J Am Oil Chem Soc (2010) 87:127–132
DOI 10.1007/s11746-009-1484-2

ORIGINAL PAPER

A Modified Amino Acid Analysis Using PITC Derivatization for Soybeans with Accurate Determination of Cysteine and Half-Cystine

Prachuab Kwanyuen · Joseph W. Burton

Received: 12 May 2009 / Revised: 23 September 2009 / Accepted: 4 October 2009 / Published online: 27 October 2009

Abstract Breeding efforts to change the amino acid profile of seed protein and the assessment of genetic variation for amino acid composition among soybean germplasm resources have been hampered by lack of a rapid and inexpensive method for amino acid determination. A modified procedure presented here is partly based on a gas-phase hydrolysis and precolumn derivatization for HPLC analysis. The procedure accurately measured cysteine and half-cystine in samples. The method was also proven to be accurate using a reference protein with known amino acid composition. It is reliable and can be automated for daily analysis with a large number of samples. The method was also tested with soybean seeds harvested from a two-replicate multi-location soybean field experiment. It was found that no soybean cultivar by location interaction was significant for any of the amino acids which demonstrate that amino acid compositions were generally stable across a range of environments, and that the repeatability of the measurement itself was high. The error associated with the determination was also low as demonstrated by the analysis of variance.

Keywords Amino acid analysis · Soybean · *Glycine max* L. Merr · Cysteine · Half-cystine

P. Kwanyuen (✉) · J. W. Burton
USDA-ARS, Soybean and Nitrogen Fixation Research Unit, Crop Science Department, North Carolina State University, 3127 Ligon Street, Raleigh, NC 27607, USA
e-mail: Prachuab.Kwanyuen@ars.usda.gov

Introduction

Soybean (*Glycine max* L. Merr.) protein meal is a major component of swine and poultry diets. Soy protein is generally considered to be a very high quality protein, but some diets require that the protein meal be amended with synthetic amino acids. Those which are deficient in soybean protein are the sulfur-containing amino acids, methionine and cystine. These have long been targets for soybean protein improvement through plant breeding and/or genetic engineering. Increasing lysine and tryptophan concentrations in soy protein would also improve soy protein quality.

Soybean breeding efforts to change the amino acid profile of seed protein has been hampered by the expense in time and money required to measure amino acids that include sulfur-containing amino acids, methionine and cysteine. This has also prevented an adequate assessment of genetic variation for amino acid concentrations among soybean germplasm resources. Currently, in the US soybean germplasm collection there are 19,266 accessions. Methionine concentrations have been determined for 2,198 of these and there is no information about other amino acids, such as lysine (GRIN USDA-ARS, 30 September 2007). Studies of genetic variation in amino acid concentration have been limited to evaluations of relatively small sets of soybean genotypes [1–3], and good estimates of heritability and genotype by environmental interaction for individual amino acids are rare in published literature. Successful soybean breeding practice requires evaluation of large populations of plants or families that are variable for genes that control particular traits, e.g. concentration of a particular amino acid. Thus an amino acid measurement method is needed that is accurate, inexpensive, and economical in terms of time required for sample analysis.

Since soybean proteins especially the storage protein, glycinin, contains both cysteine and cystine, a dimer of two cysteine molecules joined by a disulfide bond, conventional amino acid analysis requires an additional performic acid oxidation of cysteine and half-cystine and then a complete acid hydrolysis with HCl. This is time-consuming and expensive particularly when large numbers of soybean samples generated from a breeding program are involved.

In this paper, we describe a modified method for measuring amino acids in picomole levels that meets the above requirements. It does not require performic acid oxidation of cysteine and/or cystine. The method is inexpensive, can be automated for daily HPLC analysis, and suitable for soybean and its full-fat meals and products. We have also tested this method using a reference protein with known composition and 150 samples of seeds harvested from a two-replicate multi-location soybean field experiment. Results of those analyses are also presented here.

Materials and Methods

Chemicals

Amino acid standard mixture, sequanal grade of 6 N HCl and triethylamine (TEA), phenyl isothiocyanate (PITC), and saturated phenol were from Pierce Chemicals (Rockford, IL, USA). Lysozyme, L-cystine, HPLC grade sodium acetate trihydrate ($CH_3COONa{\cdot}3H_2O$), disodium hydrogen phosphate (Na_2HPO_4) were from Sigma (St. Louis, MO, USA). Acetonitrile (HPLC grade) was from Fisher Scientific (Fair Lawn, NJ, USA). Purified water (Millipore, Bedford, MA, USA) of at least 18 MΩ cm resistivity was used for the preparation of all reagents, eluents, and buffers. All were filtered through 0.2-μm membranes.

Sample Preparation

Soybean seeds were initially ground in a coffee grinder. For convenience, each sample of 100 mg of full-fat meal was homogenized in 2.0 mL water in a 16 × 100 mm glass tube with a homogenizer equipped with a 7-mm generator for 1 min at medium speed. Five microliters of homogenate was transferred to a 6 × 50 mm glass tube that had been acid-washed with 6 N HCl for 24 h. The sample was then lyophilized. Several samples of 5 μg of egg-white lysozyme were also used as reference checks for determining accuracy and reliability of the procedure.

Gas-Phase Hydrolysis

Protein hydrolysis, precolumn derivatization, and HPLC analysis were performed according to Heinrikson and Meredith [4] with modifications. Ten lyophilized samples were placed in each evacuated glass container with 200 μL of 6 N HCl and 1% phenol in the bottom. Evacuation in 25–30 in-Hg pressure alternated with nitrogen flushing was performed at least three times and sealed. Samples were hydrolyzed at 110 °C in vacuo for 24 h. After hydrolysis, samples were dried under vacuum and stored in freezer until derivatization and HPLC analysis. Although HPLC analysis of the protein hydrolysates is normally performed in duplicate, in this experiment the hydrolysis of each sample was done in triplicate. The extra replication was used as a back-up in case of a possible contamination caused by HCl condensation in the sample tube during the hydrolysis. A 10-μL sample of amino acid standard mixture was also dried in a separate 6 × 50 mm acid-washed glass tube at the same time.

Precolumn Derivatization

Both standard and hydrolyzed samples were neutralized by adding 20 μL of a 2:2:1 mixture of ethanol:water:TEA (v/v), and mixing well with a vortex stirrer. They were dried under vacuum. Derivatization was performed by adding 20 μL of a mixture of 7:1:1:1 ethanol:water:TEA:PITC (v/v), and mixing well with a vortex stirrer. The reaction between PITC and the hydrolysate to produce phenylthiocarbamyl (PTC) amino acids was allowed to complete for 20 min at room temperature. Samples were then completely dried under vacuum and stored in a freezer.

HPLC Analysis

Phenylthiocarbamyl amino acids in each sample and standard were dissolved from the dried matrix by vortex mixing with 500 μL of 5 mM Na_2HPO_4 buffer, pH 7.4 containing 5% acetonitrile. The fluid was then filtered through a 0.2-μm membrane. Samples were reconstituted one at a time due to the PTC amino acid sensitivity to light and ambient temperature. Ten microliters of sample were injected and analyzed with an HPLC system equipped with a column heater, autosampler, variable wavelength detector, and a data acquisition software controller. The reverse-phase column used was a Pico-Tag (3.9 × 150 mm), dimethyloctadecylsilyl bonded amorphous silica, with an inline column filter. The column temperature was maintained at 38 °C. The PTC amino acids were separated and eluted by a gradient resulting from mixing eluents A and B. Eluent A consisted of 150 mM $CH_3COONa{\cdot}3H_2O$, 0.05% TEA, and 6% acetonitrile, pH 6.4. Eluent B consisted of 6:4 acetonitrile: water (v/v). Both eluents were sparged with ultra pure helium gas for 10 min before use. The flow rate was 1 mL/min throughout, and the gradient consisted of the following profiles: 100% A at start, 80% A and 20%

Table 1 Soybean cultivars and references to their origin that were field tested at five NC locations in 2005

Genotypes	References
Coker 237	Proprietary
NC-Raleigh	[5]
Stonewall	[6]
Ransom 2[a]	
Braxton	PI548695, PVP8000075
Haskell	[7]
Ransom	[8]
Black Ransom[a]	
Johnston	[9]
GaSoy 17	[10]
Howard	PI548971
Bragg	[11]
N88-480	[12]
Prichard	[13]
Jackson	[14]

Plymouth, Kinston, Clayton, Clinton, and Windblow

[a] Single plant selections from the Ransom cultivar

B at 5.5 min, 54% A and 46% B at 10 min, 100% B at 10.5–12.5 min, 100% A at 13 min. The PTC amino acids eluted from column were detected at 254 nm and recorded. The column was regenerated and equilibrated with eluent A for 10 min. A new and freshly reconstituted sample was injected, either by manual injector or autosampler/injector, and analyzed every 23 min.

Soybean Field Experiments

In 2005, 14 soybean cultivars and one germplasm (Table 1) were grown in field tests using a randomized complete block experimental design with two blocks (replications) at five North Carolina locations (Plymouth, Kinston, Clayton, Clinton, and Windblow). Cultivars were grown in three-row plots. Rows were 1 × 6 m. At maturity, seeds were harvested from the middle 5 m of the center row.

Statistical Analysis

A combined analysis of variance over locations was performed for all amino acids and for protein percentages using an ANOVA SAS procedure [15]. In the analysis, cultivars were considered to be fixed and locations random. Variation among cultivars was tested for significance using an F test, with the location × cultivar mean square as the divisor (error). The location × cultivar interaction was tested using the pooled error as the devisor.

Results and Discussion

Phenylthiocarbamyl amino acids are known to be sensitive to light and can be degraded over time. We found from our preliminary experiments that they were stable for at least 8 h at room temperature in amber sample vials. Automated HPLC analysis was then conveniently performed with an autosampler with a minimum of 16 samples. Figure 1 shows a typical elution profile of standard amino acids, each contained 500 pmol except cysteine which contained only 250 pmol. Since tryptophan can only be obtained by a separate alkaline hydrolysis and additional derivatization, it is considered to be cumbersome and is not cost-effective to measure. In addition, most proteins including soy protein

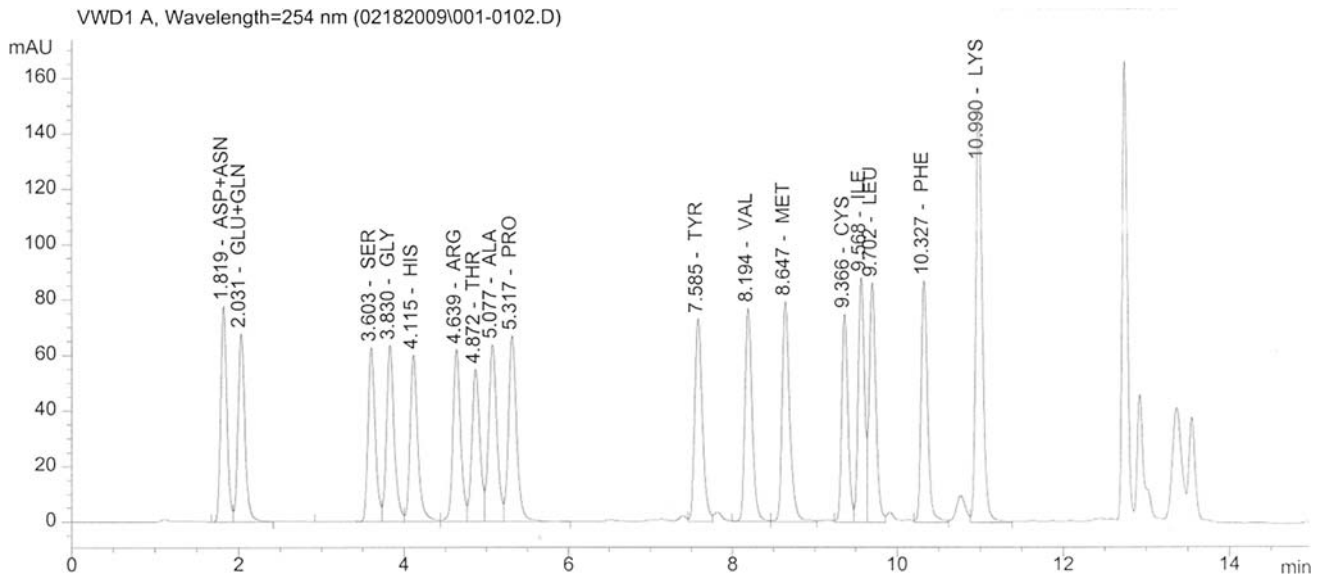

Fig. 1 Typical elution profile including retention times of standard amino acids

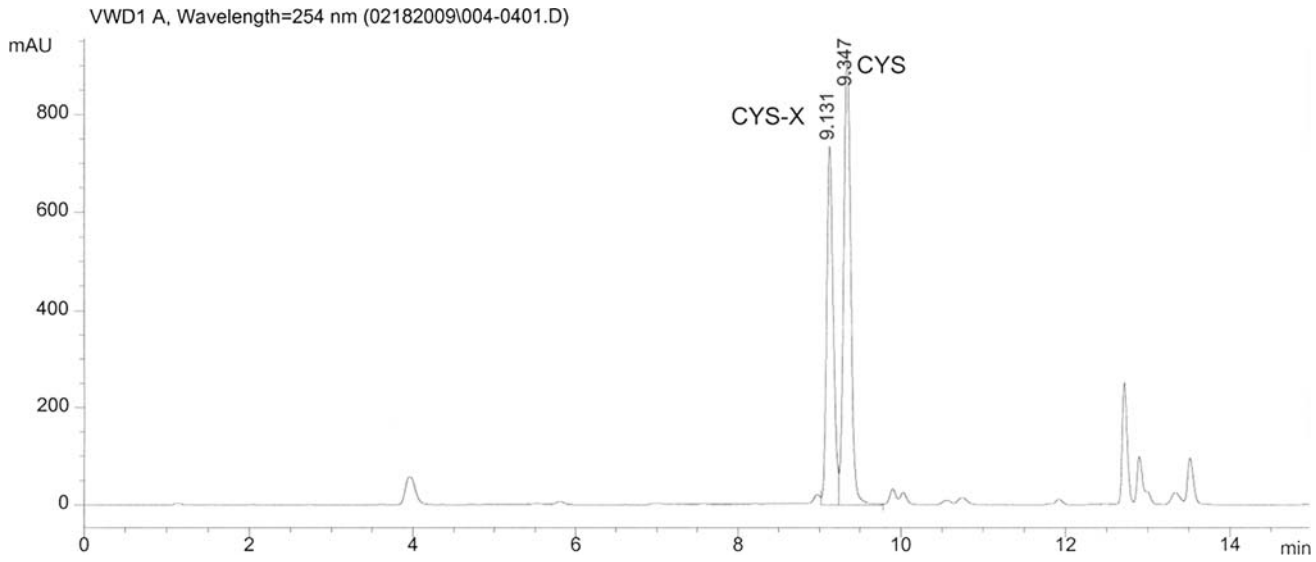

Fig. 2 Elution profile of cyteine and half-cystine (cysteine-X) when L-cystine was hydrolyzed

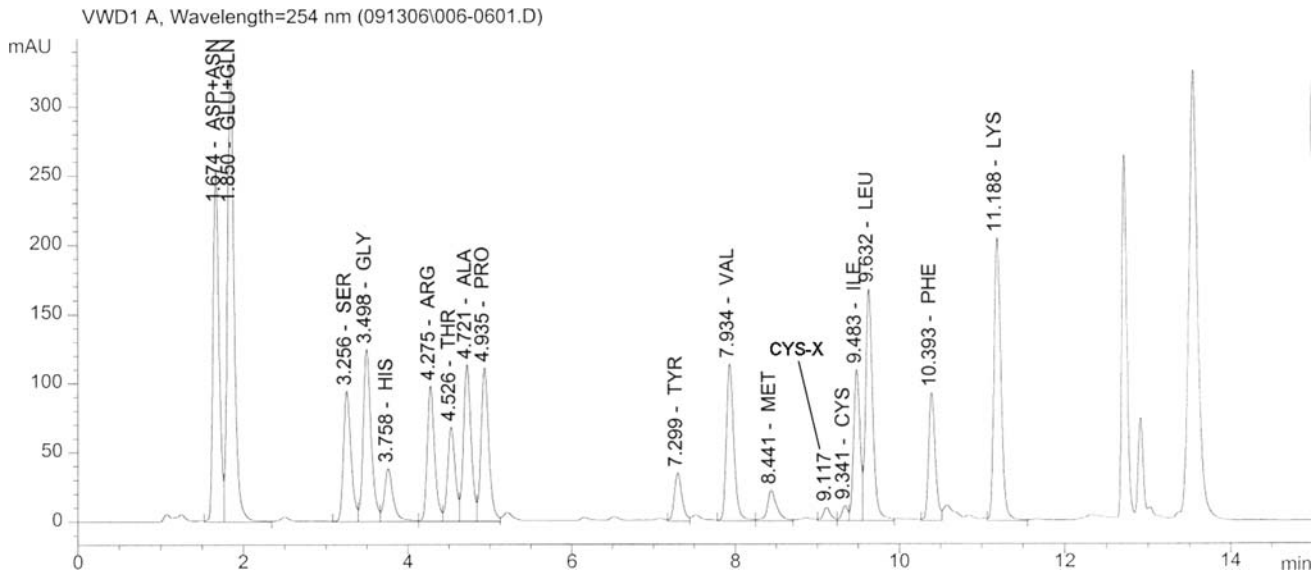

Fig. 3 Typical elution profile including retention times of amino acids from soybean samples

are known to contain very low amount of tryptophan. Therefore, tryptophan was not determined in our experiments.

Since soybean proteins especially the storage protein, glycinin, contains both cysteine and cystine, a dimer of two cysteine molecules joined by a disulfide bond, it is necessary to be able to identify and measure the half-cystine after acid hydrolysis. Cystine was subjected to hydrolysis and HPLC analysis in the same manner. Figure 2 shows the elution profile of the half-cystine, one is at a retention time corresponding to cysteine and the other a half-cystine designated as cysteine-X. This doublet is present at the same retention time in all soybean samples analyzed. These two peaks representing both cysteine and cystine were measured and therefore integrated in our amino acid determination (Fig. 3).

Accuracy and Reliability

In order to validate the accuracy and reliability of the procedure, a reference protein with known sequence, egg-white lysozyme, was subjected to this modified method. Table 2 shows the amino acid composition of lysozyme determined by this technique in comparison with numbers of residues derived from its protein sequence data.

Table 2 Amino acid composition of egg-white lysozyme determined by the modified procedure

Amino acid	Number of residues per mole	Theoretical number of residues per mole[a]
Asp + Asn	21.6	21
Glu + Gln	5.6	5
Ser	8.7	10
Gly	12.7	12
His	1.3	1
Arg	10.6	11
Thr	7.4	7
Ala	12.6	12
Pro	2.5	2
Tyr	2.9	3
Val	6	6
Met	2.2	2
Cys + Cys-X	7.7	8
Ile	5.7	6
Leu	7.8	8
Phe	3	3
Lys	5.5	6

[a] Numbers derived from known protein sequence data

We tested this modified procedure with 150 samples from 15 soybean cultivars grown at five different North Carolina locations. Differences among cultivars in mean concentrations of seed protein were highly significant. The cultivar Prichard had the highest concentration, 42.8% (dry weight basis) and N88-480, the high oil germplasm had the lowest protein content, 37.5%. The genotype by location interaction was significant. Differences among location averages were also significant. So the interaction was likely due to location differences.

The analyses have shown that statistically significant differences among cultivars were found for aspartate plus asparagine, glycine, histidine, arginine, alanine, tyrosine, and methionine (Tables 3, 4). There were no statistically significant differences among cultivars for the other amino acids. Average location differences were significant only for arginine and methionine. No cultivar by location interaction was significant for any of the amino acids which demonstrates that amino acid concentrations were generally stable across a range of environments, and that the repeatability of the measurement itself was high. The error associated with actual amino acid measurements was low as demonstrated in the analysis of variance for each amino acid by R^2 values that were generally high. Seven of the 16 R^2 values were >0.70 and seven of the remaining nine were between 0.59 and 0.69. Thus, for most amino acids, experimental error accounted for no more than 40% of the total variability among the field plots.

Even though some differences between cultivars in specific amino acids were statistically significant, the actual differences were not great. For instance, the difference in methionine content of the cultivar with the highest level and the one with the lowest level was 0.24 g/16 gN. Another amino acid in soybean that, like methionine, has been targeted for increase is lysine. The range in lysine concentration was 0.53 g/16 gN. This low level of genotypic variation in amino acid concentration is likely due to the narrow genetic base of soybean cultivars grown in the US. Only ten ancestors contribute 80% of the genes in the southern breeding programs [16]. This fact underscores the need for more research on the amino acid content of soybean protein. In particular, genetic resources outside the US cultivated germplasm pool need to be investigated. Having a reliable and efficient method for measuring

Table 3 Mean squares from the analysis of variance of seed protein (%) and amino acid (g/16 gN) concentration of 15 cultivars grown at 5 NC locations in 2005

	Degree of freedom	Protein	Asp + Asn	Glu + Gln	Ser	Gly	His	Arg	Thr
Locations	4	37.6[a]	3.44	5.25	0.18	0.07	0.07	1.00[a]	0.19
Cultivars	14	21.2[b]	0.29[a]	0.45	0.05	0.00	0.04	0.15[b]	0.04
Location × cultivar	56	1.1[a]	0.12	0.34	0.06	0.02	0.02	0.05	0.03
Pooled error	70	0.7	0.10	0.23	0.04	0.02	0.03	0.06	0.02
Average[c]		40.8	11.9	19.0	4.8	4.5	3.1	7.1	4.0

Ala	Pro	Tyr	Val	Met	Cys	Ile	Leu	Phe	Lys
0.23	0.22	0.22	0.11	0.37[a]	1.40	0.84	0.20	11.91	7.57
0.06[b]	0.04	0.07[a]	0.03	0.05[a]	0.14	0.24	0.04	0.49	0.27
0.01	0.03	0.03	0.03	0.02	0.09	0.22	0.03	0.46	0.37
0.01	0.04	0.04	0.02	0.02	0.07	0.16	0.03	0.39	0.29
4.6	5.6	2.9	5.3	1.8	1.2	4.7	8.0	5.6	6.1

[a, b] Variation is significant at the 0.05 and 0.01 probability levels, respectively

[c] Average of 15 cultivars replicated twice at five locations

Table 4 Mean protein and amino acid concentration (g/16 gN) that showed statistically significant genotypic variability among 15 soybean cultivars grown in 5 NC locations in 2005

Cultivar	Protein (%)	Asp + Asn	Gly	Ala	His	Arg	Tyr	Met
Coker 237	41.2	11.83	4.46	4.46	2.95	7.12	2.97	1.91
NC-Raleigh	38.2	11.72	4.50	4.70	3.10	7.01	3.05	1.85
Stonewall	41.5	12.21	4.57	4.54	3.17	7.09	3.04	1.84
Ransom 2	41.4	11.81	4.45	4.56	3.02	7.05	2.95	1.80
Braxton	42.4	11.94	4.50	4.48	3.02	6.97	2.87	1.77
Haskell	40.7	12.08	4.62	4.58	3.02	7.03	2.91	1.77
Ransom	41.0	11.89	4.49	4.56	3.05	7.17	2.95	1.77
Black Ransom	39.7	11.56	4.50	4.49	3.07	7.17	3.00	1.75
Johnston	42.1	11.76	4.46	4.46	3.04	7.26	2.82	1.75
GaSoy 17	40.4	11.95	4.62	4.59	3.07	6.99	2.88	1.75
Howard	41.5	12.05	4.53	4.57	2.94	6.86	2.89	1.74
Bragg	41.6	12.10	4.53	4.57	3.04	6.96	2.92	1.71
N88-480	37.5	11.89	4.65	4.74	3.17	6.82	2.72	1.68
Prichard	42.8	11.98	4.62	4.57	3.11	7.20	2.78	1.67
Jackson	40.6	12.08	4.57	4.57	3.06	7.07	2.87	1.67
$LSD_{0.05}$	1.0	0.32	0.12	0.09	0.14	0.21	0.16	0.14

amino acids in soybean protein, will move that research forward.

In general, average values of the 19 amino acids measured (Table 3) are similar to those reported in other soybean studies [1–3, 17]. Those published previously were for materials from the northern US or Canadian soybean germplasm pools. Thus, apparent small differences between those amino acid values and values reported on this paper for southern US germplasm are likely due to genetic differences although some differences due to variable amino acid analysis methods cannot be ruled out.

Conclusion

In comparison with the conventional method for determination of amino acids plus cysteine and cystine, the modified method presented above requires minimal manipulation of samples. It does not require an additional and cumbersome step of performic acid oxidation. The method is inexpensive and can be automated for HPLC analysis. It is applicable for large numbers of samples as evident in the experiment reported here. This method has the advantage of being used as a routine procedure for newly developed soybean cultivars generated from breeding programs.

Acknowledgments Financial support was provided by the United Soybean Board.

References

1. Zarkadas GG, Voldeng HD, Yu ZR, Choi VK (1999) Assessment of protein quality of nine northern adapted yellow and brown seed coated soybean cultivars by amino acid analysis. J Agric Food Chem 47:5009–5018
2. Mahmoud AA, Natarajan SS, Bennett JO, Mawhinney TP, Wiebold WJ, Krishnan HB (2006) Effects of six decades of selective breeding on soybean protein composition and quality: a biochemical and molecular analysis. J Agric Food Chem 54: 3916–3922
3. Krishnan HB, Natarajan SS, Mahmoud AA, Bennett JO, Hrishnan AH, Prasad BN (2006) Assessment of indigenous Nepalese soybean as a potential germplasm resource for improvement of protein in North American cultivars. J Agric Food Chem 54: 5489–5497
4. Heinrikson RL, Meredith SC (1984) Amino acid analysis by reverse-phase high-performance liquid chromatography: precolumn derivatization with phenylisothiocyanate. Anal Biochem 136: 65–74
5. Burton JW, Carter TE Jr, Fountain MO, Bowman DT (2006) Registration of 'NC-Raleigh' soybean. Crop Sci 13:130
6. Weaver DB (1989) Registration of 'Stonewall' soybean. Crop Sci 29:1329
7. Boerma HR, Hussey RS, Philips DV, Wood ED, Finnerty SL (1994) Registration of 'Haskell' soybean. Crop Sci 34:541
8. Brim CA, Elledge C (1973) Registration of 'Ransom' soybeans. Crop Sci 13:130
9. Burton JW, Brim CA, Young MF (1987) Registration of 'Johnston' soybean. Crop Sci 27:1093
10. Baker SH, Harris HB (1979) Registration of 'Gasoy 17' soybean. Crop Sci 19:130
11. Hinson K, Hartwig EE (1964) Bragg and Hardee soybeans. Crop Sci 4:664
12. Burton JW, Wilson RF (1994) N88–480. Soybean germplasm. Crop Sci 34:313
13. Boerma HR, Hussey RS, Phillips DV, Wood ED, Rowan GB, Finnerty SL, Griner JT (2001) Prichard soybean. Crop Sci 41:920
14. Johnson HW (1958) Registration of soybean varieties, VI. Agron J 50:690–691
15. SAS Institute Inc (1990) SAS/STAT User's guide, version 6, vol. 1, 4th edn. SAS Institute, Cary
16. Gizlice Z, Carter TE Jr, Burton JW (1994) Genetic base for North American public soybean cultivars released between 1946 and 1988. Crop Sci 34:1143–1151
17. Yazdi-Samadi B, Rinne RW, Seif RD (1977) Components of developing soybean seeds: oil, protein, sugars, starch, and organic acids, and amino acids. Agron J 69:481–486