2013

Soybean Seed Amino Acid Content QTL Detected Using the Universal Soy Linkage Panel 1.0 with 1,536 SNPs

Benjamin D. Fallen
University of Tennessee, BFALLEN@clemson.edu

Catherine N. Hatcher
Monsanto, Harrisburg, SD

Fred L. Allen
University of Tennessee, allenf@tennessee.edu

Dean A. Kopsell
University of Tennessee, dkopsell@utk.edu

Arnold M. Saxton
University of Tennessee, asaxton@utk.edu

See next page for additional authors

Journal of Plant Genome Sciences 1 (3): 68–79, 2013
doi: 10.5147/jpgs.2013.0089

Soybean Seed Amino Acid Content QTL Detected Using the Universal Soy Linkage Panel 1.0 with 1,536 SNPs

Benjamin D. Fallen[1,2*], Catherine N. Hatcher[3], Fred L. Allen[2], Dean A. Kopsell[2], Arnold M. Saxton[2], Pengyin Chen[4], Stella K. Kantartzi[5], Perry B. Cregan[6], David L. Hyten[6,7], and Vincent R. Pantalone[2]

[1] Current address: Clemson Pee Dee REC, Advanced Plant Technology Center, 2200 Pocket Road, Florence, SC 29506, USA; [2] University of Tennessee, Department of Plant Sciences, 2431 Joe Johnson Dr., Knoxville, TN 37996, USA; [3] Monsanto, 140 W. Industrial Drive, Harrisburg, SD 57032, USA; [4] University of Arkansas, Department of Crop, Soil, and Environmental Sciences, Fayetteville, AR 72701, USA; [5] Southern Illinois University, Department of Plant, Soil Science and Agricultural Systems, 1205 Lincoln Drive, Carbondale, IL 62901, USA; [6] Soybean Genomics and Improvement Laboratory, Beltsville Agricultural Research Center – West, USDA, ARS, Beltsville, MD 20705, USA; [7] Current address: DuPont Pioneer, 8305 NW 62nd Ave., PO Box 7060, Johnston, IA 50131-7060, USA.

Received: August 26, 2013 / Accepted: September 14, 2013

Abstract

Soybean [*Glycine max* (L.) Merr.] is the primary source of meal used in animal feed in the U.S. However, few studies have been conducted to evaluate genomic regions controlling amino acid composition is soybean. Designing soybean seed compositions that will benefit animal production is essential. The objective of this study was to identify genomic regions controlling essential and non-essential amino acid composition in soybean seed proteins. To achieve this objective, 282 $F_{5:9}$ recombinant inbred lines (RILs) developed from a cross of Essex × Williams 82 were used. Ground soybean seed samples were analyzed for amino acids and statistically significant differences ($p < 0.05$) were found among genotypes in the population for all amino acid concentrations. The Universal Soy Linkage Panel (USLP) 1.0 of 1,536 single nucleotide polymorphism (SNP) DNA markers were used to genotype the 282 RILs and identify 480 useful genetic markers. The software R/qtl was used to identify candidate quantitative trait loci (QTL), which were validated using R/MQM. A total of ten QTL were detected on chromosomes 5, 7, 9, 10, 13 and 20 that explained 5 to 14% of the total phenotypic variation for a particular amino acid. Using SNPs from the USLP 1.0 to detect QTL for amino acids in soybean provides additional information to select genotypes with enhanced amino acid profiles that will benefit animal production.

Keywords: Soybean meal, essential and non-essential amino acids, QTL analysis.

Abbreviations: PCA, principal component analysis; QTL, quantitative trait loci; RIL, recombinant inbred line; SNP, single nucleotide polymorphism; ETREC, East Tennessee Research and Extension Center; Ala, Alanine; Arg, Arginine; Asn, Asparagine; Asp, Aspartic acid; Cys, Cysteine; Glu, Glutamine; Gln, Glutamic acid; Gly, Glycine; His, Histidine; Iso, Isoleucine; Leu, Leucine; Lys, Lysine; Met, Methionine; Phe, Phenylalanine; Pro, Proline; Ser, Serine; Thr, Threonine; Try, Tryptophan; Tyr, Tyrosine; Val, Valine; Chr, Chromosome; TARS, Tropical Agricultural Research Station.

Introduction

Soybean [*Glycine max* (L.) Merr.] is currently used as the primary source of animal feed because of its cost effectiveness and protein content (Wilson, 2012). In the U.S. in 2012, soybean meal accounted for 68% of protein meal consumption (Soy States 2012, http://www.soystats.com/2012/Default-frames.htm). Today, the market demands are changing and if soybean meal is to remain the primary source of protein meal in animal feed it is necessary that the composition of the meal

* Corresponding author: bfallen@clemson.edu

be improved. To improve the composition of soybean meal the amino acid profile must be conformed to fit different dietary needs. Amino acids are classified as essential or nonessential depending on the animals' dietary needs. An essential amino acid is required by the animal and must be consumed daily. Non-essential amino acids can be produced and stored by the animal. However, both are important components of a healthy diet (Ufaz and Galili, 2008). Since, soybean meal is primarily used in the poultry and swine industry the essential amino acids for each are reported. The 10 essential amino acids for swine (*Sus scrofa domesticus*) are: Phe, Val, Thr, Met, Arg, Try, His, Iso, Leu and Lys (Boisen 2003). For poultry (*Gallus domesticus*) the essential amino acids are: Met, Lys, Thr, Try, Iso, Arg and Val (Baker, 2003).

Presently, very few papers are available on the genetic analysis of amino acid composition in soybean. The only QTL for amino acids on SoyBase are from Panthee et al. (2006a) and (2006b) (Grant et al., 2010; http://soybase.org/). In both studies 101 F_6-derieved recombinant inbred lines (RILs) developed from a cross of 'N87-984-16' x 'TN93-99' were screened using a total of 94 polymorphic simple sequence repeat (SSR) molecular genetic markers. Panthee et al. (2006b) identified QTL associated with Cys (chr 1, 13, and 18), Met (chr 13, 18, and 7), and Met+Cys (chr 13 and 7) concentration. Panthee et al. (2006a) also identified genomic regions associated with Ala (chr 5, 13, 14 and 19), Arg (chr 2), Asp (chr 1, 8, 9, and 20), Glu (chr 1, 2, 7, 10, 16 and 19), Gly (chr 2, 8, 9 and 19), His (chr 16), Ile (chr 1, 13 and 19), Leu (chr 1, 2, 3, 7, 17 and 19), Lys (chr 1, 15, and 18), Phe (chr 2, 9, 16 and 19), Pro (chr 18 and 19), Ser (chr 9, 7, and 19), Thr (chr 2, 5, 9, and 19), Trp (chr 1, 2, 6, 18 and 20), Tyr (chr 2, 9, 15 and 19) and Val (chr 2, 13 and 19). From the genetic mapping population used by Panthee et al. (2006a, 2006b) 'TN04-5321' was developed and released as a soybean germplasm line with significantly elevated sulfur containing amino acid levels (Panthee and Pantalone 2006). This was the first soybean line registered specifically for improved amino acid concentration.

A more recent study conducted by Warrington (2011), screened 140 F_5-derived RILs from a 'Benning' × 'Danbaekkong' cross with 98 SSR markers and 323 single nucleotide polymorphism (SNP) DNA markers. Warrington (2011) reported QTLs for Lys (chr 8 and 20), Thr (chr 9, 17, and 20), Met (chr 6, 9, 10, and 20) and for Cys (chr 10). Only a QTL detected for Thr was reported on the same chromosome (chr 9) by Panthee et al. (2006a) and Warrington (2011). However, the QTL were > 30 cM apart.

In order to efficiently develop soybean cultivars with improved amino acid profiles, the genetic basis of amino acid composition of the seed should be explored thereby allowing for marker assisted selection (MAS) of desired amino acids for improved protein quality. The objective of this study was to use the 1,536 SNP DNA markers of the Universal Soy Linkage Panel 1.0 (USLP 1.0) (Hyten et al., 2010) in an analysis of a soybean mapping population segregating for seed amino acid composition to identify genomic regions controlling essential and non-essential amino acid composition in soybean seed.

Materials and Methods

Population Development

A RIL population was created from a cross of the cultivars 'Essex' x 'Williams 82'. Essex originated from the cross 'Lee' × 'S5-7075' at the Virginia Agricultural Experiment Station and was released in 1972 (Smith and Camper, 1973). Essex is characterized as having purple flowers, gray pubescence, a group V maturity, average protein, oil, height and yield and is moderately susceptible to sudden death syndrome (SDS) (Fusarium solani f. sp. glycines nee F.virguliforme Aoki; Lightfoot et al 2005; Yesudas et al., 2013). Williams 82 was developed by the USDA-ARS and the Illinois Agricultural Experiment Station by combining four individual BC_6F_3 plants selected after a series of backcrosses to 'Williams' to transfer the *Rps1k* gene from Kingwa (Bernard and Cremeens, 1988). The *Rps1k* allele confers resistances to certain races of Phytophthora sojae, which causes phytophthora root rot. Williams 82 is characterized as having white flowers, tawny pubescence, a group III maturity, average seed protein and oil content, resistance to phytophthora root rot and moderate resistance to SDS (Gibson, 1994). Williams 82 has contributed to the genetic background of many northern U.S. cultivars and Essex has contributed to the genetic background of many southern U. S. cultivars and elite breeding lines (Sneller, 2004; Gizlice et al., 1996). A population formed from these diverse parents should reflect a broad measure of the range of amino acids available in elite U.S. soybean cultivars. Therefore, QTL detected in this population are likely to be segregating in a wide range of North American breeding programs.

The initial crosses for the Essex × Williams 82 population were made at the East Tennessee Research and Extension Center (ETREC) in Knoxville, TN in the summer of 2005. In the fall of 2005, the F_1 seeds obtained from the Essex × Williams 82 cross were harvested and grown in Isabela, PR at the USDA-ARS Tropical Agricultural Research Station (TARS). The population was advanced from the F_2 to the F_5 generation through single seed descent (Brim 1966). The F_2 generation was grown at ETREC in 2006 and the F_3 generation was grown at ETREC in 2007. The F_4 and F_5 generations were grown at the TARS location in the winter of 2007/2008 and the spring of 2008, respectively. In the summer of 2008, 284 individual $F_{5:6}$ RILs were planted in 3.1m single plant rows at ETREC. From each row, leaf tissue was collected for DNA extraction and agronomic data was recorded.

Experimental Field Procedures

In 2009, yield trials were conducted using the $F_{5:7}$ recombinant inbred lines. Three population subsets: early (94 genotypes, four checks and the two parents), mid (94 genotypes, four checks and the two parents) and late (94 genotypes, four checks and the two parents) were planted in two 6.1m row plots in a randomized complete block design replicated three times in Knoxville, TN, Harrisburg, IL and Stuttgart, AR. Checks were assigned by maturity group. In the early test 'IA4004', LD00-2817P, LD00-3309 and 'Macon' were used as checks. In the mid test TN05-4008, TN06-189, TN06-196 and '5002T' were

used as checks. In the late test JTN-5203, 'Osage', '5002T' and '5601T' were used as checks. At maturity, plant height was taken as an estimation of the distance from the soil surface to the tip of the main stem. Lodging was scored on a scale from 1-5; with 1 being all the plants in the plot erect and 5 being all the plants in a plot prostrate. Maturity was recorded when 95% of the pods achieved their mature color. Seed yield was estimated after the plots had been end trimmed to 4.9m in length. Seed yield was obtained by an onboard seed spectrometer (Almaco, Nevada, IA) and was adjusted to a 13% moisture basis. Seed size was taken as the weight in grams from a random 100 seed sample.

Laboratory Procedures

Sample Preparation for Amino Acid Composition via NIR Analysis

Approximately 20 g of soybean seed collected from plot samples were ground in a water-cooled Knifetec 1095 Sample Mill (FOSS Tecator, S-26321, Hogana, Sweden) for 20s. This produced soybean flour that was uniform in particle size. The samples were analyzed using a FOSS 6500 near infrared spectrometer (NIR). A dehumidifier was used throughout the analysis to reduce the humidity to 40%, and room temperature was maintained at approximately 20°C.

Initially the NIR was warmed up for 2h after turning on the lamp. Auto diagnostics were run for instrument response, wavelength accuracy and NIR repeatability. Ground soybean samples were scanned to obtain the predicted concentrations of oil and protein, and 18 amino acids Ala, Arg, Asp, Cys, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val using ISIscan (System II version 2.80 software (FOSS, State College, PA). The instrument was left on for the whole period of analyses, and diagnostics was performed every day until the scanning was finished. Each amino acid sample was expressed as a percentage of overall crude protein content to report values as grams of the amino acid per kilogram of crude protein. The NIRS prediction equation was developed using modified partial least squares (Shenk and Westerhaus, 1991) with the Infrasoft International NIRS 3 ver 3.0 software program (ISI, Port Matilda, Mass.) Equations were developed at the University of Minnesota as described in Panthee et al. (2006a).

Genotypic Data

A leaf was collected from 10 $F_{5:6}$ plants from each F5:6¬ RIL grown at ETREC in the summer of 2008 and DNA was extracted from the combined 10 leaf sample and processed to contain 50µl of DNA at a 200 ng/µl concentration. The samples were then sent to the Soybean Genomics and Improvement Laboratory at the USDA-ARS in Beltsville, MD, where a total of 1,536 SNP markers were assayed on each RIL genotype using the USLP 1.0 (Hyten et al. 2010), which employs the Illumina GoldenGate® assay and the resulting fluorescence data were analyzed on the Illumina BeadStation 500G (Illumina, San Diego, CA) (Hyten et al., 2008).

Statistical Analyses

Analysis of variance and LSD mean separation were conducted in SAS using PROC MIXED (SAS ver. 9.1.3, Cary, NC) to test for significant genotype differences among RIL for amino acid concentrations. Genotype was considered a fixed effect and all other factors were considered random. In addition, phenotypic data was analyzed using PROC MIXED to estimate genetic, environment and genetic by environment variances. Genotype, replication and environment were all considered random effects. For both PROC MIXED statements maturity date within environment was included as a covariate in the model to reduce its influence on phenotypic values of all traits. Relationships among the 18 amino acids were analyzed using PROC CORR and principal component analysis was performed using PRINCOMP in SAS version 9.1.3 (SAS ver. 9.1.3, Cary, NC). Restricted maximum likelihood analysis (REML) was used to estimate variance components for calculating heritability estimates. The REML estimation was performed by including METHOD=REML as an option in the PROC MIXED statement. Heritability was estimated to determine the fraction of phenotypic variation among individuals that was due to genetic differences. A broad sense estimate of heritability of the amino acid concentrations in the population was calculated on an entry mean basis (Nyquist, 1991) using the equation presented by (Panthee et al., 2006b). This estimate primarily includes additive effects because inbred lines ($F_{5:9}$) were used. Thus, the estimate functionally provides a narrow sense heritability estimate.

Marker order, position and composite interval mapping (CIM) were completed using R/qtl (Broman and Sen, 2009). For CIM the threshold of significance for each marker was established using an experiment wise Type 1 error rate of P = 0.05, determined using 1000 permutations. In addition, Multiple-QTL Mapping (MQM) was used to confirm QTL found by R/qtl (Broman and Sen, 2009). MQM combines generalized linear model regression with CIM, which allows markers to be used as cofactors. Markers were selected as cofactors by using multiple regression and backwards elimination and only the markers with a LOD score above 2.0 were used as cofactors. QTL analysis was performed across each environment and maturity group.

Results and Discussion

Variation in Amino Acid Composition

There were statistically significant differences ($p<0.001$) among the RILs for all essential and non-essential amino acids tested in this study (Table 1). Little variation in amino acid concentrations was observed across environments and maturity groups (Table 2). Deficiency in sulfur-containing amino acids, Met and Cys is the main limitation of soy protein for animal feed (Ufaz and Galili, 2008). In this study the difference between the population mean and the population maximum for Cys was 2.6g kg^{-1} crude protein (Figure 1), representing a 14% increase at the upper extreme (Table 1); similarly the difference between the population mean and the population maximum for Met was 1.5g kg^{-1} crude protein (Figure 2), representing a 8% increase

Table 1. Estimates and significance of genetic (g), environment (e), and genotype by environment (gxe) variance components; means, minimum and maximum of 282 F5:9-derived RILs of Essex x Williams 82; as well as heritability estimates and parental means for essential and non-essential amino acid concentrations in soybean seed grown in Knoxville, TN, Stuttgart, AR, and Harrisburg, IL in 2009.

Trait	Var (g)	Var (e)	Var (gxe)	Pop. Min	Pop. Mean	Pop Max.	Essex	Williams82	$LSD_{0.05}$	h^2 (%)
					(g kg^{-1} crude protein)					
Essential amino acids										
Ile	0.05	0.01	0.09	51.8	53.7	55.5	54.7	54.2	0.7	72.5
Leu	0.58	0.13	0.79	74.8	82.1	89.4	85.8	80.6	1.7	65.2
Lys	1.71	3.80	2.55	44.1	65.3	73.3	68.5	52.3	2.4	39.7
Met	0.04	0.04	0.02	16.2	17.9	19.3	16.4	17.8	0.2	67.7
Phe	0.01	0.06	0.16	55.5	57.8	60.1	59.2	58.5	0.9	74.2
Thr	0.32	0.16	0.34	43.8	49.5	54.98	50.6	49.6	0.9	63.3
Trp	0.12	0.11	0.07	7.4	12.7	14.7	12.8	11.8	0.2	71.3
Tyr	0.14	0.01	0.19	41.5	45.8	49.3	46.7	46.1	0.7	70.9
Val	0.33	0.75	0.76	59.6	66.3	74.9	67.5	68.3	1.7	63.7
His	0.46	2.99	1.74	29.1	39.1	62.2	39.4	56.1	1.7	31.2
Non-essential amino acids										
Ala	0.08	0.02	0.06	52.4	56.7	62.2	57.7	57.9	1.2	55.3
Arg	0.54	0.47	0.77	75.6	85.9	91.31	88.7	83.1	1.7	72.1
Asp	0.43	2.59	0.54	125.8	131.3	143.2	134.5	124.9	2.1	69.8
Cys	0.15	0.04	0.14	15.3	18.1	20.7	15.6	18.5	0.5	63.2
Glu	1.43	0.12	2.24	167.6	186.2	193.5	192.1	177.2	4.7	64.7
Gly	1.15	3.89	2.71	58.9	67.3	81.6	68.1	79.3	2.6	47.8
Pro	0.13	0.20	0.18	56.5	61.8	66.3	63.3	61.5	1.2	68.4
Ser	1.04	1.63	1.08	53.9	64.1	74.4	66.2	62.4	2.1	61.5

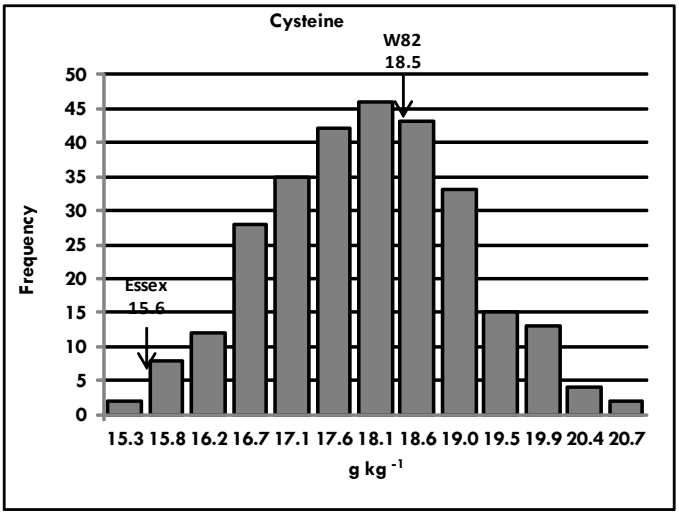

Figure 1. Histogram of cysteine concentrations (g kg^{-1} crude protein) from 282 $F_{5:9}$-derived RILs of Essex x Williams 82 grown in Knoxville, TN, Stuttgart, AR, and Harrisburg, IL in 2009.

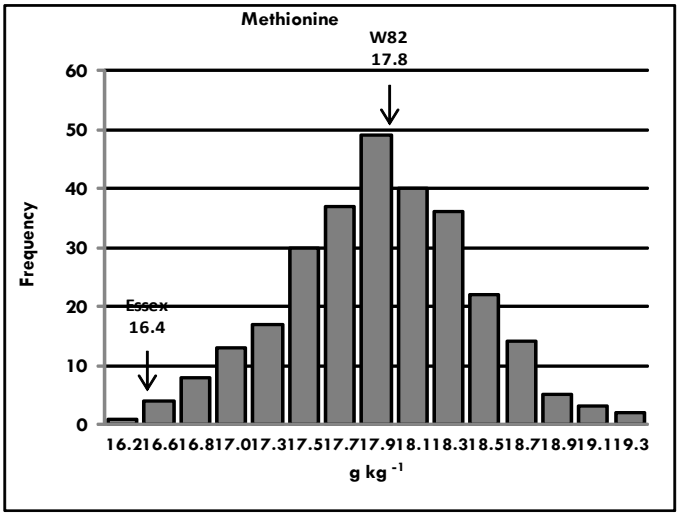

Figure 2. Histogram of methionine concentrations (g kg^{-1} crude protein) from 282 $F_{5:9}$-derived RILs of Essex x Williams 82 grown in Knoxville, TN, Stuttgart, AR, and Harrisburg, IL in 2009.

at the upper extreme (Table 1). For Cys the maximum variation averaged across environments and maturity groups was 1.1g kg^{-1} crude protein for the RIL population, 2.3g kg^{-1} crude protein for the parents (Essex and Williams 82) and 2.2g kg^{-1} crude protein for the checks. The average maximum variation for Met averaged across environments and maturity groups was 0.8g kg^{-1} crude protein for the RIL population, 1.7g kg^{-1} crude protein for the parents (Essex and Williams 82) and 1.1g kg^{-1} crude protein for the checks (Table 2). The modest amount of variation (Table 1) and the stability of the amino acid concentrations among the RILs across environments and maturity groups (Table 2) suggest that modest genetic gains can be made in soybean, including genetic gains for Cys and Met. For Cys and Met it has been reported that only a slight increase (~0.5g kg^{-1} crude pro-

Table 2. Mean amino acid concentration (g kg-1 crude protein) of 282 F5:9-derived RILs from Essex x Williams 82 that were grown in Knoxville, TN, Stuttgart, AR, and Harrisburg, IL in 2009.

Mat	Loc	Line	†Ala	†Arg	†Asp	†Cys	†Glu	†Gly	†His	†Pro	†Ser	‡Leu	‡Lys	‡Ile	‡Met	‡Phe	‡Thr	‡Trp	‡Tyr	‡Val
			g kg^{-1} crude protein																	
Early	AR	CK§	57.8	85.9	131.4	17.0	189.4	70.6	42.1	62.4	64.5	83.0	59.5	53.7	17.4	57.7	47.9	12.9	45.2	66.1
		PR¶	58.8	86.0	133.2	17.7	182.7	75.5	49.4	63.5	67.1	81.1	50.3	53.8	17.6	57.4	49.1	12.7	45.7	68.5
		RIL	58.4	88.3	134.5	18.0	186.6	72.3	42.6	63.7	67.2	82.5	61.4	54.7	18.4	58.7	50.0	12.8	46.4	69.3
	IL	CK§	57.4	86.7	132.4	18.2	199.8	64.3	33.9	62.3	62.4	86.1	73.1	54.9	17.6	59.2	50.1	13.1	46.4	64.1
		PR¶	55.7	87.7	131.7	17.4	195.0	63.2	34.5	61.7	61.1	84.3	70.6	54.5	17.1	58.8	48.0	12.6	45.4	64.9
		RIL	57.2	88.3	132.2	18.3	190.9	66.2	36.3	62.6	63.9	83.5	70.1	54.9	18.0	58.9	50.7	12.8	46.8	66.8
	TN	CK§	56.5	86.3	131.5	17.7	188.2	67.3	38.3	62.0	64.0	81.8	64.3	53.7	17.7	57.9	49.0	12.7	45.4	66.2
		PR¶	55.8	86.5	133.0	17.9	191.4	65.9	37.8	60.9	63.5	83.1	66.3	53.7	17.8	58.4	48.8	12.7	45.2	65.7
		RIL	56.9	86.3	132.0	17.8	186.0	68.3	39.2	62.0	64.7	81.5	63.9	53.8	17.9	58.0	49.5	12.8	45.9	67.0
Mid	AR	CK§	57.3	84.5	135.3	19.2	198.6	65.2	37.5	62.4	63.6	86.7	71.1	54.3	18.0	59.5	48.9	11.5	45.5	65.1
		PR¶	58.9	86.1	137.4	19.0	188.3	73.1	48.4	63.0	65.5	83.9	56.5	54.7	18.3	59.5	50.4	12.2	46.7	69.6
		RIL	57.8	86.1	134.6	18.7	189.3	69.4	41.8	63.1	65.1	83.9	64.4	54.5	18.2	59.2	50.1	12.3	46.7	68.1
	IL	CK§	57.3	88.1	132.5	18.5	192.9	65.5	35.0	62.8	64.3	84.4	72.3	54.5	18.1	59.0	50.8	12.3	46.3	66.0
		PR¶	59.3	88.6	133.3	18.4	189.1	69.7	38.6	64.1	66.5	83.6	69.2	55.4	18.4	59.3	52.8	12.4	48.5	68.7
		RIL	56.9	87.3	130.7	18.1	188.2	65.8	37.3	62.0	63.4	82.9	69.2	54.3	17.8	58.1	50.5	12.4	46.4	66.3
	TN	CK§	57.0	85.3	130.8	18.0	189.6	67.1	35.6	62.5	65.2	82.6	68.6	53.0	18.0	57.6	49.5	13.2	44.7	65.3
		PR¶	58.9	87.2	134.6	18.5	185.9	72.3	40.5	64.2	68.7	82.2	63.7	54.5	18.7	58.6	51.3	13.2	46.8	69.1
		RIL	58.1	86.8	133.6	18.4	187.3	70.1	39.6	63.2	66.3	82.7	65.1	54.3	18.3	58.4	50.8	13.1	46.3	68.0
Late	AR	CK§	56.6	85.7	131.7	18.7	189.5	66.2	37.2	62.3	64.8	83.1	67.9	53.1	18.1	57.7	49.0	12.0	45.0	64.7
		PR¶	56.5	81.6	138.2	19.7	194.5	67.1	50.8	60.9	60.8	85.9	53.6	54.3	17.6	59.5	48.1	10.9	45.1	67.0
		RIL	56.8	85.3	134.2	18.8	191.1	66.4	40.6	62.0	63.7	84.2	65.7	53.9	18.1	58.9	49.2	11.9	45.8	66.4
	IL	CK§	55.9	88.4	129.8	17.6	187.0	65.0	34.7	61.5	63.9	82.1	70.3	53.7	17.8	57.7	49.7	12.0	45.5	65.6
		PR¶	56.2	88.0	133.1	18.4	193.0	63.5	37.0	61.6	61.5	84.8	71.0	55.1	17.8	59.2	50.2	12.6	46.4	66.0
		RIL	55.9	87.6	129.9	17.7	186.5	64.6	36.6	61.2	62.5	82.2	68.8	54.0	17.6	57.9	49.7	12.2	46.0	65.9
	TN	CK§	58.9	86.6	132.3	18.5	188.5	71.0	39.4	63.6	67.6	82.6	66.9	54.1	18.5	58.2	51.1	13.0	46.5	67.9
		PR¶	58.7	86.2	134.7	18.7	186.1	72.4	42.9	62.9	67.3	82.5	61.1	54.0	18.8	58.2	50.5	12.5	46.0	69.5
		RIL	58.7	87.2	134.3	18.4	190.0	70.6	42.0	63.2	66.3	83.6	64.8	54.6	18.4	59.0	50.9	13.0	46.8	68.8

†Essential amino acids; ‡Non-essential amino acids; §Checks assigned by maturity group. Early 'IA4004', LD00-2817P, LD00-3309 and 'Macon'. Mid TN05-4008, TN06-189, TN06-196 and '5002T'. Late JTN-5203, 'Osage', '5002T' and '5601T'; ¶Parents (Essex and Williams 82).

Table 3. Simple phenotypic correlation coefficients between amino acids and agronomic traits in soybean seed in 282 F5:9-derived RILs of Essex x Williams 82 grown in Knoxville, TN, Stuttgart, AR, and Harrisburg, IL in 2009.

	Ala	Arg	Asp	Cys	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val	Protein	Maturity	Height
Arg	0.78																				
Asp	0.85	0.82																			
Cys	0.63	0.42	0.77																		
Glu	0.41	0.63	0.76	0.49																	
Gly	0.96	0.69	0.72	0.52	0.20																
His	0.84	0.57	0.79	0.70	0.43	0.78															
Iso	0.85	0.92	0.94	0.59	0.73	0.72	0.73														
Leu	0.57	0.71	0.88	0.63	0.95	0.37	0.59	0.84													
Lys	0.41	ns	ns	-0.23	0.31	-0.53	-0.61	ns	0.17												
Met	0.92	0.78	0.88	0.76	0.45	0.86	0.79	0.84	0.63	-0.29											
Phe	0.8	0.86	0.97	0.67	0.80	0.66	0.72	0.96	0.90	ns	0.88										
Pro	0.96	0.87	0.87	0.59	0.49	0.91	0.77	0.91	0.63	-0.27	0.91	0.86									
Ser	0.93	0.74	0.72	0.52	0.21	0.96	0.72	0.73	0.38	-0.43	0.89	0.67	0.92								
Thr	0.93	0.75	0.73	0.55	0.21	0.93	0.74	0.77	0.39	-0.39	0.88	0.70	0.91	0.93							
Trp	0.51	0.54	0.41	0.05	0.37	0.51	0.39	0.51	0.33	-0.16	0.41	0.44	0.53	0.47	0.42						
Tyr	0.92	0.81	0.79	0.52	0.36	0.89	0.75	0.85	0.53	-0.30	0.86	0.80	0.93	0.89	0.94	0.45					
Val	0.97	0.84	0.86	0.59	0.43	0.92	0.80	0.89	0.59	-0.32	0.91	0.84	0.97	0.90	0.94	0.50	0.94				
Protein	0.67	0.53	0.81	0.56	0.72	0.32	0.53	0.37	0.41	-0.31	0.47	0.34	0.58	0.62	0.29	0.43	0.63	0.38			
Maturity	ns	0.11	ns	ns	0.10	ns	ns	ns	0.16	ns	ns	ns	ns	ns	0.12	ns	ns	ns	0.15		
Height	ns	0.16	ns	0.14	0.17	0.23	-0.18	0.18	0.12	0.22	0.16	ns	0.19	0.12	0.22	0.28	0.24	0.12	0.25	0.40	
Yield	-0.24	-0.20	-0.23	-0.18	-0.16	ns	-0.18	ns	-0.26	-0.19	-0.17	-0.16	-0.21	-0.16	-0.10	-0.17	ns	-0.21	-0.32	0.33	ns

All values were significant at $p < 0.01$.

Table 4. Principal components obtained using amino acid concentrations in soybean seed of 282 F5:9-derived RILs of Essex x Williams 82 grown in Knoxville, TN, Stuttgart, AR, and Harrisburg, IL in 2009.

Principal Component	Eigenvalue	Difference	Proportion	Cumulative
1	12.7	10.25	0.6733	0.6733
2	2.53	1.09	0.133	0.8066
3	1.43	0.60	0.606	0.8821

Table 5. Eigenvectors for principal components obtained using amino acid concentrations of soybean seed in 282 F5:9-derived RILs of Essex x Williams 82 grown in Knoxville, TN, Stuttgart, AR, and Harrisburg, IL in 2009.

Amino Acid	Principal Component 1	Principal Component 2	Principal Component 3
Tyr	0.25	-0.11	0.03
Val	0.27	-0.09	0.04
Asp	0.26	0.17	-0.08
Cys‡	0.19	0.08	-0.43‡
Glu†	0.16	0.47†	0.05
Gly	0.24	-0.26	0.04
His	0.23	-0.11	-0.19
Iso	0.26	0.15	0.08
Leu†	0.20	0.40†	-0.05
Lys†	-0.07	0.49†	0.18
Met	0.26	-0.05	-0.09
Phe	0.25	0.22	-0.01
Pro	0.27	-0.05	0.06
Ser	0.24	-0.22	0.03
Thr	0.25	-0.21	0.01
Trp‡	0.14	-0.02	0.50‡
Ala	0.27	-0.13	0.01
Arg	0.24	0.12	0.24

† Signifies the amino acids that define principal component 2
‡ Signifies the amino acids that define principal component 3

tein) can lead to significant improvements in poultry and swine diets (Baker, 2003; Boisen, 2003). Panthee et al. (2006a) reported a positive correlation between amino acid composition and total protein content, except for Lys and Arg. In this study, as total protein content increased or decreased with environmental factors so did the individual amino acid concentrations tested, except for Lys. Lys had a negative correlation with total protein content (Table 3). These results along with the findings from other studies (Panthee et al., 2006a; Warrington, 2011; Carlson, 2011) suggest an environmental change may effect total protein content and therefore change the total amino acid concentration. However, the overall proportion of amino acid composition remains the same.

Correlation between Amino Acids and Agronomic Traits

To examine the relationship among 18 amino acids in soybean, phenotypic correlations were determined using PROC CORR in SAS version 9.1.3 (SAS Institute, 2003). Almost all the amino acids were positively correlated (r = 0.17 to 0.97) (Table 3). However, Lys was shown to have a weak to moderately negative correlation with ten amino acids and a weak to moderately positive correlation with three amino acids. Lys had a weak negative correlation with Cys (r = -0.23), Met (r = -0.29), Pro (r = -0.27), Trp (r = -0.16) and Tyr (r = -0.30). Lysine had a moderately negative correlation with Gly (r = -0.53), His (r = -0.61), Ser (r = -0.43), Thr (r = -0.39), Val (r = -0.32) and Protein (-0.31) (Table 3). In addition, Lys had a weak positive relationship with Leu (r = 0.17) and a moderately positive relationship with Ala (r = 0.41) and Glu (r = 0.31) (Table 3). Panthee et al (2006a) reported moderately negative correlations between Lys and six of the same amino acids reported in that study and weak negative correlation between Lys and total protein. Panthee et al. (2006a) reported Lys had a moderately negative relationship with Gly (r = -0.56), Pro (r = -0.29), Ser (r = -0.36), Thr (r = -0.46), Tyr (r = -0.52) and Val (r = -0.55). Lys is essential in the swine and poultry diet, as well as many other animal diets (Baker 2003; Boisen 2003). Breeding for increased Lys may be difficult due to the inverse relationship with total protein and other essential amino acids.

As mentioned earlier, a major limitation of soy proteins is their deficiency of sulfur-containing amino acids, Met and Cys. This deficiency results in the use of either synthetic or natural supplementary ingredients to fulfill the requirement of Met in soy based animal feed. However, this process can result in leaching and bacterial degradation leading to formation of undesirable volatile sulfides (George and de Lumen, 1991). In this study a strong positive correlation was seen between Met and Cys (r = 0.76) (Table 3). A moderate to strong positive correlation was also seen between Met, Cys and all other amino acids reported in this study (r = 0.45 to 0.92) except for a weak negative correlation between Lys and Cys (r = -0.23) and Lys and Met (r = -0.29) (Table 3). Panthee et al. (2006b) reported a moderate positive correlation between Cys and Met (r = 0.41). A moderate positive correlation was reported in their study between Cys, Arg, Phe, His, Trp, Thr and Ser and a moderate positive correlation was reported between Met, Arg, Pro, Phe, His, and Trp. The only amino acid they found in both the swine and poultry diet that had a negative correlation with Cys and Met was Val (r = -0.22 and r = -0.05, respectively) (Panthee et al. 2006b). These results suggest increasing Cys and Met content in soybean will not adversely affect other amino acids concentrations needed in swine and poultry diets.

In this study Arg (r = 0.11), Glu (r = 0.10), Leu (r = 0.16) and Thr (r = 0.12) had a weak positive relationship with maturity (Table 3). All other amino acids did not have a statistically significant relationship with maturity. Though maturity did have a statistically significant relationship with four amino acids, the correlations between amino acid concentration and maturity tended to be relatively small in comparison to correlations between maturity and other agronomic characteristics such as yield

Table 6. Quantitative trait loci for essential and non-essential amino acid concentration in soybean seed identified using R/qtl and R/MQM in 282 F5:9-derived RILs of Essex x Williams 82 grown in Knoxville, TN, Stuttgart, AR, and Harrisburg, IL in 2009.

MARKER(S)	TRAIT	CHR	MLG	LOC (cM)	95% CI‡	ENV§	LOD	R^2 (%)	ADDITIVE EFFECT†
ss107923612	Ala	5	A1	145.54	141.23-149.63	AR, IL	3.21	5.5	0.02 (E)
ss107917837	Ala	13	F	4.86	2.45-7.26	ALL	3.96	6.9	0.03 (W)
ss107920654/ss107924336	Ala	13	F	40.69	35.91-49.95	AR, TN	3.81	6.6	0.02 (W)
ss107917837	Arg	13	F	4.86	2.45-7.26	ALL	3.14	5.7	0.05 (W)
ss107912633/ss107918763	Arg	13	F	21.51	15.45-27.35	AR, TN	3.46	6.1	0.06 (W)
ss107920654/ss107924336	Arg	13	F	40.69	35.91-49.95	AR, TN	4.49	7.9	0.06 (W)
ss107928831/ss107926274	Asp	7	M	71.31	63.21-79.53	AR	2.97	5.5	0.06 (W)
ss107913002	Asp	9	K	62.54	52.33-71.22	ALL	3.75	5	0.07 (E)
ss107912627	Asp	9	K	86.91	78.23-95.65	IL, TN	2.92	7.4	0.06 (E)
ss107917837	Asp	13	F	4.86	2.45-7.26	ALL	3.07	5.5	0.07 (W)
ss107920654/ss107924336	Asp	13	F	40.69	35.91-49.95	AR, TN	3.46	6.1	0.07 (W)
ss107929220/ss107914151	Cys	20	I	133.42	123.23-142.75	ALL	2.94	6	0.01 (W)
ss107924237	Glu	19	L	116.09	105.16-123.59	ALL	3.32	13.8	0.08 (E)
ss107917837	Gly	13	F	4.86	2.45-7.26	ALL	3.72	6.4	0.05 (W)
ss107920654/ss107924336	Gly	13	F	40.69	35.91-49.95	AR, TN	2.95	5	0.05 (W)
ss107920438/ss107912744/ ss107919004	His	10	O	110.18	103.52-118.11	ALL	3.98	7.4	0.04 (E)
ss107917837	Iso	13	F	4.86	2.45-7.26	ALL	3.57	6.3	0.03 (W)
ss107912633/ss107918763	Iso	13	F	21.51	15.45-27.35	AR, TN	3.18	5.4	0.02 (W)
ss107920654/ss107924336	Iso	13	F	40.69	35.91-49.95	AR, TN	5.01	8.9	0.04 (W)
ss107913002	Leu	9	K	62.54	52.33-71.22	ALL	5.57	5.7	0.04 (E)
ss107912627	Leu	9	K	86.91	78.23-95.65	IL, TN	4.08	10.7	0.03 (E)
ss107917837	Met	13	F	4.86	2.45-7.26	ALL	2.97	5.2	0.01 (W)

†Additive effect refers to the quantitative change in amino acid composition that is associated with either (E) Essex or (W) Williams82.
‡95% Confidence interval of QTL position.
§Environment(s) in which QTLs were found to be statistically significant.

Table 6. Continued.

MARKER(S)	TRAIT	CHR	MLG	LOC (cM)	95% CI‡	ENV	LOD	R2 (%)	ADDITIVE EFFECT†
ss107917837	Phe	13	F	4.86	2.45-7.26	ALL	3.16	5.7	0.03 (W)
ss107912633/ss107918763	Phe	13	F	21.51	15.45-27.35	AR, TN	3.33	5.7	0.03 (W)
ss107920654/ss107924336	Phe	13	F	40.69	35.91-49.95	AR, TN	4.23	7.5	0.03 (W)
ss107917837	Pro	13	F	4.86	2.45-7.26	ALL	3.9	7	0.04 (W)
ss107912633/ss107918763	Pro	13	F	21.51	15.45-27.35	AR, TN	3.52	6	0.03 (W)
ss107920654/ss107924336	Pro	13	F	40.69	35.91-49.95	AR, TN	5.42	9.5	0.04 (W)
ss107917837	Ser	13	F	4.86	2.45-7.26	ALL	3.67	6.5	0.04 (W)
ss107912657/ss107913658	Ser	13	F	21.51	15.45-27.35	AR, TN	3.27	5.6	0.04 (W)
ss107920654/ss107924336	Ser	13	F	40.69	35.91-49.95	AR, TN	3.07	5.2	0.04 (W)
ss107917837	Thr	13	F	4.86	2.45-7.26	ALL	3.72	6.6	0.02 (W)
ss107920654/ss107924336	Trp	13	F	40.69	35.91-49.95	AR, TN	3.81	6.5	0.02 (W)
ss107920438/ss107912744/ ss107919004	Tyr	10	O	110.18	103.52-118.11	ALL	3.1	5.7	0.02 (E)
ss107917837	Tyr	13	F	4.86	2.45-7.26	ALL	3.75	6.7	0.02 (W)
ss107912657/ss107913658	Tyr	13	F	21.51	15.45-27.35	AR, TN	3.02	5	0.02 (W)
ss107920654/ss107924336	Tyr	13	F	40.69	35.91-49.95	AR, TN	3.46	6	0.02 (W)
ss107923612	Val	5	A1	145.54	141.23-149.63	AR, IL	3.38	6	0.05 (E)
ss107917837	Val	13	F	4.86	2.45-7.26	ALL	4.18	7.4	0.05 (W)
ss107912633/ss107918763	Val	13	F	21.51	15.45-27.35	AR, TN	2.97	5	0.04 (W)
ss107920654/ss107924336	Val	13	F	40.69	35.91-49.95	AR, TN	4.5	7.9	0.05 (W)

†Additive effect refers to the quantitative change in amino acid composition that is associated with either (E) Essex or (W) Williams82.
‡95% Confidence interval of QTL position.
§Environment(s) in which QTLs were found to be statistically significant.

and height. Carlson (2011) reported similar results where maturity had only a weak correlation with amino acid concentration. These results suggest maturity does not have a significantly large effect on breeding and selection for amino acid composition.

Principal Component Analysis

Though the metabolic pathways for the biosynthesis of amino acids are well understood, literature regarding the elucidation of genetic control of variation of amino acid content in soybean is limited. To further understand the relationship of amino acids in soybean, a principal component analysis (PCA) was conducted on all 18 amino acids. Using PCA, 18 amino acids were reduced to 3 principal components that explained 88.2% of the observed phenotypic variation (Tables 4, 5). Almost all amino acid concentrations contributed to PC1 and all the amino acids in PC1 had a moderately positive correlation with protein (Tables 3, 5). This suggests selection of any one the amino acids comprising PC1 will result in an increase in all amino acids in PC1 and an increase in total protein. This was also observed by Panthee et al. (2006a) who reported the same amino acids identified in PC1 had a positive correlation with total protein content. The only two amino acids to have a negative correlation with total protein content were Lys and Arg (Panthee et al., 2006a). In this study Lys was the only amino acid to have a moderately negative correlation with total protein and did not contribute to PC1. Glu, Lys and Leu concentrations mainly contributed to PC2 and Cys and Trp concentrations mainly contributed to PC3 (Table 5). The amino acid pathway was examined to understand why these amino acids had a greater effect on amino acid composition than just breeding for the amino acids in PC1 and to relate these findings to the biochemistry of amino acid production. Amino acids are classified into five families: the serine-glycine family (which also includes Cys) derived from 3-phosphoglycerate, the family of aromatic amino acids (which includes Tyr, Phe, and Trp) derived from phosphoenolpyruvate, the alanine-valine-leucine family derived from pyruvate, the aspartate family (which includes Thr, Lys, Met and Ile) derived from oxaloacetate, and the glutamate family (which includes Glu, Pro, Arg and His) derived from alpha-ketoglutarate (Taiz and Zeiger 2006). PC2 contained one amino acid from each of the last three families and PC3 contained one amino acid from each of the first two families (Table 5). The results from this study suggest breeding and selection for the amino acids in PC2 and PC3 would have the greatest impact on altering amino acid composition in soybean. Lys, Glu and Cys have all been reported to have a significant effect on amino acid production in plants. The negative correlation between Lys and other amino acids and total protein content has already been discussed (Carlson, 2011; Warrington, 2011; Panthee et al., 2006a). Also, the importance of the sulfur containing amino acid Cys has been discussed and will be mentioned again later (Panthee et al., 2006b; Brosnan and Brosnan, 2006). Glu has been reported to provide the source of nitrogen to synthesize several amino acids in plants and may also have roles in homeostasis (Forde and Lea, 2007). Further evaluation is needed to determine which amino acids have the greatest effect on amino acid composition and total protein content. An improved understanding of plant amino acid pathways would make it possible to engineer increased amino acid content not only using classical plant breeding, but also transgenic approaches.

The potential of using PCA has been shown to be a useful tool for exploring multiple trait data and multi-trait selection because trait associations and trait profiles of the genotypes can be displayed in a table or graphically using biplots. Yan et al. (2008) demonstrated how PCA can be used for selecting potential cultivars and for parent selection in plant breeding programs. Also, Yan et al. (2005) demonstrated how PCA can be used for QTL identification and marker-based selection. No other studies were available to compare the results of PC2 and PC3.

Amino Acid QTL Detection

In total ten QTL were found to be significant using both MQM and CIM (Table 6). Each QTL explained between 5%-14% of the total phenotypic variation (R2) for a particular amino acid (Table 6). A QTL detected on chromosome 5 was associated with Ala and Val and was linked to molecular marker ss107923612, which explained 5.5 and 6.0% of the total phenotypic variation (R2), for those two amino acids, respectively. A QTL linked to molecular markers ss107928831 and ss107926274 on chromosome 7 was detected for Asp (R2=5.5%). Two QTL for Asp and Leu were detected on chromosome 9. Marker ss107912627 on chromosome 9 was found to be linked to a major QTL (R2 =10.5%) for Leu. Three molecular markers (ss107920438/ss107912744/ss107919004) were linked to a QTL associated with His and Tyr on chromosome 10, explaining 7.4 and 5.7% of the phenotypic variation, respectively. Three QTL were detected on chromosome 13 that were associated with 12 amino acids, explaining 5-9.5% of the total phenotypic variation for an individual amino acid. On chromosome 19, ss107917837 was linked to a QTL associated with Glu that had an R2 of 13.8%. A QTL linked to ss107929220 and ss107914151 on chromosome 20 was associated with Cys (R^2 = 6.0%).

Lys, Thr, Met, and Trp are the most important amino acids in swine diets (Boisen 2003), whereas for young poultry Lys, Trp, Arg, Thr, and Val are the most important (Baker 2003). Four of the minor QTL reported in this study are associated with amino acids that are essential in chicken diets and two minor QTL reported are essential to swine diets. Panthee et al. (2006a) reported a QTL for Gly and Thr linked to Satt518 (46.4 cM) roughly 20 cM from a QTL we detected on chromosome 9 near marker ss107913002 (62.54 cM), which was linked to Asp and Leu. Warrington (2011) detected a QTL associated with Thr linked to BARC-048619 (79.06 cM) and Met linked to BARC-042449 (77.4 cM) on chromosome 9. Those are within ~10 cM of the two markers reported in this study on chromosome 9 (86.91 cM) associated with Asp and Leu.

Another QTL detected by Warrington (2011) associated with Met linked to Satt592 on chromosome 10 (91.4 cM) was within 20 cM of the QTL linked to markers ss107920438, ss107912744 and ss107919004 on chromosome 10 (110.18 cM) in our study. In addition, Panthee et al. (2006a) reported a QTL on chromosome 13 linked to Satt252 (16.0 cM) only 5 cM away from markers ss107912657 and ss107913658 (21.51 cM). Marker Satt252 was associated with Cys, Ile, Met and Val (Panthee et al. 2006a). In this study marker ss107912657 and ss107913658 were associ-

ated with Arg, Iso, Phe, Pro, Ser, Tyr and Val.

QTL Associated with Met and Cys

In this study one QTL was reported to be associated with Cys and one QTL was associated with Met (Table 6). The QTL, linked to marker ss107917837, associated with Met was located on chromosome 13 and was associated with eleven other amino acids. Reinprecht et al. (2006) detected a seed protein QTL associated with marker Satt569 (2.35 cM) on chromosome 13 only ~2 cM away from marker ss107917837 (4.86 cM). Brummer et al. (1997) detected a QTL associated with seed protein linked to marker K002_1 (46.3 cM) on chromosome 13. This marker was ~6 cm from markers ss107920654 and ss107924336 (40.69 cM) reported in this study linked to a QTL associated with ten amino acids. Based on the close proximity of the QTL for Cys and Met discovered in this study to the protein QTL found in previous studies suggests an increase in either Cys, Met or both my require an increase in total protein content. This may be due to the fact Met is the initiating amino acid that initiates the synthesis of almost all eukaryotic proteins. In addition, Cys is known to play an important role in protein structure and in protein-folding pathways because of its ability to form disulfide bonds (Brosnan and Brosnan, 2006).

Conclusion

The proximity of the markers reported in this study and in previous studies indicates that some of the same QTL may have been detected in several studies. Selection for only a few of these QTL may greatly enhance genetic gains. Moreover, three genomic regions on chromosome 13 (4.89, 21.51, 40.69 cM) were found to control multiple amino acids. Two of these regions were very close to previously reported QTL associated with seed protein content. This suggests that some of the QTL reported for seed protein content in soybean may also be involved in determining protein quality.

Also, in this study new QTL for improving amino acid composition in soybean were discovered that do not coincide with previously reported QTL. Through selection of these new amino acid QTL and the previously reported QTL, soybean lines with improved amino acid profiles could be developed to help meet industry demands. The results from this study are intended to provide a basis for future research in soybean amino acid composition, which could provide valuable benefits to the animal feed industry.

Acknowledgements

The authors would like to thank the United Soybean Board, the Tennessee Soybean Promotion Board, the Tennessee Agricultural Experiment Station, the Arkansas Agricultural Experiment Station, Southern Illinois University and the Soybean Genomics Lab, USDA-ARS, Beltsville, MD for making this research possible. Their support is greatly appreciated.

References

Baker DH (2003) Ideal amino acid patterns for broiler chicks. In: JPF D'Mello (2nd edn) Amino Acids in Animal Nutrition. Cambridge, MA, pp. 223-235.

Bernard RL and CR Cremeens (1988) Registration of 'Williams 82' Soybean. Crop Sci 28:1027-1028.

Boisen S (2003) Ideal dietary amino acid profiles for pigs. In: JPF D'Mello (2nd edn) Amino Acids in Animal Nutrition. Cambridge, MA, pp. 157-168.

Brim CA (1996) A modified pedigree method of selection in soybeans. Crop Sci 6:220.

Broman KW and S Sen (2009) A guide to QTL mapping with R/qtl. Springer, NY.

Brosnan JT and ME Brosnan (2006) The sulfur-containing amino acids: an overview. J Nutr 136:16365-16405.

Brummer EC, GL Graef, JH Orf, JR Wilcox, and RC Shoemaker (1997) Mapping QTL for seed protein and oil content in eight soybean populations. Crop Sci 37:370-378.

Carlson CM (2011) Genetic control of protein and amino acid content in soybean determined in two genetically connected populations. Dissertation. North Carolina State University.

Forde BG and PJ Lea (2007) Glutamate in plants: metabolism, regulation and signaling. J Exp Botany 58:2339-2358.

Gizlice Z, TE Carter Jr, TM Gerig, and JW Burton (1996) Genetic diversity patterns in North American public soybean cultivars based on coefficient of parentage. Crop Sci 36:753–765.

Gibson PT, RS Suttner, VN Nijti, and O Myers Jr (1994) Soybean varietal response to sudden death syndrome. In: Wilkinson, D (Ed) Proceedings of the 24th soybean seed research conference. Chicago, IL, pp. 20-40.

George AA and BO Lumen (1991) A novel methionine-rich protein in soybean seed: Identification, amino acid composition, and n-terminal sequence. J Agric Food Chem 39:224-227.

Grant D, RT Nelson, SB Cannon, and RC Shoemaker (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. Nucleic Acid Res 38:D843-D846.

Hyten DL, Q Song, Y Zhu, IY Choi, MS Yoon, JE Specht, LK Matukumalli, RL Nelson, RC Shoemaker, ND Young, and PB Cregan (2008) High-throughput genotyping with the GoldenGate assay in the complex genome of soybean. Theor Appl Genet 116: 945-952.

Hyten DL, IY Choi, Q Song, JE Specht, TE Carter Jr, RC Shoemaker, LK Matukumalli, and PB Cregan (2010) A high density integrated genetic linkage map of soybean and the development of a 1536 universal soy linkage panel for quantitative trait locus mapping. Crop Sci 50:960-968.

Lightfoot DA, VN Njiti, PT Gibson, MA Kassem, MJ Iqbal, and K Meksem (2005) Registration of the Essex by Forrest recombinant inbred line mapping population. Crop Sci 45:1678-1681.

Nyquist WE (1991) Estimation of heritability and prediction of selection response in plant populations. Crit Rev Plant Sci 10: 235–322.

Panthee DR and VR Pantalone (2006) Registration of soybean germplasm lines TN03-350 and TN04-5321 with improved protein concentration and quality. Crop Sci 46:2328-2329.

Panthee DR, VR Pantalone, AM Saxton, DR West, CE Sams (2006a) Genomic regions associated with amino acid composition in soybean. Crop Sci 17:79-89.

Panthee DR, VR Pantalone, AM Saxton, DR West, CE Sams, JH Orf, AS Killiam (2006b) Quantitative trait loci controlling sulfur containing amino acids, methionine and cysteine, in soybean seeds. Theor Appl Gen 112:546-553.

Reinprecht Y, V Poysa, K Yu, I Rajcan, G Ablett, and K Pauls (2006)

Seed and agronomic QTL in low linolenic acid, lipoxygenase-free soybean (*Glycine max* (L.) Merrill) germplasm. Gen 49:1510-1527.
Shenk JS and MO Westerhaus (1991) Population definition, sample selection and calibration procedures for near infrared reflectance spectroscopy. Crop Sci 31:469-474.
Sneller CH (1994) Pedigree analysis of elite soybean lines. Crop Sci 34:1515–1522.
Smith TJ and HM Camper (1973) Registration of 'Essex' Soybean. Crop Sci 13:495-495.
Taiz L and E Zeiger (2006) Amino acid biosynthesis. In: Plant Physiology (4th edn.) Sunderland, MA, pp. 296-297.
Ufaz S and G Galili (2008) Improving the content of essential amino acids in crop plants: goals and opportunities. Plant Physiol 147: 954-961.
Warrington CV (2011) QTL mapping and optimum resource allocation for enhancing amino acid content in soybean. Dissertation. University of Georgia.
Richard WF (2012) Impact of Research on the Economic Outlook for Soybeans and Soybean Products in Global Markets. In: Wilson RF (ed) Designing Soybeans for 21st Century Markets. AOCS, pp. 3-28.
Yan W and JA Fregeau-Reid (2008) Breeding line selection based on multiple traits. Crop Sci 48:417-423.
Yan W, NA Tinker, and D Falk (2005) QTL identification, mega-environment classification and strategy development for marker-based selection using biplots. J Crop Imp 14:299-324.

Journal of Plant Genome Sciences - ISSN 1949-1360. Published By Atlas Publishing, LP (www.atlas-publishing.org)