Analysis of Single Soybean Seeds for Oil and Protein

Lori L. Huskey[a], **H.E. Snyder**[*,a] and **E.E. Gbur**[b]
[a]Department of Food Science and [b]Agricultural Statistics Laboratory, University of Arkansas, Fayetteville, AR 72703

Microprocedures were developed to analyze single soybean seeds for protein, oil, and moisture. Ten mature Forrest soybean plants were taken from a yield plot and sampled, so that seeds were taken from the top, middle, and bottom areas. Also, seeds were taken from distal and proximal racemes, from distal and proximal pods on a raceme, and from distal and proximal seeds within a pod. The range in protein for the 241 seeds analyzed was 32 to 51% (dry weight basis) with a standard deviation of 2.96%. The range in oil was 16.5 to 25.5% (dry weight basis) with a standard deviation of 1.84%. The middle area had significantly more oil and significantly less protein than the top and bottom areas. Also, significant differences in protein and oil were found between plants. There was no significant difference in protein or oil due to position of a raceme, position of a pod on a raceme, or position of a seed within a pod. The standard deviation found for protein and oil can be used to estimate the sample size needed to achieve a certain degree of accuracy in protein and oil analyses.

KEY WORDS: Goldfisch extraction procedure, oil, protein, sample size, soybean.

Recently the Federal Grain Inspection Service of the United States Department of Agriculture (USDA), the American Soybean Association, and others have called attention to the possible usefulness of protein and oil content of soybeans as quality criteria for marketing soybeans (1-4). To improve the ability to control protein and oil content in soybeans, it would be useful to have information on the factors that influence variability in protein and oil content.

Regional differences in protein content have been identified in the United States with the northwest soybean region having less protein than the south or southeast (5). Within regions and from year to year, differences in oil and protein have been noted (6,7). Temperature (8,9) and moisture (10) are factors that influence protein and oil content during soybean growth.

Position of the seed on the soybean plant has been investigated (11) with the result that seeds from the top half of the plant had more protein and less oil than seeds from the bottom half.

The only data on analysis of single seeds that we are aware of came from oil analysis by wide line nuclear magnetic resonance (NMR) (12). The authors were interested in calibrating NMR for oil analysis and found a range of 12% in oil content for single seeds. No data were given on the cultivars or sources of these seeds.

With the development of a method for oil analysis by rapid equilibrium of finely ground flour with solvent (7,13), the possibility existed of analyzing single soybean seeds for protein, oil, and moisture. This analysis coupled with a study of the variability of protein and oil with position of the seed on the plant could provide information useful for improved control of protein and oil.

Herein we report the development of micromethods for analyzing protein, oil, and moisture in single soybean seeds. In addition, the methods were used to analyze seeds located at different positions on uniformly grown Forrest soybean plants.

MATERIALS AND METHODS

Soybean seed selection. Soybean seeds (Forrest cultivar) were sampled from a yield plot at the Agricultural Experiment Station, University of Arkansas, Fayetteville, Arkansas. Five rows of the yield plot were chosen at random and two plants were taken from each row to the laboratory for sampling.

The number of nodes on each plant was determined, and the plant was divided into thirds so that an equal number of nodes were in the top, middle, and bottom areas. Two branches were selected from each area. The proximal and distal pods were removed from the proximal and distal racemes of each branch. The proximal and distal seeds from the pods were analyzed.

In the top area, no distal racemes were obtained. If there was only one seed in a pod, it was labeled as proximal. Only seeds with a weight of at least 100 mg were taken to have sufficient material to analyze.

Seed moisture and seed weight. Single seeds were dried at 130°C for 3 hr, and moisture was determined using AOCS method Ac 2-41 (14). Seed dry weights were determined simultaneously.

Seed grinding. Individual dried seeds were placed in the metal cylinder of a Wig-L-Bug amalgamator (Crescent Dental Mfg., Lyons, IL) with a single steel ball. Total grinding time was 25 sec with 5 sec of grinding interspersed with 5 sec of rest to prevent overheating and agglomerating. The resulting flour was brushed through a 100-mesh sieve, and portions of the sieved flour were used for moisture, oil, and protein analyses. To minimize static charges and blinding of the sieves, the dried seeds were allowed to equilibrate with room moisture for approximately 5 days before grinding.

Flour moisture. Cellulose acetate capsules (Cahn Instruments, Cerritos, CA) were used as containers. Approximately 20 mg of 100 mesh flour was placed in the preweighed capsules. Flour plus capsule were weighed with a Cahn 300 Microbalance to the nearest μg. Uncovered samples were dried at 130°C for 1 hr, covered, and cooled in a diseccator before reweighing.

Total oil analysis. A procedure was developed for determining total oil in 20 mg samples based on the rapid equilibrium extraction (7,13). Approximately 20 mg of 100 mesh flour was weighed to the nearest 0.1 mg and placed in a 4 mL glass vial with a permeable septum lid insert. Hexane (2 mL) was added to the flour, and the slurry was stirred for 30 min using a small magnetic stir bar. The slurry was filtered through a 0.8 μm Millipore filter into a second vial using a 10

*To whom correspondence should be addressed at 272 Young Avenue.

mL syringe. Precautions were taken during all transfers to minimize loss of hexane by evaporation.

The flour free miscella was sampled by withdrawing 400 μL with a gas tight syringe and transferring to a preweighed aluminum dish positioned on a hot plate. Three samples were taken for each extraction. The hexane was evaporated on the hot plate, sample pans were cooled in air, and the remaining oil was weighed to the nearest μg.

The rapid equilibrium procedure was compared with a 4-hr Goldfisch procedure using AOCS method Ac 3-44 (15).

Protein analysis. Protein was determined by a micro-Kjeldahl procedure. The sample size was approximately 20 mg, and digestion was done in sulfuric acid with a mercuric oxide: sodium sulfate catalyst. Distillation was with a LabConCo microdistillation unit and distillate was trapped in 0.25% boric acid. The titration was done with an autoburrette with an endpoint of pH 4.7, speed 40, delay 20 sec, and proportional band 0.2.

Data analysis. Simple descriptive statistics including means, standard deviations, and correlations were calculated for percentage of oil, protein, and moisture and for seed weight. In addition, the effect of plants, area of the plant (top, middle, bottom), branch (proximal, distal), raceme (proximal, distal), and pod (proximal, distal) on percentage of oil and protein as well as on seed weight were analyzed using analysis of variance.

RESULTS AND DISCUSSION

To test the newly developed microprocedures a sample of 100-mesh soybean flour was analyzed for total oil, protein, and moisture by conventional procedures and by the microprocedures. The total oil content was 0.04% lower when compared to a 4-hr Goldfisch extraction, moisture was 0.07% lower compared to 2-g samples dried in metal pans, and there was no difference in protein content by the two analysis procedures.

The total number of single seeds analyzed was 241. The mean value for total oil was 21.90% (dry basis) with a range of 16.53 at 25.51%. The standard deviation for total oil analysis was 1.84%. Figure 1 shows the distribution of values for total oil. Values were definitely skewed toward higher oil content.

The mean value for protein in the 241 seeds was 40.78% (dry basis) with a range of 32.04 to 51.21%. The standard deviation was 2.96%. Figure 2 shows the distribution of values for protein with no obvious skewing.

We expected a range of values for total oil and protein in single seeds but the extent of the range was surprising. The only report on single soybean seed analysis that we are aware of is for total oil measured by NMR (12). Collins *et al.* (12) found a range of 12% oil content, but they chose seeds to analyze based on a wide range of oil content to determine linearity of NMR response over a wide range. Our choice of seeds was from the same cultivar grown in a yield plot, so that soil, moisture, and temperature were all reasonably uniform. To our knowledge the data presented herein are the only analyses of single soybean seeds for total protein.

The standard deviations found for analysis of total oil and protein can be used to estimate sample size needed to detect a certain allowable limit for mean error for a given level of confidence. For example, a formula (16) relating the allowable limit for mean error (L) to standard deviation σ, and sample size, n, is:

$$L = \frac{2\sigma}{\sqrt{n}}$$

The coefficient 2 represents an approximate t value for 95% confidence level and a large sample size. For detecting a mean error limit of 0.1% protein at a confidence level of 95% and $\sigma = 2.96$, the sample size should be 3,505 seeds. For Forrest soybeans with an average seed weight of 100 mg, this means a sample size of 350 g would be needed. If seed size were larger, a larger sample could be needed. The same kind of calculation done for total oil (mean error limit of 0.1 with 95% confidence level) gives a sample size of 1,354 seeds or 135 g.

The sample size recommended by Official Methods of the American Oil Chemists' Society is 60 g for oil and protein (15). For the variability found in Forrest

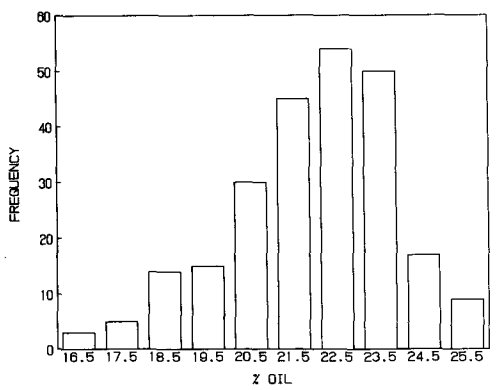

FIG. 1. The frequency distribution for oil content (dry basis) of 241 single soybean seeds from the cultivar Forrest. One percent increments used.

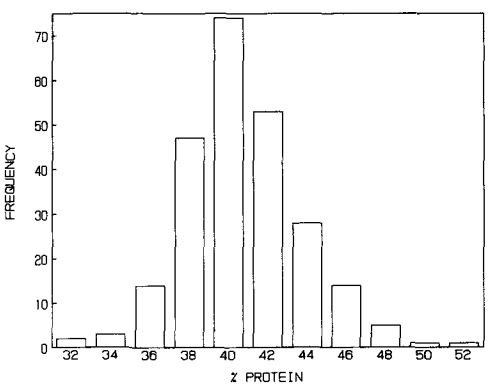

FIG. 2. The frequency distribution for protein content (dry basis) of 241 single soybean seeds from the cultivar Forrest. Two percent increments used.

TABLE 1

Effect of Plant Area on Oil, Protein (dry basis) and Weight of Individual Soybean Seeds

Area	Oil (%)[1]	Protein (%)[2]	Weight (mg)[3]
Top	21.49^b	41.50^a	128.55^a
Middle	22.66^a	39.92^b	117.40^b
Bottom	21.50^b	41.01^a	115.09^b

[1]Means in the same column with different letters are significantly different at 1% level (LSD = 0.73).
[2]Means in the same column with different letters are significantly different at 5% level (LSD = 0.91).
[3]Means in the same column with different letters are significantly different at 1% level (LSD = 9.10).

TABLE 2

Oil and Protein Contents (% dry basis) for Plants Calculated from Individual Seed Data

Plant	Oil[1]	Protein[1]
895B	23.22^a	40.32^c
37B	22.82^a	42.22^a
895A	22.47^{ab}	38.50^d
37A	22.43^{ab}	40.54^c
250A	21.85^{bc}	42.88^a
250B	21.59^c	40.57^c
1561A	21.34^c	40.93^{bc}
1561B	21.34^c	40.01^c
1450A	21.34^c	40.38^c
1450B	20.34^d	42.15^{ab}

[1]Means in the same column with different letters differ significantly at the 5% level (LSD = 0.83 for oil, 1.50 for protein).

soybeans, a 60-g sample would give a mean error level of 0.24% for protein and 0.15% for oil at a 95% confidence level.

There was a significant negative correlation between protein and oil in single seeds with a correlation coefficient of −0.408. This confirms, for single seeds, the often-made observation for larger samples that protein and oil are negatively correlated in soybeans. Because of the large sample sizes, both protein and oil content showed significant positive correlations with moisture content and with seed weight although the correlation coefficients were small (protein:moisture 0.18; protein:weight 0.28; oil:moisture 0.14; oil:weight 0.17).

The single seed analyses were done in part to investigate what effect position of the seed on the plant had on the protein and oil content. Table 1 shows that dividing the plant into three areas (top, middle, and bottom) did have a significant effect. Seeds from the middle third of the plants had significantly more oil and significantly less protein than seeds from the top or bottom of the plant.

We also investigated the effect of proximal or distal racemes, proximal or distal pods on the racemes, and proximal or distal seeds in the pods. These positions showed no significant differences for either protein or oil.

Oil and protein contents were compared in ten plants based on means of the 20 to 30 seeds analyzed from each plant (Table 2). Plants did show significant differences from each other in protein and oil content. The combined oil and protein ranged from 65% for plant 37B to 61% for plant 895A.

While the main effects of area and plant showed very significant differences, there was also a highly significant interaction between plant and area. The interaction indicated that the differences between top, middle and bottom areas were not the same from plant to plant. Also, the differences between plants varied depending on the area analyses.

The significance and possible utility of the different oil and protein content found in the middle area compared to top and bottom areas remain to be seen. The results from single seed analysis of soybeans provide a basis for making rational judgments about the sample size needed for oil and protein analyses. Based on our data for the cultivar Forrest, the American Oil Chemists' Society recommendation of a 60-g sample for protein seems low.

The source of the highly significant difference in oil and protein content for individual plants would be useful to know. Although there is an inverse correlation between oil and protein content, one plant of the ten sampled had reasonably high oil (22.82%) and protein (42.22%).

REFERENCES

1. Anonymous, *Federal Register 54*:778 (1989).
2. Anonymous, *Ibid. 54*:33702 (1989).
3. Guinn, J.M., *Oil Mill Gazeteer 94(6)*:26 (1988).
4. Hurburgh, C.R., *Cereal Foods World 33*:503 (1988).
5. Breene, W.M., S. Lin, L. Hardman and J. Orf, *J. Am. Oil Chem. Soc. 65*:1927 (1988).
6. Caviness, C.E., *Rice Journal 77(3)*:20 (1974).
7. Clark, P.K., and H.E. Snyder, *J. Am. Oil Chem. Soc. 66*:1316 (1989).
8. Howell, R.W., and J.L. Cartter, *Agronomy J. 45*:526 (1953).
9. Howell, R.W., and J.L. Cartter, *Ibid. 50*:664 (1958).
10. Rose, I.A., *Aust. J. Agric. Res. 39*:163 (1988).
11. Collins, F.I., and J.C. Cartter, *Agronomy J. 48*:216 (1956).
12. Collins, F.I., D.E. Alexander, R.C. Rodgers and L. Silvela, *J. Am. Oil Chem. Soc. 44*:708 (1967).
13. Snyder, H.E., G. Sheu, H.G. Brown, P. Clark and K.L. Wiese, *Ibid. 65*:255 (1988).
14. *American Oil Chemists' Society, Official and Tentative Methods, Method Ac 2-41, edited by R.C. Walker, AOCS, Champaign, IL, 1985.*
15. American Oil Chemists' Society, *Ibid.*, Method AC 3-44, edited by R.C. Walker, AOCS, Champaign, IL, 1985.
16 Snedecor, G.W., and W.G. Cochran, *Statistical Methods*, 86th edn., Iowa State University Press, Ames, IA, 1989, p. 52.

[Received January 13, 1990; accepted June 9, 1990]