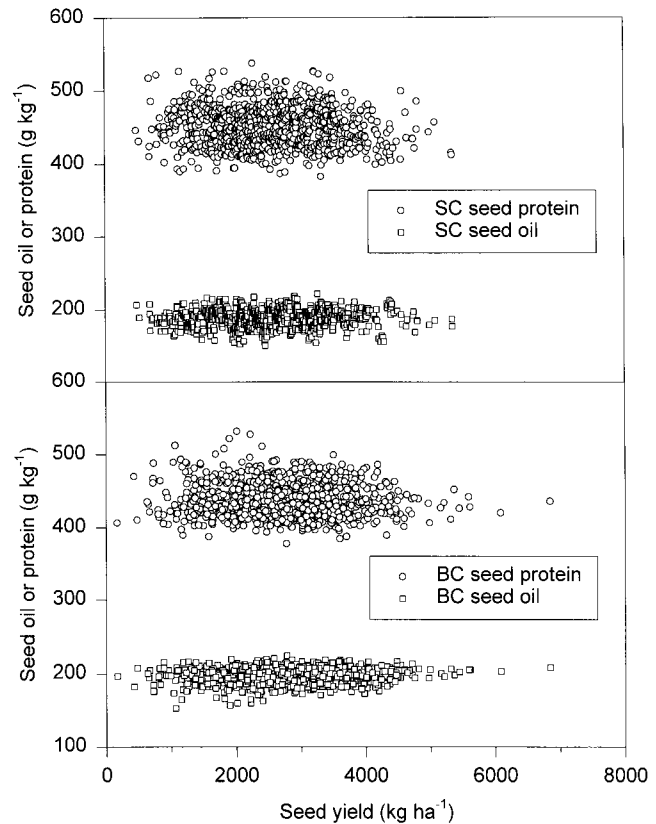
Genotype	Seed yield kg ha ⁻¹	Days to maturity	Seed		Protein yield kg ha ⁻¹
			Protein	Oil	
			g kg ⁻¹		
Maple Glen	3039	115	427	207	1297
AC Proteina	2995	113	478	172	1433
SC lines					
X3585-139-3-B	2243	110	462	176	1036
X3585-69-1-B	2389	112	487	166	1164
X3585-136-4-B	2607	113	459	177	1196
X3585-6-6-B	2656	116	491	169	1303
X3585-116-4-B	2758	116	480	170	1323
X3585-116-3-B	2869	118	501	161	1436
Mean	2587	114	480	170	1241
BC lines					
X3659-B-B-79-8	2451	110	458	186	1122
X3659-B-B-96-4	2475	112	458	189	1135
X3659-B-B-79-9	2757	112	458	191	1261
X3659-B-B-79-7	2599	113	456	190	1186
X3659-B-B-96-7	2617	114	478	171	1250
X3659-B-B-96-6	2869	114	457	183	1312
X3659-B-B-96-3	2968	114	460	181	1364
X3659-B-B-79-6	2534	114	471	178	1193
X3659-B-B-86-9	2475	116	461	187	1141
Mean	2660	114	468	179	1245
SC vs. BC, <i>P</i> value†	0.11	0.0001	0.0001	0.0001	0.18
LSD (0.05)‡	201	1.2	9	5	88

† Contrast comparing single and back cross derived lines.

‡ LSD (*P* = 0.05) value to compare any two entries within a column.

elevated seed protein content, however, none of these lines out yielded Maple Glen or the recently released high protein cultivar AC Proteina. Generally, single cross lines had higher seed protein and lower seed yield while an additional cross to the high yield parent, Maple Glen, resulted in higher seed yield and lower seed protein.

It does not appear that a back cross is necessary to develop high-yield, high-protein lines with these parents. This contrasts with the findings of previous studies (Thorne and Fehr, 1970; Hartwig and Hinson, 1972; Wehrmann et al., 1987), where a first or further back cross was found to be superior to single cross breeding schemes. The high protein source used in this study may

**Fig. 2. Seed oil vs. seed protein content for 42 single cross and 42 back cross derived soybean line tested in two trials on the Central Experimental Farm, Ottawa in 1996.****Fig. 3. Seed yield vs. seed protein and oil content for randomly selected lines from single cross (SC) and back cross (BC) derived soybean line grown in single-row plots on the Central Experimental Farm, Ottawa in 1994.**

account for the lack of agreement with previous reports. Our high-protein source, AC Proteus (Voldeng et al., 1996a), was a BC₂-derived line developed from a back crossing program which used 'Maple Arrow' as the recurrent parent (Table 1). AC Proteus maybe more adapted and contain more favorable alleles than is usual in studies of introgression of high protein. In our case, therefore, the use of further back crossing was less beneficial for the development of high protein high yield lines.

In the randomly selected lines tested in 1994 (Fig. 2), seed protein was negatively correlated with seed yield in both single cross ($r = -0.07$, $P = 0.010$, $n = 886$) and back cross derived lines ($r = -0.08$, $P = 0.016$, $n = 800$); however, the large number of lines may be largely

Table 4. Estimates of genetic variance and broad sense heritability for seed yield and seed protein in single cross and back cross soybean populations from 84 lines grown at two locations on the Central Experimental Farm, Ottawa in 1996 and 15 lines grown at six locations in Eastern Canada in 1997.

	Single cross population		Back cross population	
	1996	1997	1996	1997
Seed yield				
Genetic variance	11 630	45 880	18 650	28 050
Broad sense (h^2)	0.26	0.33	0.33	0.27
Seed protein				
Genetic variance	1.83	2.92	1.16	0.541
Broad sense (h^2)	0.77	0.74	0.65	0.40

responsible for results slightly but significantly different from zero. Using the 42 lines from each cross tested in 1996 (Fig. 1), we observed no significant relationship between seed yield and seed protein (single cross $r = -0.21$, $P = 0.176$; back cross $r = -0.06$, $P = 0.7$) or oil (single cross $r = -0.05$, $P = 0.8$; back cross $r = -0.02$, $P = 0.9$). The usual inverse relationship between seed protein and oil (Burton, 1984) was observed in these two populations (Fig. 3; single cross $r = -0.84$, $P = 0.0001$; back cross $r = -0.86$, $P = 0.0001$). It appears that, in these two populations, alleles which do not exhibit the usual pleiotropic effect of low seed yield and high seed protein content are present.

It may be hypothesized that the low association observed between seed yield and protein in this study results from a lack of genetic variance for yield in these crosses. This does not seem to be the case. Estimates of heritability for seed yield and protein (Table 4) are within the range reported by Burton (1987) even though the lines in these populations used for heritability estimates were selected for high protein and yield and were not random lines. This selection would be expected to result in lower estimates of genetic variance and heritability in comparison to the use of random lines.

In summary, the parents used to develop these populations or lines derived from these populations may be useful parents for breeding programs with breeding objectives which include high seed protein content and high yield. This study also indicates that a single cross breeding scheme may be as effective, in developing high-yield high-protein lines, as a breeding scheme with an additional back cross to a high-yielding parent when the high-protein parent is improved germplasm.

ACKNOWLEDGMENTS

We would like to thank Kemptville College, University of Guelph; Novartis Seeds; First Lines Seeds; Semences Prograin; and the Ministère de l'agriculture, des pêcheries et de l'alimentation du Québec for providing trial sites in 1997. We would like to thank Ron Guillemette, Andrew Bird, John Bowman and Gerry Petznick for excellent technical assistance.

REFERENCES

- Burton, J.W. 1984. 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.
- Burton, J.W. 1987. Quantitative genetics: Results relevant to soybean breeding. p. 211–247. *In* J.R. Wilcox (ed.) Soybeans: Improvement, production, and uses. 2nd ed. Agron. Monogr. 16. ASA, CSSA, and SSSA, Madison, WI.
- Dudley, J.W. 1987. Theory for transfer of alleles. *Crop Sci.* 22:631–637.
- Hartwig, E.E., and K. Hinson. 1972. Association between chemical composition of seed and seed yield of soybeans. *Crop Sci.* 12: 829–830.
- Lin, C.S., and G. Poushinsky. 1985. A modified augmented design (type 2) for rectangular plots. *Can. J. Plant Sci.* 65:743–749.
- McNiven, M.A., P.H. Robinson, and J.A. MacLeod. 1994. Evaluation of a new high protein variety of soybeans as a source of protein and energy for dairy cows. *J. Dairy Sci.* 77:2605–2613.
- Thorne, J.C., and W.R. Fehr. 1970. Incorporation of high-protein, exotic germplasm into soybean populations by 2- and 3-way crosses. *Crop Sci.* 10:652–655.
- Voldeng, H.D., R.J.D. Guillemette, D.A. Leonard, and E.R. Cober. 1996a. AC Proteus soybean. *Can. J. Plant Sci.* 76:153–154.
- Voldeng, H.D., R.J.D. Guillemette, D.A. Leonard, and E.R. Cober. 1996b. Maple Glen soybean. *Can. J. Plant Sci.* 76:475–476.
- 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., and J.F. Cavins. 1995. Backcrossing high seed protein to a soybean cultivar. *Crop Sci.* 35:1036–1041.