Korean J. Crop Sci. 45(4): 227~231(2000)

Identification of Quantitative Trait Loci Associated with Seed Size and Weight in Soybean

Hong-Sik Kim*†, Suk-Ha Lee, Keum-Yong Park*** and Yeong-Ho Lee*****

**Wheat and Barley Research Division, National Crop Experiment Station, RDA, Suwon 441-100, Korea*
***Dept. of Agronomy, College of Agri. and Life Sci., Seoul National University, Suwon 441-744, Korea*
****Upland Crop Division, National Crop Experiment Station, RDA, Suwon 441-100, Korea*

ABSTRACT: Small seed size is one of the major traits of soybean cultivars for sprouts with regard to high sprout yield. This study was conducted to identify quantitative trait loci (QTL) for seed size and weight in a set of F_6 seeds of 89 lines derived from a cross between 'Pureunkong', a soybean cultivar developed for sprouts and 'Jinpumkong 2', a soybean cultivar with no beany taste in seed due to the lack of lipoxygenases. The genetic map of 25 linkage groups with a total of 98 markers including RFLP, RAPD, SSR and classical markers was constructed from this F_5-derived population and was used for QTL analysis. 'Pureunkong' was significantly smaller ($P<0.01$) than 'Jinpumkong 2' in seed size and seed weight. Genetic variation was detected and transgressive segregation was common in the population for these traits. Seven DNA markers including opT14-1600 in LG A2, opF02-400 in LG B2, Satt100, opC09-700, opG04-730 and opQ11-650 in LG C2, and opY07-1100 & 1000 in LG(unknown) were significantly associated and accounted for 4.7 to 10.9% and 5.1 to 10.1% of the phenotypic variation in seed size and seed weight, respectively. 'Pureunkong' alleles increased seed size and seed weight at the all four significant marker loci on the LG C2. These marker loci in LG C2 were closely linked and were presumed to be a single QTL. Overall, at least three independent QTLs from 3 linkage groups (A2, B2, and C2) were putatively involved in the control of seed size and seed weight.

Keywords : soybean, sprout, quantitative trait loci, seed size, seed weight, genetic linkage map, RFLP, RAPD, SSR.

Soybean sprout is one of the popular soybean food products extensively used in Korea. As the dietary and functional effects of soybean sprout are scientifically proved these days, its consumption is expected to increase. In particular, highly qualified domestic soybean sprouts produced by chemical-free cultivation are preferred by the consumers along with a rise in the standard of living.

For stable production of soybean sprouts with good quality, it is important to develop new soybean varieties with superior quality for sprouts as well as efficient cultivation methods. Even if the criteria of soybean peculiar to soybean sprouts are not laid down, soybean cultivars should have several agronomic traits such as small seed size (<12 g per 100 seeds), high yield potential (>250 kg/10a), resistance to lodging and pest, long germinability ($\geq$ 1 year) and good sensory quality. Small seed size is primarily preferred because it is closely associated with high sprout yield. Soybean cultivars with small seed size were also superior to those with large seed size in major traits for soybean sprouts, such as water uptake, hypocotyl elongation, seed germination rate, seed germinability, and sprout yield (Park *et al.*, 1994; Kim *et al.*, 1994). Calero *et al.* (1981) reported that small seeds tended to have rapid water uptake due to higher percentage by weight of seed coat and large, rounded pores than large seeds in soybean. They also found that small elongated pores and high density of waxy material embedded in the epidermis of large seeds were associated with low absorption of water.

In general, soybean sprout-related traits are quantitatively inherited. It is not easy to evaluate and select appropriate breeding lines for those traits, especially in the early growth stages. Also, these ordinary tasks are sometimes time-consuming and costly. With introduction of molecular genetic maps and DNA markers as biotechnological aids to conventional breeding, marker-assisted selection strategy is recently suggested to improve those traits of interest. In soybean, analyses of quantitative trait loci (QTL) associated with end-use quality-related seed traits (Diers *et al.*, 1992a; Mansur *et al.*, 1993), lodging and maturity (Keim *et al.*, 1990; Lee *et al.*, 1996) were conducted. Lee *et al.* (1999) identified several QTLs associated with traits of soybean for sprout in the 83 F_2-derived lines from a cross between 'Pureunkong' and 'Jinpumkong 2'. In their studies (Lee *et al.*, 1999), some RFLP markers in the linkage group B1 were associated with all three traits including hypocotyl length, abnormal seedling rate, and sprout yield, providing genetic basis for the correlations among these traits. Linkage analyses

†Corresponding author: (Phone) +82-31-290-6680 (E-mail) kimhs@nces.go.kr
<Received April 6, 2000>

of soybean disease resistance controlled by one or more genes have also been conducted for indirect selection by DNA markers (Kim and Diers, 2000; Yu *et al.*, 1994; Diers *et al.*, 1992b).

In our study, we aimed at identification of QTLs associated with seed size/weight and studies of their genetic basis in the Korean soybean genetic background.

MATERIALS AND METHODS

Eighty-nine F_5 lines were developed from a cross between 'Pureunkong', which was released for soybean sprout, and 'Jinpumkong 2' which had no beany taste in seed due to lack of lipoxygenase 1, 2 and 3. Ten seeds of each line were planted in the 30 cm plastic pots and randomly thinned to 5 heathy plants per pot in the R1 growth stage (Fehr *et al.*, 1971). Bed soil in the pots was prepared by mixing loess, compost, and sand with the ratio of 2:2:1. Plants were grown in the greenhouse and seeds were bulk-harvested for each pot.

Twenty F_6 seeds randomly selected for each line were measured and calculated for the mean values for seed length, seed width, and seed thickness of individual lines. This process was repeated three times for each line. Soybean seed size was determined by the following formula:

Seed size=(seed length × seed width × seed thickness)/3

The collected data were used for statistical analysis. For seed weight, 100 seeds were randomly selected and measured three times for each line. Analysis of variance was conducted on the data using PROC GLM of SAS (SAS Institute, 1985).

Genetic linkage map derived from the F_5 population developed from a cross between 'Pureunkong' and 'Jinpumkong 2' (Kim *et al.*, 2000) was used for analysis of QTLs associated with seed size and weight. Significant associations between traits and markers on linkage groups were tested with one-factor analysis of variance using PROC GLM of SAS.

RESULTS AND DISCUSSION

A genetic linkage map was constructed with 113 RFLP (7), RAPD(79), SSR(24), and morphological markers(3) (Kim *et al.*, 2000). The map defined 807.4 cM of the soybean genome comprising 25 linkage groups with 98 markers. Fifteen polymorphic markers remained unlinked. Additional SSR markers are being evaluated for polymorphism between two soybean mapping parents, 'Pureunkong' and 'Jinpumkong 2'. The current genetic map is expected to be more saturated and extended along with introduction of new polymorphic markers.

A soybean cultivar developed for sprouts, 'Pureunkong' was significantly smaller ($P<0.01$) than the other parental cultivar, 'Jinpumkong 2' in seed size (Table 1). They differed on average by 1.5 mm in seed length, 1.3 mm in seed width, and 1.1 mm in seed thickness. The seed size values of the progeny lines in the population were normally distributed ($P>0.01$). They had ranges of 6.1~8.5 mm in seed length, 5.8~7.5 mm in seed width, and 4.5~6.6 mm in seed thickness. As shown in Fig. 1, some progeny lines were smaller than 'Pureunkong' or larger than 'Jinpumkong 2' in seed size.

As for seed weight (100-seed wt), similar trends were observed in the parents and population lines. 'Pureunkong' and 'Jinpumkong 2' significantly differed by 11.9 g for 100-seed wt. Normal distribution was observed in the progenies ranging from 10.1 to 27.8 g (Fig. 1).

One factor analysis of variance conducted for each marker revealed that seven DNA markers were significantly associated with seed size (Table 2). In particular, a SSR marker, Satt100, showed a highly significant association with seed size ($P<0.01$). All of these seven markers were also significantly ($P<0.05$) associated with 100-seed wt. (Table 3). These markers accounted for 4.7 to 10.9% and 5.1 to 10.1% of the phenotypic variations in seed size and seed weight, respectively.

At the significant marker loci on the linkage groups including A2 and B2 except for LG C2, 'Pureunkong' alleles reduced seed size and seed weight (Tables 2 and 3). On

Table 1. Means and ranges of seed size and seed weight for parents and a population of 89 F_6 progeny lines.

	Seed Size				100-Seed Wt
	Length (L)	Width (W)	Thickness (T)	(L × W × T)/3	
Parents	(mm)	(mm)	(mm)		(g)
Pureunkong	6.67	6.08	5.32	72.15	13.60
Jinpumkong 2	8.19	7.41	6.43	130.69	25.48
Population					
Range	6.13~8.46	5.76~7.55	4.50~6.61	53.58~136.17	10.07~27.77
Mean	7.30	6.68	5.58	92.01	18.23
$LSD_{0.05}$	0.17	0.15	0.16	5.94	1.26

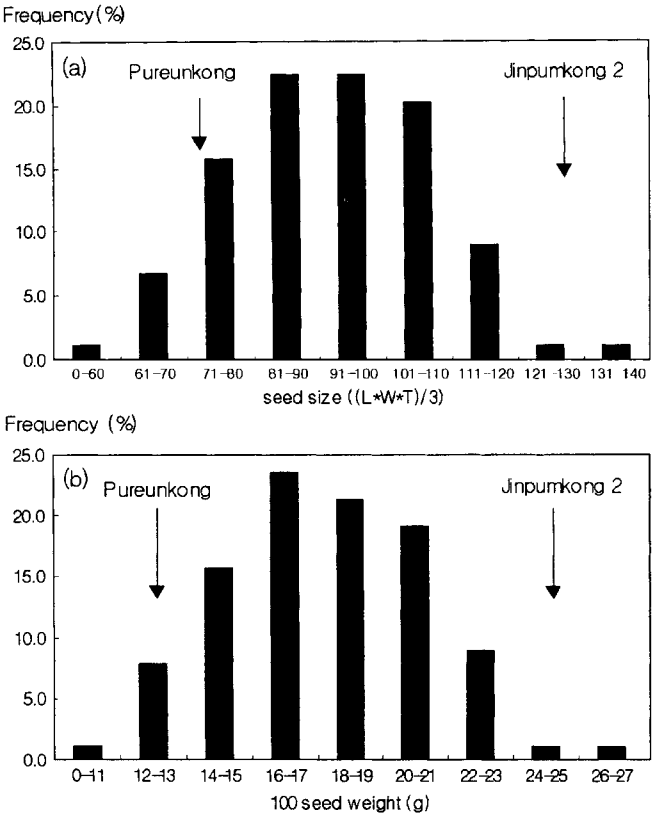

Fig. 1. Frequency distribution of the mean (a) seed size ((length × width × thickness)/3) and (b) seed weight of F_6 lines and parents.

the other hand, Pureunkong alleles increased seed size and seed weight at the all four significant marker loci on the LG C2. As shown in Fig. 2, these four marker loci including Satt100, opC09-700, opG04-730, and opQ11-650 were closely linked and were presumed to be a single QTL associated with seed size and weight. Overall, at least three independent QTLs from 3 linkage groups (A2, B2, and C2) were putatively involved in these traits.

Table 3. DNA markers significantly associated with 100-seed weight at the 0.05 probability level.

Marker	LG	Prob.†	R^2(%)	Marker genotypic means		
				PP	PJ	JJ
Satt100	8 (C2)	*	10.1	18.6	21.1	17.6
$opY07_{1100\ \&\ 1000}$	20 (?)	*	7.8	17.9	14.4	18.8
				P_	JJ	
$opT14_{1600}$	2 (A2)	*	5.3	17.7	19.1	
$opF02_{400}$	6 (B2)	*	5.4	17.5	19.0	
$opC09_{700}$	8 (C2)	*	6.3	19.1	17.5	
$opG04_{730}$	8 (C2)	*	5.1	19.1	17.6	
$opQ11_{650}$	8 (C2)	*	5.2	19.0	17.6	

†*P<0.05

In the previous studies using F_2-derived, F_4-derived and F_7-derived populations with different genetic background, a few QTLs associated with seed weight were identified in at least 14 linkage groups in the USDA soybean linkage map, such as A1, A2, C1, C2, D2, E, F, G, J, K, L, M, P, and R (Mansur *et al.*, 1996; Mian *et al.*, 1996). This suggests that a few number of genes are involved in the control of seed weight and they distribute widely over the soybean genome. Most QTLs except for B031-1n identified in these studies explained 5 to 11% of the total variation of seed weight. The locus B031-1n on the LG G was identified as the most significant QTL with R^2 value of 22% in the 'Young' × 'PI416937' population (Mian *et al.*, 1996).

It is quite possible that 'Pureunkong', the small seeded parent might possess positive genes capable of increasing

Table 2. DNA markers significantly associated with seed size ((L × W × T)/3) at the 0.05 probability level.

Marker	LG	Prob.†	R^2(%)	Marker genotypic means‡		
				PP	PJ	JJ
Satt100	8 (C2)	**	10.9	94.3	106.6	88.6
$opY07_{1100\&1000}$	20 (?)	*	7.4	90.2	74.1	95.1
				P_	JJ	
$opT14_{1600}$	2 (A2)	*	5.8	89.2	96.7	
$opF02_{400}$	6 (B2)	*	4.7	88.8	95.4	
$opC09_{700}$	8 (C2)	*	7.1	96.6	88.4	
$opG04_{730}$	8 (C2)	*	6.0	96.4	88.9	
$opQ11_{650}$	8 (C2)	*	6.0	96.2	88.7	

†*P<0.05; **P<0.01

‡ PP designates homozygous 'Pureunkong' class; PJ designates segregating class; JJ designates homozygous 'Jinpumkong 2' class; P_ designates segregating and homozygous 'Pureunkong' classes.

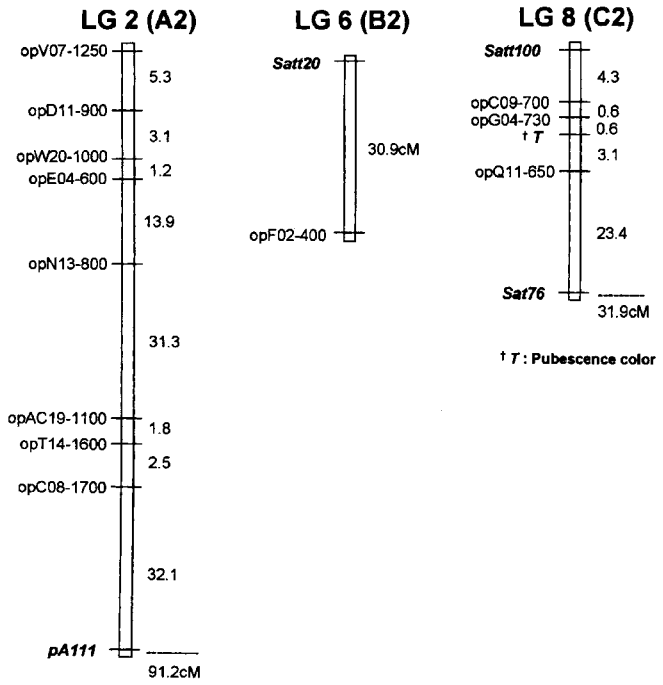

Fig. 2. Linkage groups of putative DNA markers significantly associated with QTL of seed size and seed weight of soybean. This map was constructed from segregation of 89 F_5 lines of the 'Pureunkong' × 'Jinpumkong 2' cross. DNA markers represented with bold type are ankering markers for individual linkage groups in the USDA map. DNA marker loci significantly associated with seed size are identical with those for seed weight (see Tables 2 & 3).

seed size. It is generally known that a phenotype is conditioned by several genes having either positive or negative alleles for most quantitative traits. This has been shown in other studies for various traits and explained in part why transgressive segregation was observed for those traits. For example, Xiao *et al.* (1996) has shown that wild rice germplasm that has very low yield can have genes that contribute to the yield increase of cultivars. This means that of all the alleles for high yield, the wild parent had a few that are missing in the high yielding elite parent. Overall, the high yielding parent has much more positive alleles that increase yield than the wild parent.

In the case of soybean seed weight, similar results of QTLs for larger seed size have already been reported from the small seeded parent (Mansur *et al.*, 1996; Mian *et al.*, 1996). Interestingly, Mian *et al.* (1996) also identified a significant marker locus (A635-1) associated with seed weight on LG C2, where the average 100-seed wt of homozygous class of small seed parent ('Young') was greater than that of large seed parent ('PI416937'). About 6% of the total variation was explained by this locus for seed weight. According to the integrated genetic map published by Cregan *et al.* (1999), this RFLP locus was located 21.2 cM away from Satt100, the SSR marker locus significantly associated with seed size and weight in our study.

One of the major concerns in this study is that those marker alleles on the LG C2 from the small seeded parent had such a large effect. Seed size and seed weight are quantitatively inherited and many genes control these traits in soybean population. We supposed that the genes associated with these marker alleles from the small seeded parent are minor for large seed size. If that is the case, it would not be expected that the homozygous classes for any marker were quite different in seed size and that R-square values for these markers on the LG C2 should be large.

The accurate mapping of QTLs can not be often achieved since some loci might be segregating in the recombinant F_5 or F_6 population. In addition, we analyzed genetic marker loci one-at-a-time at a false positive rate of α=0.05. Many markers are tested in this one-point ANOVA in which it is more likely that type I error may be committed over the genome. For example, we could not exclude the possibility of type I error as for the significant marker opT14-1600 because its two closely linked markers (opAC19-1100 and opC08-1700) in the LG2 (A2) were not significant. More progeny lines should be required in this population to obtain more accurate estimates of phenotypic effects and the genetic location of QTLs.

As a matter of fact, our work reported here may be the preliminary phase of QTL mapping. The map we used is not a complete representative of soybean genome. It covers only the small portion of the genome as compared to the currently published frame map (Cregan *et al.*, 1999). If the genetic map derived from 'Pureunkong' × 'Jinpumkong 2' is more saturated with DNA markers and recombinant inbred line population is analyzed, it would be possible to identify putatively major gene(s) with large effect on controlling seed size and weight. It would be also more likely to find out the major QTLs associated with seed size if the mapping population developed from the parents with more remarkable contrast was analyzed than the parents used here for a target trait.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. Seok-Dong Kim, General Director of the National Honam Agricultural Experiment Station, RDA who was the Director of the Upland Crop Division, NCES for his support to complete postdoctoral research project in 1998.

REFERENCES

Calero, E., S. H. West, and K. Hinson. 1981. Water absorption of

soybean seeds and associated causal factors. *Crop Sci.* 21 : 926-933.

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

Diers, B. W., L. M. Mansur, J. Imsande, R. C. Shoemaker. 1992b. Mapping phytophthora resistance loci in soybean with restriction fragment length polymorphism markers. *Crop Sci.* 32 : 377-383.

Fehr, W. R., C. E. Caviness, D. T. Burmood, and J. S. Pennington. 1971. Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. *Crop Sci.* 11 : 929-931.

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

Kim, H.S. and B.W. Diers. 2000. Inheritance of partial resistance to Sclerotinia stem rot in soybean. *Crop Sci.* 40 : 55-61.

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

Kim, Y. H., S. D. Kim, and E. H. Hong. 1994. Characteristics of soy sprouts cultivated with soybeans for sprout. RDA *J. Agri. Sci.* 36(2) : 107-112.

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.

Lee, S. H. M. A. Bailey, M. A. R. Mian, E. R. Shipe, D. A. Ashley, W. A. Parrott, R. S. Hussey, H. R. Boerma. 1996. Identification of quantitative trait loci for plant height, lodging, and maturity in a soybean population segregating for growth habit. *Theor. Appl. Genet.* 92 : 516-523.

Mansur, L. M., J. H. Orf, K. Chase, T. Jarvik, P. B. Cregan, and K.G. Lark. 1996. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. *Crop Sci.* 36 : 1327-1336.

Mansur, L. M., K. G. Lark, H. Cross, 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.

Mian, M. A. R., M. A. Bailey, J. P. Tamulonis, E. R. Shipe, T. E. Jr. Carter, W. A. Parrott, D. A. Ashley, R. S. Hussey, H. R. Boerma. 1996. Molecular markers associated with seed weight in two soybean populations. *Theor. Appl. Genet.* 93 : 1011-1016.

Park, K. Y., S. D. Kim, and Y. H. Ryu. 1994. Water uptake, cotyledon damage after imbibition and hypocotyl elongation in soybean with different seed size and color. *Korean J. Crop Sci.* 39(4) : 331-338.

SAS Institute. 1985. SAS User's guide: Statistics 5th ed. SAS Inst., Cary, NC.

Xiao, J., S. Grandillo, S. N. Ahn, S. R. McCouch, S. D. Tanksley, J. Li, and L. Yuan. Genes from wild rice improve yield. *Nature.* 384 : 223-224.

Yu, Y. G., M. A. Saghai-Maroof, G. R. Buss, P. J. Maughan, and S.A. Tolin. 1994. RFLP and microsatellite mapping of gene for soybean mosaic virus resistance. *Phytopathology.* 84 : 60-64.