Korean J. Crop Sci. 49(5) 423~428(2004)

Identification of Quantitative Trait Loci Associated with Isoflavone Contents in Soybean Seed

Myung Sik Kim*‡, Min Jung Park*, Jung Gyu Hwang*, Soo Ho Jo*, Mi Suk Ko*, Ill Min Chung‡, and Jong Il Chung*†**

*Department of Agronomy, Research Institute of Life Sci., Gyeongsang National University, Chinju 660-701, Korea
**Department of Applied Life Science, Konkuk University, Seoul 143-701, Korea

ABSTRACT : Soybean seeds contain high amounts of isoflavones that display biological effects and isoflavone content of soybean seed can vary by year, environment, and genotype. Objective of this study was to identify quantitative trait loci that underlie isoflavone content in soybean seeds. The study involved 85 F_2 populations derived from Korean soybean cultivar 'Kwangkyo' and wild type soybean 'IT182305' for QTL analysis associated with isoflavone content. Isoflavone content of seeds was determined by HPLC. The genetic map of 33 linkage groups with 207 markers was constructed. The linkage map spanned 2,607.5 cM across all 33 linkage groups. The average linkage distance between pair of markers among all linkage groups was 12.6 cM in Kosambi map units. Isoflavone content in F_2 generations varied in a fashion that suggested a continuous, polygenic inheritance. Eleven markers (4 RAPD, 3 SSR, 4 AFLP) were significantly associated with isoflavone content. Only two markers, Satt419 and CTCGAG3 had F-tests that were significant at $P<0.01$ in F_2 generation for isoflavone content. Interval mapping using the F_2 data revealed only two putative QTLs for isoflavone content. The peak QTL region on linkage group 3, which was near OPAG03c, explained 14% variation for isoflavone content. The peak QTL region on linkage group 5, which was located near OPN14 accounted for 35.3% variation for isoflavone content. Using both MapMaker-QTL ($LOD \geq 2.0$) and single-factor analysis ($P \leq 0.05$), one marker, CTCGAG3 in linkage group 3 was associated with QTLs for isoflavone content. This information would then be used in identification of QTLs for isoflavone content with precision

Keywords: soybean, isoflavone, genetic map, molecular marker, QTL

Soybean is widely used as a protein source both for humans and domesticated animals. Soybean contains several biologically active components (e.g., isoflavones, saponins, phytic acids, phytosterols, phenolic acids,..). Isoflavones are found in soybeans in large amounts. The beneficial effects of isoflavones are reduction on risk of several types of cancer (Barnes *et al.*, 1999), of cardiovascular disease (Tikkanen *et al.*, 1998), of osteoporosis (Erdman & Potter, 1997), and attenuation of menopause symptoms (Kurzer, 2000).

Isoflavone content and composition of soybean seeds is very different according to the morphological part of the seed (Kudou *et al.*, 1991) and genotypes and the environment conditions during seed development. Wang & Murphy (1994) observed the difference of isoflavone content in variety (8 American and 3 Japanese varieties), crop year (1989-1991), and location Eldridge & Kwolek (1983) reported total isoflavone of soybeans varied from 116 to 309 mg/g within variety and varied from 46 to 195 mg/g with the same variety grown in different locations. Yang & Chung (2001) reported effect of crop year and genotype on isoflavone content using sixty Korean local cultivars. Meksem *et al.* (2001) observed that the heritability estimates were 79 % for daidzein, 22 % for genistein, and 88 % for glycitein. Lucimara *et al.* (2004) reported that the broad sense heritabilities for the contents of the various isoflavone forms were higher than 90 %. These results indicate that isoflavone content of soybean seeds is controlled as quantitative trait.

Traits of agronomic and economic importance are controlled by polygenes or quantitative trait loci (QTLs). The phenotype of a quantitative character is a consequence of both genetic and environmental sources. Thus, in a polygenic system, an individual gene substitution may well contribute a relatively small effect to the phenotype of the character (Mather & Jinks, 1971). Molecular markers based upon DNA polymorphism have greatly simplified the genetic analysis of quantitative traits, providing a reliable and extensive framework of qualitative markers to which quantitative trait loci (QTLs) can be linked. Much effort has been put into the genetic and genomic analysis of soybeans. DNA markers allowed the construction of genetic linkage maps of soybean (Keim *et al.*, 1990, 1997; Shoemaker & Specht, 1995; Cregan *et al.*, 1999; Kim *et al.*, 2000, 2003; Yamanaka *et al.*, 2002), QTL mapping (Diers *et al.*, 1992a; Mansur *et al.*, 1993; Lee *et al.*, 1996, 1999; Chung *et al.*, 2003). In isoflavone content of soybean seeds, four genomic

†Corresponding author: (Phone) +82-55-751-5421 (E-mail) jongil@gshp.gsnu.ac.kr <Received November 22, 2004>
‡These two authors contributed equally to this work

regions were found to be significantly associated with QTLs controlling isoflavone content using 150 polymorphic SSR markers (Meksem *et al.*, 2001). Information on the association between genetic markers and isoflavone content should help breeders construct beneficial allelic combinations and accelerate the development of high isoflavone soybean cultivars. The purpose of this study was to identify various genetic markers that are linked to loci conditioning variation in isoflavone production.

MATERIALS AND METHODS

Plant materials and DNA marker analysis

Mapping population was derived from a mating between soybean cultivar 'Kwangkyo' (female parent) and wild type soybean 'IT182305' (male parent). F_1 hybrids were made in the greenhouse and F_2 seeds were planted in the field on May, 2001. Young leaves were collected from the 85 individual F_2 plants and parent plants. Genomic DNA was extracted from finely ground leaf tissue by means of a modified CTAB procedure (Saghai Maroof *et al.*, 1984) At maturity, individual F_2 plants were individually harvested and the harvested seeds were air-dried to a constant seed moisture. The procedures of RAPD, SSR, and AFLP marker analysis have been described elsewhere (Kim *et al.*, 2003)

Isoflavone analysis

The ground soybean seeds (2 g) including seed coat from both parents and 85 individual F_2 harvested seeds were mixed with 2 mL of 0.1 N HCl and 10 mL of acetonitrile, stirred for 2 hrs at room temperature, and filtered through Whatman No. 42 filter paper. The filtrate was taken to dryness under vacuum a temperature below 30 °C. The dried material was redissolved in 10 mL of 80 % HPLC grade MeOH in distilled water. A aliquot of sample was filtered through a 0.45 m filter unit and analyzed by HPLC Authentic standards of daidzin, daidzein, genstin, and genstein were obtained from commercial source (Aldrich Chemical Co.). The glycitein was also obtained with some modifications by the method of Wang & Murphy (1994). The instrumentations for HPLC analysis were applied by the method of Wang & Murphy (1994). A linear HPLC gradient was employed: solvent A was 0.1 % glacial acetic acid in distilled H_2O, and solvent B was 0.1 % glacial acetic acid in acetonitrile. Following injection of 20 L of sample, solvent B was increased from 15 % to 35 % over 50 min and then held at 35 % for 10 min. The solvent flow rate was 1 mL/min. The HPLC system used Yon-Lin company and the column employed YMC-AM 303 (ODS 4.52 × 50 mm). The UV detector wavelength was 254 nm. Standard compounds were chromatographed alone and as mixture. HPLC standard chromatogram which was mixed five isoflavones standard was gained. Relative retention times for the major peaks in the extracts were determined for each gradient by dividing the compound's retention time by the retention time of an internal reference.

Linkage map and QTL mapping

All molecular marker data were subjected to Chi-square analysis to test the goodness of fit for observed to expected ratios (1 : 2 : 1 for codominant or 3 : 1 for dominant markers). Only these loci fitting (P = 0.05) the Chi-square test were used in mapping procedures. The software program Mapmaker/Exp 3.0b was used to detect probable linkages at a LOD value of 4 0 and maximum distance 50 cM. After markers were assigned to a linkage group, compare and ripple commands were used to develop linkage groups and established most likely gene order. Genetic distance was described by centimorgan (cM) using Kosambi (1944) mapping function.

Single-factor analysis of variance (ANOVA) was performed for each marker locus to discern the effect of its alleles on isoflavone content (PROC GLM, SAS). For the codominant markers (SSRs and several RAPDs), three phenotypic classes were distinguishable (i.e., homozygous A or B and heterozygous H), whereas for the dominant markers (most RAPDs), only two phenotypic classes were distinguishable (i.e., D vs B when allele B was a recessive null, or C vs A when allele A was the recessive null) A simple F-test was used to determine if the least square, isoflavone content means for the marker types distinguished at a given locus were statistically different. An F-test significance level of $P<0.01$ was chosen for declaring that a marker was linked to a isoflavone content QTL. An interval mapping technique (Lander & Botstein, 1989) was also performed using the computer program MAPMAKER-QTL v. 1.1 (Lincoln & Lander, 1992). This QTL analysis was performed on the mean values. A LOD score of 2.0 was chosen as the threshold to be used to declare the presence of a QTL in any given

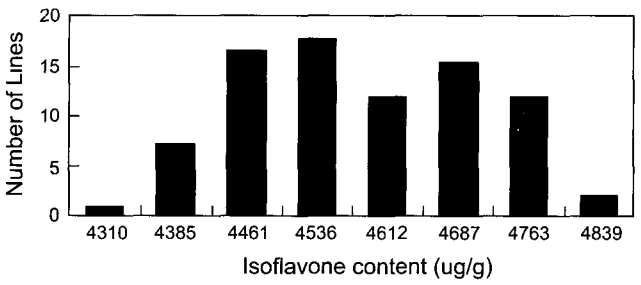

Fig. 1. Frequency distribution of 85 F_2 for isoflavone content in soybean seed Mean isoflavone content of the parent is 4,755 μg/g for Kwangkyo and 3,935 μg/g for IT182305.

the interval between two adjacent markers.

RESULTS AND DISCUSSION

The frequency distributions for isoflavone content in the 85 F_2 plants are shown in Fig. 1. Isoflavone content ranged from 3,927 to 4,827 μg/g in the F_2 generation.

The means of the F_2 generations were 4,536 μg/g. There was transgressive segregation for isoflavone content. Isoflavone content in F_2 generation varied in a fashion that sug-

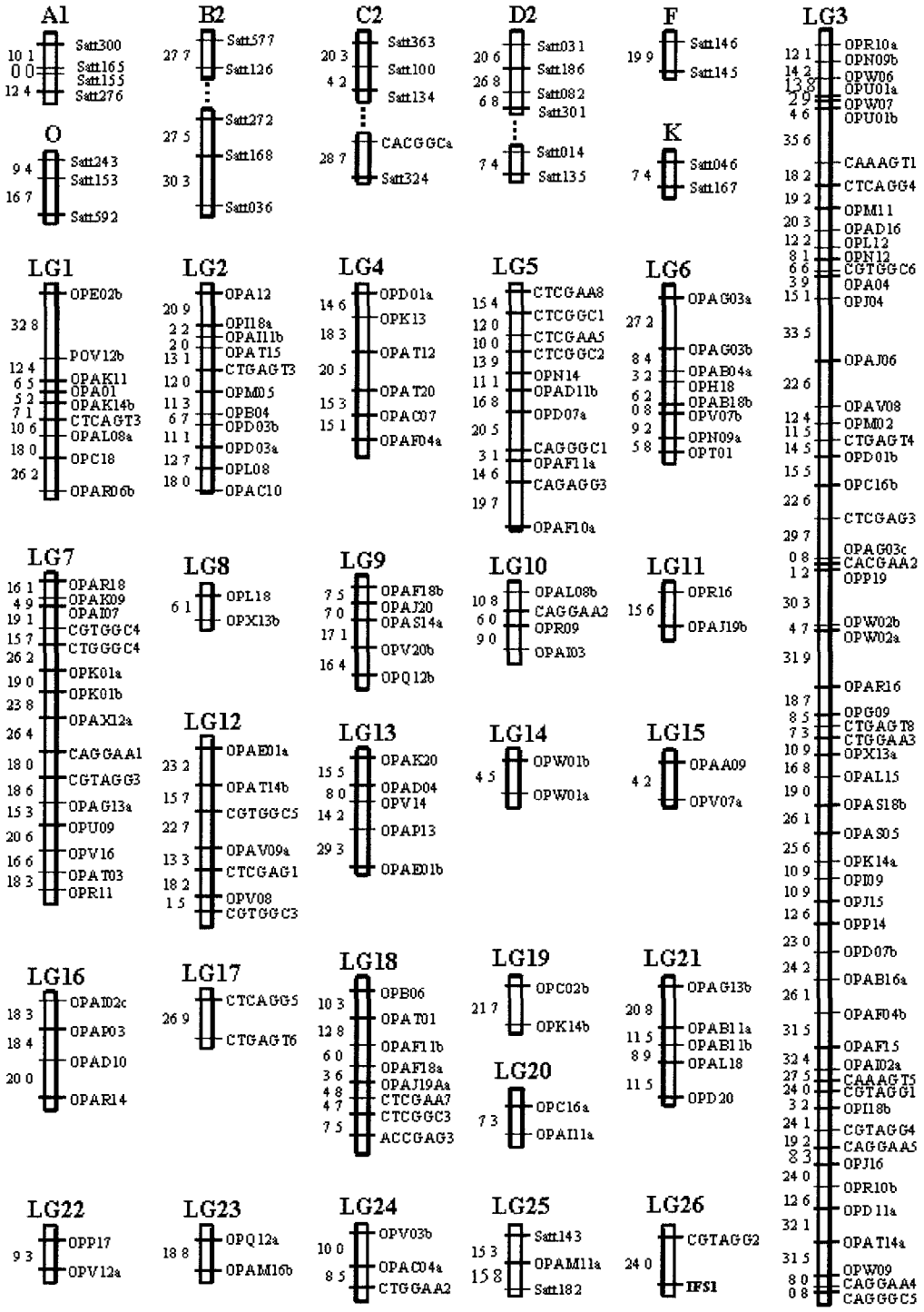

Fig. 2. Soybean linkage map with 207 markers constructed from a segregating F_2 population derived from the mating of Kwangkyo and IT182305 The linkage group names are listed at the top of each group, with map distances shown on the left and marker positions on the right.

gested a continuous, polygenic inheritance. This observation is consistent with an additive gene model of many loci whose alleles have small effects. This result is consistent with previous reports (Meksem *et al.*, 2001).

Of the 1,000 RAPD primers tested on the two parents, only 242 markers produced amplicons that were scorable as either codominant 1 : 2 : 1 segregations or dominant 3 : 1 segregations. Of the 100 SSR primers tested, only 54 markers produced a scorable codominant 1 : 2 : 1 segregation pattern. 83 AFLP markers among 49 combiantions of selective primers tested produced. Total 379 markers (242 RAPD, 54 SSR, and 83 AFLP) were obtained in 85 F_2 mapping population. Total 342 markers (219 RAPD, 50 SSR , and 73 AFLP) were subjected to Chi-square analysis to test the goodness of fit for the observed to expected ratio (1 : 2 : 1 for codominant or 3 : 1 for dominant markers in F_2 population). These markers were used in construction for genetic map. Of the 342 markers, 207 markers were found to be genetically linked. These markers coalesced into 33 linkage groups (Fig. 2). The linkage map spanned 2,607.5 cM across all 33 linkage groups. The average linkage distance between pair of markers among all linkage groups was 12.6 cM in Kosambi map units. By SSR marker homology, A1, B2, C2, D2, F, K, and O was matched with MLG (Cregan *et al.*, 1999). Another linkage groups have been designated LG1-LG26. Compared to previous map (Kim *et al.*, 2003), 31 markers were added and were advanced.

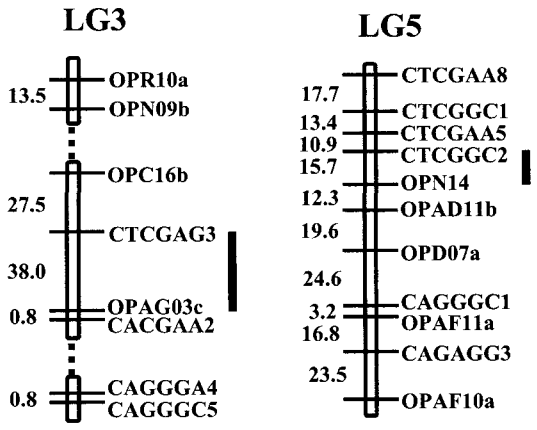

Fig 3. Two linkage groups identified as having putative isoflavone content QTLs in a population derived from a Kwangkyo × IT182305 mating. Map distance (cM) is shown on the left and markers are shown on the right of each linkage group. A solid bar on the right denotes a QTL region exceeding LOD 2.0 as detected by interval mapping.

The results of single-factor analysis between marker and isoflavone content are presented in Table 1.

Eleven markers (4 RAPD, 3 SSR, 4 AFLP) were significantly associated with isoflavone content. Only two markers, Satt419 and CTCGAG3 had F-tests that were significant at $P<0.01$ in F_2 generation for isoflavone content. Interval mapping using the F_2 data revealed only two putative QTLs for isoflavone content (Fig. 3). The peak QTL region on

Table 1. RAPD, SSR, and AFLP loci putatively associated with isoflavone content in the F_2 generation as detected by a single-factor analysis of variance.

Marker	LG	F-value	Pr > F	Means			R^2
				A	B	H	
Satt419	I	6.52	0.0024	4610.3	4453.3	4450.6	0.139
Satt077	D1a+Q	3 28	0 0427	4510 7	4628.7	4521.5	0.075
Satt167	K	3 88	0.0245	4490.1	4487.5	4578.4	0 087
OPM02	3	4 36	0.0399	4476.0	4555.7		0 051
OPAB16a	3	6.18	0.0150	4487.0	4568.9		0.070
CAAAGT5	3	4.13	0.0454	4499.5	4566 1		0.048
CTCGAG3	3	11.58	0.0010	4457.4	4572 4		0 124
CTCGAA6		6.63	0.0118	4498.3	4581.2		0 075
CTCGAA8	5	4.61	0.048	4441.9	4549.6		0 053
OPAE01b	13	5.32	0.0236	4509 7	4588 1		0.061
OPAI13a		5 26	0.0243	4578.6	4503.8		0.060

Table 2. Location and magnitude of QTLs related to isoflavone content in soybean seeds, as determined by interval mapping analysis

Trait	Group	Marker interval	QTL position	LOD peak	Variation explained
Isoflavone	LG3	CTCGAG3-OPAG03c	38.0	2.56	14.0 %
	LG5	CTCGGC2-OPN14	15.7	2 96	35.3 %

linkage group 3, which was near OPAG03c, explained 14% variation for isoflavone content.

The peak QTL region on linkage group 5, which was located near OPN14 accounted for 35.3 % variation for isoflavone content (Table 2).

Using both MapMaker-QTL (LOD 2.0) and single-factor analysis (P 0.05), one marker, CTCGAG3 in linkage group 3 was associated with QTLs for isoflavone content. Meksem *et al.* (2001) reported that molecular linkage group B1 (for glycitein content), N (for glycitein and daidzein), and A1 (for daidzein) contained a major QTLs for each isoflavone content. These QTLs were not detected in our population. This information would then be used in identification of QTLs for isoflavone content with precision

ACKNOWLEDGEMENTS

This work was supported by a grant (CG3122) from Functional Genomics Center (CFGC), 21C Frontier R & D Project, Ministry of Science & Technology.

REFERENCES

Barnes, S., H Kim, and J Xu 1999. Soy the prevention and treatment of chronic disease. Annals of the Brazilian Soybean Congress, pp 265-308. I Brazilian Soybean Congress, Londrina, PR, Brazil.

Chung, J., H. L. Babka, G L. Graef, P. E. Staswick, D. J. Lee, P. B. Cregan, R C Shoemaker, and J. E. Specht. 2003. The seed protein, oil, and yield QTL on soybean linkage Group I Crop Sci. 43 : 1053-1067

Cregan, P. B , T Jarvik, A L. Bush, R. C Shoemaker, K. G Lark, A. L. Kahler, N. Kaya, T. T VanToai, D G Lohnes, J. Chung, and J. E. Specht. 1999. An integrated genetic linkage map of the soybean genome. Crop Sci. 39 . 1464-1490.

Diers, B W., R. Keim, W R Fehr, and R C. Shoemaker 1992 RFLP analysis of soybean seed protein and oil content Theor Appl Genet 83 : 608-612

Eldridge, A C and W. F Kwolek. 1983 Soybean isoflavons: effect of environment and variety on composition. J. Agric. Food Chem 31 : 394-396

Erdman , J. W. Jr. and S. M Potter. 1997. Soy and bone health. The Soy Connection 5(2) 1-4

Kosambi, D. D. 1944. The estimation of map distance from recombination values. Ann. Eugen. 12 · 172-175.

Keim, P, B. W. Diers, T C Olson, and R C Shoemaker 1990 RFLP mapping in soybean. Association between marker loci and variation in quantitative traits Genetics 126 · 735-742

Keim, P., J. M. Schupp, S. E. Travis, K. Clayton, T. Zhu, L. Shi, A. Ferreira, and D. M. Webb 1997 A high-density soybean genetic map based on AFLP markers Crop Sci. 37 537-543.

Kim, H S., S. H. Lee, and Y. H. Lee. 2000. A genetic linkage map of soybean with RFLP, RAPD, SSR, and morphological markers Kor J. Crop Sci. Vol. 45(2) . 123-127.

Kim, H. S., S. H. Lee, K. Y. Park, and Y. H. Lee. 2000. Identification of quantitative trait loci associated with seed size and weight in soybean. Kor. J. Crop Sci. 45(4) : 227-231.

Kim, M. S , Y. J. Cho, D. J. Park, S. J. Han, J. H. Oh, J. G. Hwang, M. S. Ko, and J. I. Chung. 2003. Construction of genetic linkage map for Korean soybean genotypes using molecular markers Korean J. Crop Sci 48(4) : 297-302

Kudou, S , Y, Fleury, D. Welti, D Magnolato, T Uchida, K Kitamura, and K. Okubo. 1991. Malonyl isoflavone glycosides in soybean seeds (*Glycine max* MERRILL). Agric Biol Chem 55 . 2227-2233.

Kurzer, M S 2000. Hormonal effects of soy isoflavones. studies in premenopausal and postmenopausal women, J. Nutr. 130 : 660S-661S.

Lander E. S and D. Botstein. 1989. Mapping mendelian factors underlying quantitative traits using RFLP likage maps. Genetics 121 185-199

Lincoln, S., M. Daly, and E. Lander. 1992 Mapping genes controlling quantitative traits with MAPMAKER/QTL Whitehead Institutue Technical Report. Ed 2

Lee, S. H , M A. Bailey, M. A R Mian, T E Carter, Jr , D A. Ashley, R. S Hussey, W. A Parrott, and H R Boerma. 1996 Molecular markers associated with soybean plant height, lodging, and maturity across locations. Crop Sci. 36(3) . 728-735.

Lee, S H., K Y. Park, H S Lee, and H R. Boerma. 1999. Identification of quantitative trait loci associated with traits of soybean for sprout Korean J Crop Sci 44(2) : 166-170.

Lucimara, C., N. D Piovesan, L. K Naoe, I. C Jose, J. M. S Viana, M A Moreira, and E. G. D. Barros. 2004. Genetic parameters relating isoflavone and protein content in soybean seeds. Euphytica 138 · 55-60

Mather, K. and J L Jinks. 1971. Bimmetrical Genetics. Chapman and Hall, London

Mansur, L. M., K. G. Lark, H. Kross, and A. Oliveira. 1993 Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). Theor. Appl. Genet 86 · 907-913.

Meksem, K , V N Njiti, W J Banz, M. J. Iqbal, M M. Kassem, D. L. Hyten, J. Yuang, T. A Winters, and D. A. Lightfoot 2001 Genomic regions that underline soybean seed isoflavone content J. of Biomedicine and Biotechnology 1(1) : 38-44.

Saghai-Maroof, M. A, K. M Soliman, R A. Jorgensen, and R W Allard. 1984 Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci. 81 · 8014-8018

Shoemaker, R C and J E. Specht 1995 Integration of the soybean molecular and classical genetic linkage groups Crop Sci. 35 : 436-446.

Tikkanen, M. J., K Wahala, S Ojala, V Wihma, and H. Adlercreutz. 1998. Effect of soybean phytoestrogen intake on low density lipoprotein oxidation resistance. Proc Nat Acad Sci. U.S A. 95 : 3106-3110.

Wang, H J. and P. A. Murphy 1994. Isoflavone content in commercial soybean foods. J Agric Food Chem 42: 1666-1673

Yang K J and I. M. Chung. 2001. Yearly and genotypic variations in seed isoflavone content of local soybean cultivars. Korean J. Crop Sci 46(2) 139-144.

Yamanaka, N , S. Ninomiya, M Hoshi, Y. Tsubokura, M Yano, Y. Nagamura, T. Sasaki, and K Harada. 2001. An informative linkage map of soybean reveals QTLs for flowering time, leaflet morphology, and regions of segregation distortion DNA Research 8 61-72