Plant Res. 1(2) 81~91(1998)

Genotype Fingerprinting, Differentiation and Association between Morphological Traits and SSR Loci of Soybean Landraces

Qijian Song[1,2], Ik Young Choi[1], Nam Kee Heo[3] and Nam Soo Kim[1]
[1]Division of Applied Plant Sciences, College of Agriculture and Life Sciences, Kangwon National University, Chunchon, Korea, 200-701
[2]Department of Plant Genetics and Statistics, Nanjing Agricultural University, Nanjing, 210095, the People's Republic of China
[3]Kangwon Provincial ATA, Chunchon, Korea, 200-150

ABSTRACT

Fifty-nine Korean soybean (*Glycine max* L. Merr.) landrace accessions were tested for genotype fingerprinting, differentiation and association between morphological traits and SSR profile. Using 8 SSR loci, 59 varieties were divided into 55 groups, and only 4 pairs of varieties were not uniquely identified. The resolving power of SSR for soybean genotyping was much higher than that of the morphological traits that were studied. Identification efficiency also differed among SSR loci. Those loci with higher numbers of alleles distinguished varieties more effectively. Genetic differentiation values of the soybean landraces varied from 0.57 to 0.82 with a mean of 0.68. The number of alleles detected by the 8 loci ranged from 3 to 8, and the effective number of alleles ranged from 2.3 to 5.1. In a study of the association of SSR alleles with morphological traits, some alleles seemed to be related with some specific morphological traits. Comparison of two kinds of dendrograms which were derived from SSR markers and quantitative traits indicated that the dendrograms were not consistent. Considering the correlation between single SSR locus and qualitative traits governed by major genes, the data suggest that alleles of microsatellite loci be more closely related to some traits determined by major genes than those determined by minor genes.

Key words: soybean, microsatellites, genotype identification, differentiation

INTRODUCTION

Collecting and maintaining different genotypes are necessary for preserving valuable genetic resources for breeding purposes. Thousands of new varieties or lines are introduced into plant germplasm or breeding institutes around the world each year. Establishing a genetic identity for each introduction is a high priority. In soybeans and most other species, morphological traits such as flowers, leaf, seed coat and pubescence colors, and flowering and maturity dates, etc., have been used for identification. As more new cultivars and lines are collected, it becomes difficult to distinguish them from those already in the collection, mainly because of the paucity of available morphological markers. In such a situation, many duplicate introductions might be maintained. Moreover, the Plant Variety Protection Act requires that breeders demonstrate that a new cultivar has unique characters differing from varieties already released. The use of molecular markers provides one solution to this problem by providing marker allele profiles for distinguishing between cultivars (Rongwen et al. 1995).

Frequent use of outstanding lines as parents in breeding programs results in a small number of genotypes being recycled, thus narrowing the genetic base. This is especially true in soybeans. Gizlice et al. (1994) found that 28 ancestors and 7 first progenies contributed 95% of the genes

found in current North American soybean cultivars and that cultivars Lincoln and Harosoy in the North, Lee and its full sib (D49-2491) in the South contributed nearly 40% of the genes to these newly developed varieties. Delanney et al.(1983) and Hymowitz and Bernard(1991) also reported similar results. The narrow genetic base increases potential vulnerability to pests, diseases and environmental stress. Better knowledge of the genetic diversity of potential soybean breeding materials could help to develop optimal plant breeding programs balancing the maintenance of genetic diversity in order to sustain long term selection progress and make use of high performance.

Many scientists have tried to construct core germplasm collections based on genetic diversity of DNA profiles in order to make full use of germplasm and to save time and energy required to evaluate entire collections. From the breeding point of view, if diversity based on DNA profiles is associated with the phenotypic variability of genotypes, this information would be more valuable. If the preserved DNA diversity based upon DNA profile is not consistent with the phenotypic variability of the varieties, breeders would need to use the diversity estimates with caution. However, an assessment of the relationship between diversity based upon DNA profiles and phenotypic diversity is not available in soybean. In recent years, numerous work is undertaken to define the relation between coefficient of coancestry, pedigree or species origin information from molecular markers. The associations being reported on different crops varied from strong (Messmer et al. 1993, Smith et al. 1990, Akkaya et al. 1992) to moderate or low (Graner et al. 1994, Keim et al. 1992, Plaschke 1995, Powell et al. 1996, Tinker et al. 1993). Even if a strong relationship between phenotypic variability and molecular marker variations is found, it could not be used as direct proof of association between variability of morphological traits and genetic diversity based on molecular markers, due to selection and genetic drift.

The objectives of the present study are to detect and measure SSR variation in Korean soybean landraces, determine the association between genetic diversity based on SSR information and genetic variability based on morphological traits, and to assess the associations of morphological traits with single and multi-locus SSR information.

MATERIALS AND METHODS

Soybean genotypes and morphological traits observation

Fifty-nine soybean landraces from Korea were surveyed. The field experiments were carried out at the Experiment Farm, Rural Development Agency, Kangwon Branch from 1992 to 1994. Each plot consisted of 3 rows, each of which

Table 1. Characteristics of the eight microsatellite loci used in this research.

Locus	Repeat sequence	PCR primer sequence	Reference†	$T_{ann.}$ ℃ used	Gel type
SATT5	$(TAA)_{21}$	TATCCTAGAGAAGAACTAAAAAA GTCGATTAGGCTTGAAATA	1	55	Agarose
SATT2	$(AAT)_{16}$	AATAATGTGGAAACTAAATGG TAATGTGCCTATCCTTGTCTT	1	55	Agarose
SOYPRP1	$(TAT)_{20}$	CGAAGAGGTACGTGCCAAATT GTTAGAAAACTCCGCCCACAC	2	58	Agarose
SOYHSP176	$(AT)_{15}$	TGTGGGCCACAAAACGTATAG CGTACGTTCTAGCTAGTCTTC	2	58	PAGE
SATT1	$(ATT)_{24}$	CTGGTGGACTATTGATACGACC AACTGCGAAGATACTACCCTCC	1	55	PAGE
SOYAC7L	$(CT)_{16}$	CCCTATCCCTCTTTTCTCGTG AATTGGAATTGAGGTTGGGAC	2	58	PAGE
SOYSC514	$(AT)_{14}$	AGTCGTAGTCTAGCTACATGAC CAGTGGAATATGTGAAGCAATG	2	58	PAGE
SAT43	$(AT)_{20}$	AAATTCTGTTCATTGTCCGTC CATTTTAATATCCCGAGTAGG	1	45	PAGE

†Reference: 1, Rongwen et al. (1995); 2, Morgante et al. (1994), and Powell et al. (1996)

was 4 meters long. Twelve quantitative traits were measured; flowering date, maturity date, lodging score, SMV susceptible score, SMV-N susceptible scores, plant height, number of branches, number of pods per plant, 100-seed weight, plot yield, protein content, and oil content. Each landrace was phenotypically classified relative to five qualitative traits which included leaf color, pod color, flower color, pubescence color, and coat color.

DNA isolation, amplification, electrophoresis, and silver staining

The protocol for extracting genome DNA was described by Dellaporta et al. (1983) with slight modification. Eight primer pairs, which were reported by Morgante et al. (1994), Rongwen et al.(1995), Powell et al. (1996), were chosen for the analysis. Microsatellite designation, microsatellite composition, primer sequences, annealing temperature, and gel types are as shown in Table 1. PCR amplification was carried out in a TC1 Thermocycler (Perkin Elmer Co.) in a volume of 25μl which contained 40 ng of template genomic DNA, 6.25μM each primer, 2.25 μM dNTPs, 1.4 unit of thermostable DNA polymerase F-500L dynazme and 2.5μl 10 × storage buffer. Reactions were first heated at 95℃ for 5 min, amplified for 30 cycles. Each cycle consisted of 1 min at the denaturing temperature of 94℃, 1 min at the annealing temperature specific for each primer pair (Table 2), and 45 sec at 72℃. After the last cycle, the reactions were held at 72℃ for 10 min. Amplification products were separated either on 3.5% agarose gels containing 2% Metaphor and 1.5% SeaKem LE, or on DNA sequence gels containing 6% polyacrylamide, 7 M urea and 1x TBE. Gel types for each primer pair were as specified in Table 1. DNA fragments on polyacrylamide gel were visualized by silver staining, based on the method described in Promega Technical Manual for the Silver Sequence DNA Sequencing System. After being dried, polyacrylamide gels were exposed to a Promega Silver Sequence TM APC Film.

Data analysis

Alleles of each genotype were coded by digits from 1, 2, 3, · · · to n. If the alleles of different genotypes were common, the codes were the same, otherwise, they were different. Single and joint primer pairs were used for grouping through pair-wise comparisons between genotypes. Genotypes with the same codes for all alleles at a given locus were placed in the same group. The number of groups was established and genotypes in each group were recorded.

Genetic diversity (measured by number of effective alleles) and population differentiation were estimated by Gregorius' s algorithms (1986, 1987): $\upsilon = (\Sigma pi^2)^{-1}$ and $\delta_T = N/(N-1) \cdot (1-\Sigma pi^2)$, respectively. Ratio of groups to sample size was estimated by using NG/N, where N is the sample size, p is the frequency of the ith allele, NG is the number of groups which have been divided, υ is the effective number of alleles.

For quantitative traits, the t-test was used to evaluate the difference between phenotypic means of the different groups. A chi-square independence test was used to evaluate the association between alleles of each given SSR locus and the qualitative traits. Fifty-nine genotypes were allocated

Table 2. Genetic diversity, number of effective alleles, and number of groups identified using alleles of these loci.

Locus	No. of effective alleles	No. of alleles Observed	Differentiation	Groups	Ratios(%)
SATT5	3.5	4	0.73	4	6.8
SATT2	2.8	4	0.68	4	6.8
SOYPRP1	2.5	3	0.61	3	5.1
SOYHSP176	2.3	4	0.57	4	6.8
SATT1	4.6	8	0.80	8	13.6
SOYAC7L	2.6	4	0.63	4	6.8
SOYSC514	2.9	5	0.66	5	8.5
SAT43	5.1	8	0.82	8	13.6
Mean	3.2	5	0.68	5	8.5

to the appropriate cell of the two dimensional contingency table defined by SSR alleles and phenotypic classes of qualitative traits. The total number of genotypes in each cell was used as the basis for analysis. Joint analysis combining the information from all 8 SSR loci was carried out by calculating similarity index S_{xy} between genotypes x and y, i.e. $S_{xy}=2N_{xy}/(N_x+N_y)$, where Nxy refers to the number of alleles in common between genotypes x and y. N_x and N_y indicate the total number of alleles for genotypes x and y, respectively. The average linkage (UPGMA) algorithm was chosen as a clustering method. Based on the cluster trees, three and four subclusters were unambiguously identified. Frequencies of qualitative trait alleles and the means of quantitative traits in each subcluster were calculated. Mean differences between populations for the morphological traits were assessed for their statistical significance using either the chi-square or t-test. In order to study the association between SSR and morphological information, dendrograms based on 12 quantitative phenotypic traits were also studied. Distance index between genotypes x and y was calculated by $d_{xy}^2=\Sigma(P_{kx}-P_{ky})^2$, where P_{ki} and P_{kj} were observation values for genotype i and j, $k=1 \cdots t$, t refers to the number of traits. Dendrograms based on morphological traits were also generated using the UPGMA algorithm. A dissimilarity matrix of 59 genotypes, based on 8 SSR loci, was calculated using $D_{xy}=1-S_{xy}$. Product moment correlation, a measure of goodness of fit between two matrices of distance was calculated using the method of Mantel (1967) with NTSYS-PC software (Rohlf 1992).

RESULTS AND DISCUSSION

Genotype identification using different microsatellite combinations: Eight microsatellite loci were used to reveal the effectiveness of SSR length polymorphism for the identification of Korean soybean landraces. As shown in Tables 2 and 3, a single SSR locus was capable of dividing the 59 varieties into as few as 3 or as many as 8 groups. On average, five groups were detected per locus. When two SSR loci were jointly used, the 59 varieties could be divided into as few as 13 to as many as 28 groups. On average the two-locus combination yielded 16.1 groups, with an average ratio of identification of 27.4%. Of the 8 loci (Table 3), SAT43 and SATT1 were more effective in identification when single locus was used. The identification efficiency was also quite different among loci combinations. When two SSR loci were combined, the SATT1 and SAT43 loci allowed the landrace population to divided into 28 groups in contrast with 13 groups divided by the SOYPRP1 and SOYHSP176 loci. When three loci, SATT1, SOYSC514 and SAT43, were combined, the population was divided into 42 groups while other low efficiency three-locus combinations could only divide the population into 23 groups. Likewise, great differences were observed among 4-, 5-, 6-, and 7-loci combinations. From Tables 2 and 3, it was relatively easy to establish that the more effective locus or loci combinations were those with the higher numbers of effective alleles. Loci or locus combinations with higher numbers of effective alleles

Table 3. Differentiation, number of groups identified using different loci combinations

Number of loci combined	Differentiation	Groups (%)	Ratio of Identification	Best Combinations†
one	0.684	4.8(3-8)	8.5	5 or 8
two	0.895	16.1(13-28)	27.4	5,8
three	0.962	31.2(23-42)	52.8	5,7,8
four	0.984	42.8(34-49)	72.5	3,5,7,8 or 2,3,5,8
five	0.993	49.4(43-53)	83.7	2,3,5,7,8 or 1,2,5,6,8
six	0.996	52.9(49-55)	89.6	1,2,4,5,6,7
seven	0.997	54.5(54-55)	92.4	1,2,4,5,6,7,8 or 1,2,3,4,5,6,7 or 1,2,3,5,6,7,8 or 2,3,4,5,6,7,8
eight	0.998	55	93.2	

† Combinations: 1, SATT5; 2, SATT2; 3, SOYPRP1; 4, SOYHSP176; 5, SATT1; 6, SOYAC7L; 7, SOYSC514; 8, SAT43.

Table 4. Phenotypes of the two cultivars placed into each of four groups by an 8-locus SSR analysis.

Group	Entry	Leaf Color[a]	Pod Color[b]	Coat Color[c]	Flower Color[d]	Pubescence Color[e]
1	KWS 9	G	L.B	M.C	P	G
	KWS 12	G	D.B	B	P	T
2	KWS 19	G	Br	B	W	T
	KWS 24	G	Br	B	W	T
3	KWS 32	D.G	D.B	B/G	P	T
	KWS 36	D.G	D.B	B/G	P	T
4	KWS 48	D.G	D.B	B(G)	P	T
	KWS 49	D.G	D.B	B(G)	P	T

[a]Leaf color: G, green; D.G, dark green.
[b]Pod color: Br, brown; L.B, light brown; D.B, dark brown.
[c]Coat color: B, black; G, green; M.C, mixed color; B/G, black/green; B(G), black seed coat with green cotyledon.
[d]Flower color: P, purple; W, white.
[e]Pubescence color: G, gray; T, tawny.

could distinguish varieties more effectively. The number of groups that were established as well as the number of individuals that were identified increased rapidly as more loci were combined for identification. When eight loci were involved, the 59 landraces were divided into 55 groups, which allowed 93.2% of the varieties to be uniquely identified. Only KWS9 vs KWS12, KWS19 vs KWS24, KWS32 vs KWS36, KWS48 vs KWS49 could not be distinguished by the 8-locus SSR analysis. These four pairs were further examined for their leaf color, pod color, seed coat color, flower color and pubescence color(Table 4). Only the KWS9 and KWS12 pair could be distinguished using these data.

Soybean leaf color, pod color, seed coat color, flower color and pubescence color have long been and are still among the common morphological traits used for the identification of soybean landraces. The efficiency of these traits in comparison to SSR marker was also assessed. When one trait was used alone, the 59 landraces were divided on average into 3.4 groups. The actual number of such groups ranged from 2 to 7 because some traits (flower and pubescence colors) had only two contrasting phenotypes (i.e., alleles). The landraces could be further divided into 8.5 (range from 3 to 7), 15.8 (range from 7 to 25), 24.6 (range from 17 to 32) and 34 groups, when two, three, four or five traits, respectively, were combined for identification. Although the number of groups into which the landrace were placed varied with the combinations of morphological traits, the mean group number and maximum group number were much lower than that obtained with SSR markers. High resolving power of microsatellite DNA markers for soybean genotype identification was also demonstrated by Rongwen et al. (1995) using 7 loci. Among 96 diverse soybean genotypes in their study, only two genotypes had identical SSR allelic profiles and these had very similar pedigrees. Generally, the variabilities of the morphological traits observed among landrace of soybean were more obvious than among released modern cultivars. The SSR information was expected to be relatively more informative when it was used for modern cultivar identification in comparison with morphological traits.

Genetic differentiation and diversity among Korean soybean landraces: Genetic differentiation based on a one-locus SSR analysis of Korean soybean landraces varied from 0.57 to 0.82 with a mean of 0.68. Rongwen et al. (1995) selected a group of 56 soybean cultivars, 23 plant introductions, and 12 Chinese landraces, along with 5 wild soybean accessions to represent a diversity of soybean genotypes. Calculations from the data presented indicated that gene differentiation values associated with a four-locus combinations of SSR markers were extremely high, ranging from 0.72 to 0.86 with a mean value of 0.81. When the subset of genotypes consisted of cultivars developed in North American soybean breeding

programs, the genetic differentiation ranged from 0.54 to 0.80, with a mean of 0.66. Morgante et al. (1994) selected 61 soybean germplasm accessions (14*G. soja*, 15 PIs, and 32 ancestral lines or cultivars) to screen SSRPs (Simple Sequence Repeat Polymorphisms) using 7 microsatellites. The differentiation values calculated on the basis of three-locus SSR marker combinations ranged from 0.84 to 0.9, with a mean of 0.88 for the *G. soja* group, from 0.23 to 0.66, with a mean of 0.45 for plant introductions and ancestral lines, and from 0.39 to 0.44, with a mean of 0.42 for public cultivars. In comparison, the differentiation value (for 3-locus or 4-locus combinations) for the Korean landraces was similar to that of soybean non-ancestral cultivars developed in North American soybean breeding programs reported by Rongwen et al. (1995), but higher than that of PI and public cultivar populations, and lower than that of wild soybeans as reported by Morgante et al (1994).

The number of alleles detected by 8 loci varied from 3 to 8 among the 59 landraces. However, what is important in using markers in genetic diversity studies is the effective number of alleles which is a measure that takes into account both the number of alleles per locus and the frequencies of those alleles in the population. Among the 59 landraces, the effective number of alleles ranged from 2.3 to 5.1. Similar results were also observed by Powell et al.(1996) in *G. max* by studying 37 *G. max* and *G. soja* accessions collected throughout the native geographical ranges of these two species, and a subset of 12 genotypes(core genepool) representing 92% of total allelic diversity identified at 115 RFLP loci for over 700 lines including 134 public lines, 88 plant introductions, 45 commercial cultivars and 427 breeding lines or other germplasms. The number of alleles identified by these authors for the SOYRPRP1, SOYHSP176, SOYAC7L and SOYSC514 in *G. max* were 5, 3, 2, 5 respectively. The number of alleles we identified were 3, 4, 4, 5, respectively (Table 2). Additional alleles were found for the SOYHSP176, SOYAC7L loci in the Korean landraces, but fewer alleles were observed for SOYPRP1, while the number of alleles was the same for the SOYSC514 locus in both studies.

On average, SSR loci might be preferable to RFLP when measuring of genetic differentiation in soybean. Calculation from the data presented by Keim et al. (1989) indicated that the 17 most informative RFLP probes in a group of 58 diverse soybean and wild soybean genotypes had a genetic differentiation value of 0.56. The remaining 16 probes had values less than 0.51. This contrasted to the high level of polymorphism detected by SSR loci in soybeans by Powell et al. (1996) and in our research. Morgante at al (1994) also carried out comparative studies between SSRs and RFLPs. The differentiation values for SSR and RFLP were 0.52 and 0.38 for 19 lines in common, 0.62 and 0.36 for all 61 lines respectively.

Association of SSR alleles with quantitative traits and qualitative traits: The 59 landrace varieties were divided into 4, 4, 3, 4, 8, 4, 5 and 8 groups by single-locus analysis using SSRs SATT5, SATT2, SOYPRP1, SOYHSP176, SATT1, SOYAC7L, SOYSC514 and SAT43, respectively, (Table 2). The chi-square values of 5 qualitative traits with alleles of different loci and means of 12 quantitative

Table 5. Chi-square tests of independence between(or among) alleles at each SSR locus.

Loci	Leaf Color	Pod Color	Coat Color	Flower Color	Pubescence Color
SATT5	4.13	4.83	19.30	9.67*	1.34
SATT2	7.11	12.17	28.48	7.20	10.17*
SOYPRP1	10.89*	3.30	16.40	2.85	0.30
SOYHSP176	1.21	11.40**	24.95*	2.11	1.66
SATT1	11.76	12.69	31.14	9.62	4.88
SOYAC7L	6.53	25.13**	18.23	9.31*	0.10
SOYSC514	18.83*	16.00*	39.22*	2.02	12.83*
SAT43	12.80	20.34	39.94	7.20	2.65

*Significant at 5% probability. **Significant at 1% probability.

Table 6. Means of quantitative traits in each subgroup divided separately or jointly by microsatellite allele.

A	N	Flower (days)	Maturity (days)	Height (cm)	No. Branch	lodeing	No. pod	100wt (g)	Yield (kg/10a)	SMV score	SMV-N score	Protein (%)	Oil (%)
SATT5													
1	19	70.74	133.50	87.63	3.87	2.13	40.70	32.54	193.05	3.53	0.95	38.01	17.18
2	12	73.17	140.96	108.42	4.46	2.46	41.81	37.52	208.92	2.42	0.92	37.95	17.32
3	21	69.76	134.12	81.86	4.04	2.48	45.32	30.19	119.14	2.81	0.81	38.97	16.63
4	7	65.07	125.50	77.93	3.47	2.50	40.20	31.63	183.07	3.29	1.29	39.47	18.23
SATT2													
1	29	70.79	135.43	94.90	3.95	2.64	39.03	35.00	191.78	3.28	1.00	37.78	17.26
2	9	71.50	134.06	80.28	4.03	2.06	43.83	27.69	209.78	2.89	0.78	39.40	16.10
3	18	68.83	132.64	82.44	3.96	2.11	43.79	32.26	196.83	2.89	1.06	39.06	17.38
4	3	69.00	133.83	90.50	4.70	2.17	64.79	26.40	215.33	1.67	0.00	39.67	17.63
SOYPRP1													
1	38	73.47	138.78	95.78	4.17	2.39	41.33	34.78	200.33	3.11	0.83	38.11	17.54
2	31	68.76	132.63	87.40	3.92	2.58	44.32	31.23	135.45	3.13	1.10	39.10	16.75
3	10	68.85	131.35	79.70	3.99	1.65	39.02	32.97	197.35	2.50	0.60	37.42	17.61
SOYHSP176													
1	35	70.51	134.77	92.16	3.98	2.41	39.80	34.74	192.06	3.31	1.03	38.21	17.52
2	15	70.07	133.80	86.13	3.84	2.50	45.45	28.61	206.53	2.47	0.93	39.43	16.55
3	8	69.25	133.13	81.31	4.25	2.00	48.16	31.20	200.56	2.75	0.63	38.11	16.58
SATT1													
1	22	68.50	133.18	81.89	3.71	2.07	41.71	33.45	187.89	2.82	0.32	38.64	16.88
2	7	67.57	126.64	77.21	3.59	1.86	38.37	33.24	189.57	3.14	1.10	39.09	17.00
3	3	74.17	133.17	73.83	2.60	1.67	38.83	24.50	214.00	4.33	1.67	40.00	16.33
4	11	72.95	137.05	86.50	4.58	2.18	44.05	34.04	204.82	2.64	1.09	38.08	16.55
5	8	70.63	139.19	108.31	4.55	3.31	42.94	30.84	211.06	3.25	2.38	38.39	17.96
6	3	71.33	137.67	91.50	4.13	4.00	40.50	36.23	206.83	4.33	1.33	37.63	18.17
7	3	71.00	134.50	95.00	4.50	2.50	58.60	27.13	201.83	2.33	0.00	38.53	16.73
8	2	72.75	134.75	89.75	4.55	3.00	39.95	35.30	184.25	3.00	0.00	37.00	20.60
SOYAC7L													
1	7	73.14	136.29	92.71	4.19	1.86	35.91	38.59	184.71	2.86	0.71	38.33	17.71
2	7	67.64	128.07	67.07	2.93	2.00	39.24	27.66	192.57	3.57	1.29	41.07	17.17
3	33	70.58	136.02	92.32	4.15	2.45	43.83	33.36	198.98	3.00	0.82	37.84	17.35
4	12	69.00	132.00	88.79	4.14	2.63	44.64	29.93	202.58	2.83	1.17	38.98	16.20
SOYSC514													
1	5	67.90	132.30	79.40	3.84	1.90	39.60	37.98	198.00	3.40	0.80	36.90	17.00
2	7	69.64	134.14	87.43	4.57	1.93	48.76	32.43	197.64	2.86	0.14	38.14	16.49
3	32	71.50	134.70	89.98	3.98	2.45	41.74	31.31	199.36	2.88	1.47	39.06	17.24
4	7	66.07	132.36	90.86	3.67	3.00	37.10	33.90	179.93	3.86	0.43	38.11	18.21
5	8	70.63	135.69	88.25	3.99	2.13	46.69	33.49	203.25	2.75	0.00	37.99	16.40
SAT43													
1	1	77.00	137.00	108.00	4.50	3.00	45.40	30.20	139.50	3.00	3.00	41.20	16.20
2	3	68.88	136.38	95.00	4.53	3.25	51.68	25.93	220.88	3.00	2.00	39.57	16.45
3	12	68.63	131.08	77.67	3.83	1.50	40.16	33.52	188.58	2.50	0.58	38.45	17.65
4	19	69.63	133.84	88.13	4.16	2.42	40.70	33.49	196.97	3.16	0.68	38.64	17.32
5	8	71.31	136.00	102.94	4.18	3.19	41.02	33.34	193.13	4.25	1.88	38.06	16.24
6	8	72.88	137.44	98.00	4.29	2.94	44.35	33.45	208.69	2.00	0.88	37.58	18.04
7	6	70.92	134.17	75.75	3.02	1.42	44.75	32.93	191.17	3.33	0.83	39.33	16.48
8	2	70.75	134.50	91.00	3.70	2.50	47.35	24.90	194.00	3.00	0.00	38.65	15.65
Three subclusters based on the dendrogram Fig. 1 of microsatellite clusters													
1	43	70.84	134.98	90.45	3.93	2.44	41.26	33.32	193.01	3.14	1.05	38.50	17.43
2	9	69.28	133.72	83.17	4.30	4.89	44.98	33.12	204.28	2.67	0.67	38.02	16.82
3	7	69.50	133.21	88.57	4.20	2.64	46.40	28.16	214.50	2.71	0.71	39.13	15.99

*Significant at 5% probability; **significant at 1% probability.
The number to the right of the asterisk(s) was the subgroup which differed significantly from the other subgroups.
A, subgroup number; N, number of genotypes in the subgroup.

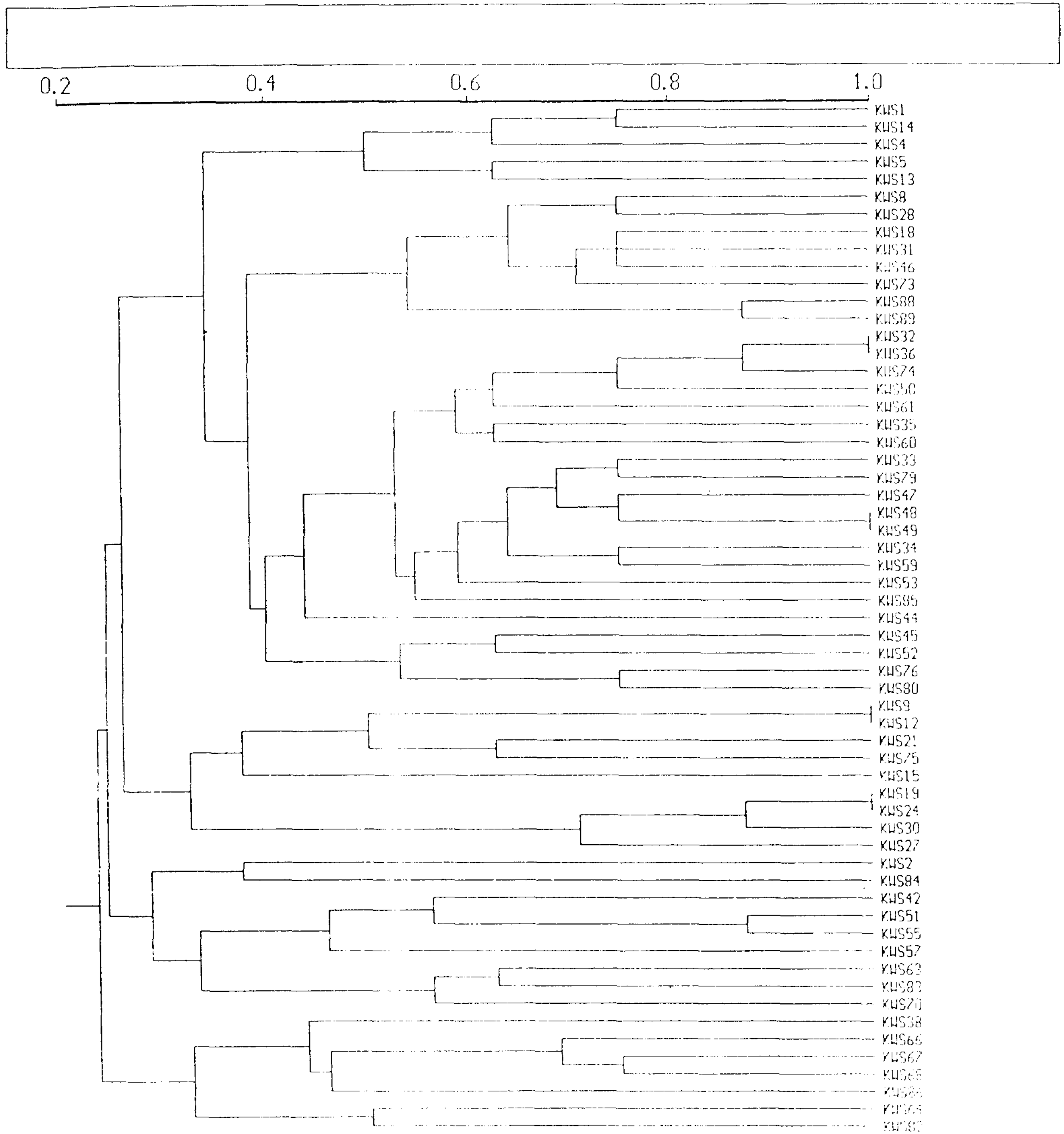

Fig. 1. Dendrogram of 59 soybean landrace based on 8 SSR profiles.

traits in each group were calculated and are shown in Tables 5 and 6. Contingency tests for independence indicated that there were strong interactions and associations among the 4 alleles of SATT5 and the two alleles of the flower color locus, among 4 alleles of SATT2 and two alleles of the pubescence color locus, among the 3 alleles of SOYPRP1 and the three alleles of the leaf color locus, among the 4 alleles of SOYHSP176 and the three alleles of the pod color locus as well as the five alleles of the seed coat color lcous, among the 4 alleles of SOYAC7L and the pod color alleles as well as the flower color alleles, and among the 5 alleles of SOYSC514 and alleles at all qualitative traits except for the flower color.

The group means for quantitative traits were calculated based on the values of genotypes in each group corresponding to the different alleles at each loci (Table 6). A t-test of the difference between means in different groups showed that alleles for such loci as SATT5, SOYPRP1, SOYHSP176, SOYAC7L, SOYSC514 and SAT43 were not associated with the quantitative traits being investigated. However, associations between SATT2 and number of pods, between SATT1 and number of branches, 100-seed weight, as well as oil content were found.

The high frequency of association between each SSR locus and one or several traits of qualitative characters as well as a relatively high frequency of association between each SSR locus and one or several quantitative traits indicated that SSR loci might be useful in integrating linkage maps, and in QTL detection and measurement or for germplasm screening in soybeans. Since the above association was based on a single locus, analysis incorporating information from all 8 SSR loci was further carried out through cluster analysis (Fig. 1). With a base line similarity index of 0.22, the 59 genotypes clustered into 3 groups, each with 43, 9, and 7 genotypes, respectively. As shown at the bottom of Table 6, quantitative trait differences among three subclusters were not significant. Similar results were obtained when 59 genotypes were divided into 4 subclusters if similarity index was set at 0.28. Means of 12 quantitative traits between 4 subclusters were not significant except for the subclusters 1 and 3 in studies of the maturity trait (data not shown). It would be expected that dendrograms based on multi-microsatellite information would not be consistent with that based on phenotypic values of quantitative traits. This was confirmed by calculating the values of product-moment correlation between dissimilarity matrices computed for the eight SSR markers and distance matrices computed for the 12 quantitative traits. The correlation coefficient was 0.320. Therefore, genetic diversity based on combined SSR information might not be consistent with phenotypic variability. Combined SSR data probably would not be useful for assigning soybean landraces to different diversity pools for quantitative traits without first testing the correlation between microsatellite loci information and morphological traits. Low correlations were

Table 7. Frequencies of qualitative traits in each subclusters based on dendrogram of microsatellite information.

Sub		Leaf color[a]				Pod color[b]			Seed coat color[c]					Flower color[d]		Pubescence color[e]	
class	N	G	L.G	D.G	Br	L.B	D.B	B	Br	Y	M.C	B/G	B(G)	P	W	T	G
1	43	0.56	0.07	0.37	0.14	0.26	0.61*3	0.37	0.00	0.05	0.07	0.07	0.33	0.93	0.07	0.79**2	0.21**2
2	9	0.56	0.22	0.22	0.33	0.22	0.44	0.56	0.22**1	0.11	0.11	0.00	0.00	1.00	0.00	0.33**3	0.67**3
3	7	0.43	0.14	0.43	0.14	0.71*1	0.14	0.86*1	0.00	0.00	0.14	0.00	0.00	1.00	0.00	1.00	0.00

*, ** significant difference between subclusters at the 5% and 1% probability levels, respectively.
The number to the right of the asterisks identifies the subclusters which differed significantly from the other subclusters.
SC, subcluster; N, number of genotypes in each subcluster.
[a]Leaf color: G, green; L.G. light green; D.G, dark green; Br, brown.
[b]Pod color: Br, brown; L.B, light brown; D.B, dark brown.
[c]Coat color: B, black; G, green; Br, brown; M.C, mixed color; B/G, black/green; B(G), black seed coat with green cotyledon.
[d]Flower color: P, purple; W, white.
[e]Pubescence color: G, gray; T, tawny.

also reported by Graner et al. (1994), who evaluated the association between coefficient of coancestry and genetic similarity based on RFLP analysis of 24 winter and 24 spring barley cultivars from the European barley germplasm, and by Powell et al. (1996), who studied allelic variability based on pedigree and molecular data in soybeans.

The expression of qualitative traits such as light brown, dark brown pod color, black seed coat colour, and brown seed and pubescence colors were significantly different between subclusters (Table 7). Considering the correlation between single loci and qualitative traits, it might be that alleles of microsatellites were more related to some traits determined by major genes than traits determined by minor genes. Since expression of quantitative traits was complicated by the presence of many alleles, with each allele having a small effect, the expression of each such character more easily affected by environment than the expression of qualitative traits.

ACKNOWLEDGEMENT

The authors thank KMOE for a financial support of Qijian Song to work in Korea. The authors also appreciate the painstaking proof-reading by Drs. Ramesh Bhambhani and John Kuspira. Part of this work was carried out with a genetic engineering grant from Korea Ministry of Education to NSK at the period of 1996, 1997.

REFERENCES

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

Delanney, X., Rodgers, D. M., Palmer, R. G. 1983. Relative genetic contribution among ancestral lines to North America soybean cultivars. Crop Sci. 23:944-949.

Dellaporta, S. L., Wood, J., Hicks, J. B. 1983. A simple and rapid method for preparation. Version II. Plant Mol. Bio. Rep. 1: 19-21.

Gizlice, Z., Carter Jr., T. E., Burton, J. W. 1994. Genetic base for North American public soybean cultivars released between 1947 and 1988. Crop Sci. 34:1143-1151.

Graner, A., Ludwig, W. F., Melchinger, A. E. 1994. Relationship among European barley germplasm: II Comparison of RFLP and pedigree data. Crop Sci. 34:1199-1205.

Gregorius, H. R. 1987. The relationship between the concepts of genetic diversity and differentiation. Theor. Appl. Genet. 74:397-401.

Gregorius, H. R., Roberds, J. H. 1986. Measurement of genetical differentiation among sub-populations. Theor. Appl. Genet. 71:826-834.

Hymowitz, T., Bernard, R. L., 1991. Origin of the soybean and germplasm introduction and development in North America. In: Shands HL and Wiesner LE(ed.) Use of plant introduction in cultivar development. Part 1. CSSA Spec. Publ. 17, CSSA, Madison, WI, pp147-164.

Keim, P., Shoemaker, R. C., Palmer, R. G. 1989. Restriction fragment length polymorphism diversity in soybean. Theor. Appl. Genet. 77:786-792.

Keim, P., Beavis, W., Schupp, J., Freestone, R. 1992. Evaluation of soybean RFLP marker diversity in adapted germplasm. Theor. Appl. Genet. 85:205-212.

Lark, K. G., Weisemann, J. M., Matthews, B. F., Palmer, R., Chase, K., Macalma, T. 1993. A genetic map of soybean (*Glycine max* L.) using an intraspecific cross of two cultivars: "Minsoy" and "Noir 1". Theor. Appl. Genet. 86:901-906.

Mantel, N. A. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res., 27:209-220.

Messmer, M. M., Melchinger, A. E., Herrmann, and Boppermaier, J. 1993 Relationship among early European maize inbreds: II. Comparison of pedigree and RFLP data. Crop Sci. 33:944-950.

Morgante, M., Rafalski, A., Biddle, P., Tingey, S., Olivieri, A. M. 1994. Genetic mapping and variability of seven soybean simple sequence repeat loci. Genome 37:763-769.

Plaschke, J., Ganal, M. W., Roder, M. S., 1995. Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theor. Appl. Genet. 91:1001-1007.

Powell, W., Morgante, M., Doyle, J. J., McNicol, J. W., Tingey, S. V., Rafalski, A. J. 1996. Gene pool variation in Genus *Glycine* subgenus *Soja* revealed by polymorphic nuclear and chloroplast microsatellites. Genetics 144:793-803.

Rohlf, S. B. 1992. NTSYS-pc numerical taxonomy and multivariate analysis system, version 1.70, Applied Biostatistics Inc., New York.

Rongwen, J., Akkaya, M. S., Bhagwat, A. A., Lavi, U., Cregan, P.B. 1995. The use of microsatellite DNA markers for soybean genotype identification. Theor. Appl. Genet. 90:43-48.

Smith, O. S., Smith, J. S. C., Bowen, S. L., Tenborg, R. A., Wall, S. J. 1990. Similarities among a group of elite maize inbreds as measured by pedigree, F1 grain yield, heterosis, and RFLPs. Theor. Appl. Genet. 80:833-840.

Tinker, N. A., Fortin, M. G., Mather, D. E. 1993. Random amplified polymorphic DNA and pedigree relationships in spring barley. Theor. Appl. Genet. 85:976-984.