Korean J. Crop Sci. 48(5): 413~418(2003)

Genetic Variation in Sprout-related Traits and Microsatellite DNA Loci of Soybean

Kyujung Van*, Moon Young Kim*, Jae-Gyun Gwag*, Kyung-Geun Bae, Young Jin Oh***, Kyong Ho Kim***, Ho Ki Park***, and Suk-Ha Lee*†**

**School of Plant Science, Seoul National University, Seoul 151-742, Korea*
***Pulmuone R&D Center, Seoul 120-600, Korea*
****National Honam Agricultal Experiment Station, RDA, Iksan 570-080, Korea*

ABSTRACT: Genetic diversity and soybean sprout-related traits were evaluated in a total of 72 soybean accessions (60 *Glycine max*, 7 *Glycine soja*, and 5 *Glycine gracilis*). 100-seed weight (SW) was greatly varied and ranged from 3.2 g to 32.3 g in 72 soybean accessions. Positive correlation was observed between GR and hypocotyl length (HL), whereas negative correlation was observed between SW and hypocotyl diameter (HD). Re-evaluation by discarding two soybean genotypes characterized with low GR indicated that much higher correlation of sprout yield (SY) with HD and SW. Based on the principal component analysis (PCA) for sprout-related traits, 57 accessions were classified. Soybean genotypes with better traits for sprout, such as small size of seeds and high SY, were characterized with high PCA 1 and PCA 2 values. The seed size in second is small but showed low GR and SY, whereas the third has large seed, high GR and more than 400% SY. In genetic similarity analysis using 60 SSR marker genotyping, 72 accessions were classified into three major and several minor groups. Nine of twelve accessions that were identified as the representatives of soybean for sprout based on PCA were in a group by the SSR marker analysis, indicating the SSR marker selection of parental genotypes for soybean sprout improvement program.

Keywords: genetic variation, microsatellite, sprout-related traits, soybean

Soybean gets attention because it contains many beneficial components to human health and even prevents certain diseases like cancer (Naim *et al.*, 1976; Messina and Barnes, 1991; Boué *et al.*, 2003). Also, the isoflavones daidzein and genistein present as high concentrations in soybean are responsible for many health benefits of soybean (Messina and Barnes, 1991; Barnes, 1997). So, foods made of soybean such as tofu, soymilk, fermented products and sprouts, are popularly used in not only Asia but also throughout the world. Especially, soybean sprout is a popular food type and a year-round vegetable because it is excellent source of minerals as well as vitamins (Wijeratne and Nelson, 1986).

Selection of soybean variety for sprout is an important key factor for growth of soybean sprout. Small seed size, less than 12 g per 100-seed weight, is a favorable condition because it gives higher germination rate and leads to higher sprout yield (Lee *et al.*, 1999). Also, hypocotyl length is considered to be the other significant factor for increasing soybean sprout yield depending on genotypes (Kim, 1981; Kim *et al.*, 1994). However, breeding program for improving soybean sprout yield has not developed well, although significant genetic traits associated with soybean sprout was observed. Multiple gene controls by traits, time-consuming and complex quality determination for soybean sprout leaded to slower development of new soybean genotypes for sprouts (Lee *et al.*, 1999).

The identification of quantitative trait loci (QTL) would be helpful for improving soybean sprout yield. Lee et al. (1999) were used 92 restriction fragment length polymorphism (RFLP) loci and two morphological markers (*W1* and *T*) associated with hypocotyls length, abnormal seedling rate and sprout yield in 100 F_2-derived lines from a cross of PureunkongxJinpumkong 2. Four markers (A802n on Linkage group (LG) B1, A069 on LG E, Cr321 on LG F and A235 on LG G) were associated with 50-seed weight and Bng 119, K455n and K418n were linked to QTLs conditioning abnormal seedling rate. Also, sprout yield showed association with four loci, L154 (LG G) for Pureunkong and A089 (LG B1), A688n (LG K) and B046 (LG L) for Jinpumkong 2 (Lee *et al.*, 2001).

Due to low polymorphism of RFLP markers among soybean genotypes (Akkaya *et al.*, 1995; Cregan *et al.*, 1999), simple sequence repeats (SSRs) markers were used for detection of polymorphisms among multiple populations to create new version of the genetic frame map (Akkaya *et al.*, 1992; Maughan *et al.*, 1995). Thus, SSRs in USDA soybean consensus map with all 20 common LGs (Cregan *et al.*, 1999) might be helpful for identification of QTL related to soybean sprout yield.

In this study, germination rate (GR), hypocotyls length

†Corresponding author: (Phone) +82-2-880-4545 (E-mail) sukhalee@snu.ac.kr
<Received November 9, 2003>

(HL), root length (RL), hypocotyls diameter (HD), 100-seed weight (SW), sprout yield (SY) were evaluated as characters of soybean for sprouts with 57 genotypes including domestic and wild soybean. And, we also tested 72 genotypes with SSR markers for clarifying relationship between genetic variation and characterization of soybean for sprouts.

MATERIALS AND METHODS

Plant materials and evaluation of sprout-related traits

A total of 72 genotypes were selected for this study, including collections of domestic and wild species from Korea and even from other parts of the world (Table 1). Wild soybean species as well as soybean genotypes with large seed size were included in this study to compare sprout-related traits better.

For evaluating sprout characters, 40 seeds each from 57 genotypes were kept with three replications at 5°C in dark for germination after measurement of SW. GR were recorded and decayed seeds were excluded for SW measurement averaged over three replications. HL, RL and HD were evaluated by calipers and also averaged over three replications. SY was calculated as % of fresh weight over SW.

DNA extraction

After healthy and young leaves were harvested three weeks after germination, genomic DNA was isolated by CTAB method (Gelvin *et al.*, 1995). A Hoechst dye-based protocol was used for determination of DNA concentration by a fluorescence spectrophotometer (Model F-4500, Hitachi Ltd., Ibaragi, Japan). The DNA solutions were adjusted to 25 ng/ml with Tris-EDTA buffer (pH 8.0) and stored at -20°C until use.

SSR marker analysis

Among SSR markers from SoyBase website (http://soybase.agron.iastate.edu/), 60 markers were randomly selected for investigating polymorphisms. The PCR reaction was conducted in a volume of 10 μl, containing 50 ng of DNA template, 5 μM each primer, 250 μM of each dNTP, 1× GeneAmp PCR buffer plus 15 mM $MgCl_2$ and 0.5 U of AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA). PCR reactions were initiated with 32 cycles of 94°C for 25 sec, 47°C for 25 sec and 68°C for 25 sec followed by 30 min of final extension at 72°C.

After amplification on a PTC-100 programmable thermal controller (MJ Research, Watertown, MA, USA), the PCR products were confirmed by gel electrophoresis on 1.4% ethidium bromide stained agarose. ABI PRISM 377 DNA sequencer (Applied Biosystems, Foster City, CA, USA) was used for accurate estimate of amplified product size. GeneScan 3.1 (Applied Biosystems, Foster City, CA, USA) and Genotyper 2.0 softwares (Applied Biosystems, Foster City, CA, USA) were also used for gel image analysis and accurate characterization of the alleles, respectively.

Data analysis

Variance and correlation analysis for soybean sprout-related traits as well as principal component analysis (PCA) were analyzed by SAS Program (SAS institute, Inc., Cary, NC, 1990). To confirm genetic variation of SSR markers, polymorphic information content (PIC) value was determined (Botstein *et al.* 1980). Each soybean genotype was given 1 or 0 depending on presence or absence of each band, respectively. And, genetic relationship between microsatellite scores and sprout-related traits were analyzed by simple matching coefficients using NTSYS-pc program (Rohlf, 2000). Based on Jaccardís similarity/dissimilarity coefficients, dendrogram was created from 72 soybean genotypes with unweighted pair group method (UPGMA) (Sokal and Michener, 1958).

RESULTS AND DISCUSSION

Characterizations of sprout-related traits

In this study, sprout-related traits were evaluated with 72 soybean varieties including accessions for sprout and regular purpose and wild genotypes. A total of 72 genotypes were started with investigating sprout-related traits, including 60 *Glycine max*, 7 *G. soja*, and 5 *G. gracilis*. Due to incomplete GR caused by bad seed quality, only 57 of them were finally used for measurements (Table 1). Several characters for sprout-related traits were investigated. Average of GR was 90.6%. 100% of GR was observed in 14 among 57 genotypes whereas hard seed of domestic species collected from Yanggujaerae showed only 5.3% of GR. Generally, higher than 85% of GR was observed among 57 accessions (Table 1). About 2/3 of accessions had less than 15g of SW and higher than 700% in SY were recorded in 16 accessions. Also, large genotypic variation was observed in HL, RL and SW (Table 1). Our results showed phenotypic characters related to sprout were significantly different among 72 genotypes (Table 1), indicating the large genotypic variation in soybean sprout-related traits for improving soybean genotypes for sprout.

Correlations among sprout-related traits

GR showed significant correlation among sprout-related

Table 1. Sprout-related traits of 57 soybean genotypes.

Genotypes	GR[a] (%)	HL (cm)	RL (cm)	HD (mm)	SW (g)	SY (%)
Banjaebae 1	63	9.1	7.4	2.8	14.0	358
Bukwangkong	92	7.1	6.9	2.7	16.4	504
Camp	96	7.7	7.2	2.6	9.4	633
Dawonkong	100	8.5	8.5	2.8	9.5	770
Eunhakong	96	7.5	5.1	2.8	12.9	571
Hannamkong	89	9.0	6.5	2.6	11.5	572
Iksan 10	99	8.3	9.2	2.8	9.7	758
Iksan 15	96	8.6	7.6	2.8	10.1	755
Iksan 18	97	8.2	5.7	3.0	10.7	745
Iksan 2	99	7.1	7.4	2.8	13.2	593
Iksanjarae-3	99	7.8	9.5	2.6	11.2	670
Iksannamulkong	100	7.0	7.4	3.0	15.7	590
Il-2	97	10.1	8.7	2.9	11.0	732
Iri 3	87	9.1	7.9	2.7	10.1	571
IT021845	92	8.9	6.1	2.3	8.2	684
IT157613	44	8.6	9.3	2.3	5.8	383
IT182305	93	7.5	6.0	2.1	4.5	659
IT184204	13	6.9	6.4	1.2	3.2	176
IT194539	100	9.3	9.1	2.5	10.0	695
Jangyupkong	99	8.7	8.1	3.2	21.7	492
Jejujaerae-1	91	9.6	9.7	2.5	8.8	715
Jinju 1	99	9.4	7.7	2.7	9.8	815
Jinpumkong 2	83	5.9	8.2	3.2	21.4	409
Junnam 24	100	7.6	9.0	2.8	9.7	701
Junnamjaerae-1	93	8.1	8.7	2.8	10.1	727
Kangjinjaerae	100	8.5	10.2	2.6	10.7	752
Keungnamjaerae-1	99	8.2	7.6	2.7	9.7	786
Kochangjaerae-5	99	9.6	7.7	2.4	8.7	751
Koheungjaerae	99	10.0	8.1	2.8	9.6	803
Kumjungkong 1	96	7.7	8.5	3.2	25.5	463
Kureajaerae-26	96	11.4	7.7	2.9	12.7	721
Kwangahnkong	99	7.3	9.0	3.1	18.7	504
Mallikong	100	10.0	10.3	2.9	20.4	575
Milyang 68	97	5.4	7.5	3.6	25.8	444
Milyang 78	100	7.7	8.8	3.0	24.4	436
Milyangjaerae	99	4.9	10.2	3.6	33.8	370
Muhankong	97	7.9	9.9	3.2	22.0	438
Myungjunamulkong	100	7.7	8.7	2.8	15.1	623
Namhaekong	95	7.0	7.4	3.0	18.1	481
NOD1-3	85	8.1	7.3	2.7	15.2	461
Paldokong	99	7.5	9.6	3.0	12.9	625
Peking	52	8.6	8.7	2.6	10.5	344
PI458082	96	8.7	7.9	2.6	10.2	679
PI82264	97	8.5	9.7	2.6	7.8	726
PI85505	100	9.6	7.8	2.4	8.2	733
Pungsannamulkong	95	7.4	6.7	2.9	13.2	630
Pureunkong	100	8.4	11.1	3.0	14.6	559
Samnamkong	92	7.0	8.0	3.2	22.1	473
Sinpaldalkong	100	6.9	10.1	3.0	20.5	532
Sinpaldalkong 2	100	9.1	8.0	2.8	20.5	504
Sobaekkong	96	6.6	8.4	2.9	12.1	642
SS2-2	89	7.2	6.7	2.9	18.3	369
SS2-4	69	7.0	5.4	2.5	14.4	373
Suwon 157	96	8.0	10.6	2.9	23.5	418
Taekwangkong	100	7.7	9.1	3.2	25.3	436
Williams 82	100	7.6	7.2	2.9	17.3	507
Yanggujaerae	5	6.8	1.8	1.1	4.6	141
Min.	5	4.9	1.8	1.1	3.2	141
Max.	100	11.4	11.1	3.6	33.8	815
Average	91	8.1	8.1	2.8	14.1	571

[a]GR=germination rate; HL=hypocotyls length; RL=root length; HD=hypocotyl diameter; SW=100-seed weight; SY=sprout yield.

Table 2. Correlation among sprout-related traits of soybean.

	GR[a]	HL	RL	HD	SW	SY
GR	1.000					
HL	0.148ns	1.000				
RL	0.473***	0.152ns	1.000			
HD	0.713***	–0.197ns	0.497***	1.000		
SW	0.355**	–0.423***	0.339**	0.775***	1.000	
SY	0.661***	0.521***	0.234ns	0.122ns	–0.399**	1.000

[a]See Table 1 for the abbreviation. ***Significant at 0.001 level. **Significant at 0.01 level. ns=nonsignificant.

traits except HL ($P<0.01$ or $P<0.001$) (Table 2). HL was negatively correlated with SW (r=-0.423, $P<0.001$), whereas HL and SY showed positively significant correlation (r=0.521, $P<0.001$). RL was significantly correlated with GR (r=0.473, $P<0.001$), HD (r=0.497, $P<0.001$) and SW (r=0.339, $P<0.01$), but there was no significant correlation between RL and SY. At $P<0.001$, positive significant correlations were found between HD and GR (r=0.713), RL (r=0.497), SW (r=0.775). SW showed positive correlation with GR (r=0.355, $P<0.001$), RL (r=0.339, $P<0.001$) and HD (r=0.775, $P<0.001$), but negative correlation with HL (r=-0.423, $P<0.001$) and SY (r=-0.399, $P<0.001$). Additionally, SY was found to be positively correlated with GR (r=0.661, $P<0.001$) and HL (r=0.521, $P<0.001$) and negatively with SW (r=-0.399, $P<0.01$).

Correlation analysis was repeated after exclusion of two accessions, showing less than 5% of GR that was caused by seed hardness, undesirable trait for soybean sprout (data not shown). Non-significant correlation between SY and HD was turned to negative correlation at $P<0.01$. Also, SY was negatively correlated with SW at $P<0.001$ instead of $P<0.01$. It was suggested that small seed size, high GR and long HL were traits for determining SY of soybean sprout (Table 2).

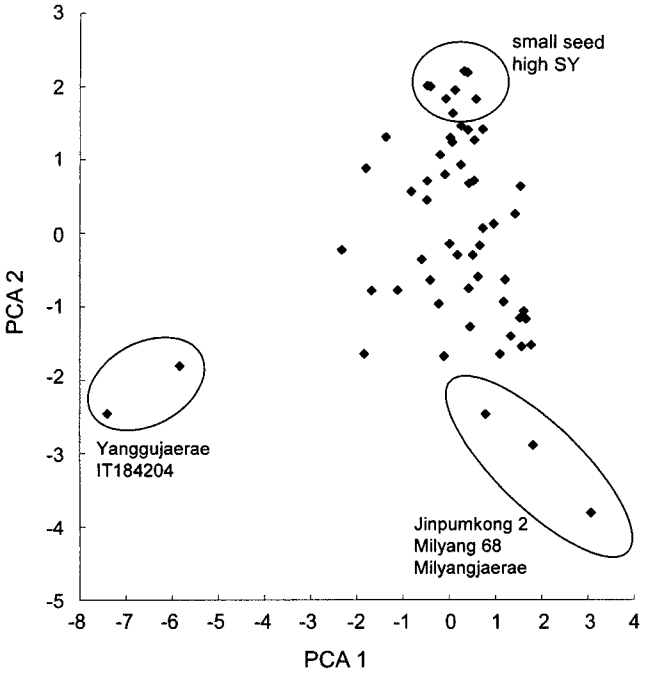

Fig. 1. Grouping of soybean genotypes using principal component analysis (PCA) based on sprout-related traits.

Grouping of soybean genotypes by sprout-related traits

PCA was performed for sprout-related traits of 57 soybean genotypes (Fig. 1). Yanggujaerae and IT184204 with small seed size and yield due to hard seed were placed in the lower value of PCA 1 and PCA 2. Jinpumkong 2, Milyang 68 and Milyangjaerae characterized with large seed and short HL were placed in high value of PCA 1, but low value of PCA 2. Soybean genotypes with high GR, good HL, small seed size and high SY were shown in high value of PCA 1 and PCA 2 (Fig. 1). A total of 12 accessions (Dawonkong, Iksan 10, Kangjinjaerae, Il-2, Jinju 1, PI82264, PI85505, IT194539, Kureajaerae-26, Kochangjaerae-5, Jejujaerae-1 and Iksan 15) with high PCA 1 and PCA 2 were thought to be the representatives of soybean for sprout (Fig. 1).

SSR marker and cluster analyses

Microsatellites are a power tool for studying genetic diversity. SSR markers were used to survey for DNA variation in this study. 60 SSR markers were used for calculating PIC value, measuring genetic variation of markers. It would be very efficient for estimating genetic diversity, if PIC value of each marker was closer to 1. Higher than 0.6 in PIC value was shown in 51 out of 60 markers, indicating high efficiency of SSR markers for measurement of genetic diversity (data not shown).

Based on amplified fragments by SSR markers, a dendrogram was created by the simple matching coefficient method. The dendrogram classified 72 accessions into three major and several minor groups (Fig. 2). A total of 9 out of 12 accessions that were evaluated as the representatives for soybean sprout based on phenotypic data were grouped

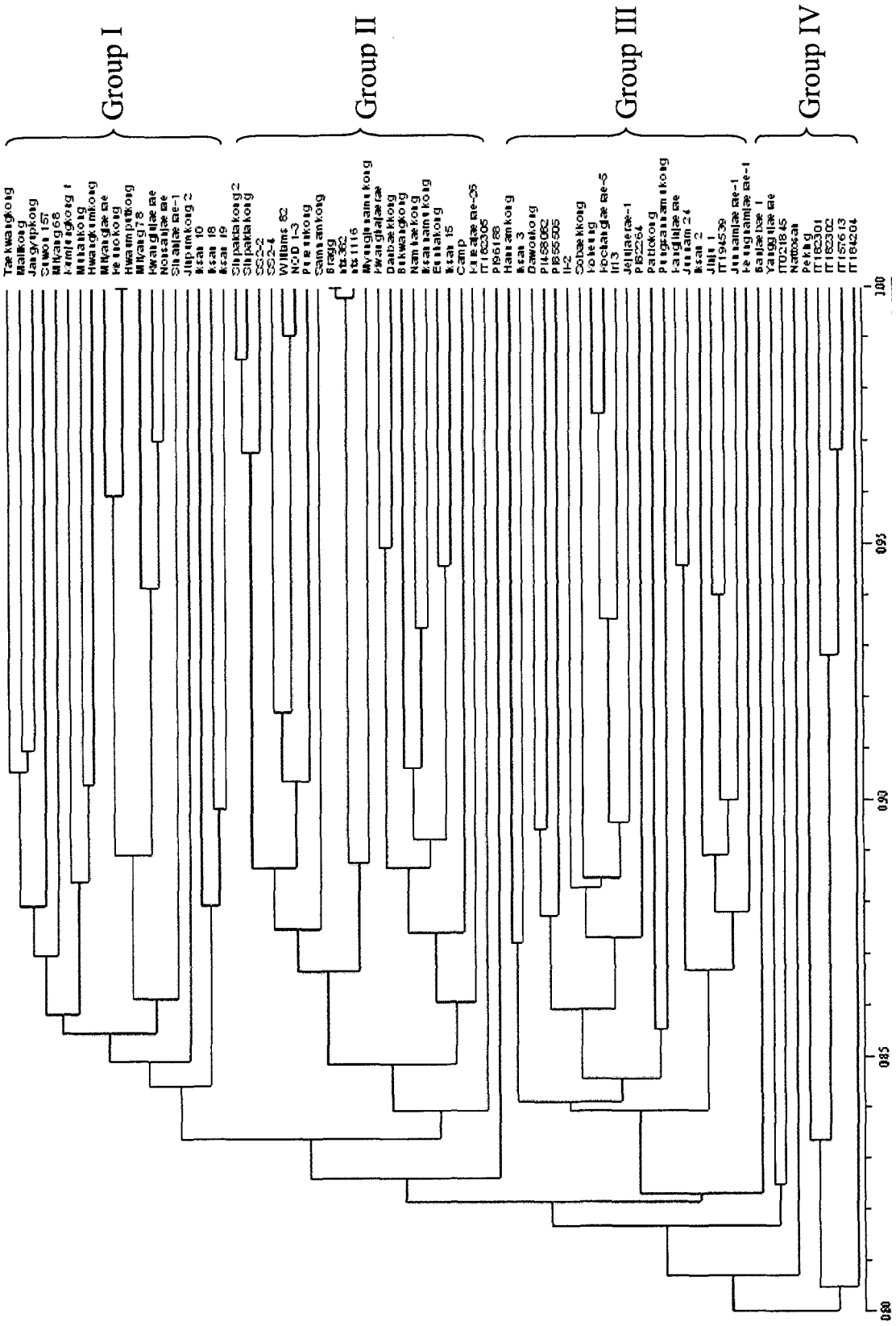

Fig. 2. Dendrogram of 72 soybean genotypes produced by using UPGMA based on simple matching coefficient computed with 60 SSR markers. These genotypes were classified by four different groups.

together in Group III (Fig. 2). Thus, morphological traits for sprout could be evaluated fairly by microsatellite information. Therefore, these SSR markers could be used for selection of parents for crossing population.

ACKNOWLEDGEMENTS

This work was supported by a grant from Crop Functional Genomic Center, 21C Frontier R & D Project, Ministry of Science & Technology in Korea. We thank National Instrumentation Center for Environmental Management at Seoul National University in Korea.

REFERENCES

Akkaya, M.S., A.A. Bhagwat and P.B. Cregan, 1992. Length polymorphism of simple sequence repeat DNA in soybean. Genetics. 132 : 1131-1139.

Akkaya, M.S., R.C. Shoemaker, J.E. Specht, A.A. Bhagwat and P.B. Cregan. 1995. Integration of simple sequence repeat (SSR) DNA markers into a soybean linkage map. Crop Sci 35 : 1439-1445.

Barnes, S. 1997. The chemopreventive properties of soy isoflavonoids in animal models of breast cancer. Breast Cancer Res Treatment 46 : 169-179.

Boué, S.M., T.E. Wiese, S. Nehls, M.E. Burow, S. Elliott, C.H. Carter-Wientjes, B.Y. Shih, J.A. McLachlan and T. E. Cleveland. 2003. Evalution of the estrogenic effects of legume extracts containing phytoestrigens. J Agric Food Chem 51 : 2193-2199.

Botstein, D., R.L. White, M. Skolnick and R.W. Davis, 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphism. Am J of Hum Genet 32: 314-331.

Cregan, P.B., T. Javik, 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.

Gelvin, S.B. and Schilperoort, R.A. 1995. Plant Molecular Biology Manual. Norwell, MA: Kluwer Acedemic Publishers.

Kim, K.H. 1981. Studies on the growing characteristics of soybean sprout. Korean J Food Sci Technol. 13 : 247-251.

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 : 107-112.

Kwon, S.H., K.H. Lim and J.R. Kim. 1972. Studies on diversity of seed weight in Korean soybean land races and wild soybean. *Korean J Breed.* 4 : 70-74.

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 : 166-170.

Lee, S.H., K.Y. Park, H.S. Lee and H.R. Boerma, 2001. Genetic mapping of QTLs conditioning soybean sprout yield and quality. *Theor Appl Genet.* 103 : 702-709.

Maughan, P.J., M.A. Saghai Maroof and R.G. Buss, 1995. Microsatellite and amplified sequence length polymorphism in cultivars and wild soybean. Genome. 38 : 715-723.

Messina, M. and S. Barnes, 1991. The role of soy products in reducing risk of cancer. *J. Natl Cancer Inst.* 83 : 2971-2875.

Naim, M, B. Gestetner, A. Bondi and Y. Birk, 1976. Antioxidative and activities of soybean isoflavones. *J. Agric Food Chem.* 24 : 1174-1177.

Rolf, F.J. 2000. NTSYS-pc Numerical taxonomy and multivariate analysis system, Version 2.0. Applied Biostatistics. New York.

Sokal, R.R. and C.D. Michener, 1958. A statistical method evaluating systematic relationship. Univ. of Kansas Sci. Bull. 38 : 1409-1438.

Wijeratne, W.B. and A.L. Nelson, 1986. Proceeding for food legume improvement for Asia farming system. pp. 183-192.