Korean J. Crop Sci. 46(4): 334~340(2001)

Evaluation of Genetic Diversity among Soybean Genotypes Using SSR and SNP

P. Tanya*, P. Srinives*, T. Toojinda*, A. Vanavichit*, Bo-Keun Ha, Jeong-Suk Bae**, Jung-Kyung Moon*** and Suk-Ha Lee**†**

**Dept. of Agronomy, Kasetsart Univ., Kamphaeng Saen, Nakhon Pathom 73140, Thailand*
***School of Plant Science, Seoul National Univ., Suwon 441-744, Korea*
****National Crop Experiment Station, RDA, Suwon 441-100, Korea*

ABSTRACT: Two different types of molecular markers, simple sequence repeat (SSR) and single nucleotide polymorphism (SNP), were used to measure genetic diversity among five Korean, eight Thai, and three wild soybeans. For SSR analysis, a total of 20 markers were surveyed to detect polymorphisms. For SNP analysis, four primers were designed from consensus sequence regions on disease resistance protein homolog genes, and used to amplify the genomic region. The PCR products were sequenced. A number of polymorphic SSR and SNP bands were scored on all genotypes and their genetic similarity was measured. Clustering analysis was performed independently on both types of markers. Clustering based on SSR markers separated the genotypes into three main groups originated from Korea, Thailand, and wild soybeans. On the other hand, two main groups were classified using SNP analysis. It seemed that SSR was more informative than SNP in this study. This may be due to the fact that SNP was surveyed on the smaller genomic region than SSR. Grouping based on the combined data of both markers revealed similar results to that of SNP rather than that of SSR. This might be due to the fact that more loci from SNP were considered to measure genetic relatedness than those from the SSR.

Keywords **:** soybean, simple sequence repeat, SSR, single nucleotide polymorphism, SNP

The genus *Glycine* consists of two subgenera, i.e. *Glycine* and *Soja*. *Glycine max* (L.) Merr. and *G. soja* Sieb. & Zucc. belong to *Soja* subgenus. *G. max* is the cultivated soybean while *G. soja* is its wild progenitor. Soybean (*Glycine max* (L.) Merrill) is an important crop as a source of protein and oil. Charles and Morse (1923) reported that the origin of soybean is China, Manchuria and Korea.

Molecular markers have become new tools to save time and labor in identifying genotypes in most plant breeding steps, including introduction, selection and certification. Among them, simple sequence repeat (SSR), also known as microsatellite or short tandem repeat (STR) or simple sequence length polymorphisms (SSLP) is a repeated nucleotide sequence of 2-7 base pair units. Repeatability of fragment resulted from slippage in replication (Schlötterer and Tautz, 1992) and unequal crossover (Smith, 1976). The SSR technique uses PCR to amplify DNA fragments by repeated cycles of DNA denaturing, annealing and extension using DNA polymerase enzyme. The resulting DNA was separated by banding on gel (Akkaya *et al.*, 1992). This method has many advantages such as rapid, reliable (Diwan and Cregan, 1997), abundance (Lagercrantz *et al.*, 1993), co-dominant (Akkaya *et al.*, 1992), high heterozygosity (Powell *et al.*, 1996), and high polymorphism (Akkaya *et al.*, 1995).

SSR has been used in the construction of genetic linkage map of human (Gyapay *et al.*, 1994), mouse (Dietrich *et al.*, 1994), fruit fly (Schlötterer *et al.*, 1997), soybean (Shoemaker *et al.*, 1992; Morgante, 1994; Akkaya *et al.*, 1995; Maughan *et al.*, 1995), rice (Temnykh *et al.*, 2000), corn (Taramino and Tingey, 1996) and sorghum (Taramino *et al.*, 1997), and also is used in DNA fingerprint of soybean (Yanagisawa *et al.*, 1994; Diwan and Cregan, 1997), in tag DNA of coconut (Teulat *et al.*, 2000), in pedigree analysis of soybean (Diwan and Cregan, 1997) and barley (Russell *et al.*, 1997), in genetic identification in barley (Russell *et al.*, 1997), in genetic diversity of rice (Cho *et al.*, 2000), barley (Struss and Plieske, 1998), and wheat (Prasad *et al.*, 2000).

SNP is the single base variation between two otherwise identifical DNA sequence. Brookes (1999) reported that SNP was single base pair positions in genomic DNA at which different sequence alternatives (alleles) exist in normal individuals of natural populations. SNP was the powerful tool and was the most frequent type of variation in the human genome (Wang *et al.*, 1998). This technology allows greater number of tests to be run at a significantly lower cost than existing technology. This technology has application in genome sequencing, SNP identification and typing, screening for genetically-linked diseases, identification of genetic

†Corresponding author: (Phone) +82-31-290-2305 (E-mail) sukhalee@snu.ac.kr
<Received August 31, 2001>

drug targets, screening for individuals for potential drug side effects, gene cloning, screening potential tissue donors, screening cancer cells for genes conferring chemotherapy resistance, typing pathogenic and antibiotic resistant bacteria, forensics and pharmacogenomics. This technique has more advantages than markers such as isozymes, RFLP, RAPD, or SSR. These advantages include very large number of polymorphic loci which distributed throughout the genome, marker present within coding regions, introns and regions that flank genes, simple and unambiguous assay techniques, high levels of polymorphism in the population, stable Mendelian inheritance and low levels of spontaneous mutation (Brown, 1999)

Soybean genotypes grown in Thailand are more adapted to short-day-length than those in Korea, indicating that there is a great genetic variability among soybean genotypes grown in Thailand and Korea. The objective of this study was to evaluate genetic relationship among soybean genotypes from diverse genetic bases, i.e. from Korean, Thailand, and wild soybean (*G. soja*) using SSR and especially SNP.

MATERIALS AND METHODS

Plant materials

Eight Thai, 5 Korean and 3 wild soybeans were used in this study (Table 1). Several parameters were evaluated such as flower color, stem color, pubescence color, leaf shape, plant height, number of nodes, and color of seed coat in the experimental field of Seoul National University, Suwon, Korea during the summer of 2000.

DNA isolation

Total genomic DNA was isolated from unfolded leaflets and kept in lyophillizer at -50°C. DNA was extracted using modified protocols of Keim *et al.* (1988). Five to six pieces of lyophilizer were taken into 96-well plate containing a single steel bead ball. By the use of a reciprocal saw, the soybean tissue was pulverized, then 700 ul CTAB was added into each well and taken to a shaker-incubator at 225 rpm, and 65°C for 35 min. Chloroform: isoamyl (24 : 1) was added at 600ul into the plate and taken into the shaker-incubator at 24°C for 10 min, then centrifuged in Eppendrof 5804 at 4,000 rpm for 30 min. Cold isopropanol was added at 420ul into each well of new 96-well plate and the supernatant was transfered 3 times. In each time, 200 ul of the supernatant was taken to new plate and brought to refrigerator for 15 min. Plate was spun in Eppendrof 5804 at 4,000 rpm for 20 min. The supernatant was added with 500 ul 70% ethanol to each well of plate and kept on shaker-incubator at 24°C for 10-15 min then centrifuged one more time. Supernatant was poured and incubated at 37°C for 15 min then resuspended each well by 300 ul TE buffer, shaken at 225 rpm for 1 day and kept in refrigerator. DNA concentration was measured using F-4500 Fluorescence Spectrophotometer (Hitachi, Ltd., Tokyo, Japan).

Table 1. List of soybean varieties used in evaluation of their genetic relationship.

Soybean genotypes	Origin	Scienfic name
Kumjung 2, Sowon, Jangkyung, Keumkang, Songhak	Korea	Glycine max
IT161471, IT182307, IT184222	Korea	Glycine soja
SJ1, SJ2, SJ5, Sukhothai 1 (ST1), Chiang Mai 1 (CM1), KUSL20004, Chakkrabhandhu 1 (CKP1), Rajamangkala 1 (RM1)	Thailand	Glycine max

Table 2. Primers used in the SSR analysis.

Locus	Fluorescent label	Linkage group†	Locus	Fluorescent label	Linkage group
Satt 175	6-FAM	M	Satt 196	HEX	K
Satt 414	6-FAM	J	Satt 253	HEX	H
Satt 556	6-FAM	B2	Satt 146	NED	F
Satt 590	6-FAM	M	Satt 177	NED	A2
Satt 141	6-FAM	D1b+W	Satt 530	NED	N
Satt 545	6-FAM	A1	Satt 143	NED	L
Satt 192	6-FAM	H	Satt 445	NED	O
Satt 269	6-FAM	F	Satt 038	NED	G
Satt 187	HEX	A2	Satt 294	NED	C1
Satt 167	HEX	K	Satt 271	NED	D1b+W

†USDA linkage group (Cregan *et al.* 1999).

Simple Sequence Repeat (SSR)

Twenty of the fluorescent 5-end forward primers (Table 2) were labeled with either 6-FAM (blue), NED (yellow) and HEX (green) (PE-ABI, Foster City, CA). Twenty SSR markers were selected due to their high polymorphism (personal comm. Perry Cregan) DNA amplification was proceeded in genomic DNA at the concentration of 10ng/ul, 10x buffer, 2.5 mM of each nucleotide, 5 units of *Taq* DNA polymerase, 20 mM $MgCl_2$ and 5 pM Primer Mix, with 32 cycles of 25 s of denaturation at 94°C, 25 s of annealing at 46°C and 25 s of extension at 68°C on PCT-100™ Thermal Controller (MJ research, Watertown, Mass). Then PCR product was examined in 2 % agarose gel.

Two ul of each PCR product of 6-FAM, NED, HEX which had different allele size was loaded in the same gel lane. Then 4 ul of cocktail consisted of 110ul formamide deionized, 55 ul loading buffer, 15 ul Genscan 500 (500XL) were added to 1.5 ul of each combined PCR product. Combined PCR product was loaded and separated in ABI Prism® 377 (ABPEC, Foster City, CA), analyzed allele size by GeneScan® Analysis software, version 2.1.1(ABPEC, Foster City, CA) and Genotyper® software, version 2.0 (ABPEC, Foster City, CA).

Single nucleotide polymorphism (SNP)

Primers were designed from consensus sequence regions on disease resistance protein homolog genes such as *N* gene of tobacco, *L6* gene of flax, *RPM1* and *RPS2* genes of arabidopsis, *Xa21* gene of rice, and *Pto* and *Cf-9* genes of tomato (Table 3). Genomic DNA was amplified with designed primers. Alleles of different sizes obtained by amplification of SRGA5, SRGA6, SRGA7, and SRGA8 were cloned and sequenced after carrying out amplification reactions. The cloning process was based on the T/A cloning method (Clark 1988; Mead *et al.* 1999). The ligation reaction was carried out at 14°C overnight in 10 ul with 4 Units of T4 DNA ligase (Clonetech Lab., Inc.), 50 ng of the plasmid vector pT-Adv (Clontech Lab., Inc.) in 1× reaction buffer (6 mM Tris-HCl pH 7.6, 6 mM $MgCl_2$, 5 mM NaCl, 2 mM of DTT, 1 mg/ml BSA, 7 mM ß-mercaptoethanol, 1 mM spermidine, and 0.1 mM ATP). The ligation mixture was transformed into *E. coli* TOP10F competent cells and the recombinant clones were screened for inserts on LB plates with color selection. The plasmid DNA was isolated from white colonies and grown in LB broth. Plasmids were prepared according to the standard protocol using Qiagen columns (Qiagen, Hiden, Germany) from 5 ml culture. The DNA was sequenced on an ABI prism BigDye terminators (Perkin-Elmer Biosystems, Foster City, CA) according to the manufacturers instructions (Perkin-Elmer Biosystems, Foster City, CA). The sequences of the amplified and cloned products were aligned using the CLUSTAL W (1.60) multiple alignment package (Thompson *et al.* 1994).

Data analysis

Allelic polymorphic information content was calculated using the formula $PIC_i = 1-\Sigma p_{ij}^2$, where p_{ij} is the frequency of the j allele for marker i (Anderson, 1992). The data were used to analyze genetic diversity from scoring the bands of each SSR loci and sequence data by 0 = absent and 1 = present indicating the corresponding band and nucleotide. Each band was then treated as a unit character and pair-wise genetic similarity coefficients among cultivars were quantified based on Jaccard's coefficient = a/(n-d) using NTSYS-pc software (Rohlf, 1998). The cluster method was carried out by the unweighted pair-group method with arithmetic mean (UPGMA) (Sokal & Michener, 1958).

RESULTS

Morphological variation

The importance of genetic diversity relies on the variant parameters used in separating different characters. In this study, the qualitative characters were stem color (SC), flower color (FC), pubescence color (PC), pubescence density (PD), and leaf shape (LS), some quantitative characters were average height (AH) and average node (AN) (Table 4). Three Korean varieties (Jangkyung, Keumkang and Songhak) had green stem, others were purple stem. Since the gene controlling colors of stem and flower is the same locus it is predictable that when the stem of soybean is green the flower is white, and when the stem is purple the flower is also purple. Two Thai varieties Sukhothai (ST 1) and

Table 3. Primers selected from consensus sequence regions on disease resistance protein homolog genes used in SNP detection.

Primer code	Forward primer (5' → 3')	Reverse primer (5' → 3')
SRGA5	TGCTAGAAAAGTCTATGAAG	TCAATCATTTCTTTGCACAA
SRGA6	AGCCAAAGCCATCTACAGT	AACTACATTTCTTGCAAGT
SRGA7	AGTTTATAATTCCATTGCT	CCGAAGCATAAGTTGCTG
SRGA8	AGCGAGAGTTGTATTTAAG	AGCCACTTTTGACAACTGC

Table 4. Morphological traits of 16 soybean accessions used in this study.

No.	Varieties name	%G	SC	FC	PC	LS	PD	AH (cm)	AN (cm)
1	SJ 1	60	Purple	Purple	Brown	Oval	Sparse	13	5
2	SJ 2	60	Purple	Purple	Brown	Lanceolate	Sparse	13	5
3	SJ 5	56	Purple	Purple	Brown	Oval	Sparse	14	4
4	ST 1	82	Purple	Purple	White	Lanceolate	Sparse	11	5
5	CM 1	48	Purple	Purple	White	Oval	Sparse	12	4
6	KUSL 20004	76	Purple	Purple	Brown	Oval	Sparse	9	4
7	RM 1	57	Purple	Purple	Brown	Oval	Dense	13	5
8	CKP 1	34	Purple	Purple	Brown	Oval	Sparse	8	2
9	Kumjung 2	76	Purple	Purple	Brown	Oval	Dense	11	3
10	Sowon	83	Purple	Purple	White	Lanceolate	Normal	8	3
11	Jangkyung	86	Green	White	White	Oval	Normal	10	3
12	Keumkang	60	Green	White	Brown	Oval	Normal	8	3
13	Songhak	72	Green	White	White	Oval	Normal	11	4
14	IT 161471	62	Purple	Purple	Brown	Rhomboidlanceolate	Sparse	5	2
15	IT 182307	10	Purple	Purple	Brown	Rhomboidlanceolate	Sparse	8	3
16	IT 184222	10	Purple	purple	Brown	Oval	Sparse	4	2

%G=% germination; SC=stem color; FC=flower color; PC=pubescence color; LS=leaf shape; PD=pubescence density; AH=average height; AN=average node (AH and AN were recorded 30 days after planting).

Chiang Mai (CM 1) and the three Korean varieties were white pubescence, while the others were brown pubescence at main stem. Rajamangkala 1 (RM 1) and Kumjung 2 had dense pubescence. ST1 was the best in germination, while IT 182307 and IT 184222 gave the least germination, probably due to high hard seed percentage. They also had different leaf shape (rhomboid-lanceolate) while the other varieties were either lanceolate or ovate.

Automated sizing of SSR and SNP

For DNA amplification, twenty microsatellite fluorescent 5-end forward primers were used as summarized in Table 5. The total of 149 alleles were detected from 20 primers. The number of alleles per locus ranged from 3 to 9, with an average of 6.65 alleles. The maximum number of alleles were obtained from Satt 445, Satt 167 and Satt 590. Satt 271 produced as low as 3 alleles per locus, while Satt 187 produced 4 alleles, Satt 038, Satt 143 and Satt 196 produced 5 alleles each, Satt141, Satt 177 and Satt 192 gave 6 alleles each, Satt 146, Satt 253. Satt 269, Satt 294, Satt 530 and Satt 556 gave 7 alleles each, Satt 175, Satt 414, and Satt 545 gave 8 alleles each, Satt 167, Satt 445 and Satt 590 produced 9 alleles each. Gene diversity or PIC values were found ranging from 0.48 to 0.85 across all 20 SSR loci with the mean value of 0.72.

After sequencing, the fragment of DNA products from the primers SRGA5, SRGA6, SRGA7 and SRGA8 had the

Table 5. Allele number, size range and gene diversity of 20 fluorescent labeled SSR loci in 8 Thai, 5 Korean and 3 wild soybeans.

Locus	Fluorescent label	Linkage Group	Cultivar genotypes		
			Allele size range (bp)	No. of alleles	PIC
Satt 175	6-FAM	M	143-183	8	0.70
Satt 414	6-FAM	J	266-313	8	0.80
Satt 556	6-FAM	B2	163-211	7	0.73
Satt 590	6-FAM	M	263-340	9	0.80
Satt 141	6-FAM	D1b+W	148-201	6	0.65
Satt 545	6-FAM	A1	155-203	8	0.81
Satt 192	6-FAM	H	234-264	6	0.68
Satt 269	6-FAM	F	160-345	7	0.70
Satt 187	HEX	A2	243-280	4	0.62
Satt 167	HEX	K	235-273	9	0.85
Satt 196	HEX	K	178-205	5	0.70
Satt 253	HEX	H	130-175	7	0.71
Satt 146	NED	F	287-320	7	0.70
Satt 177	NED	A2	105-131	6	0.75
Satt 530	NED	N	215-241	7	0.83
Satt 143	NED	L	235-276	5	0.69
Satt 445	NED	O	162-228	9	0.79
Satt 038	NED	G	157-184	5	0.64
Satt 294	NED	C1	252-296	7	0.80
Satt 271	NED	D1b+W	112-121	3	0.48
Mean				6.65	0.72

```
1   CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
2   CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
3   CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
4   CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
5   CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
6   CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
7   CATTAGAGAGGTGCCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
8   CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
9   CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
10  CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
11  CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
12  CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
13  CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
14  CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCG
15  CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
16  CATTAGAGAGGTGTCTAAGACAAATGGCTTAGTTCATATTCAAAAGGAACTTTCTAATCT
```

Fig. 1. Aligment of the base sequence using CLUSTAL W of 8 Thai, 5 Korean and 3 wild soybeans; 1 = SJ 1, 2 = SJ 2, 3 = SJ 5, 4 = ST 1, 5 = CM 1, 6 = KUSL 20004, 7 = RM 1, 8 = CKP 1, 9 = Kumjung 2, 10 = Sowon, 11 = Jangkyung, 12 = Keumkang, 13 = Songhak, 14 = IT161471, 15 = IT82307, and 16 = IT184222.

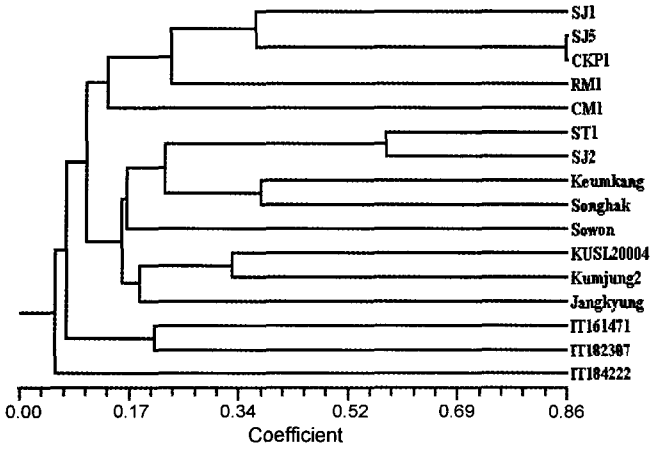

Fig 2. A dendrogram of soybean varieties created by genetic similarity based on the binary data from SSR analysis.

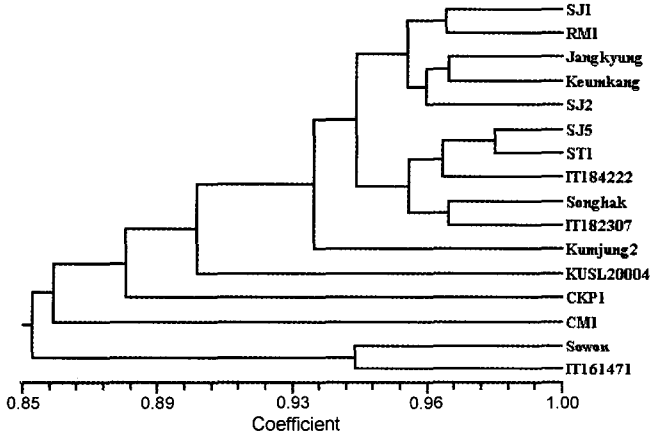

Fig. 3. A dendrogram of soybean varieties created by genetic similarity based on the binary data from SNP analysis.

sizes of 400, 400, 400 and 312 bp, respectively. They were aligned using CLUSTAL W and presented in Fig. 1. The number of polymorphisms in SNP were 263 positions that appeared to have different base pairs.

Clustering of the soybean genotypes

The dendrograms prepared through cluster analysis using SSR and SNP are shown in Fig. 2 and 3. For SSR marker, the genotypes can be grouped into four clusters (Fig. 2), cluster I with 5 genotypes (SJ 1, SJ 5, CKP 1, RM 1 and CM 1), cluster II with 7 genotypes (ST 1, SJ 2, Keumkang, Songhak, Sowon, KUSL 20004, Kumjung 2 and Jangkyung), cluster III with 2 wild genotypes (IT 161471 and IT 182307), and cluster IV contains a solitary wild genotype, IT 184222. SNP separated these soybeans into eight cluster (Fig. 3), cluster I with 5 genotypes (SJ 1, SJ 2, Jangkyung, Keum-

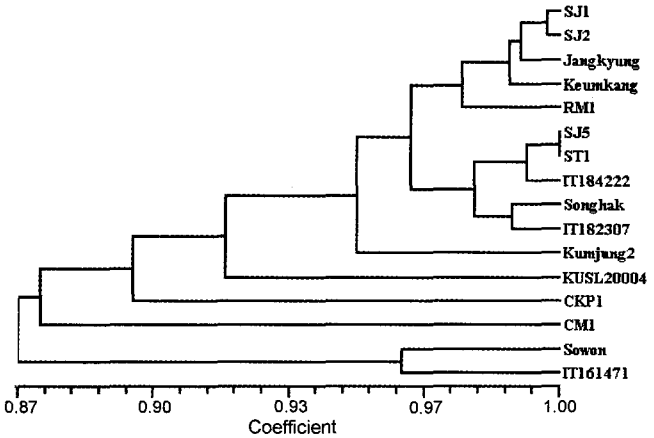

Fig. 4. A dendrogram of soybean varieties created by genetic similarity based on integration of binary data from SSR and SNP markers.

kang and RM 1), cluster II with 5 genotypes (SJ 5, ST 1, IT 184222, Songhak and IT 182307), while group III to VIII were clustered based on one genotype each of the rests. Integration of the binary data from both markers revealed that the new grouping followed the trend of SNP clustering (Fig. 4).

Discussion

The automated sizing of fluorescent-labeled marker alleles taken from ABI Prism® 377 DNA sequencing was discovered to rapidly and precisely assay diversity of the soybean genotypes with less use of polyacrylamide gel, since each gel can employ 5 or 6 primers depending on allelic range. Then the Genotyper® software can be used in sizing and visualization.

In this study, 20 SSR primers were used to amplify repeated regions, while SNP utilized the primers made from the consensus sequence regions on disease resistance protein homolog genes and cloned it for specific regions. SNP considered the difference in base pair position and thus gave more number of loci and polymorphism than the SSR that used number of primers as the number of loci and polymorphism. With only 4 SNP primers, 1412 loci and 263 polymorphisms were detected while SSR used 20 primers which gave rise to 20 loci and 20 polymorphisms. After the binary data (0,1) were scored and analyzed by NTSYS program, the SSR dendrogram could separate the Thai, Korean and wild soybeans more clearly than the SNP dendrogram. This is because SNP was surveyed on the smaller region and specific only to the disease resistance protein homolog genes. However, integration of SSR and SNP showed that the dendrogram was closer to that of the SNP because the latter has more scored data (1674) as can pared to the binary data of SSR (149).

ACKNOWLEDGEMENT

This research was supported by Seoul National University, Kasetsart University, and Thailand Research Fund.

REFERENCES

Akkaya, M. S., A. A. Bhagwat and P. B. Cregan. 1992. Length polymorphisms 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 DNA markers into a soybean linkage map. *Crop Sci.* 35 : 1439-1445.

Anderson, J. A. 1992. Optimizing parental selection for genetic linkage maps. *Genome* 36 : 181-186.

Brown, S. M. 1999. Snapping up SNP. *Biotechniques* 26(6) : 1090-1094.

Brookes, A. J. 1999. The essence of SNP. *Gene* 234 : 177-186.

Charles. V. P., and W. J. Morse. 1923. The Soybean, 2nd ed., McGraw-Hill Book Co., USA.

Cho, Y. G., T. Ishii, S. Temnykh, X. Chen, L. Iipovich, S. R. McCouch, W. D. Park, N. Ayres and S. Cartinhour. 2000. Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 100 : 713-722.

Clark, J. M. 1988. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerase. *Nucleic Acids Res.* 16 : 9677-9686.

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.

Dietrich, W. F., J. Miller, R. G. Steen, M. Merchant, D. Damron, R. Nahf, A. Gross, D. C. Joyce, M. Wessel, R. D. Dredge, A. Marquis, L. D. Stein, N. Goodman, D. C. Page and E. S. Lander. 1994. A genetic map of the mouse with 4,006 simple sequence length polymorphisms. *Nature Genet.* 7 : 220-245.

Diwan, N. and P. B.Cregan. 1997. Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theor. Appl. Genet.* 95 : 723-733.

Gyapay, G., J. Morissette, A. Viganal, C. Dib, C. Fizames, P. Millassean, S. Marc, G. Bernardi, M. Lathroppand and J. Weissenbach. 1994. The 1993-94 Genethon human genetic linkage map. Nat. Genet. 7 : 246-249.

Jaccard, P. 1908 Nouvelles recherches sur la distribution florale. Bull Soc. *Vaudoise Sci. Nat.* 44 : 223-270.

Keim, P., T. C. Olson, and R. C. Shoemaker. 1988. A rapid protocol for isolating soybean DNA. *Soybean Genet. Newsl.* 15 : 150-154.

Lagercrantz, U., H. Ellegran, and L. Anderson. 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Res.* 21 : 1111-1115.

Maughan, P. J., M. A. S. Maroof and G. R. Buss. 1995. Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean. *Genome* 38 :715-723.

Mead, D. A., N. K. Pey, C. Herrnstadt, R. A. Marcil and L.M . Smith. 1991. A universal method for the direct cloning of PCR amplified nucleic acid. *Bio/Technology* 9 : 657-663.

Morgante, M. 1994. Genetic mapping and variability of seven soybean simple sequence repeat loci. *Genome* 37 :763-769.

Powell, W., M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey, and A. Rafalski. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol. Bre.* 2 : 225-238.

Prasad, M., R. K. Varshney, J. K. Roy, H. S Balyan and P. K. Gupta. 2000. The use of microsatellite for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. *Theor. Appl. Genet.* 100 : 584-592.

Rohlf, F. J. 1998. Numerical taxonomy and multivariate analysis system, NTSYS-pc program manual version 2.0. Applied Bio-

statistics Inc., 3 Heritage Lane, Setauket, NY 11733.

Russell, J., J. Fuller, G. Young, B. Thomas, G. Taramino, M. Macaulay, R. Waugh and W. Powell. 1997. Discriminating between barley genotypes using microsatellites markers. *Genome* 40 : 442-450.

Schlötterer, C. and D. Tautz. 1992 Slippage synthesis of simple sequence DNA. *Nuleic Acids Res.* 20 : 211-215.

Schlötterer, C., C. Volgl and D. Tautz. 1997. Polymorphism and locus-specific effects on polymorphism at microsatellite loci in natural *Drosophila melanogaster* populations. *Genetics* 146 : 309-320.

Shoemaker, R. C., R. D. Guffy, L. L. Lorenzen and J. E. Specht. 1992. Molecular genetic mapping of soybean: map utilization. Crop Sci. 32 : 1091-1098.

Smith, G. P. 1976. Evolution of repeated DNA sequences by unequal crossover. *Science.* 191 : 528-535.

Sokal R. R. and C. D. Michener. 1958. A Statistic Method for Evaluating Systematic Relationships. *Univ. Kans. Sci. Bull.* 28 : 1409-1438.

Struss, D. and J. Plieske. 1998. The use of microsatelillite markers for detection of genetic diversity in barley populations. *Theor. Appl. Genet.* 97 : 308-315.

Taramino, G., R. Tarchini and S. Ferrario. 1997. Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor. Theor. Appl. Genet.* 95 : 66-72.

Taramino, G. and S. Tingey. 1996. Simple sequence repeats for germplasm analysis and mapping in maize. *Genome* 39 : 277-287.

Temnykh, S., W. D. Park, N. Ayres, S. Cartinhour, N. Hauch, L. Lipovich, Y. G. Cho, T. Ishii and S. R. McCouch. 2000. Mapping and genome organization of microsatellite sequences in rice (*Orysa sativa* L.). *Theor. Appl. Genet.* 100 : 697-712.

Teulat, B., C. Aldam, R. Trehin, P. Lebrun, J.H.A. Barker, G. M. Arnold, A. Karp, L. Baudouin and F. Rognon. 2000. An analysis of genetic diversity in coconut (*Cocos nucifera*) populations from across the geographic range using sequence-tagged microsatellites (SSRs) and AFLPs. *Theor. Appl. Genet.* 100 : 764-771.

Thompson. J. D., D. G. Higgins and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22 : 4673-4680.

Wang, D. G., J. B. Fan, C. J Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, E. Hubbell, E. Robinson, M. Mittmann, M. S. Morris, N. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T. J. Huddon, R. Lipshutz, M. Chee and E. S. Lander. 1998. Large-scale identification, mapping, and genotyping of single nucleotide polymorphisms in the human genome. *Science* 208 : 1077-1082.

Yanagisawa, T., M. Hayashi, A. Hirai and K. Harada. 1994. DNA fingerprinting in soybean (*Glycine max* (L). Merrill) with oligonucleotide probes for simple repetitive sequences. *Euphytica* 80 : 129-136.