Theor Appl Genet (2006) 112: 546–553
DOI 10.1007/s00122-005-0161-6

ORIGINAL PAPER

D. R. Panthee · V. R. Pantalone · C. E. Sams
A. M. Saxton · D. R. West · J. H. Orf · A. S. Killam

Quantitative trait loci controlling sulfur containing amino acids, methionine and cysteine, in soybean seeds

Received: 27 April 2005 / Accepted: 13 November 2005 / Published online: 10 December 2005

Abstract Soybean [*Glycine max* (L.) Merr.] is the single largest source of protein in animal feed. However, a major limitation of soy proteins is their deficiency in sulfur-containing amino acids, methionine (Met) and cysteine (Cys). The objective of this study was to identify quantitative trait loci (QTL) associated with Met and Cys concentration in soybean seed. To achieve this objective, 101 F_6-derived recombinant inbred lines (RIL) from a population developed from a cross of N87-984-16 × TN93-99 were used. Ground soybean seed samples were analyzed for Met and Cys concentration using a near infrared spectroscopy instrument. Data were analyzed using SAS software and QTL Cartographer. RIL differed ($P < 0.01$) in Met and Cys concentrations, with a range of 5.1–7.3 (g kg^{-1} seed dry weight) for Cys and 4.4–8.8 (g kg^{-1} seed dry weight) for Met. Heritability estimates on an entry mean basis were 0.14 and 0.57 for Cys and Met, respectively. A total of 94 polymorphic simple sequence repeat molecular genetic markers were screened in the RIL. Single factor ANOVA was used to identify candidate QTL, which were confirmed by composite interval mapping using QTL Cartographer. Four QTL linked to molecular markers Satt235, Satt252, Satt427 and Satt436 distributed on three molecular linkage groups (MLG) D1a, F and G were associated with Cys and three QTL linked to molecular markers Satt252, Satt564 and Satt590 distributed on MLG F, G and M were associated with Met concentration in soybean seed. QTL associated with Met and Cys in soybean seed will provide important information to breeders targeting improvements in the nutritional quality of soybean.

Communicated by S. J. Knapp

D. R. Panthee · V. R. Pantalone (✉) · C. E. Sams · D. R. West
Department of Plant Sciences, University of Tennessee,
2431 Joe Johnson Dr, 252 Ellington Plant Sciences Bldg,
Knoxville, TN 37996, USA
E-mail: vpantalo@utk.edu
Tel.: +1-865-9748801
Fax: +1-865-9741947

A. M. Saxton
Department of Animal Science, University of Tennessee,
2505 River Dr., 208c Brehm Animal Sciences,
Knoxville, TN 37996, USA

J. H. Orf · A. S. Killam
Department of Agronomy and Plant Genetics,
University of Minnesota, St. Paul, MN 55108, USA

Introduction

Soybean [*Glycine max* (L.) Merr.] is a major source of protein for humans and livestock throughout the world. It is estimated that more than 70% of the world's total protein meal comes from soybean (Soytech Inc. 2004). A major function of proteins in nutrition is to supply adequate amounts of required amino acids (Friedman and Brandon 2001). Although soybean protein is relatively rich in most of the essential amino acids, the concentration of the sulfur containing amino acids methionine (Met) and cysteine (Cys) is approximately half that of egg protein, which is the standard nutritional reference protein (George and de Lumen 1991) identified by the Food and Agriculture Organization (FAO) of the United Nations. Because of this deficiency, either synthetic or natural supplementary ingredients are utilized to fulfill the requirement of Met in soy based animal feed. However, Met supplementation has possible problems such as leaching during processing and bacterial degradation leading to formation of undesirable volatile sulfides (George and de Lumen 1991).

Methionine occupies a central position in cellular metabolism in which the processes of protein synthesis, methyl group transfer, polyamine and ethylene syntheses are interconnected (Ravanel et al. 1998). Among these pathways, the synthesis of proteins is the only pathway consuming the entire Met molecule. Since Met is involved in all these physiological and biochemical processes, an ample supply of Met in animal feed is needed. The FAO recommended a combined requirement of Cys

and Met as 3.5% of the total protein (George and de Lumen 1991). In regular soybean cultivars, average Met and Cys content is approximately 0.65 g/100 g of dry matter each (Friedman and Brandon 2001). Efforts are needed to develop soybeans with enhanced Met and/or Cys to overcome the nutritional limitations of existing cultivars.

The amount of supplemental Met required to balance a deficient ration for animals is about 4–10 g per day per animal. Sulfur containing amino acids are extremely important in wool producing animals to produce high quality wool (Liu and Masters 2003). Natural sources of Met (such as corn gluten meal or animal proteins) supply Met at a cost of about US$ 0.03 g^{-1}. Synthetic rumen-protected Met products typically supply Met at about US$ 0.02 g^{-1}. It is estimated that the commercial value of a 10% improvement in the present level of Met would be about US$ 3.5 t^{-1} (Clarke and Wiseman 2000). The poultry and swine industry would derive considerable savings if an improved cultivar of soybean could be developed with enhanced S-containing amino acids. Some people, whose primary source of food is root crops e.g., cassava (*Manihot esculenta* Crantz), suffer from deficiency of S-containing amino acids. If a genetic solution could be provided in soybean, it may also help to solve this deficiency problem.

Efforts made in the past to address low Met include conventional plant breeding, biotechnological approaches, and nutrient amendments. One of the successful efforts was the development of mutants with increased concentration of Met and Cys by the use of 0.5% Ethylmethanesulfonate (EMS) (Imsande 2001). In this study, a mutant line was developed with both Met and Cys concentrations of 1.85% of total protein. Another research group was able to increase the Met concentration in maize inbreds by 17% over their recurrent parent via backcrossing (Olsen et al. 2003).

Muntz et al. (1998) have discussed in detail the approaches adopted to increase the concentration of essential amino acids including Met and Cys by means of biotechnology. The most promising approach adopted was the introduction of foreign gene(s) encoding for proteins with extremely high Met level. Examples of such approaches are the gene encoding for the 2S albumin of Brazil nut (*Bertholettia excelsa*) (Altenbach et al. 1987) and sunflower (*Helianthus annus*) (Kortt and Caldwell 1990). The feasibility of this strategy was shown by transforming tobacco (Altenbach et al. 1989). The transformation approach was also successfully used in soybean (Townsend and Thomas 1994). The Met concentration in transgenic soybean was nearly double that of the wild type. However, the Brazil nut protein proved to be allergenic making the transgenic beans unusable (Nordlee et al. 1996). Despite all of these efforts, not a single commercial cultivar of soybean with the FAO standard total sulfur containing amino acids has been developed to date.

A nutrient amendment approach to improve sulfur containing amino acids included regulation of N and S in soil (Imsande and Schmidt 1998; Sexton et al. 1998). They assessed the seed quality of soybean with respect to 7S and 11S fractions of seed storage protein because 11S contains more Met concentration than 7S. However, this is not a sustainable approach.

Development of a cultivar with increased levels of S-containing amino acids in a short time would be an alternative to address this problem. Marker-assisted selection (MAS) is promising for this purpose, as has been used to improve various traits in several crops including soybean (Orf et al. 2004). However, mapping quantitative trait loci (QTL) or identifying a closely linked marker is a prerequisite for this purpose. The objective of this study was to identify QTL associated with Met and Cys concentration in soybean seeds.

Materials and methods

Plant material

A total of 101 F_6-derived recombinant inbred lines (RIL) were developed from a cross of N87-984-16 (P1) × TN93-99 (P2). The N87-984-16 parent is one of two F_8-derived near isogenic sister lines, whose blend constitutes the high protein commercial cultivar 'Prolina' (Burton et al. 1999). The TN93-99 parent is a high yielding line, developed by the Tennessee Agricultural Experiment Station, and is registered as germplasm GP-280 (Pantalone et al. 2003). Crossing was done during summer of 1998. F_1 seeds were harvested in October 1998 and F_1 single plants were grown in Costa Rica during winter 1998. Generations were advanced via single seed descent until F_5 in Costa Rica and $F_{5:6}$ seeds were obtained in May 2000. About 300 $F_{5:6}$ single plants were grown at the East Tennessee Research and Education Center in Knoxville, TN as an RIL population and 101 random plants of Tennessee adapted maturity were selected as a source population for this study. All 101 lines and parents were planted in 6 m length two-row plots with three replications in a randomized complete block design (RCBD) at the Knoxville Plant Science Farm (KPSF) in 2001. All the RIL, the two parents, as well as checks (Hutcheson, 5002T and 5601T) were planted in an RCBD with three replications at three locations: the Knoxville and Holston units of the East Tennessee Research and Education Center and the Ames Plantation Research and Education Center near Grand Junction, TN in May 2002 and 2003. Each line was planted in a four-row plot of 6 m length with a spacing of 75 cm between rows.

Sample preparation and amino acid analysis

Approximately 20 g of soybean seeds were ground in a water-cooled Knifetec 1095 Sample Mill (FOSS Tecator, S-26321, Hogana, Sweden) for 20 s. This setting produced soybean flour with a uniform particle size.

The near infrared spectroscopy (NIRS) instrument (NIRS 6500, FOSS North America) was warmed up for 2 h after turning on the lamp and auto diagnostics were run. Diagnostics tests ensured that the instrument passed three different tests for instrument response, wavelength accuracy, and NIRS repeatability. A room dehumidifier was used throughout the analysis, setting the humidity to 40%; room temperature was approximately 20°C. The ground soybean samples were scanned using Winisi II 1.5 software. The instrument was left on for the entire period of analysis, and diagnostics were performed every day until scanning of all samples was completed. The scanning produced the predicted levels of Met and Cys as concentration of total seed on a dry weight basis. We also calculated total sulfur containing amino acids (Met + Cys) and their ratio (Met:Cys).

DNA extraction and polymerase chain reaction

DNA was extracted from the RIL and parental lines utilizing Qiagen Plant Easy DNA Extraction Kit (Qiagen Hilden, Germany). PCR consisted of 7.4 µl of ddH_2O, 1 µl of 10× PCR Buffer, 1 µl of 2 mM dNTPs mixture (Pharmacia, Piscataway, NJ, USA), 0.5 µl of 20 µM forward and reverse primer, 0.1 µl of 5 U/µl^{-1} Klentaq (Ab Peptides Inc., St Louis, MO, USA) and 2 µl of 20 ng µl^{-1} template DNA. The PCR was performed in a 96-well MBS Hybaid thermocycler (Hybaid, Franklin, MA, USA). PCR conditions were (a) 94°C for 5 min followed by 35 cycles at 94°C for denaturation for 25 s, (b) 47°C for annealing for 30 s, (c) 72°C for 25 s for extension, and one last cycle at 72°C for final extension for 5 min. Parents were screened for a total of 585 $(ATT)_n$ simple sequence repeat (SSR) genetic markers (Cregan et al. 1999), of which 138 were found polymorphic. The sequence information for the markers is publicly available from the USDA website http://www.soybase.org, which was verified on 20 October, 2005. A total of 94 scorable and polymorphic SSR markers were used in QTL analyses.

DNA gel electrophoresis

A 6% non-denaturing polyacrylamide gel electrophoresis (PAGE) used to separate the PCR product consisted of 6% bis-acrylamide, 0.5% TBE buffer, 0.07% APS and 0.035% TEMED. Two microliter of loading buffer (6×) were added to PCR products and a 10 µl sample was loaded on the gel. The running buffer was 0.5× TBE. The gel was run at constant 300 V for 3 h. A fan was used to keep the glass plates cool during gel running. Ethidium bromide (50 µl, 10 mg ml^{-1}) was added to the running buffer in order to visualize the bands under exposure to UV light. Bands were scored using 1 to represent P_1 (N87-984-16), 2 to represent heterozygote, and 3 to represent P_2 (TN93-99) alleles for each primer locus.

Data analysis

Phenotypic data on Met and Cys were analyzed using MIXED procedure of SAS software to determine genotypic differences among the RIL (SAS Institute Inc. 2002). Location and replication were considered as two blocking factors in the model. Heritability of the trait in the population was estimated on an entry mean basis (Nyquist 1991) as follows:

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \left(\frac{\sigma_{ge}^2}{e}\right) + \left(\frac{\sigma^2}{re}\right)},$$

where h^2 represents narrow-sense heritability, σ_g^2 is genotypic variance, σ_{ge}^2 is genotype × environment variance, σ^2 is error variance, r is number of replications and e is the number of environments.

REML was used to estimate variance components for calculating heritability estimates. Phenotypic correlations were determined using the CORR procedure (SAS Institute Inc. 2002) and genetic correlations were determined using the following formula (Falconer and Mackay 1996; Kearsey and Pooni 1996):

$$r_G = \frac{\mathrm{Cov}_{xy}}{\sqrt{\sigma_x^2 \sigma_y^2}},$$

where r_G represents genetic correlation, x represents one trait, y represents a second trait and σ^2 is genetic variance. Cross products (Cov) were generated using the MANOVA option of the PROC GLM procedure. Associations between molecular data and least squares means of NIRS predicted Met and Cys were detected by analyzing the data using PROC GLM (SAS Institute Inc. 2002). Heterozygotes were excluded from the analysis because the RIL population was F_6-derived, and our interest was to detect heritable alleles from pure lines. QTL position and distance were estimated using Mapmaker/Exp 3.0 (Lincoln et al. 1993) and the information from Mapmaker was then used in QTL Cartographer (Wang et al. 2003) to confirm the QTL by composite interval mapping (CIM). We used the standard model *Zmapqtl* 6 in the CIM procedure with a 10 cM window size and a 2 cM walking speed. We also performed QTL × Environment interaction using least squares means for Cys and Met from six environments by the multiple trait analysis (MTA) method of QTL Cartographer choosing hypothesis 14 (Wang et al. 2003). We performed 1,000 permutations on Cys, Met and Met + Cys in each environment and in the combined data from six environments to establish empirical LOD thresholds at the 5% probability level (Churchill and Doerge 1994). However, QTL with LOD ≥ 2.0 are also reported as has been done in other papers (Concibido et al. 2003; Njiti et al. 2002;

Table 1 Descriptive statistics (seed dry weight basis) for cysteine (Cys), methionine (Met), total sulfur containing amino acids (Met + Cys) and total seed protein concentration averaged over six environments in a soybean RIL population derived from N87-984-16 × TN93-99

Trait	Minimum	Maximum	Mean	P1	P2	$LSD_{0.05}$	h^2
Cysteine (g kg^{-1})	5.1	7.3	6.2	7.0	6.0	0.3	0.14
Methionine (g kg^{-1})	4.4	8.8	6.4	6.8	6.8	0.7	0.57
Cys + Met (g kg^{-1})	9.7	15.9	12.6	13.8	12.8	0.7	0.42
Protein (g kg^{-1})	399.7	434.5	416.1	431.0	401.0	3.3	0.66

P1 N87-984-16, *P2* TN93-99, *LSD* Least significant difference at the 0.05 probability level, h^2 Heritability on an entry mean basis

Rector et al. 1998), in order to identify potential QTL for further research.

Results and discussion

Methionine and cysteine phenotype

A significant difference ($P < 0.01$) was found among genotypes in the population for Met and Cys concentration. Descriptive statistics (g kg^{-1} seed dry weight) for Met, Cys and a combination of Met + Cys are presented in Table 1. The frequency distributions for Met, Cys and Met + Cys are presented in Fig. 1. RIL means ranged from 5.1 to 7.3, and from 4.4 to 8.8 (g kg^{-1} seed dry weight) for Cys and Met, respectively. The combination of total S-containing amino acids (Met + Cys) ranged from 9.7 to 15.9 (g kg^{-1} seed dry weight) among the RIL. The upper end of these ranges represents a reasonably high level of S-containing amino acid concentration in soybean. The average concentrations of Met and Cys in soybean reported by Friedman and Brandon (2001) are 0.65% each of dry matter, which is 6.5 g kg^{-1} seed dry weight each, making the total S-containing amino acids 13.0 g kg^{-1} of dry matter The combination of Met + Cys was 1.28% of dry matter (or 12.8 g kg^{-1} seed dry weight) in a study by Fontaine et al. (2001). Kwanyuen et al. (1997) analyzed the amino acid composition in selected *Glycine soja*, a wild relative of cultivated soybean germplasms and found that the concentration of Met and Cys ranged from 0.7 to 1.1% (7–11 g kg^{-1}dry weight) and 0.4 to 0.8% (4–8 g kg^{-1}seed dry weight), respectively. Thus, on the basis of these past observations, the average Met and Cys levels found in our population were high enough to enable QTL detection for these S-containing amino acids.

The genetic variance and breeders' ability to select from observed phenotypes for Met and Cys should be interpreted with respect to heritability. A moderate heritability was found for Met (0.57) in data combined from six environments, whereas heritability was low (0.14) for Cys. Quantitative traits like Met and Cys are influenced by the growing environments to a great extent. Whitehead et al. (1989) concluded from a similar type of study that a high level of heritability observed for amino acids indicated that phenotypes observed are likely to have similar observations in future generations. Selection from among high performing lines is more likely to achieve significant genetic gains when heritabilities are high. Unlike qualitative traits, it is difficult to make significant response to selection in quantitative traits if the heritability is low.

According to the FAO standard, the combination of Met + Cys should be ≥3.5% of protein (approximately 14.0 g kg^{-1} seed dry weight) to make soybean protein equivalent to egg protein, a reference protein (Clarke and Wiseman 2000). In the present study, the combined

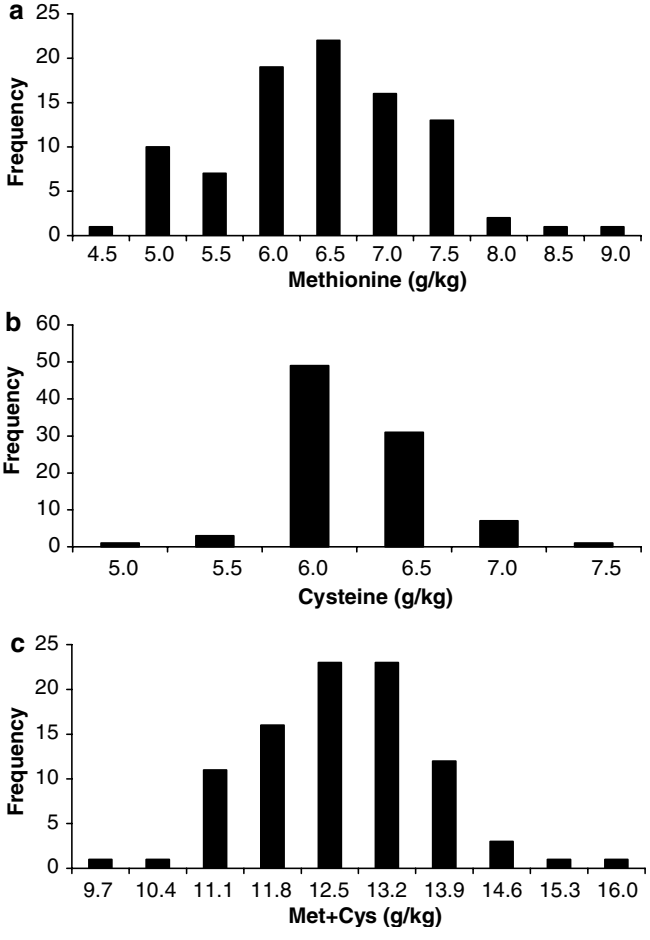

Fig. 1 Distribution of **a** methionine **b** cysteine, and **c** methionine + cysteine (g kg^{-1} seed dry matter) in an F_6-derived RIL population of soybean developed from a cross of N87-984-16 × TN93-99

average Met + Cys was 12.6 g kg^{-1}, ranging from 9.7 to 15.9 g kg^{-1} seed dry weight, thus a few RILs existed which were greater than the FAO standard. Generally, both Met and Cys were higher in this population, but Cys was much higher than in other populations. One of the parents (N87-984-16) constitutes the high protein commercial cultivar 'Prolina', which contains not only high protein but also exhibits superior gelling properties of its storage protein, indicating that it may contain more S-containing amino acids (Kwanyuen et al. 1998; Luck et al. 2001). Our choice of N87-984-16 as one of the parents provided us with an opportunity to generate lines with higher S-containing amino acids.

In our population, a negative phenotypic correlation between protein and Cys ($r = -0.48$) and no correlation between protein and Met (Table 2) indicated that high protein genotypes per se may not have high S-containing amino acids, which is in agreement with past findings (Whitehead et al. 1989). Therefore, it is necessary to assess the lines for S-containing amino acids in a breeding program aiming to improve protein quality.

The population was grouped into two subsets of relative plant maturity for ease of field production and to minimize measurement errors. We labeled the subsets as maturity 'A' and 'B' whose RIL differed by a range of approximately 14 days in maturity. It is interesting to note that there are six clusters in the scatter diagram of total S-containing amino acid (Met + Cys) (Fig. 2). Evaluation of individual lines and Met and Cys concentration revealed that the earlier maturing genotypes (maturity A) produced higher S-containing amino acids in all three locations than maturity B. There was a negative phenotypic correlation between maturity and total S-containing amino acid content ($r = -0.68$) in this population. To our knowledge, this is the first report of such an association in soybean.

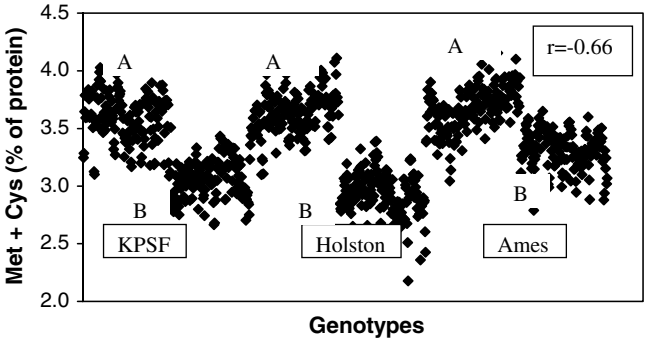

Fig. 2 Scatter diagram of total S-containing amino acid (% of protein) in an F_6-derived RIL soybean population from N87-984-16 × TN93-99. Earlier maturity subset of the population is indicated by '*A*' and later maturity subset of population is indicated by '*B*' in all three experiment sites, which differ by 14 days on average

Quantitative trait loci analysis

We initially identified candidate QTL for Cys and Met on the basis of single factor ANOVA and then confirmed by CIM using QTL Cartographer (Wang et al. 2003). CIM analysis detected four QTL near molecular markers Satt235, Satt252, Satt427 and Satt436 significantly associated with Cys concentration in this population. These QTL were distributed on three molecular linkage groups (MLG) D1a, F and G (Table 3). All these QTL showed association with Cys in multiple environments indicating that they may be an important QTL. However, MTA of QTL Cartographer confirmed that QTL near Satt252 and Satt235 were stable across the environments. Phenotypic variation explained by individual QTL ranged from 8.5 to 12.6% (Table 3). Alleles for increased Cys in three of the four loci were contributed through N87-984-16 whereas that near Satt436 was contributed through TN93-99. There were four additional QTL associated with Cys near Satt002, Satt185, Satt201 and Satt590 markers on MLG D2, E and M, respectively, but all of them were significant only in one environment (data not shown); hence they may be environmentally sensitive QTL (Brummer et al. 1997).

In our population, genetic gains for Cys were attainable by both parents but the magnitude of contributions was higher for N87-984-16. This is as expected because N87-984-16 was thought to be rich in S-containing amino acids (Kwanyuen et al. 1998; Luck et al. 2001). Moreover, N87-984-16 had some unusual biochemistry in its β-conglycinin (7S) fraction of protein in the sense that it had almost five times more Cys concentration (Kwanyuen et al. 1998), which is presumably because of sulfur-bonding. That may explain why more Cys was contributed via N87-984-16 alleles. Moreover, the distribution of Cys was skewed towards the higher

Table 2 Genetic and phenotypic correlation coefficient among Cys, Met and other traits in an F_6-derived RIL population of soybean from N87-984-16 × TN93-99

Trait	Cysteine		Methionine		Met + Cys	
	Genetic	Phenotypic	Genetic	Phenotypic	Genetic	Phenotypic
Methionine	0.73***	0.44***	–	–	–	–
Met + Cys	0.97***	0.91***	0.86***	0.76***	–	–
Met:Cys	−1.00***	−0.93***	−0.64***	−0.150 NS	−0.94***	−0.73***
Protein	−0.64***	−0.48***	−0.250**	−0.130 NS	−0.55***	−0.40***

NS Nonsignificant
*, **, *** Significant at 0.05, 0.01 and 0.001 probability levels, respectively

Table 3 QTL detected by composite interval mapping for cysteine, methionine, and Cys + Met in F_6-derived RIL population of soybean from N87-984-16 × TN93-99

Trait	Environment	MLG	Marker	QTL position(cM)	LOD score	Additive effect	R^2 (%)
Cysteine	Holston (2002–2003)	D1a	Satt436	11.57	2.46[c]	−0.02	8.5
	Combined (2002–2003)	F	Satt252	16.01	2.22[b]	0.01	11.0
	Combined (2002–2003)	G	Satt235	25.36	2.78[a,b]	0.01	12.6
	Ames (2002–2003)	G	Satt427	47.17	2.29 [c]	0.01	12.1
Methionine	KPSF (2002–2003)	F	Satt252	16.01	2.84[a,c]	0.47	15.2
	Combined (2003)	G	Satt564	56.80	2.59[a,b]	−0.54	18.8
	Combined (2002–2003)	M	Satt590	12.01	2.42[c]	−0.04	22.9
Met + Cys	Holston (2002–2003)	D2	Satt002	23.22	2.10[d]	−0.04	7.6
	Combined (2002–2003)	F	Satt252	16.01	2.45 [d]	0.04	12.3
	Ames (2002–2003)	M	Satt590	12.01	2.43 [d]	−0.05	17.6

MLG Molecular linkage group, *LOD* Log of odds ratio, *Additive* Additive genetic effect with respect to P1 (N87-984-16), *KPSF* Knoxville Plant Science Farm of the East Tennessee Research and Education Center at Knoxville, TN, *Holston* Holston Unit of the East Tennessee Research and Education Center at Knoxville, TN, *Ames* Ames Plantation Research and Education Center near Grand Junction, TN

[a]Significant QTL at empirical threshold levels determined by 1000 permutations

[b]QTL × Environment interaction analyzed by multi trait analysis detected QTL as stable across environments

[c]QTL × Environment interaction analyzed by multi trait analysis detected QTL as unstable across environments

[d]Significant at LOD = 2.0, but not at empirical threshold determined by 1,000 permutations

tail (Fig. 1b). This may result in higher type I or type II errors for Cys QTL detected in this population.

There were three QTL near molecular markers Satt252, Satt564 and Satt590 significantly associated with Met concentration in this population, which were distributed on MLG F, G and M, respectively (Table 3, Fig. 3). Among these, QTL near Satt564 and Satt590 were associated with Met concentration in multiple environments. However, MTA of QTL Cartographer revealed that the genomic region linked to Satt564 was stable across the environments. Another QTL near Satt252 was detected only in one location for both the years. The QTL near Satt002 was significant only in one environment indicating that this may be an environmentally sensitive QTL. Interestingly, unlike for Cys, three of the four loci (including the QTL near Satt002) had alleles for increased Met through TN93-99 (Table 3). Phenotypic variation explained by an individual QTL ranged from 15.2 to 22.9% indicating that all the QTL detected were major QTL (Falconer and Mackay 1996).

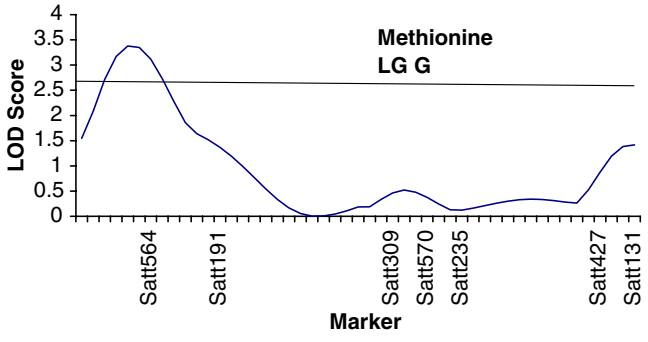

Fig. 3 LOD score plot on MLG G for Methionine QTL (Holston 2002) in an F_6-derived soybean population of N87-984-16 × TN93-99 indicated a major Met QTL upstream of Satt564 on MLG G

It is interesting to note that most of the QTL associated with Met were also associated with Cys in this population although some of them appear to be environmentally sensitive. It should be noted that Satt564 (a marker linked to a Met QTL) located on LG G was within 6 cM distance from Satt427 (a marker linked to a Cys QTL) in the composite molecular linkage map. An explanation for this observation could be that in the biochemical pathway of Met biosynthesis, Cys is the intermediate product in the process of assimilating sulfur (Matthews 1999; Saito 1999). With one amino acid being dependent on the concentration of others, it is possible to detect the same QTL for both Met and Cys.

With this in mind, an estimate of the total sulfur containing amino acid was derived by adding Met and Cys in this population and QTL analysis was performed on that sum. We detected three QTL in multiple environments near Satt002, Satt252 and Satt590 markers associated with total S-containing amino acids (Table 3). Furthermore, two QTLs near Satt185 and Satt564 were detected in only one environment, indicating that they may be environmentally sensitive QTL (data not shown). All the QTL were distributed among five MLG: D2, E, F, G and M (Table 3). The significant QTL associated with total S-containing amino acids were the same as those detected either for Cys or Met or both, as expected.

We know that amino acids including Met and Cys are components of protein. Therefore, QTL associated with Met and Cys are expected to be associated with protein. In this population, we previously identified some QTL associated with seed nitrogen accumulation at various reproductive growth stages (Panthee et al. 2004a), which is the major component of storage protein. It is interesting to note that not all QTL associated with Met and Cys were the same as those detected for seed nitrogen concentration, although some of them were located in the same linkage groups. We found that there was no correlation ($r = -0.13$) between Met and protein

whereas the correlation was moderately negative between Cys and protein ($r = -0.48$) (Table 2). Thus, despite knowledge that amino acids are the components of protein, there are some different genomic regions which govern total protein concentration in soybean.

Several researchers have mapped QTL for protein in soybean, which are close to the QTL detected in our study or from the same MLG. A protein QTL near Satt077 has been reported (Csanadi et al. 2001), which is within 7 cM downstream of molecular marker Satt436 on MLG D1a in the composite molecular linkage map of soybean. However, the closest protein QTL (near K002_1) reported from MLG F (Brummer et al. 1997) is about 31 cM downstream of Satt252. Molecular marker Satt235 was found to be associated with seed nitrogen accumulation processes during reproductive stages of soybean (Panthee et al. 2004a). Two protein QTL linked to molecular markers A816_1 and A890_1 have been reported from MLG G (Brummer et al. 1997), which are about 16 cM downstream of molecular marker Satt427 which we detected to be associated with a Cys QTL in the present study. A QTL associated with the glycinin fraction of storage protein, which is believed to be rich in S-containing amino acids, was detected between Satt564 and Satt191 from MLG G (Panthee et al. 2004b). Although a QTL associated with protein fractions has been reported from MLG D2 (Panthee et al. 2004b), no protein QTL close to marker Satt002 has been reported. A protein QTL linked with B174_1 from MLG E has been reported (Brummer et al. 1997), which is approximately 13 cM upstream of Satt185. Molecular marker Satt201 has been reported as associated with seed nitrogen accumulation processes in soybean during reproductive stages (Panthee et al. 2004a), and was found to be associated with an environmentally sensitive Cys QTL in the present study. Therefore, all MLG (D1a, E, F, G and M) which contained Cys and/or Met QTL have already been reported to have protein QTL. Thus, our data provide additional information that may be important for future genetic research targeting the improvement of soybean protein quality.

In quantitative traits, genetic gain or accumulation of favorable alleles in the progeny is possible from both of the parents. Favorable alleles are distributed in the population, which can be brought together through breeding efforts to make significant genetic gains. The QTL for Met and Cys that we report provide a first step to utilize a MAS program to achieve genetic gains.

Although there are several papers on QTL mapping for other traits, there is very limited information on QTL mapping for amino acids in any species. Loudelt et al. (2003) mapped QTL for amino acids in *Arabidopsis* and Wang et al. (2001b) mapped elongation factor involved in lysine biosynthesis in *Zea mays*. Wang et al. (2001a) also mapped a QTL involved in determining the free amino acid content in *Z. mays*. QTL detected for Met and Cys in the present study have potentially important applications in improving protein quality in soybean by MAS. The QTL detected in the present population need to be verified in other populations to confirm, for example, whether conserved alleles for Met and Cys exist among important breeding population parents, or whether specific lines would better facilitate MAS.

Acknowledgements Funding from the United Soybean Board in support of objectives of the Better Bean Initiative, and from the Tennessee Soybean Promotion Board is greatly appreciated. Continued support by the Tennessee Agriculture Experiment Station is also highly appreciated. We would like to express our thanks to Dr. Madeleine Spencer, Dr. David L. Hyten, Jr. and to Debbie Landau-Ellis for their contributions to this project.

References

Altenbach SB, Pearson KW, Leung FW, Sun SS (1987) Cloning and sequence analysis of a cdna encoding a brazil nut protein exceptionally rich in methionine. Plant Mol Biol 8:239–250

Altenbach SB, Pearson KW, Meeker G, Staraci LC, Sun SS (1989) Enhancement of the methionine content of seed proteins by the expression of a chimeric gene encoding a methionine-rich protein in transgenic plants. Plant Mol Biol 13:513–522

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

Burton JW, Carter TE, Wilson RF (1999) Registration of 'prolina' soybean. Crop Sci 39:294–295

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Clarke EJ, Wiseman J (2000) Developments in plant breeding for improved nutritional quality of soybeans I. Protein and amino acids content. J Agric Sci 134:111–124

Concibido VC, La Vallee B, McLaird P, Pineda N, Meyer J, Hummel L, Yang J, Wu K, Delannay X (2003) Introgression of a quantitative trait locus for yield from glycine soja into commercial soybean cultivars. Theor Appl Genet 106:575–582

Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J (1999) An integrated genetic linkage map of the soybean genome. Crop Sci 39:1464–1490

Csanadi G, Vollmann J, Stift G, Lelley T (2001) Seed quality qtls identified in a molecular map of early maturing soybean. Theor Appl Genet 103:912–919

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th edn. Longman Group Ltd, London

Fontaine J, Horr J, Schirmer B (2001) Near-infrared reflectance spectroscopy enables the fast and accurate prediction of the essential amino acid contents in soy, rapeseed meal, sunflower meal, peas, fish meal, meat meal products, and poultry meal. J Agric Food Chem 49:57–66

Friedman M, Brandon DL (2001) Nutritional and health benefits of soy proteins. J Agric Food Chem 49:1069–1086

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

Imsande J (2001) Selection of soybean mutants with increased concentrations of seed methionine and cysteine. Crop Sci 41:510–515

Imsande J, Schmidt JM (1998) Effect of n source during soybean pod filling on nitrogen and sulfur assimilation and remobilization. Plant Soil 202:41–47

Kearsey MJ, Pooni HS (1996) The genetical analysis of quantitative traits, 1st edn. Chapman and Hall, New York

Kortt AA, Caldwell JB (1990) Low molecular weight albumins from sunflower seed: identification of a methionine-rich. Phytochemistry 29:2805–2810

Kwanyuen P, Wilson RF, Burton JW (1998) Soybean protein quality. In: Koseoglu SS et al (eds) Oilseeds and edible oils processing, vol I. AOCS Press, Champaign, IL, pp 284–289
Kwanyuen P, Pantalone VR, Burton JW, Wilson RF (1997) A new approach to genetic alteration of soybean protein composition and quality. J Am Oil Chem Soc 74:983–987
Lincoln SE, Daly MJ, Lander ES (1993) Constructing genetic linkage maps with mapmaker version 3.: a tutorial and reference manual, 3rd edn. Whitehead Institute for Biomedical Research, Cambridge, MA
Liu SM, Masters DG (2003) Amino acid utilization for wool production. In: D'Mello JPF (ed) Amino acids in animal nutrition, 2nd edn. CABI Publishing, Cambridge, MA, pp 309–328
Loudelt O, Chaillou S, Merigout P, Talbotec J, Daniel-Vedele F (2003) Quantitative trait loci analysis of nitrogen use efficiency in *Arabidopsis*. Plant Physiol 131:345–358
Luck P, Lanier T, Daubert C, Wilson RF, Kwanyuen P (2001) Functionality and visoelastic behaviour of prolina soybean isolate. In: Wilson RF (ed) Oilseed processing and utilization. AOCS Press, Champaign, IL, pp 197–202
Matthews BF (1999) Lysine, threonine, and methionine biosynthesis. In: Singh BK (ed) Plant amino acids: biochemistry and biotechnology. Marcel Dekker Inc, New York, pp 205–225
Muntz K, Christov V, Saalbach G, Saalbach I, Waddel D, Pickardt T, Schieder O, Wustenhagen T (1998) Genetic engineering for high methionine grain legumes. Nahrung 42:125–127
Njiti VN, Meksem K, Iqbal MJ, Johnson JE, Kassem MA, Zobrist KF, Kilo VY, Lightfoot DA (2002) Common loci underlie field resistance to soybean sudden death syndrome in forrest, pyramid, essex, and douglas. Theor Appl Genet 104:294–300
Nordlee JA, Taylor SL, Twonsend JA, Thomas LA, Bush RK (1996) Identification of a brazil-nut allergen in transgenic soybeans. New Engl J Medicine 334:688–692
Nyquist WE (1991) Estimation of heritability and prediction of selection response in plant populations. Crit Rev Plant Sci 10:235–322
Olsen MS, Krone TL, Phillips RL (2003) Bsss53 as a donor source for increased whole-kernel methionine in maize: selection and evaluation of high-methionine inbreds and hybrids. Crop Sci 43:1634–1642
Orf JH, Diers BW, Boerma HR (2004) Genetic improvement: conventional and molecular-based strategies. In: Boerma HR, Specht JE (eds) Soybeans: improvement, production, and uses, 3rd edn. American Society of Agronomy, Madison, pp 417–450
Pantalone VR, Allen FL, Landau-Ellis D (2003) Registration of TN93–99 soybean germplasm. Crop Sci 43:1137
Panthee DR, Pantalone VR, Sams CE, Saxton AM, West DR, Rayford WE (2004a) Genomic regions governing soybean seed nitrogen accumulation. J Am Oil Chem Soc 81:77–81
Panthee DR, Kwanyuen P, Sams CE, West DR, Saxton AM, Pantalone VR (2004b) Quantitative trait loci for beta-conglycinin (7 s) and glycinin (11 s) fractions of soybean storage protein. J Am Oil Chem Soci 81:1005–1012
Ravanel S, Gakiere B, Job D, Douce R (1998) The specific features of methionine biosynthesis and metabolism in plants. Proc Nat Acad Sci USA 95:7805–7812
Rector BG, All JN, Parrott WA, Boerma HR (1998) Identification of molecular markers linked to quantitative trait loci for soybean resistance to corn earworm. Theor Appl Genet 96:786–790
Saito K (1999) Biosynthesis of cysteine. In: Singh BK (ed) Plant amino acids: Biochemistry and biotechnology. Marcel Dekker Inc, New York, pp 267–291
SAS Institute Inc (2002) Sas guide for personal computers, 9th edn. SAS Institute, Cary
Sexton PJ, Paek NC, Shibles RM (1998) Effects of nitrogen source and timing of sulfur deficiency on seed yield and expression of 11s and 7s seed storage proteins of soybean. Field Crops Res 59:1–8
Soytech Inc (2004) Oilseed statistics. Soytech Inc., Bar Harbor
Townsend JA, Thomas LA (1994) Factors which influence the *Agrobacterium*-mediated transformation of soybean. J Cell Biochem Suppl 18A:X1–014
Wang S, Basten CJ, Zeng ZB (2003) Windows qtl cartographer. Release V2.0. North Carolina State University, Raleigh, NC
Wang X, Stumpf DK, Larkins BA (2001a) Aspartate kinase 2. A candidate gene of a quantitative trait locus influencing free amino acid content in maize endosperm. Plant Physiol 125:1778–1787
Wang X, Won Y, Kim CS, Larkins BA (2001b) Quantitative trait locus mapping of loci influencing elongation factor 1a content in maize endosperm. Plant Physiol 125:1271–1282
Whitehead TF, Sapra VT, Rao DR (1989) Sulfur-containing amino acids and protein content in selected soybean progenies. Soybean Genet Newsl 16:74–80