Theor Appl Genet (2009) 118:1455–1463
DOI 10.1007/s00122-009-0994-5

ORIGINAL PAPER

Identification of QTL underlying isoflavone contents in soybean seeds among multiple environments

Guoliang Zeng · Dongmei Li · Yingpeng Han · Weili Teng · Jian Wang · Liquan Qiu · Wenbin Li

Received: 9 July 2008 / Accepted: 13 February 2009 / Published online: 6 March 2009
© Springer-Verlag 2009

Abstract Soybean isoflavones are valued in certain medicines, cosmetics, foods and feeds. Selection for high-isoflavone content in seeds along with agronomic traits is a goal of many soybean breeders. The aim of the study was to identify the quantitative trait loci (QTL) underlying seed isoflavone content in soybean among seven environments in China. A cross was made between 'Zhongdou 27', a soybean cultivar with higher mean isoflavone content in the seven environments (daidzein, DZ, 1,865 μg g^{-1}; genistein, GT, 1,614 μg g^{-1}; glycitein, GC, 311 μg g^{-1} and total isoflavone, TI, 3,791 μg g^{-1}) and 'Jiunong 20', a soybean cultivar with lower isoflavone content (DZ, 844 μg g^{-1}; GT, 1,046 μg g^{-1}; GC, 193 μg g^{-1} and TI, 2,061 μg g^{-1}). Through single-seed-descent, 130 F_5-derived F_6 recombinant inbred lines were advanced. A total of 99 simple-sequence repeat markers were used to construct a genetic linkage map. Seed isoflavone contents were analyzed using high-performance liquid chromatography for multiple years and locations (Harbin in 2005, 2006 and 2007, Hulan in 2006 and 2007, and Suihua in 2006 and 2007). Three QTL were associated with DZ content, four with GT content, three with GC content, and five with TI content. For all QTL detected the beneficial allele was from Zhongdou 27. QTL were located on three (DZ), three (GC), four (GT) and five (TI) molecular linkage groups (LG). A novel QTL was detected with marker Satt144 on LG F that was associated with DZ ($0.0014 > P > 0.0001$, $5\% < R^2 < 11\%$; $254 <$ DZ < 552 μg g^{-1}), GT ($0.0027 > P > 0.0001$; $4\% < R^2 < 9\%$; $262 <$ GT < 391 μg g^{-1}), and TI ($0.0011 > P > 0.0001$; $4\% < R^2 < 15\%$; $195 <$ TI < 871 μg g^{-1}) across the various environments. A previously reported QTL on LG M detected by Satt540 was associated with TI across four environments and TI mean ($0.0022 > P > 0.0001$; $3\% < R^2 < 8\%$; $182 <$ TI < 334 μg g^{-1}) in China. Because both beneficial alleles were from Zhongdou 27, it was concluded that these two QTL would have the greatest potential value for marker-assisted selection for high-isoflavone content in soybean seed in China.

Keywords QTL · SSR · MAS · Soybean · Isoflavone

Introduction

Soybean seeds are a rich source of isoflavones with approximately 1,000–3,000 μg g^{-1} of seed mass (Wang and Murphy 1994). The isoflavones are found in the following forms in different organs, daidzein (DZ), genistein (GT), glycitein (GC), genistin, daidzin, glycitin, 6″-*O*-acetylgenistin, 6″-*O*-acetyldaidzin, 6″-*O*-acetylglycitin, 6″-*O*malonylgenistin, 6″-*O*-malonyldaidzin and 6″-*O*malonylglycitin (Kudou et al. 1991; Griffith and Collison 2001). In soybean, DZ, GT and GC are synthesized via the phenylpropanoid pathway and stored in the vacuole as glucosyl and malonylglucose conjugates (Graham and Graham 1991; Kudou et al. 1991). The pathway to DZ branches from the phenylpropanoid pathway, following the chalcone synthase

G. Zeng, D. Li and Y. Han have equal contributions to the paper.

Communicated by D. Lightfoot.

G. Zeng · D. Li · Y. Han · W. Teng · J. Wang · W. Li (✉)
Soybean Research Institute (Key Laboratory of Soybean Biology in Chinese Ministry of Education), Northeast Agricultural University, 150030 Harbin, China
e-mail: wenbinli@neau.edu.cn

L. Qiu
Institute of Crop Science, Chinese Academy of Agriculture, 100081 Beijing, China

catalyzed reaction through a legume-specific enzyme, chalcone reductase. GT synthesis is not yet clearly defined but is likely derived from isoliquiritigenin, an intermediate of DZ synthesis (Latunde-Dada et al. 2001). GC synthesis shares the naringenin intermediate with the flavonoid–anthocyanin branch of the phenylpropanoid pathway. In all cases, the unique aryl migration reaction to create the isoflavones is mediated by isoflavone synthase (Primomo et al. 2005). Many studies showed that isoflavones have important functions in soybean, such as promoting nodulation by rhizobia (Kosslak et al. 1987), changing or adjusting the category and cluster of microorganism around soybean root (Lozovaya et al. 2004), and protecting soybeans from pests and pathogenic microbes (Morris et al.1991; Benhamou and Nicole 1999). Moreover, human nutritional studies have suggested that isoflavones might play an important role in the prevention and treatment of a number of chronic diseases including certain forms of cancers, osteoporosis and heart disease and for their ability to relieve menopause symptoms (Munro et al. 2003). Therefore, soybean cultivars with increased isoflavone content in the seed would be desirable for breeding programs.

Selection for high-seed isoflavone content in the seed is a major goal of some soybean breeders. Traditionally, plant improvement has relied on phenotypic selection of populations from crosses between commercial cultivars or experimental lines (Stuber et al. 1992). However, selection for high-seed isoflavone based on phenotype was complicated by genotype × environment interaction (GE) that significantly influenced isoflavone accumulation (Eldridge and Kwolek 1983; Wang and Murphy 1994; Carrao-Panizzi and Kitamura 1995; Hoeck et al. 2000; Lee et al. 2003; Primomo et al. 2005). Hence, selection for soybean cultivars with high seed isoflavone contents in seed required evaluation in multiple environments over several years, which is expensive, time-consuming, and labor intensive. Molecular markers offer a faster and more accurate approach to breeding, because selection could be based on genotype rather than solely on phenotype. The use of molecular markers for indirect selection of important agronomic traits, or marker-assisted selection (MAS) could improve the efficiency of traditional plant breeding. Some aspects of plant breeding that can be improved by MAS include; identification and elimination of undesirable individuals in the early stages of selection; identification of individuals prior to flowering when backcrossing genes governing the favorable expression of quantitative traits into adapted genotypes; and facilitation of selection for several traits simultaneously (Allen 1994). MAS could potentially improve selection of traits that have low heritability using markers with high heritability.

Several studies have attempted to map QTL associated with seed isoflavone content in soybean seed. QTL associated with this trait have been reported on LG A1, B1, B2, D1a_Q, H, K, and N in the mapping population **'**Essex**'** × **'**Forrest**'** (Njiti et al. 1999; Meksem et al. 2001; Kassem et al. 2004, 2006). Primomo et al. (2005) analyzed individual and total isoflavone in soybean seed using a set of RILs from a cross 'AC756' × 'RCAT Angora' in two locations of Canada, and found that nine genomic regions (LG A1, C2, D1a,_Q, F, G, H, J, K, and M) were associated with individual and total isoflavone content. All these studies focused on the North American soybean germplasm (Njiti et al. 1999; Meksem et al. 2001; Kassem et al. 2004; Primomo et al. 2005). The QTL associated with isoflavone content in soybean seeds in Chinese germplasm have not been reported to date.

The objective of the present study was to identify additional QTL associated with individual and total isoflavone using Zhongdou 27 × Jiunong 20 RILs in multiple environments using SSR markers.

Materials and methods

Plant materials

The mapping population, of 130 $F_{5:6}$ recombinant inbred lines (RILs), was advanced by single-seed-descent from the cross between Zhongdou 27 (developed by the Chinese Academy of Agriculture, Beijing, China) and Jiunong 20 (developed by Jinlin Academy of Agriculture, Jilin, China). Zhongdou 27 contained high individual and total isoflavone (TI) contents in seed. Jiunong 20 had low individual and TI contents in seed.

Field experiment

Recombinant inbred lines were grown together with the parents at Harbin (fine-mesic chernozen soil) in 2005, 2006 and 2007, at Hulan (fine-mesic chernozen soil) in 2006 and 2007, and at Suihua (fine-mesic chernozen soil) in 2006 and 2007. Seeds were planted in rows 3 m long, 0.65 m apart and with a space of 6 cm between two plants. Three replicates were used with a randomized complete block design. Each plot of a single genotype provided 20 plants as seed donors that were later used to analyze individual and total isoflavone content.

Isoflavone extraction and high-performance liquid chromatography (HPLC) analysis

Isoflavone concentrations were determined using HPLC as described by Vyn et al. (2002). One hundred grams of soybean flour was used for each extraction with 10 ml of 80% (v/v) ethanol in 10-ml Falcon tube that was slowly vortexed

for 1 h and then left overnight. The extraction was hydrolyzed by 2 ml HCl solvent. The extraction was filtered through PTFE membrane (Sartorius, 0.45 μm, Germany) and 10 μl of the filtrate was used to detect the isoflavone content by HPLC (Dionex ASI-100, USA). The following condition of HPLC was employed; solvent A was double-distilled water (ddH_2O) and solvent B was methanol (chromatography purity); solvent A:B ratio was 50:50%; the solvent flow rate was 1.0 ml/min and the temperature of column was kept at 50°C; A Dionex UV-170 detector was monitored at 254 nm; UV spectra were recorded, and area responses were integrated by software Dionex 2.0. Quantification of the isoflavones was carried out by reference to an external standard.

SSR analyses

Total DNA of the parents and each RIL were isolated from freeze-dried leaf tissue by CTAB method (Doyle and Doyle 1990). SSR analysis was performed with the primers developed by Cregan et al. (1999). PCR was performed in 20 μl containing 2 μl genomic DNA (25 ng/μl), 1.5 μl $MgCl_2$ (25 mM), 0.3 μl dNTP mixtures (10 mM), 2 μl 10× PCR buffer, 2 μl SSR primer (2 μM), 0.2 μl *Taq* polymerase (10 units/μl), 12 μl double-distilled water. The amplification temperature profiles were 2 min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s at 47°C, 30 s at 72°C, then 5 min at 72°C. After amplification, the PCR products were mixed with loading buffer (2.5 mg/ml bromophenol blue, 2.5 mg/ml diphenylamine blue, 10 mM EDTA, 95% (v/v) formamide), and denatured for 5 min at 94°C and then kept on ice for 5 min. The denatured PCR products were separated on 6% (w/v) denaturing polyacrylamide gel and visualized by silver straining (Trigizano and Caetano-Anolles 1998).

Data analysis

The frequency distribution, broad-sense heritabilities, and statistic parameters for RILs were analyzed using the SAS procedures (PROC. Shaprio-Wilk tests, line regression, and GLM.SAS). A linkage map was constructed by MAPMAKER/EXP version 3.0 (Lander et al. 1987) following methods described by Primomo et al. (2005). The commands "group", "map", "sequence", "lod table", "try", and "compare" were used for building the linkage groups. The error detection ratio was set at 1%. The Haldane mapping function was used with a minimum LOD score of 3.0 and a maximum distance of 50 cM.

Quantitative trait loci were identified by single-factor analysis of variance (ANOVA) (PROC. GLM. SAS) as described by Primomo et al. (2005), based on individual environment values and their means. Locus main effects were considered for linear models if they were significant at $P \leq 0.01$. Significant loci on the same LG were tested by two-factor ANOVA without interactions. If both loci were significant at $P \leq 0.05$ in the two-factor model, they would both be considered for linear models. Otherwise, the locus with the larger individual R^2 value was chosen to represent the effect of the putative QTL on the LG. The nomenclature of the QTL included four parts following the recommendations of the Soybean Germpasm Coordination Committee. For example, QTIM_1, Q, TI, M and 1 represent QTL, trait (total isoflavone, TI), linkage group name and QTL order in the linkage group, respectively.

GT (genotype by trait) biplot methodology (Yan 2001) was employed to analyze the interaction between QTL and environments or individual and total isoflavone means, based on the formula: $T_{ij} - T_j/S_j = \lambda_1\zeta_{i1}\tau_{j1} + \lambda_2\zeta_{i2}\tau_{j2} + \varepsilon_{ij}$ where T_{ij} was the mean value of QTL i for environment j; T_j was the mean value of environment j over all QTL, Sj was the standard deviation of environment j among QTL mean i; ζ_{i1} and ζ_{i2} were the PC1 (first principle component) and PC2 (second principle component) scores, respectively, for QTL mean i; τ_{j1} and τ_{j2} were the PC1 and PC2 scores, respectively, for environment j; and ε_{ij} was the residual of the model associated with QTL i, challenged with environment j.

Results

Phenotypic analysis of individual and total isoflavone content

The ranges of seed DZ, GT, GC, TI contents among the 130 RILs across seven different environments and individual or total isoflavone means were large (Table 1). The differences between the two parents were significant for individual and total isoflavone contents and their means across all seven environments. The trait mean values for Zhongdou 27 were higher than those of Jiunong 20 except for Hulan location in 2006. In contrast, the variation of trait values for 130 RILs across various environments and individual or total isoflavone means were not significant, with low coefficient of variation values (<0.5). Both skewness and kurtosis values for these traits were <1.0 in the different environments and individual or total isoflavone means, suggesting that the segregation of these traits fit a normal distribution model. Broad-sense heritabilities of 130 RILs across various environments and individual or total isoflavone means (Table 1) were relatively low (<0.5).

Linkage analysis

A total of 606 SSR markers were used to detect polymorphisms between the two parents. One hundred twenty-five

Table 1 Range, average, standard deviation, coefficient of variation, skewness, kurtosis, and broad-sense heritabilities for seed isoflavone contents of RILs across multiple environments

Trait	Environment	Parents		RIL						
		Zhongdou27[e]	Jiunong20[e]	Range[e]	Mean[e]	SD[f]	CV[g]	Skewness	Kurtosis	BSH[i]
DZ[a]	05Harbin	1,335.18	678.88	520.43–1,545.329	823.84	332.61	0.40	0.69	0.72	0.42
	06Harbin	1,424.12	840.23	623.27–1,545.30	1,293.69	412.17	0.32	0.5	−0.24	0.53
	06Hulan	2,065.84	1,002.05	876.45–2,345.22	1,287.55	418.67	0.33	0.33	−0.11	0.39
	06Suihua	1,763.2	1,024.79	956.23–1,845.32	1,394.57	360.18	0.26	0.51	0.51	0.44
	07Harbin	2,245.44	899.32	795.34–2,286.04	1,922.65	398.54	0.21	0.42	−0.26	0.53
	07Hulan	1,989.32	773.45	874.56–2,045.34	1,974.34	400.32	0.20	−0.34	−0.05	0.38
	07Suihua	2,234.56	689.45	673.23–2,476.45	2,036.34	428.17	0.21	0.27	0.4	0.49
DZ Mean[e, h]		1,865.38	844.02	759.93–2,012.72	1,533.28	436.78	0.28	0.1	−0.37	0.47
GT[b]	05Harbin	1,839.55	1,052.71	934.71–1,958.82	1,076.74	370.88	0.34	1.3	0.23	0.42
	06 Harbin	1,735.61	1,092.08	945.34–1,903.69	1,403.6	388.92	0.28	0.07	−0.08	0.31
	06Hulan	903.01	612.9	603.23–1,023.32	932.74	297.51	0.32	0.06	−0.2	0.57
	06Suihua	1,666.34	1,255.1	1,025.56–1,765.56	1,143.6	372.34	0.33	0.42	0.12	0.4
	07Harbin	1,823.41	1,043.2	987.43–1,976.85	1,782.67	385.34	0.22	0.3	0.44	0.31
	07Hulan	1,567.23	1,034.45	926.56–1,635.34	1,487.67	345.56	0.23	0.87	0.23	0.46
	07Suihua	1,763.44	1,234.45	1,164.34–1,867.34	1,752.47	353.21	0.20	0.5	−0.53	0.43
GT Mean[e, h]		1,614.08	1,046.41	941.02–1,732.99	1,304.51	382.98	0.29	0.9	0.43	0.52
GC[c]	05Harbin	462.37	189.08	174.45–453.23	359.45	24.33	0.07	0.97	0.42	0.38
	06 Harbin	261.73	193.9	149.20–289.32	223.14	41.17	0.18	0.8	1.3	0.41
	06Hulan	176.83	216.28	142.67–254.43	214.76	41.14	0.19	0.93	0.94	0.54
	06Suihua	184.68	132.67	131.34–204.54	170.52	39.74	0.23	0.4	−0.13	0.44
	07Harbin	326.45	203.34	189.23–340.43	283.45	35.34	0.12	0.55	0.73	0.47
	07Hulan	367.36	204.46	192.23–396.39	314.34	39.45	0.13	0.3	0.61	0.54
	07Suihua	401.23	212.46	179.45–436.67	356.97	23.67	0.07	−0.22	0.52	0.45
GC Mean[e, h]		311.52	193.17	165.93–339.28	274.67	33.76	0.12	0.8	0.57	0.48
TI[d]	05Harbin	3,637.1	1,920.67	1,834.45–4,987.23	2,754.08	770.87	0.28	0.59	1	0.35
	06 Harbin	3,423.63	1,974.03	1,845.28–4,062.48	2,900.36	718.46	0.25	−0.12	−0.39	0.46
	06Hulan	3,145.08	1,831.23	1,623.59–3,989.21	2,357.04	618.29	0.26	0.08	−0.42	0.5
	06Suihua	3,614.23	2,412.57	2,231.53–4,065.38	2,867.7	699.01	0.24	0.11	0.23	0.37
	07Harbin	4,395.29	2,145.86	1,986.59–4,476.23	4,003.24	705.34	0.18	0.12	0.22	0.45
	07Hulan	3,923.1	2,012.36	1,893.82–4,045.97	3,683.56	736.34	0.20	−0.23	−0.17	0.4
	07Suihua	4,399.21	2,136.36	1,986.45–4,496.34	3,754	745.39	0.20	0.22	−0.09	0.36
TI Mean[e, h]		3,791.09	2,061.87	1,914.53–4,303.27	3,188.57	739.59	0.23	0.52	0.34	0.49

[a] Daidzein

[b] Genistein

[c] Glycitein

[d] Total isoflavones

[e] μg g^{-1}

[f] Standard deviation

[g] Coefficient of variation

[h] Mean of all data from seven different environments

[i] Broad-sense heritabilities

of them (20%) were polymorphic among RI lines. The map including 99 SSR markers on 20 LGs of the map by Cregan et al. (1999) and Song et al. (2004). The map developed here encompassed about 2,020 cM with mean distance of 20.4 cM between markers (data not shown). The remaining 26 markers (Satt334, Sat_074, Satt569, Sat_105, Satt587,

Satt124, Satt247, Satt264, Sat_117, Satt309, Sat_040, Satt341, Sct_067, Satt470, Satt437, Satt274, Satt157, Satt271, Satt125, Sat_084, Satt530, Satt345, Satt262, Satt358, Satt241 and Satt576) were not linked to any other marker but could be used to detect QTL by single-factor ANOVA. The order of the SSR markers on map within most LGs was similar to the public soybean genetic map (Cregan et al. 1999 and Song et al. 2004).

QTL associated with individual and total isoflavone content

Quantitative trait loci associated with individual and total isoflavone contents were identified by ANOVA (Table 2). For all QTL detected the beneficial allele was from Zhongdou 27. Three QTL, QDZF_1 (Satt144), QDZI_1 (Satt587) and QDZK_1 (Satt124) were associated with DZ, located in LGs F, I and K,respectively. Of them, QDZF_1 explained 5.7–10.4% of the phenotypic variation at three locations in 3 years. QDZI_1 explained 4.0–11.3% of the phenotypic variation at three locations in 3 years. QDZK_1 explained 3.8–8.5% of the phenotypic variation at two locations in 1 year.

Four QTL underlying GT were detected and mapped onto four LGs (Table 2), explained 3.8–18.9% of the phenotypic variation at three locations in 3 years. Of them, QGTA2_1 (Sat_040), QGTF_1 (Satt144), and QGTM_1 (Satt540) could be identified in five, six, and three environments, respectively. However, QGC2_1 (Satt460) was detected in only one environment.

Three QTL, QGCI_1 (Satt330), QGCM_1 (Satt540), and QGCO_1 (Satt592) that were associated with GC were identified on LGs I, M, and O, respectively. The phenotypic variation ranged from 2.9 to 10.4% at three locations in 3 years (Table 2). Of them, QGCI_1 (Satt330) and QGCM_1 (Satt540) could be identified in four and five environments, respectively. However, QGCO_1 was detected in only one environment.

Five QTL underlying TI were identified and mapped onto five LGs (Table 2), explained 3.6–14.4% of the phenotypic variation. QTID2_1 (Satt208), QTIF_1 (Satt144), QTIG_1 (Satt288), QTIO_1 (Satt241), and QGTM_1 (Satt540) could be identified in two, six, three, four, and four environments, respectively.

QGCM_1 associated with GC across five environments and the GC mean; QGTM_1 associated with GT across three environments; QTIM_1 associated with TI across four environments and TI mean were linked with the same SSR marker Satt540 in linkage group M. QDZF_1 associated with DZ across seven environments and DZ mean; QGTF_1 associated with GT across six environments and GT mean; QTIF_1 associated with TI across six environments and TI mean, were linked with the same SSR marker Satt144 in linkage group F.

Analyses of QTL by environment interaction

GT biplot analysis (Yan 2001) of all QTL with individual and total seed isoflavone contents against seven environments and individual or total isoflavone means showed that these QTL jointly explained 64% of the total variation of isoflavones, more that the estimate of broad-sense heritability had predicted. Therefore, the performance of different QTL in each environment was evaluated. When QTL QDZK_1, QDZF_1, QTIF_1, QGTF_1, QGCI_1, and QGTA2_1 were set as the corner QTL, seven different environments and individual or total isoflavone mean fell in the sector in which QTIF_1 were the best QTL for five environments (at Harbin in 2006 and 2007, at Hulan in 2006 and 2007, and at Suihua in 2007) and TI mean. QDZF_1 was the best QTL at Suihua in 2006 and DZ mean. QGTF_1 was the best QTL at Harbin in 2005, GC, and GT mean.

Discussion

Soybean seed isoflavones have many uses in foods, medicines, cosmetics, and animal farming (Brouns 2002). The demand for soybean isoflavone in the international market has increased year by year (Brouns 2002). Thus, the improvement of seed isoflavone content in soybean cultivar is increasingly emphasized by breeders. Zhongdou 27 (3,791.09 μg g^{-1}) was proved to have high-isoflavone content in China for many years. Because isoflavone content in soybean was difficult to evaluate by phenotype, increasing the genotype selection intensity by MAS will lead to improved selection gain.

Wang and Murphy (1994) reported that total isoflavone content changed from 1,176 to 3,309 μg g^{-1} across locations within the same year to single soybean cultivar. Choi et al. (1996) reported that total isoflavone content ranged from 458 to 2,317 μg g^{-1} among cultivars in a single year. The TI values in China corresponded well to the previous studies (Table 1). Lee et al. (2003) determined that the interactions between environmental factors, such as year × location and genotype × environment, were the primary sources of variation in isoflavone contents in soybean seeds. These studies showed that the main genotypic effects of total and individual isoflavone were large enough for effective cultivar improvement.

Out of the 99, SSR markers used for QTL analysis, three QTL associated with DZ, four QTL associated with GT, three QTL associated with GC, and five QTL associated with TI were mapped onto three, four, three, and five LGs, respectively. These QTL explained 3–18.9% of phenotypic variation for individual and total isoflavone in different environments, most of the variation was <10%. The low

Table 2 QTL associated with individual and total seed isoflavone content using one-way ANOVA

Trait	QTL	LG	Marker	Environment	R^2 (%)[e]	*P*	Allelic means (μg g^{-1}) ± SEM[f]	
							'Zhongdou 27'	'Jiunong 20'
DZ[a]	QDZF_1	F	Satt144	05Harbin	8.54	0.0001	1,554 ± 46	1,186 ± 45
				06Harbin	9.35	0.0001	1,580 ± 47	1,194 ± 38
				06Hulan	5.74	0.0014	1,500 ± 36	1,246 ± 46
				06Suihua	10.43	<0.0001	1,589 ± 45	1,037 ± 44
				07Harbin	8.66	0.0004	1,500 ± 46	1,187 ± 38
				07Hulan	6.39	0.0001	1,534 ± 45	1,250 ± 44
				07Suihua	8.67	0.0001	1,498 ± 41	1,175 ± 38
				DZ Mean	7.06	0.0001	1,500 ± 41	1,205 ± 41
	QDZI_1	I	Satt587	06Hulan	11.29	0.0003	1,498 ± 43	926 ± 36
				07Harbin	8.94	0.0012	1,420 ± 36	1,046 ± 40
				07Hulan	4.9	0.0001	1,520 ± 41	1,275 ± 36
				07Suihua	4.03	0.002	1,498 ± 35	1,265 ± 42
	QDZK_1	K	Satt124	05Harbin	8.49	0.001	1,432 ± 41	1,062 ± 43
				06Harbin	7.43	<0.0001	1,486 ± 47	1,193 ± 31
				07Hulan	7.39	0.0009	1,534 ± 40	1,239 ± 42
				07Suihua	3.82	0.0043	1,629 ± 22	1,430 ± 45
				DZ Mean	4.76	0.0006	1,546 ± 41	1,286 ± 41
GT[b]	QGTA2_1	A2	Sat_040	06Habin	8.59	0.0004	1,675 ± 42	1,316 ± 40
				06Hulan	7.83	0.0001	1,693 ± 41	1,380 ± 41
				07Harbin	6.99	0.0006	1,586 ± 47	1,312 ± 44
				07Hulan	7.59	<0.0001	1,608 ± 44	1,324 ± 42
				07Suihua	18.92	<0.0001	1,594 ± 41	942 ± 31
	QGTC2_1	C2	Satt460	05Harbin	3.8	0.0032	1,628 ± 30	1,434 ± 45
	QGTF_1	F	Satt144	05Habin	6.79	0.0003	1,636 ± 44	1,350 ± 40
				06Habin	7.48	<0.0001	1,586 ± 43	1,290 ± 41
				06Suihua	8.48	<0.0001	1,650 ± 43	1,282 ± 41
				07Harbin	6.86	0.002	1,590 ± 36	1,328 ± 44
				07Hulan	7.2	0.0027	1,634 ± 35	1,243 ± 46
				07Suihua	4.98	0.0013	1,548 ± 36	1,264 ± 42
				GT Mean	5.42	0.0001	1,670 ± 42	1,418 ± 43
	QGTM_1	M	Satt540	06Habin	4.39	0.0024	1,556 ± 36	1,398 ± 43
				06Hulan	6.84	<0.0001	1,564 ± 42	1,287 ± 41
				07Hulan	5.94	0.0001	1,588 ± 42	1,320 ± 39
GC[c]	QGCI_1	I	Satt330	05harbin	7.34	0.0001	340 ± 8	268 ± 7
				06hulan	5.79	<0.0001	346 ± 8	280 ± 7
				06Suihua	8.12	0.0008	332 ± 7	247 ± 5
				07Harbin	10.43	<0.0001	354 ± 8	246 ± 5
	QGCM_1	M	Satt540	05Harbin	8.45	0.0003	318 ± 8	232 ± 8
				06Hulan	3.94	0.0021	325 ± 6	278 ± 8
				06Suihua	6.88	0.0001	350 ± 8	272 ± 8
				07Habin	6.65	0.0002	368 ± 8	298 ± 7
				07Suihua	3.67	0.003	356 ± 5	310 ± 9
				GC Mean	2.96	0.0006	346 ± 7	301 ± 6
	QGCO_1	O	Satt592	05Harbin	5.24	0.0001	360 ± 8	300 ± 8
TI[d]	QTID2_1	D2	Satt208	05Harbin	11.24	<0.0001	2,704 ± 53	2,008 ± 50
				07Suihua	7.89	0.0001	2,691 ± 54	2,176 ± 53
	QTIF_1	F	Satt144	05Harbin	8.48	0.0006	2,730 ± 51	2,049 ± 47

Table 2 continued

Trait	QTL	LG	Marker	Environment	R^2 (%)[e]	*P*	Allelic means (μg g^{-1}) ± SEM[f]	
							'Zhongdou 27'	'Jiunong 20'
				06Harbin	14.37	0.0005	2,871 ± 55	2,000 ± 51
				06Hulan	5.14	0.0001	2,675 ± 52	2,480 ± 53
				07Harbin	6.43	<0.0001	2,600 ± 55	2,248 ± 54
				07Hulan	5.87	0.0011	2,700 ± 48	2,474 ± 55
				07Suihua	7.9	0.0002	2,683 ± 54	2,231 ± 50
				TI Mean	7.48	<0.0001	2,688 ± 53	2,196 ± 54
	QTIG_1	G	Satt288	06Hulan	12.39	<0.0001	2,736 ± 55	1,990 ± 55
				07Harbin	9.62	0.0003	2,680 ± 54	2,147 ± 55
				07Suihua	6.3	0.0012	2,640 ± 50	2,336 ± 55
	QTIO_1	O	Satt241	06Harbin	7	0.002	2,587 ± 54	2,226 ± 50
				06Suihua	5.36	0.0001	2,646 ± 54	2,376 ± 54
				07Hulan	6.38	0.0006	2,700 ± 50	2,380 ± 54
				07Suihua	3.9	0.0004	2,564 ± 54	2,300 ± 53
				TI Mean	5.42	0.0026	2,532 ± 47	2,258 ± 54
	QTIM_1	M	Satt540	05Harbin	5.83	<0.0001	2,648 ± 55	2,410 ± 53
				06Suihua	3.6	0.0015	2,680 ± 53	2,498 ± 55
				07Harbin	7.58	<0.0001	2,700 ± 58	2,366 ± 54
				07Hulan	5.3	0.0022	2,683 ± 50	2,390 ± 55
				TI Mean	3.7	0.0001	2,580 ± 53	2,386 ± 53

[a] Daidzein

[b] Genistein

[c] Glycitein

[d] Total isoflavones

[e] R^2 is *R*-square or the proportion of the phenotypic data explained by the marker locus

[f] SEM (standard error mean): $SD/\sqrt{N}$; where *N* was the number of each of allele

level of phenotypic variation evaluated by these QTL was indicative of the quantitative nature of isoflavone inheritance in soybean seeds, which was similar to the other studies (Njiti et al. 1999; Meksem et al. 2001; Kassem et al. 2004 and Kassem et al. 2006; Primomo et al. 2005).

Quantitative trait loci specific to one environment were also reported by other studies for different traits (Rauh et al. 2002; Price et al. 2002; Li et al. 2003). Inconsistent QTL detection across multi-environments could be due to non- or weak-expression of the QTL, QTL × environment interaction in the opposite direction to the main QTL effects, and/or epistasis. The results here demonstrated that QTL × environment interaction was an important property of many QTL underlying individual and total isoflavone content. Moreover, the interaction between QTL and environment suggested that QTL QTIF_1 (associated with TI) were the best QTL for MAS across six environments (at Harbin in 2006 and 2007, at Hulan in 2006 and 2007, and at Suihua in 2007) and the TI mean (Fig. 1). Therefore, the information of QTL × environment interaction should be considered if MAS was to be applied to the manipulation of quantitative traits.

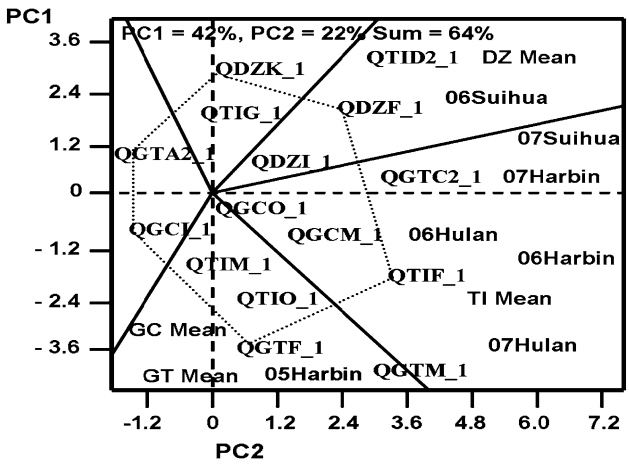

Fig. 1 GT biplot analysis for the relatedness of QTL and environment or individual and total isoflavone mean. *PC1* first principle component, *PC2* second principle component. Different environment was represented by year and location, for example 'at Harbin location in 2005' was represented by 05 Harbin

Significant epistatic interactions have been reported in soybean for other traits including height (Lark et al. 1995) and yield (Orf et al. 1999). Primomo et al. (2005) detected

impact of epistatic interactions on isoflavones in soybean seed using a set of RILs from a cross 'AC756' × 'RCAT Angora' in two locations of Canada, the results showed that 23 epistatic interactions were found and could explained 10% of phenotypic variation. In this study, significant ($P \leq 0.001$) two-way epistatic interactions were sought using EPISTACY 2.0 (Holland et al. 1997) and SAS program on individual environment values. No epistatic interactions were found in this study. It was possible that longer distance between markers made detection of epistatic interactions difficult. The detection of epistasis may benefit from mapping more SSR markers.

In this study, QGCM_1 associated with GC across five environments and GC mean, QGTM_1 associated with GT across three environments and QTIM_1 associated with TI across four environments and TI mean, were identified (Table 2). It was not surprising that these three QTL corresponded to the same SSR marker (Satt540 in LG M) detected previously by Primomo et al. (2005) using 'AC756' × 'RCAT Angora' in two locations of Canada. It should be noted that the germplasm reported by Primomo et al. (2005) was different from the plant material used here. Therefore, Satt540, associated with GC, GT and TI, was identified in both China and North American germplasm across mega-environment conditions. This suggests Satt540 was weakly influenced by genetic background and environment, and may have application in MAS.

Traits that are correlated may have loci in common manifested through linkage or pleiotropy (Aastveit and Aastveit 1993). Genetic correlations observed among isoflavones and resistance traits were investigated. Satt540 was associated with certain foliar resistances such as to aphids (Li et al. 2007) and to white mold (Guo et al. 2008). It is possible that different QTL conditioning these traits are inherited in clusters as tightly linked loci. Alternatively, isoflavones in leaves that protected soybeans from pests or pathogenic microbes (Morris et al. 1991; Benhamou and Nicole 1999) may be transported to seed. Therefore, the region detected here on LG M could represent both a major seed and leaf isoflavone content locus pleiotropic to or clustered with a pest or resistance gene locus. A fine mapping of this region could resolve this issue.

QDZF_1 associated with DZ across seven environments and DZ mean, QGTF_1 associated with GT across six environments and GT mean, QTIF_1 associated with TI across six environments and TI mean, were linked with the same SSR marker (Satt144 in LG F). Satt144 was a new finding related to seed isoflavone content QTL (Table 2). In this region of the soybean genome, a QTL for resistance to foliar herbivory by corn ear worm has been noted previously (Narvel et al. 2001).

The availability of QTL associated with individual and total isoflavones in soybean seed could facilitate MAS in breeding programs aiming to transfer high-isoflavone content from soybean cultivar Zhongdou 27 to other elite breeding lines. Knowledge of the locations of QTL controlling individual and total isoflavone in crosses will allow the design and implementation of more efficient selection schemes to develop high-isoflavone soybean cultivars. So far, progress in breeding high-isoflavone cultivars has been slow because of large environmental interactions on this trait. However, the use of the QTL identified here could improve the design of an efficient and cost-effective breeding strategy for developing high-isoflavone soybean cultivars through MAS.

Acknowledgments This study was conducted in the Key Laboratory of Soybean Biology of Chinese Education Ministry and financially supported by National 863 Projects (contract No. 2006AA10Z1F and 2006AA100104-4), National 973 Project (2004CB117203-4), National Nature Foundation Program (30671318), and Science and Technology Project of Heilongjiang Province (GA06B 101-1-3).

References

Aastveit AH, Aastveit K (1993) Effects of genotype–environment interactions on genetic correlations. Theor Appl Genet 86:1007–1013

Allen FL (1994) Usefulness of plant genome mapping to plant breeding. In: Gresshoff P (ed) Plant genome analysis. CRC Press, Boca Raton, pp 11–18

Benhamou N, Nicole M (1999) Cell biology of plant immunization against microbial infection: the potential of induced resistance in controlling plant diseases. Plant Physiol Biochem 37(10):703–719

Brouns F (2002) Soya isoflavones: a new and promising ingredient for the health food sector. Food Res Int 35(2–3):187–193

Carrao-Panizzi M, Kitamura K (1995) Isoflavone content in Brazilian soybean cultivars. Breed Sci 45:295–300

Choi JS, Kwon TW, Kim JS (1996) Isoflavone contents in some varieties of soybean. Korean Food Biotechnol 5:167–169

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

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Eldridge AC, Kwolek WF (1983) Soybean isoflavones: effect of environment and variety on composition. Agric Food Chem 31:394–396

Graham MY, Graham TL (1991) Rapid accumulation of anionic peroxidases and phenolic polymers in soybean cotyledon tissues following treatment with *Phytophthora megasperma* f. sp. *Glycinea* wall glucan. Plant Physiol 97:1445–1455

Griffith AP, Collison MW (2001) Improved methods for the extraction and analysis of isoflavones from soy-containing foods and nutritional supplements by reversed-phase high-performance liquid chromatography and liquid chromatography mass spectrometry. Chromatogr A 913:397–413

Guo XM, Wang DH, Gordon SG, Helliwell E, Smith T, Berry SA, St Martin SK, Dorrance AE (2008) Genetic mapping of QTLs underlying partial resistance to *Sclerotinia sclerotiorum* in Soybean PI 391589A and PI 391589B. Crop Sci 48:1129–1139

Hoeck JA, Fehr WR, Murphy PA, Welke GA (2000) Influence of genotype and environment on isoflavone contents of soybean. Crop Sci 40:48–51

Holland JB, Moser HS, O'Donoughue LS, Lee M (1997) QTLs and epistasis associated with vernalization in oat. Crop Sci 37:1306–1316

Kassem A, Meksem K, Njiti V, Iqbal MJ, Banz WJ, Winters TA, Wood AJ, Lightfoot DA (2004) Definition of soybean genomic regions that control seed Phytoestrogen amounts. Biotech Biomed 2:52–60

Kassem MA, Shultz J, Meksem K, Cho Y, Wood AJ, Iqbal MJ, Lightfoot DA (2006) An updated 'Essex' by 'Forrest' linkage map and first composite interval map of QTL underlying six soybean traits. Theor Appl Genet 113:1015–1026

Kosslak RM, Bookland R, Barkei J, Paaren H, Appelbaum ER (1987) Induction of *Bradyhizobium japonicum* common nod genes by isolated from *Glycine max*. Proc Natl Acad Sci USA 84:7428–7432

Kudou S, Fleury Y, Welti D, Magnolato D, Uchida T, Kitamura K, Okubo K (1991) Malonyl isoflavone glycosides in soybean seeds (*Glycine max* Merrill). Agric Bio Chem 55:2227–2233

Lander ES, Green P, Abrahamson J, Barlow A, Daly M, Lincoln S, Newburg L (1987) Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lark KG, Chase K, Adler F, Mansur LM, Orf JH (1995) Interactions between quantitative trait loci in soybean in which trait variation at one locus is conditional upon a specific allele at another. Proc Natl Acad Sci USA 92:4656–4660

Latunde-Dada AO, Cabello-Hurtado F, Czittrich N, Didierjean L, Schopfer C, Hertkorn N, Werck-Reichhart D, Ebel J (2001) Flavonoid 6-hydroxylase from soybean (*Glycine max* L.), a novel plant P-450 monooxygenase. Biol Chem 276:1688–1695

Lee SJ, Yan W, Ahn JK, Chung IM (2003) Effects of year, site, genotype and their interactions on various soybean isoflavones. Field Crops Res 81:181–192

Li ZK, Yu SB, Lafitte HR, Huang N, Courtois B, Hittalmani S, Vijayakumar CHM, Liu GF, Wang GC, Shashidhar HE, ZHuang JY, Zheng KL, Singh VP, Sidhu JS, Srivantaneeyakul S, Khush GS (2003) QTL by environment interactions in rice I. heading date and plant height. Theor Appl Genet 108:141–153

Li Y, Hill CB, Carlson SR, Diers BW, Hartman GL (2007) Soybean aphid resistance genes in the soybean cultivars Dowling and Jackson map to linkage group M. Mol Breed 19:25–34

Lozovaya VV, Lygin AV, Zernova OV LISX, Hartman GL, Widholm JM (2004) Isoflavonoid accumulation in soybean hairy roots upon treatment with *Fusarium solani*. Plant Physiol and Bioch 42(7–8):671–679

Meksem K, Njiti V, Banz WJ, Iqbal MJ, Kassem MM, Hyten D, Yuang J, Winters TA, Lightfoot DA (2001) Molecular markers of phytoestrogen content in soybeans. Biomed Biotechnol 1:38–44

Morris PF, Savard ME, Ward EWB (1991) Identification and accumulation of isoflavonoids and isoflavone glucosides in soybean leaves and hypocotyls in resistance responses to *Phytophthora megasperma* f.sp. glycinea. Physiol Mol Plant Pathol 39(3):229–244

Munro IC, Harwood M, Hlywka JJ, Stephen AM, Doull J, Flamm WG, Adlercreutz H (2003) Soy isoflavones: a safety review. Nutr Rev 61(1):1–33

Narvel JM, Walker DR, Rector BG, All JN, Parrott WA, Boerma HR (2001) A retrospective DNA marker assessment of the development of insect resistant soybean. Crop Sci 41:1931–1939

Njiti VK, Meksem Lightfoot DA, Banz WJ, Winters TA (1999) Molecular markers of phytoestrogen content in soybeans. Med Food 2:165–167

Orf JH, Chase K, Adler FR, Mansur LM, Lark KG (1999) Genetics of soybean agronomic traits: II. interaction between yield quantitative trait loci in soybean. Crop Sci 39:1652–1657

Price AH, Towhend J, Jones MP, Audebert A, Courtois B (2002) Mapping QTL associated with drought avoidance in upland rice grown in the Philippines and West Africa. Plant Mol Bio 48:683–695

Primomo VS, Poysa V, Ablett GR, Jackson CJ, Gijzen M, Rajcan I (2005) Mapping QTL for individual and total isoflavone content in soybean seeds. Crop Sci 45:2454–2464

Rauh BL, Basten C, Buckler ES IV (2002) Quantitative trait loci analysis of growth response to varying nitrogen sources in *Arabidopsis thaliana*. Theor Appl Genet 104:743–750

Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. Theor Appl Genet 109:122–128

Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:832–839

Trigizano RN, Caetano-Anolles G (1998) Laboratory exercises on DNA amplification fingerprinting for evaluating the molecular diversity of horticultural species. Hort Technol 8:413–423

Vyn TJ, Yin X, Bruulsema TW, Jackson CC, Rajcan I, Brouder SM (2002) Potassium fertilization effects on isoflavone concentrations in soybean (*Glycine max* (L.) Merr.). Agric Food Chem 50:3501–3506

Wang H, Murphy P (1994) Isoflavone composition of American and Japanese soybeans in Iowa: effects of variety, crop year, and location. Agric Food Chem 42:1674–1677

Yan W (2001) GGEbiplot-a windows application for graphical analysis of multienvironment trial data and other types of two way data. Agron J 93:1111–1117