CROP BREEDING, GENETICS & CYTOLOGY

Interrelationships among Seed Quality Attributes in Soybean

James R. Wilcox* and Richard M. Shibles

ABSTRACT

Soybean [*Glycine max* (L.) Merr.] meal is used primarily as a livestock feed. The high protein concentration and sulfur-containing amino acids in the meal contribute to its nutritional value. Oligosaccharides, including raffinose and stachyose in the meal, have detrimental effects on the nutritive value of soy meal as a livestock feed. The objective of this research was to determine the interrelationships among seed protein, oil, oligosaccharides, and S in a breeding population that varied widely in seed protein concentration. Forty-three random breeding lines that varied in seed protein concentration from 413 to 468 g kg^{-1} on a dry seed basis, were grown in replicated tests in three environments. Seed yield, protein, oil, oligosaccharides, and S concentrations were determined for entries in each replication in the three environments. Breeding lines and environments varied significantly for each of the traits measured. Concentrations of carbohydrates were not associated with seed yield. Protein increased at the expense of oil ($b = -1.560$), total carbohydrates ($b = -0.171$), and sucrose ($b = -0.151$). Sulfur increased with increasing protein ($b = 0.008$), but S/N ratios were constant across protein concentrations. Decreases in carbohydrates with increases in protein would contribute to increased nutritional value of the meal from these breeding lines. The consistent S/N ratio across the range of seed protein concentrations indicates that S-containing amino acids were not sacrificed with increases in seed protein.

SOYBEAN IS PRODUCED for oil and protein in the seed, which are the economically important components of the crop. The high concentration of protein in soy meal makes the meal a valuable livestock feed. Oligosaccharides, including raffinose and stachyose, are undesirable components of the meal because they may have a detrimental effect on the nutritive value of soy meal to animals (Liu, 1997).

Krober and Cartter (1962) evaluated seed compositional traits of soybean strains either very high or very low in seed protein concentration. Nonprotein constituents of seed of high protein samples, averaging 483 g kg^{-1} seed protein, on a dry seed basis, decreased by one-third for sugars, one-third for oil, and one-third for holocelluloses and pentosans, compared with samples averaging 391 g kg^{-1} seed protein. Among low protein samples, averaging 318 g kg^{-1} protein, oil increased from one-half to one-third, and holocellulose and pentosans increased about one-third compared with samples averaging 391 g kg^{-1} protein.

Hymowitz et al. (1972) evaluated oligosaccharides in 60 plant introductions that varied widely in seed protein and oil. They reported ranges of 2.5 to 8.2 g for sucrose, 0.1 to 0.9 g for raffinose, and 1.4 to 4.1 g for stachyose per 100 g seed. They observed a positive relationship between stachyose and seed protein ($r = 0.41**$) that they felt would pose difficulties for the soybean breeder who wanted to reduce stachyose content of the seed while maintaining high seed protein.

Hartwig et al. (1997) measured quantities of raffinose, stachyose, and sucrose among 20 soybean lines high in oil and among 20 breeding lines high in protein. Correlation coefficients between stachyose + raffinose and protein of ($r = -0.16$), and oil ($r = -0.03$), were not significant. In contrast, there was a strong inverse relationship between sucrose and protein ($r = -0.78**$) and a positive relationship between sucrose and oil ($r = 0.67**$) among all lines combined across the two groups. These relationships demonstrated the feasibility of developing soybean germplasm with high seed protein and low stachyose + raffinose in the meal.

The nutritional value of soybean meal could be improved by increasing amounts of the S-containing amino acids, methionine and cysteine. Soy protein is deficient in these amino acids and must be supplemented with other protein sources, or with synthetic methionine, when soy meal is used as the primary source of protein for humans and for monogastric animals. Glycinin (11S) and β-conglycinin (7S) are the two main classes of multisubunit seed storage proteins and account for ≈70% of total soybean seed protein (Meinke et al., 1981). Glycinin is a well-balanced protein with 3.0 to 4.5% of its amino acid residues consisting of cysteine and methionine (Nielsen et al., 1989; Fukushima, 1991), but β-conglycinin is very deficient in S-amino acids. Only 1% of its amino acid residues contain S (Harada et al., 1989; Sebastiani et al., 1990), with one of its three subunits, the β-subunit, having no S-amino acids at all (Coates et al., 1985). In hydroponic nutrition studies in which 'Harper' soybean was grown on various compositions of N during seed filling, Paek et al. (1997) learned that total protein concentration of seed could be increased 4.5 to 5.0%, from 369 to 420 g kg^{-1} in one experimental run and from 410 to 455 g kg^{-1} in the second, by substitution of NH$_3$–N for NO$_3$ in the growth medium. Storage proteins were increased by ≈4% in both runs, but the increase in storage protein was entirely because of an increase in β-conglycinin, in particular of the S-devoid, β-subunit of β-conglycinin. Thus, protein quality de-

J.R. Wilcox, USDA-ARS, Crop Production and Pest Control Research and Dep. of Agronomy, Purdue Univ., W. Lafayette, IN 47907-1150; R.M. Shibles, Prof. Emeritus, Dep. of Agronomy, Iowa State Univ., Ames, IA 50011-1010. Journal paper no. 16160 of the Purdue Univ. Agric. Res. Programs. Received 3 Dec. 1999. *Corresponding author (jwilcox@purdue.edu).

clined with increases in protein concentration. Paek et al. (1997) concluded that breeding efforts to improve soybean seed protein should not focus entirely on concentration. Potentially, soy protein quality could decline as lines with greater protein concentration are developed.

Information reported to date compares quantities of oligosaccharides among groups of soybean lines that differ widely in seed protein and oil concentration. The objective of this study was to determine the interrelationships among seed protein, oil, oligosaccharides, and S among random progenies from a cross between parents that differed in seed protein and oil.

MATERIALS AND METHODS

Random selections that differed in seed protein concentration among progenies of the cross C1834 × CX1314-37 were used in the study. The maternal parent C1834, with 397 g kg^{-1} seed protein and 213 g kg^{-1} seed oil, is a selection from the cross C1678 × 'Resnik' (McBlain et al., 1990); C1678 is a selection from the cross 'Hobbit' (Cooper et al., 1991) × 'Lakota' (Bahrenfus and Fehr, 1984). The breeding line CX1314-37, with 475 g kg^{-1} protein and 184 g kg^{-1} oil, is a selection from the cross CX1038-63 × HC84-553-1(Wilcox and Zhang, 1997). The F_2 through F_4 generations from the cross were advanced by single seed descent at Puerto Rico and at W. Lafayette, IN. Random $F_{4:5}$ lines were grown at W. Lafayette, IN. Forty-three individual $F_{4:6}$ lines in maturity Group III were harvested and evaluated in two-replicate tests at W. Lafayette, IN, in 1996 and $F_{4:7}$ lines in 1997 and at Ames, IA, in 1997. The soil at W. Lafayette was a fine-silty, mixed, mesic Typic Haplaquoll and at Ames a fine-loamy, mixed mesic Hapludoll. Four-row plots were used at W. Lafayette and three-row plots at Ames. The row spacing was 0.61 m at W. Lafayette and 0.70 m at Ames. The row length at W. Lafayette was 4.9 m, and was end trimmed to 3.7 m just prior to harvest. At Ames, row length was 5.0 m, and end trimmed to 4.27 m prior to harvest. The seeding rate was 23 seeds per m at both locations. The center two rows of each plot were harvested for yield at W. Lafayette and the single center row at Ames. Seed yield was measured as g per plot and converted to kg ha^{-1} prior to analyses.

A 25 g sample of seed from each replication at each environment was analyzed for seed protein, oil, total carbohydrate, and stachyose + raffinose by near-infrared reflectance at the National Center for Agricultural Utilization Research at Peoria, IL. Sucrose was determined as total carbohydrate minus raffinose + stachyose. Near-infrared reflectance spectroscopy is a demonstrated effective method of determining concentrations of sugar in plant materials (Giangiacomo et al., 1981; W. Rayford, USDA-ARS, 1999, personal communication). Sulfur was determined by inductively-coupled plasma emission spectrometry as described in Sexton et al. (1998).

Analyses of variance were computed on the data in which lines were considered random and environments fixed. Regression analyses were computed to evaluate relationships among specific traits.

RESULTS AND DISCUSSION

There was significant variability among lines and environments for each of the traits measured (Table 1). Interactions of lines × environments were significant for seed yield, protein, oil, and stachyose + raffinose,

Table 1. Mean squares for 43 soybean lines evaluated in three environments and for the interactions of lines × environments.

Trait	Lines	Environments	Lines × Environments
Seed yield (kg ha^{-1})	171 647.0**	41 243 282.0**	81 007.0**
Seed protein (g kg^{-1})	1 223.7**	25 817.0**	136.3**
Seed oil (g kg^{-1})	334.7**	7 854.7**	49.6**
Total carbohydrates (g kg^{-1})	68.6**	2 005.7**	18.8ns
Raffinose + stachyose (g kg^{-1})	3.3**	422.0**	2.6**
Sucrose (g kg^{-1})	62.4**	1 336.7**	15.5ns
Sulfur (g kg^{-1})	0.2**	2.4**	0.0ns

ns, ** non-significant and significant at the 0.01 probability level.

but not for total carbohydrate, sucrose, or S. Protein varied by 55 g kg^{-1} and oil by 32 g kg^{-1}, among lines (Table 2). There was greater variability relative to the mean among lines for sucrose (27%) than for raffinose + stachyose (9%) or total carbohydrate (16%). Variability among lines for S was 16%.

Regression analyses demonstrated that protein decreased by 3.9 g kg^{-1} for every 100 kg ha^{-1} increase in seed yield (Table 3). Oil increased by half this amount, 1.9 g kg^{-1} for every 100 kg ha^{-1} increase in seed yield (Table 3). These relationships among seed yield and both protein and oil are consistent with previous reports in populations where there is considerable variability among lines for protein and oil concentration (Burton, 1984). The data presented here indicate that these lines were typical of progenies from crosses between parents with typical and high levels of protein concentration.

There was a strong inverse relationship between protein and oil, with protein decreasing by 15.6 g kg^{-1} for each 10 g kg^{-1} increase in oil (Table 3). In virtually every soybean population that varies in these two traits, this inverse relationship has been reported (Burton, 1994; Wilcox, 1998).

No relationships were found between concentrations of any of the carbohydrates and seed yield in this population (Table 3). These data indicate that concentrations of either total or component carbohydrates could be altered without affecting seed yield. Carbohydrates in the seed decreased as protein increased, with the greatest decrease in sucrose (Fig. 1, Table 3). Increases in protein among these lines occurred at the expense of both oil and carbohydrates. Conversely, total carbohydrates ($b = 0.173$**) and sucrose ($b = 0.148$*) increased with increases in seed oil (Table 3). This is consistent with the strong inverse relationship between protein and oil among these lines.

Sulfur concentration in the seed increased with increasing protein (Fig. 1). This would be expected because S is a component of cysteine and methionine, two

Table 2. Means and ranges for seed yield and quality traits of 43 soybean lines averaged over three environments.

Trait	Mean	Range
Seed yield (kg ha^{-1})	2367.0	2023.0–2592.0**
Seed protein (g kg^{-1})	438.0	413.0–468.0**
Seed oil (g kg^{-1})	183.0	166.0–198.0**
Total carbohydrate (g kg^{-1})	90.0	82.0–96.0**
Raffinose + stachyose (g kg^{-1})	41.1	38.5–42.2**
Sucrose (g kg^{-1})	48.9	43.3–56.7**
Sulfur (g kg^{-1})	4.0	3.8–4.4**

** Significant at the 0.01 probability level.

Table 3. Linear regression analyses and correlation coefficients among soybean yield and quality traits for 43 breeding lines averaged over three environments.

Dependent versus independent variables	Regression mean squares	R^2	Slope (b)	Standard error of b	Correlation coefficient
Protein vs. seed yield	1387.40**	0.201	−0.039	0.012	−0.45**
Oil vs. seed yield	324.27*	0.147	0.019	0.007	0.38*
Total carbohydrate vs. yield	26.90ns	0.068	0.006	0.003	0.26ns
Stachyose + raffinose vs. yield	0.07ns	0.003	<0.001	<0.001	0.05ns
Sucrose vs. yield	24.69ns	0.069	0.005	0.003	0.26ns
Sulfur vs. yield	0.05ns	0.055	<0.001	<0.001	0.23ns
Protein vs. oil	5368.39**	0.777	−1.560	0.130	−0.88**
Total carbohydrate vs. protein	201.91**	0.508	−0.171	0.026	−0.71**
Stachyose + raffinose vs. protein	2.50*	0.104	−0.019	0.009	−0.32*
Sucrose vs. protein	157.61**	0.440	−0.151	0.027	−0.66**
Sulfur vs. protein	0.44**	0.456	0.008	0.001	0.68**

ns, *, ** non-significant and significant at the 0.05 and 0.01 probability levels, respectively.

of the amino acids in soy protein. Conversely, there was an inverse relationship between S and oil in the seed ($R^2 = 0.365*$) associated with the decrease in protein with increasing oil content. Radford et al. (1977) reported close relationships between percentage S and mg of methionine ($r = 0.83$), mg cysteine ($r = 0.83$), and mg of methionine plus cysteine ($r = 0.85$) for 13 lines of *G. max* and *G. soja* (Sieb. and Zucc.).

The ratio of S/N was regressed on seed protein to determine if the amount of S in the protein increased with increasing seed protein concentration. This would be indicative of higher amounts of S-containing amino acids, relative to other amino acids, as protein increased. No relationship was observed between the S/N ratio and seed protein ($R^2 = 0.016$). This indicated that the amount of S-containing amino acids was relatively constant regardless of the amount of protein in the seed of these breeding lines. Burton et al. (1982) reported no
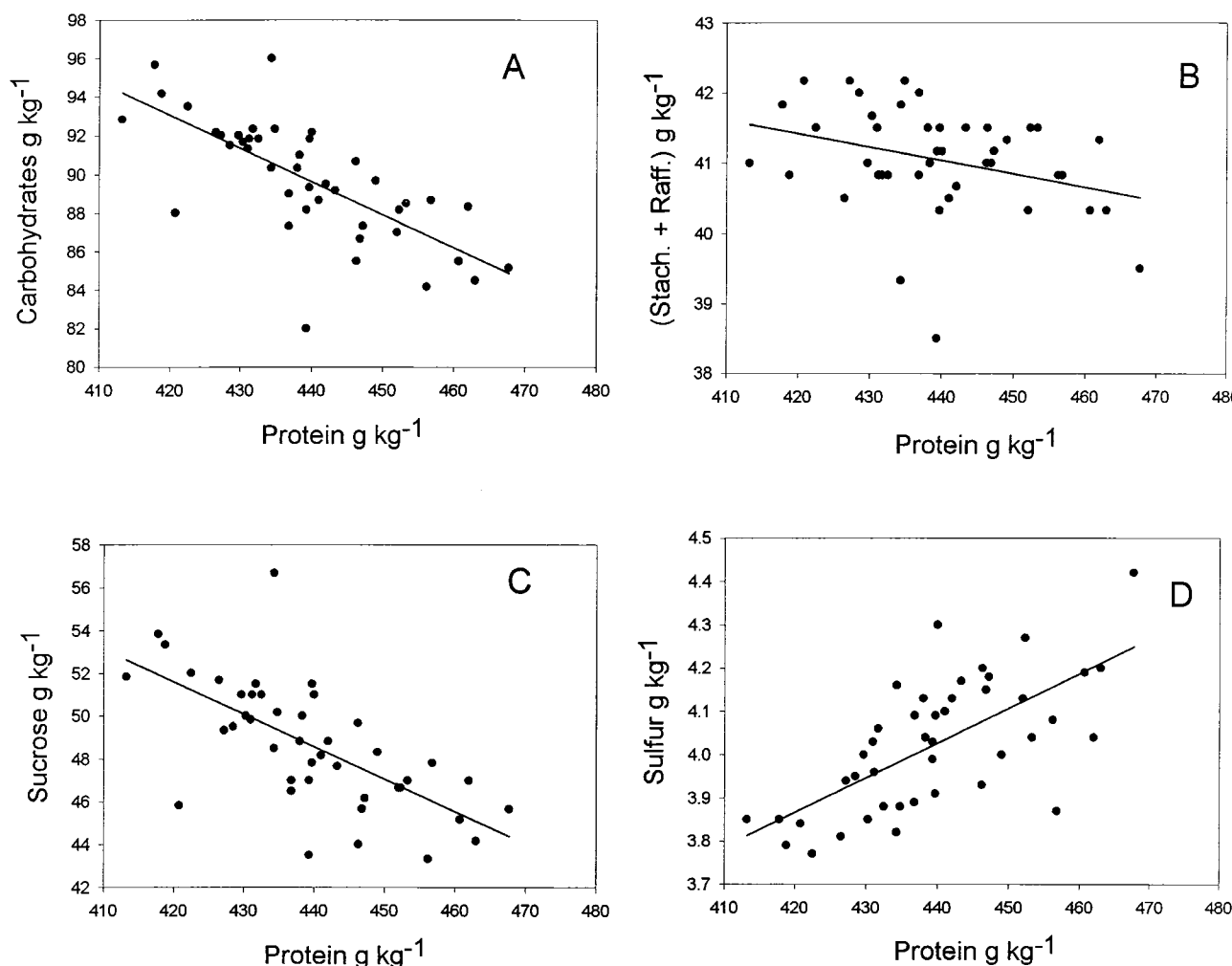
Fig. 1. Regression of total carbohydrates (A), stachyose + raffinose (B), sucrose (C), and sulfur (D) in the seed on seed protein for 43 soybean progenies averaged across three environments. Regression statistics in Table 3.

significant changes in methionine among cycles of recurrent selection for high protein where protein increased from 438 in Cycle 0 to 474 g kg^{-1} in Cycle 6 in one population and from 450 in Cycle 0 to 472 g kg^{-1} in Cycle 4 in a second population. In contrast, Krober and Cartter (1966) reported a positive association between percentage methionine in the protein and percentage protein in the seed ($r = 0.56$) among 12 soybean strains that varied from 390 to 480 g kg^{-1} seed protein.

Paek et al. (1997) found the proportion of S-poor protein to increase as protein concentration increased in a nutrition study with Harper soybean, so that protein quality declined as protein concentration increased. This suggested that it may be difficult to maintain protein quality when breeding for increased seed protein concentration. In the population of breeding lines reported in this study, soybean S assimilation was adequate to maintain S-amino acid level across a wide range of protein concentrations, as indicated by a constant S/N ratio with protein concentration. We caution, however, that seed yields were not high in this population, ranging from 2000 to 2600 kg ha^{-1}. Thus, there still remains a question whether S-amino acid concentration can be maintained while breeding for high protein concentration and simultaneously maintaining yields of ≈4000 kg ha^{-1}.

The data demonstrated that increases in seed protein in this population were at the expense of both oil and carbohydrates, particularly sucrose. The decreases in seed carbohydrates with increasing seed protein would contribute to increased nutritional value of the meal. A consistent ratio of S/N across the range of seed protein among these breeding lines indicated that S-containing amino acids were not sacrificed with increases in seed protein.

REFERENCES

Bahrenfus, J.B., and W.R. Fehr. 1984. Registration of Lakota soybean. Crop Sci. 24:384–385.

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., A.E. Purcell, and W.M. Walter, Jr. 1982. Methionine concentration in soybean protein from populations selected for increased seed protein. Crop Sci. 22:430–432.

Coates, J.B., J.B. Mederiros, V.H. Thanh, and N.C. Nielsen. 1985. Characterization of the subunits of β-conglycinin. Arch. Biochem. Biophys. 243:184–194.

Cooper, R.L., R.J. Martin, A.K. Walker, and A.F. Schmitthenner. 1991. Registration of 'Hobbit' soybean. Crop Sci. 31:231.

Fukushima, D. 1991. Recent progress of soybean protein foods: Chemistry, technology, and nutrition. Food Rev. Int. 7:323–351.

Giangiacomo, R., J.B. MaGee, G.S. Birth, and G.G. Dull. 1981. Predicting concentrations of individual sugars in dry mixtures by near-infrared spectroscopy. J. Food Sci. 46:531–534.

Harada, J.J., S.J. Barker, and R.B. Goldberg. 1989. Soybean beta-conglycinin genes are clustered in several DNA regions and are regulated by transcriptional and posttranscriptional processes. Plant Cell 1:415–425.

Hartwig, E.E., T.M. Kuo, and M.M. Kenty. 1997. Seed protein and its relationship to soluble sugars in soybean. Crop Sci. 37:770–773.

Hymowitz, T., F.I. Collins, J. Panczner, and W.M. Walker. 1972. Relationship between the content of oil, protein, and sugar in soybean seed. Agron. J. 64:613–616.

Krober, O.A., and J.L. Cartter. 1962. Quantitative interrelationships of protein and nonprotein constituents of soybean. Crop Sci. 2:171–172.

Krober, O.A., and J.L. Cartter. 1966. Relation of methionine content to protein levels in soybeans. Cereal Chem. 43:320–325.

Liu, K. 1997. Soybeans chemistry, technology, and utilization. Chapman & Hall, New York.

McBlain, B.A., R.J. Fioritto, S.K. St. Martin, A. Calip-DuBois, A.F. Schmitthenner, R.L. Cooper, and R.J. Martin. 1990. Registration of 'Resnik' soybean. Crop Sci. 30:424–425.

Meinke, D.W., J. Chen, and R.N. Beachy. 1981. Expression of storage-protein genes during soybean seed development. Planta 153:130–139.

Nielsen, N.C., C.D. Dickinson, T. Cho, V.H. Thanh, B.J. Scallon, R.L. Fischer, T.L. Sims, G.N. Drews, and R.B. Goldberg. 1989. Characterization of the glycinin family in soybean. Plant Cell 1:313–328.

Paek, N.C., J. Imsande, R.C. Shoemaker, and R. Shibles. 1997. Nutritional control of soybean seed storage protein. Crop Sci. 37:498–503.

Radford, R.L., Jr., C. Chavengsaksongkram, and T. Hymowitz. 1977. Utilization of nitrogen to sulfur ratio for evaluating sulfur-containing amino acid concentrations in seed of *Glycine max* and *G. sojae*. Crop Sci. 17:273–277.

Sebastiani, F.L., L.B. Farrell, M.A. Schuler, and R.N. Beachy. 1990. Complete sequence of a cDNA of α subunit of soybean β-conglycinin. Plant Mol. Biol. 15:197–201.

Sexton, P.J., N.C. Paek, and R. Shibles. 1998. Soybean sulfur and nitrogen balance under varying levels of available sulfur. Crop Sci. 38:975–982.

Wilcox, J.R. 1998. Increasing seed protein with eight cycles of recurrent selection. Crop Sci. 38:1536–1540.

Wilcox, J.R., and G. Zhang 1997. Relationships between seed yield and seed protein in determinate and indeterminate soybean populations. Crop Sci. 37:361–364.