Genes & Genomics (2010) 32: 299-308
DOI 10.1007/s13258-010-0014-9

RESEARCH ARTICLE

Assessment of genetic diversity and population structure in mungbean

Jae-Gyun Gwag · Anupam Dixit · Yong-Jin Park · Kyung-Ho Ma · Soon-Jae Kwon · Gyu-Taek Cho · Gi-An Lee · Sok-Young Lee · Hee-Kyoung Kang · Suk-Ha Lee

Received: 8 February 2010 / Accepted: 20 April 2010 / Published online: 31 August 2010
© The Genetics Society of Korea and Springer 2010

Abstract
This study was carried out to assess the genetic diversity and to analyze the population genetic structure for a total of 692 mungbean accessions preserved at National Agrobiodiversity Center (NAC) of the Rural Development Administration (RDA), Korea. Mungbean accessions were collected from 27 countries in nine different geographic regions, and were genotyped using 15 microsatellite markers, which were developed in our previous study. A total of 66 alleles were detected among 692 accessions at all the loci with an average of 4.4 alleles per locus. All the microsatellite loci were found to be polymorphic. The expected heterozygosity (H_E) and polymorphism information content (*PIC*) ranged from 0.081 to 0.588 (mean = 0.345) and from 0.080 to 0.544 (mean = 0.295), respectively. Of the 66 alleles, 17 (25.8%) were common (frequency range between 0.05 and 0.5), 15 (22.7%) were abundant (frequency range > 0.5), and 34 (51.5%) were rare (frequency range < 0.05). Locus GB-VR-7 provided the highest number of rare alleles(eight), followed by GB-VR-91(six) and GB-VR-113(four). Country-wide comparative study on genetic diversity showed that accessions from the USA possessed the highest genetic diversity (*PIC*) followed by Nepal, Iran, and Afghanistan. And region-wide showed that accessions from Europe possessed the highest average genetic diversity, followed by accessions from the USA, South Asia, West Asia, and Oceania. Twenty-seven countries were grouped into seven clades by phylogenetic relationship analysis, but clustering pattern did not strictly follow their geographical origin because of extensive germplasm exchange between/among countries and regions. As a result of a model-based analysis (STRUCTURE) of microsatellite data, two distinct genetic groups were identified which shared more than 75% membership with one of the two genetic groups. However the genetic group pattern did not reflect their geographical origin. The Duncan's Multiple Range Test among these two genetic groups and an admixed group, with a mean of 16 phenotypic traits, showed significant difference in 12 quantitative and qualitative traits on the basis of ANOVA. These 15 newly developed SSR markers proved to be useful as DNA markers to detect genetic variation in mungbean germplasm for reasonable management and crossbreeding purposes.

J.-G. Gwag · A. Dixit · K.-H. Ma · S.-J. Kwon · G.-T. Cho ·
G.-A. Lee · S.-Y. Lee
National Agrobiodiversity Center, National Academy of Agricultural Science, RDA, 88-20, Seodun-dong, Suwon, Kyunggi 441-707, Korea

Y.-J. Park · H.-K. Kang
College of Industrial Science, Kongju National University, Yesan, 340-802, Korea

S.-H. Lee (✉)
Department of Plant Science and Research Institute for Agriculture and Life Sciences, Seoul National University, Seoul 151-921, Korea
Plant Genomics and Breeding Institute, Seoul National University, Seoul 151-921, Korea
e-mail: sukhalee@snu.ac.kr

S.-J. Kwon
US Department of Agriculture-Agricultural Research Service, Western Regional Plant Introduction Station, 59 Johnson Hall, Washington State University, Pullman, WA 99164

Keywords Mungbean; SSRs; Genetic diversity; Population structure

Introduction

Mungbean (*Vigna radiata* L. Wilczek) is a leguminous species or pulse crop, grown principally for its protein-rich edible seeds. Mungbean seed flour is used for making soup and pizza-like ‘Bindaeddeok’ in Korea. Like soybean sprouts (Bae et

Springer

Table 1. Origin of 692 mungbean accessions used in this study and their model-based population genetic groups using 15 SSR markers.

Origin		No. of	Population genetic group		
Region	Country	accessions	1	2	admixed
Africa	Kenya	1	-	1	-
	Madagascar	1	-	-	1
	Nigeria	5	4	-	1
	Sub-total	**7**	**4**	**1**	**2**
East Asia	Korea	427	167	255	5
	Japan	5	4	1	-
	China	10	3	5	2
	Taiwan	14	7	5	2
	Sub-total	**456**	**181**	**266**	**9**
Southeast Asia	Indonesia	2	1	1	-
	Malaysia	2	2	-	-
	Philippines	50	31	16	3
	Thailand	19	10	5	4
	Vietnam	11	5	5	1
	Sub-total	**84**	**49**	**27**	**8**
Central Asia	Uzbekistan	16	1	15	-
	Afghanistan	13	4	6	3
	Sub-total	**29**	**5**	**21**	**3**
South Asia	India	29	12	7	10
	Pakistan	10	8	-	2
	Sri Lanka	1	1	-	-
	Nepal	10	5	4	1
	Sub-total	**50**	**26**	**11**	**13**
West Asia	Iran	20	11	8	1
	Turkey	8	5	2	1
	Sub-total	**28**	**16**	**10**	**2**
Europe	United Kingdom	3	1	1	1
	Netherlands	1	-	-	1
	Russia	4	1	2	1
	Sub-total	**8**	**2**	**3**	**3**
Oceania	Australia	11	5	6	-
America	USA	17	6	9	2
	Guatemala	1	1	-	-
	Nicaragua	1	1	-	-
	Sub-total	**19**	**8**	**9**	**2**
Total		**692**	**296**	**354**	**42**

al., 2004; Lee and Kim, 2004), mungbean sprouts are also an important traditional food in Korea and are an excellent source of protein, calcium, and vitamin C. Despite its usefulness, limited research efforts have been made to understand its genome structure and genetic diversity as compared with other legumes (Humphry et al., 2002). Accurate knowledge of genetic diversity and relationships among preserved germplasm collections of any crop is essential and of critical importance for establishing, managing and ensuring long-term success of appropriate crop improvement programs through breeding. The National Agrobiodiversity Center of South Korea has a collection of more than 1,000 accessions of mungbean germplasm. However, unclear understanding of genetic variability within this collection has been greatly hampered due to the lack of molecular characterization and evaluation data which represents a great challenge for successful improvement of this crop. Few reports are available on the use of molecular markers, such as RAPD and microsatellite, for estimation of genetic variability of mungbean cultivars (Betal et al., 2004; Kumar et al., 2002a, 2002b; Lakhanpaul et al., 2000), and the number of accessions was very limited.

Molecular markers can detect differences in DNA sequences and are less ambiguous than phenotypic markers. The polymerase chain reaction described by Mullis et al. (1986) is an important tool in molecular biology. In recent years, many DNA markers have been developed and have become powerful tools for detecting genetic diversity. Molecular markers provide various methods, based on DNA polymorphism, to study genetic diversity within and between populations. The choices include restriction fragment length polymorphism (RFLP) (Botstein et al., 1980), random amplified polymorphic DNA (RAPD) (Welsh and McClelland, 1990; Williams et al., 1990), amplified fragment length polymorphism (AFLP) (Vos et al., 1995), microsatellites (SSR) (Tautz, 1989), and single nucleotide polymorphism (SNP) (Gupta et al., 2001). Especially, simple sequence repeat (SSR) or microsatellite marker is one of the most useful genetic marker systems that using PCR to identify differences in microsatellite repeat units, and it is widely used because it can easily read co-dominant, multi-allelic, high polymorphism and reproduction (Litt and Luty, 1989; Weber and May, 1989).

Evaluation of genetic diversity has an important role within preservation programs for genetic resources of crop plants (Newbury and Ford-Lloyd, 1997). Knowledge of germplasm diversity using molecular markers provides a significant impact on the improvement of crop plants. For example, molecular markers have been used to identify duplicate accessions, genetic relationships, and population structures, and to determine how variation is distributed between individuals, accessions, and races (Westman and Kresovich, 1997).

The objective of this study is to assess the genetic diversity and to analyze the population structure of the 692 mungbean germplasm using newly developed microsatellite markers.

Materials and Methods

Plant materials and DNA extraction

The 692 mungbean accessions collected from 27 countries were taken from the National Agrobiodiversity Center (NAC) of Korea (Table 1). Each accession was grown in a greenhouse, and DNA was extracted from the fresh leaves of 15-day-

old seedlings using QIAGEN DNA extraction kit (QIAGEN). The relative purity and concentration of extracted DNA was estimated with the NanoDrop ND-1000 (Dupont Agricultural Genomics Laboratory, USA). The final concentration of each DNA sample was adjusted to 20 ng/μl.

Microsatellite analysis

A set of 15 newly developed primer pairs developed by Gwag et al. (2008) were used in this study. The size of polymorphic amplified products was measured accurately by the M13 tail PCR method of Schuelke (2000). Amplification reactions were carried out in a total volume of 20 μl containing 200 ng of template DNA, 1× PCR buffer, 0.2 mM of each dNTP, 1u Taq DNA polymerase, 8 pmol of each reverse and fluorescent-labeled M13 (-21) primer and 2 pmol of the forward primer with M13 (-21) tail at its 5' end. Conditions of the PCR amplification were as follows: 94℃ (3 min) then 30 cycles each at 94℃ (30 sec), 55℃ (45 sec), 72℃ (1 min) followed by 10 cycles of 94℃ (30 sec), 53℃ (45 sec), 72℃ (1 min), and a final extension at 72℃ for 10 min. Microsatellite alleles were resolved on 3130x1 Genetic Analyzer (Applied Biosystems, USA) using GENESCAN 3.7 software and sized precisely against 6-carboxy-X-rhodamine (ROX) molecular size standards using GENOTYPER 3.7 software (Applied Biosystems, USA).

Characterization of microsatellite loci and genetic diversity

Basic statistics were calculated using the genetic analysis package, PowerMarker version 1.31 (Liu and Muse, 2005), for diversity measurements of 692 mungbean accessions at each microsatellite locus, including the total number of alleles, allele frequency, rare alleles, observed heterozygosity (H_O), expected heterozygosity (H_E), and polymorphic information content (PIC). The PIC was calculated with the equation:

$$PIC=1-\sum_i P_i^2$$

where P_i is the frequency of the i-th allele in the sample examined (Anderson et al., 1993). The phylogenetic tree was constructed using the neighbor joining method and retrieved using TreeView program Version 1.6.6.

Population genetic structure analysis

For the analysis of population structure and identification of ancestral and hybrid forms, we used the model-based software program, 'STRUCTURE' (Pritchard et al., 2000). In this model, a number of populations (K) are assumed to be present, each of which is characterized by a set of allele frequencies at each locus. Individuals in the sample are assigned to populations (groups), or jointly to more populations if their genotypes indicate that they are admixed. All loci are assumed to be independent, and each K population is assumed to follow Hardy-Weinberg equilibrium. The posterior probabilities were estimated using a Markov Chain Monte Carlo (MCMC) method. The MCMC chains were run at burn-in period lengths of 100,000 at fixed iterations of five for each fixed number of population (K, 1-10). The final results were based on a burn-in period length of 100,000 and five iterations of this chain using a model allowing for admixed and correlated allele frequencies. An individual having more than 75% of its genome fraction value was assigned to a group.

Significance test for population genetic groups

Duncan's Multiple Range Test was used to clarify significance among three groups (two genetic groups and an admixed group) by population structure analysis using STRUCTURE software (Pritchard, et al., 2000) for the 16 phenotypic traits data, such as hypocotyl color (HC, 1. green, 2. greenish purple, 3. purple), seed coat color (SC, 1. yellow, 2. greenish yellow, 3. light green, 4. dark green, 5. brown, 6. mottled), luster on seed surface [LS, 1. absent (dull), 2. present (shiny)], growth habit (GH, 1. erect, 2. semi-erect, 3. spreading), days to flowering (DF, number of days from the date of sowing to the 50% of plants begin flowering), days to maturing (DM, number of days from the date of sowing to the 50% of plants begin first pod maturing), pod length (PL, cm, average of 20 pods), pod width (PW, cm, average of 20 pods), seed length (SL, mm, average of 20 seeds), seed width (SDW, mm, average of 20 seeds), plant height (PH, cm, average of 10 plants), number of first branches (BN, number of pod bearing branches originating from the leaf axils on the main stem, average of 10 plants), number of pods per plant (PNP, average of 10 plants), number of seeds per pod (SNP, average number of seeds from 20 pods), 100-seed weight (SW, g), yield per plant (YI, g). The SAS software (SAS Institute 2004) was used for Duncan's Multiple Range Test and analysis of variance (ANOVA).

Results

Genetic variations in mungbean collection

The variability at each microsatellite locus was measured in terms of the number of observed alleles (NA), observed heterozygosity (H_O), expected heterozygosity (H_E), and polymorphic information content (*PIC*) (Table 2). All 15 loci showed polymorphism, producing a total of 66 alleles among 692 mungbean accessions. The loci varied in the number of observed alleles

Table 2. Overall diversity statistics at each SSR loci for the mungbean germplasm.

SSR locus	Repeat motif	Size range(bp)	*NA*	*MAF*	H_O	H_E	*PIC*
GB-VR-7	$(CT)_7$	270 - 310	9	0.922	0.125	0.148	0.145
GB-VR-13	$(CAG)_4$	154 - 181	2	0.869	0.103	0.227	0.202
GB-VR-14	$(AAGAG)_4$	251 - 256	2	0.554	0.007	0.494	0.372
GB-VR-17	$(AG)_{10}$	145 - 163	4	0.875	0.193	0.226	0.211
GB-VR-38	$(CCG)_6$	124 - 142	5	0.926	0.127	0.137	0.129
GB-VR-77	$(GTT)_{5,}(GA)_5A(AG)_6$	301 - 313	3	0.559	0.004	0.494	0.374
GB-VR-87	$(CAA)_4$	263 - 275	2	0.723	0.012	0.401	0.321
GB-VR-91	$(GA)_{14}$	151 - 167	9	0.645	0.038	0.540	0.502
GB-VR-93	$(GAA)_{13}$	110 - 125	6	0.595	0.012	0.585	0.544
GB-VR-113	$(GA)_{10}$	147 - 241	5	0.959	0.023	0.081	0.080
GB-VR-142	$(TG)_2T(TG)_5CTC(TG)_2T(TG)_4$	224 – 260	6	0.690	0.221	0.487	0.450
GB-VR-172	$(GAA)_7$	226 - 235	4	0.814	0.007	0.304	0.259
GB-VR-180	$(CAA)_4$	253 - 295	3	0.911	0.007	0.162	0.140
GB-VR-184	$(AG)_3AA(AG)_3AA(AG)_4$	271 - 279	2	0.740	0.004	0.384	0.311
GB-VR-198	$(GA)_5AA(GA)_6$	223 - 303	4	0.528	0.003	0.503	0.381
Total			66				
Average			4.4	0.754	0.059	0.345	0.295

* *NA*: number of allele, *MAF*: major allele frequency, H_O: observed heterozygosity, H_E: expected heterozygosity, *PIC*; polymorphic information content.

from two to nine, with a mean of 4.4 alleles per locus. The size of alleles ranged from 110 to 313 bp. The two loci, GB-VR-7 and GB-VR-91, produced the highest number of alleles (nine in each locus). The lowest number of alleles (two alleles) was observed at each of the four loci, GB-VR-13, GB-VR-14, GB-VR-87, and GB-VR-184. For the majority of loci, differences in the size of alleles were observed almost exclusively by unit repeats, indicating that variations at these loci were due to differences in the number of repeat units. The H_O values ranged between 0.003 (GB-VR-198) and 0.221 (GB-VR-142) with a mean value of 0.059, and those of H_E and *PIC* ranged between 0.081 and 0.080 (GB-VR-113) and 0.585 and 0.544 (GB-VR-93), with mean values of 0.345 and 0.295, respectively. In this regard, the locus GB-VR-93 was the most informative, whereas the locus GB-VR-113 was the least informative. No correlation was observed between number of repeat units/number of alleles and genetic diversity (*PIC*). The values of the H_E were significantly higher than the corresponding H_O at all the loci. Of the 66 alleles, only 17 (25.8%) were common, with frequency range between 0.05 and 0.5, 15 (22.7%) were abundant (frequency range > 0.5), and 34 (51.5%) were rare (frequency range < 0.05). Noteworthy is the locus GB-VR-7 which provided the highest number of eight rare alleles, followed by six at GB-VR-91 and four at GB-VR-113 (Tables 3, 4). A country-wide comparative study of genetic diversity showed that accessions from the United States possessed the highest genetic diversity (*PIC*=0.334) followed by Nepal (*PIC*=0.333), Iran (*PIC*=0.317), Afghanistan (*PIC*=0.313), and Thailand (*PIC*=0.302). Region-wide comparative studies revealed that accessions from Europe possessed the highest average genetic diversity (*PIC*=0.345), followed by America (*PIC*=0.342), South Asia (*PIC*=0.313), West Asia (*PIC*=0.301), and Oceania (*PIC*=0.283). Even though a lot of accessions from the East Asia region countries were used in this study, their genetic diversities were quite low. Furthermore, accessions from Korea occupied 61.7% of the total accessions but the genetic diversity index ranked eighth among 19 countries with at least three accessions, with the value of *PIC*=0.273 (Table 5).

Cluster analysis

Twenty-seven countries were grouped into seven clades by Neighbor-Joining cluster analysis (Fig. 1). Clade 1 contained only one country, namely India; clade 2 consisted of Thailand, Turkey, and the Philippines; and clade 3 consisted of Nicaragua, Guatemala, Sri Lanka, Malaysia, Nigeria, and Japan. Korea was grouped into clade 4 with Russia, Madagascar, Uzbekistan, Afghanistan, China, and the United States. Iran and Nepal were grouped into clade 5, Vietnam and Indonesia were grouped into clade 6, and the United Kingdom, Netherlands, Kenya, Taiwan,

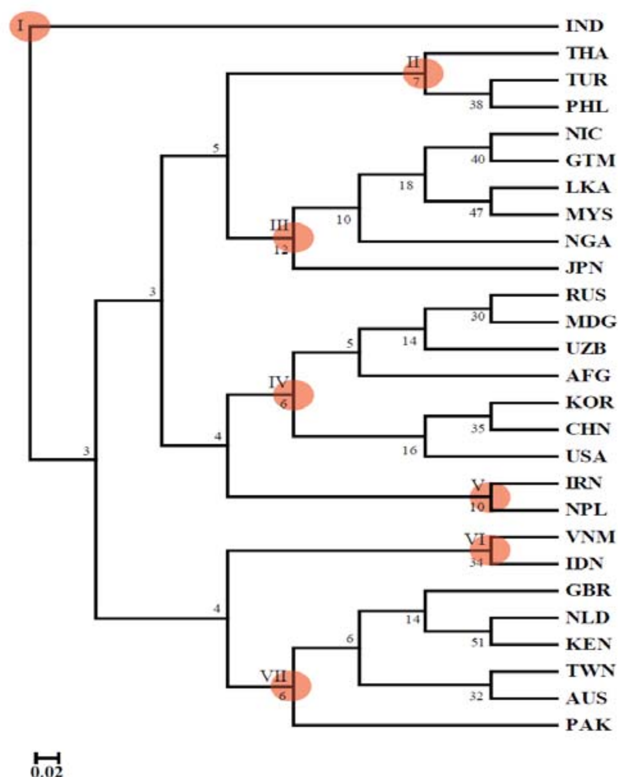

Figure 1. Phylogenetic dendrogram among 27 countries by neighbor joining analysis in mungbean germplasm.The numbers on the branches indicates bootstrap values (expressed in percentages, based on 100 replications).

Australia, and Pakistan were grouped into clade 7. Interestingly, although Korea and China belong to the same geographic region (East Asia) Taiwan and Japan, they are grouped into different clade from the two countries.

Population genetic structure

An analysis of the genetic structure and detection of ancestral and hybrid forms among 692 individuals of mungbean accessions was carried out using a model-based approach (Pritchard, et al., 2000). Five runs of STRUCTURE were done by setting the number of populations (K) from 1 to 10. For each run, burn-in time and replication number were both set to 100,000. K = 2 was found to converge well and showed comparable or higher likelihoods than K = 3~10 among runs of the program (Fig. 2). An individual having more than 75% of its genome fraction value was assigned to a group.

Of the 692 accessions, 650 (93.9%) shared > 75% membership with one of the two genetic groups and were classified as members of that genetic group. In addition to the accessions that were clearly assigned to a single genetic group, 42 accessions (6.1%) were categorized as admixtures with varying

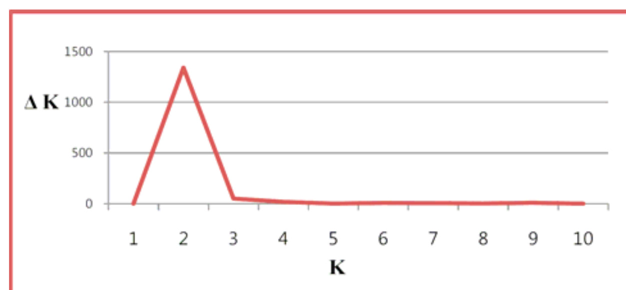

Figure 2. Magnitude of ΔK as a function of K. The modal value of this distribution is the true K or the uppermost level of structure, in this case two groups.

levels of membership shared among the two groups. The first genetic group (Group 1) includes 296 accessions with the majority (288 accessions; 97.3%) including > 90% membership of this group and with their memberships between 75.7 and 99.6%. The second genetic group (Group 2) was represented by 354 accessions, all of which showed their memberships between 78.9 and 99.6%, and majority of them (349 accessions; 98.6%) having > 90% membership of this group (Fig. 3).

Significance test for population genetic groups

Duncan's multiple range test (DMRT) was used to determine the significant difference among two genetic groups and an admixed group by model-based population structure analysis with mean of 16 morphological traits data (Table 6). Analysis of variance revealed significant difference in 12 qualitative and quantitative traits, such as hypocotyl color, luster on seed surface, growth habit, days to flowering, days to maturing, pod length, seed length, seed width, plant height, number of first branches, number of seeds per pod, and 100-seed weight. Multiple means comparisons using DMRT revealed that the luster on seed surface, days to flowering, and seed width were the traits showing significant difference for all three groups. Group 2 contained many accessions with purple hypocotyl, presence of luster on seed surface, semi-erect growth habit, long days to flowering and maturing, short seed length, narrow seed width, high plant height, and small seed compared to group 1.

Discussion

In this study, a total of 66 alleles were detected among the 692 of mungbean accessions at 15 microsatellite loci with an average of 4.4 alleles per locus which was quite low compared with other crops (Kapila et al., 2008; Lee et al., 2008) but

Table 3. Frequency distribution of different alleles in 15 microsatellite loci of 692 mungbean accessions.

Microsatellite loci	Allele (bp)	Frequency	Microsatellite loci	Allele (bp)	Frequency	Microsatellite loci	Allele (bp)	Frequency
GB-VR-7	288	0.030	GB-VR-77	319	0.441	GB-VR-113	237	0.959
GB-VR-7	292	0.004	GB-VR-77	321	0.558	GB-VR-113	239	0.013
GB-VR-7	296	0.001	GB-VR-77	331	0.001	GB-VR-113	259	0.003
GB-VR-7	298	0.003	GB-VR-87	281	0.721	GB-VR-142	242	0.145
GB-VR-7	300	0.924	GB-VR-87	293	0.279	GB-VR-142	248	0.123
GB-VR-7	308	0.002	GB-VR-91	169	0.006	GB-VR-142	250	0.008
GB-VR-7	324	0.015	GB-VR-91	171	0.186	GB-VR-142	252	0.691
GB-VR-7	326	0.018	GB-VR-91	173	0.079	GB-VR-142	254	0.001
GB-VR-7	328	0.003	GB-VR-91	175	0.649	GB-VR-142	278	0.032
GB-VR-13	171	0.133	GB-VR-91	177	0.015	GB-VR-172	244	0.001
GB-VR-13	200	0.867	GB-VR-91	179	0.006	GB-VR-172	247	0.183
GB-VR-14	269	0.440	GB-VR-91	181	0.014	GB-VR-172	250	0.815
GB-VR-14	275	0.560	GB-VR-91	183	0.043	GB-VR-172	253	0.001
GB-VR-17	163	0.018	GB-VR-91	185	0.002	GB-VR-180	271	0.913
GB-VR-17	165	0.877	GB-VR-93	110	0.162	GB-VR-180	274	0.086
GB-VR-17	167	0.012	GB-VR-93	113	0.003	GB-VR-180	312	0.001
GB-VR-17	181	0.093	GB-VR-93	116	0.056	GB-VR-184	289	0.738
GB-VR-38	124	0.071	GB-VR-93	119	0.166	GB-VR-184	297	0.262
GB-VR-38	127	0.002	GB-VR-93	122	0.598	GB-VR-198	241	0.469
GB-VR-38	133	0.001	GB-VR-93	125	0.016	GB-VR-198	243	0.527
GB-VR-38	139	0.001	GB-VR-113	165	0.010	GB-VR-198	253	0.003
GB-VR-38	142	0.925	GB-VR-113	235	0.015	GB-VR-198	321	0.001

Table 4. Distribution of alleles according to frequency range.

Frequency range	> 0.05 (rare)	0.05 ~ 0.5 (common)	0.5 < (abundant)	Total
No. of alleles	34 (51.5%)	17 (25.8%)	15 (22.7%)	66 (100%)

higher than other reports in mungbean (Kumar et al., 2002a, 2002b). Their reports revealed average of 3.6 and 3.1 alleles per locus, respectively. However, all the microsatellite loci were found to be polymorphic, and there was considerable genetic variation among mungbean accessions with a mean values of expected heterozygosity (H_E) 0.345 and polymorphic information content (*PIC*) 0.295. Majority of the microsatellite loci differed greatly in the number of alleles, ranging from two to nine. For majority of the loci, differences in size of alleles were observed exclusively by repeat units. The degree of polymorphism detected by these markers did not show any correlation with the number of alleles, H_E and *PIC* values. Some microsatellite loci producing similar number of alleles varied greatly in their H_E and *PIC* values. For example, nine alleles were detected at each of the two loci, GB-VR-7 and GB-VR-91, but there was a significant difference in their H_E and *PIC* values. This observed difference in H_E and *PIC* values may be attributed to the variation in allele frequency between the data sets of the two loci. A similar trend was also reported in rice (Queller et al., 1993).

A significantly high frequency of rare alleles (34 alleles; 51.5%) among mungbean accessions indicated that they made a greater contribution to the overall genetic diversity of the collection (Roussel et al., 2004; Yifru et al., 2006). Hence, it is important to include rare alleles for maximizing the genetic variations in the gene bank collections and to utilize them in a breeding program (Yifru et al., 2006). Genetic diversity of the USA revealed the highest among 27 countries, indicating that mungbean germplasm of the USA are more heterogeneous than other countries. Capo-chichi et al. (2003) also reported similar trend in velvetbean. Twenty-seven countries were grouped into seven clades by Neighbor-Joining cluster analysis with lower bootstrap values [6(clade 4 and 7) ~ 34(clade 6)] compared to other reports (Capo-chichi et al., 2003; Cuc et al., 2008; Pallottini et al., 2004). The bootstrap value is the measurement of confidence limits in the bootstrap branch. Higher level of polymorphism detected by SSR markers has contributed to the lower genetic similarity (Maras et al., 2008; Powell et al., 1996). Genetic diversity between countries or regions was higher than that observed among individual accession levels. It is possible to infer that because of the phylogenetic dendrogram was constructed among countries in this study, it might show lower bootstrap values. Overall clustering pattern did not strictly follow their geographic origins (country or region) because many accessions from different geographic locations were clustered together. Similar results were reported by Poehlman (1991) using phenotypic data. In his report, geographic diversity of origin was not always held in mungbean since many strains were dis-

tributed extensively with their origin frequently obscured. Besides, Malhotra et al. (1974) failed to identify genetic associations in germplasm with geographical diversity. These observations are indicative of the extensive germplasm exchanges among countries/regions from different geographical regions. According to the genetic distance among countries by neighbor-joining tree, Chinese accessions showed close relationship with Korean accessions. This might be explained by the fact that China is located on the path of mungbean introduction to Korea from its center of origin, such as India and Myanmar long ago. It could be surmised that Korea and China shared genetically similar mungbean accessions probably through germplasm exchange among the Asian Vegetable Research and Development Center (AVRDC) or other countries.

Population genetics deals with the variation of allele frequency between and within populations. The model-based method utilized a Bayesian clustering approach to probabilistically assign individuals to populations based on their genotypes and attempts to find population structure in which the population is in Hardy-Weinberg equilibrium (Barkley et al., 2006). In this study, 650 (93.9% of total accessions) out of 692 accessions were classified into two distinct genetic groups which shared more than 75% of its genome fraction. Accessions from East Asian countries, except Japan and Taiwan, were predominant in genetic group 2, and accessions

Table 5. Major allele frequency(*MAF*), observed heterozygosity(H_O), expected heterozygosity(H_E), and polymorphic information content(*PIC*) according to origin regions/countries.

Origin		No. of acc.	*MAF*	H_O	H_E	*PIC*
Region	Country					
Africa	KEN	1	0.867	-	-	0.117
	MDG	1	0.933	0.133	0.067	0.050
	NGA	5	0.823	0.053	0.248	0.207
	Sub-total	**7**	**0.768**	**0.076**	**0.326**	**0.275**
E. Asia	KOR	427	0.771	0.054	0.322	0.273
	JPN	5	0.893	0.027	0.173	0.147
	CHN	10	0.787	0.053	0.289	0.240
	TWN	14	0.738	0.049	0.320	0.266
	Sub-total	**456**	**0.769**	**0.054**	**0.325**	**0.275**
S. E. Asia	IDN	2	0.867	0.000	0.133	0.100
	MYS	2	0.900	0.133	0.117	0.091
	PHL	50	0.830	0.043	0.247	0.230
	THA	19	0.760	0.088	0.351	0.302
	VNM	11	0.790	0.043	0.281	0.238
	Sub-total	**84**	**0.807**	**0.054**	**0.292**	**0.254**
C. Asia	UZB	16	0.847	0.098	0.227	0.194
	AFG	13	0.741	0.118	0.369	0.313
	Sub-total	**29**	**0.797**	**0.107**	**0.318**	**0.276**
S. Asia	IND	29	0.748	0.073	0.345	0.297
	PAK	10	0.747	0.033	0.324	0.270
	LKA	1	1.000	0.000	0.000	0.000
	NPL	10	0.720	0.060	0.384	0.333
	Sub-total	**50**	**0.737**	**0.061**	**0.363**	**0.313**
W. Asia	IRN	20	0.742	0.100	0.370	0.317
	TUR	8	0.827	0.060	0.267	0.226
	Sub-total	**28**	**0.766**	**0.089**	**0.349**	**0.301**
Europe	GBR	3	0.811	0.022	0.256	0.200
	NLD	1	0.933	0.133	0.067	0.050
	RUS	4	0.758	0.117	0.298	0.246
	Sub-total	**8**	**0.713**	**0.083**	**0.401**	**0.345**
Oceania	AUS	11	0.736	0.055	0.330	0.283
America	USA	17	0.703	0.063	0.390	0.334
	GTM	1	0.900	0.200	0.100	0.075
	NIC	1	0.933	0.133	0.067	0.050
	Sub-total	**19**	**0.688**	**0.074**	**0.398**	**0.342**
Total(Mean)		**692**	**0.754**	**0.059**	**0.345**	**0.295**

* Full country name is shown in Table 1.

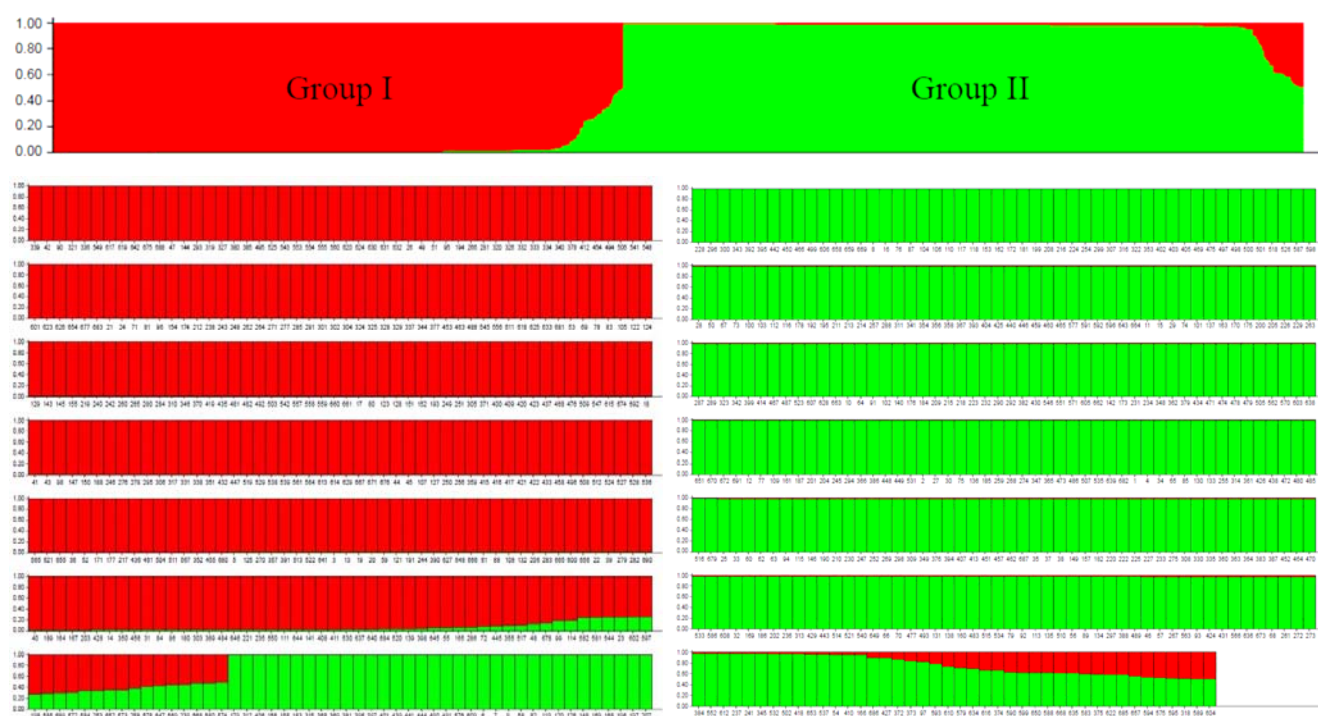

Figure 3. Bar plot of the genetic composition of 692 mungbean accessions based on 15 SSR markers generated by STRUCTURE 2.2 using the admixture model. Groups for each panel are represented by colors. Each individual bar represents each accession. The red- and green-colored bar refer to different genetic groups. The y-axis displays the estimated ancestry of membership of each individual in a particular gene pool.

from South Asia which is the center of origin of this crop, and accessions from South East Asian countries, except Vietnam, were predominant in genetic group 1. The view of Tomooka et al. (1992), who examined the variations of seed proteins in mungbean landraces from Asia, proposed the regions of protein type diversity and two dissemination pathways in mungbean was not fully supported by our results due to extensive germplasm exchanges among countries. Population determination is usually based upon geographical origin of samples or phenotypes. However, the genetic structure of population is not always reflected in the geographical proximity of individuals. Populations that are not discretely distributed can, nevertheless, be genetically structured due to unidentified barriers to gene flow. In addition, groups of individuals with different geographical locations, behavioral patterns or phenotypes are not necessarily genetically differentiated (Evanno et al., 2005). For example, migratory bats from the same breeding roost could be sampled thousands of kilometers apart in winter (Petit et al., 2001). Forty-two accessions (6.1%) did not fit into one of these two groups since they consist of a mixture of these two primary groups. Though only 42 accessions were categorized as admixed group, relatively a lot of accessions from India [10 of 29 accessions (34.5%)], Thailand [4 of 19 accessions (21.1%)], and Afghanistan [3 of 13 accessions (23.1%)] belonged to this group. On the contrary, Korean accessions belonged to this group relatively less accessions [5 of 427 accessions (1.2%)]. Moreover, any accession from Japan, Uzbekistan, and Australia was not observed in the admixed group. From these results, it is possible to infer that India, Thailand, and Afghanistan possessed relatively many introgressed forms (partial ancestry) of mungbean accessions, while Korea, Japan, Uzbekistan and Australia possessed many ancestral forms. These introgressed forms of accessions probably had a complex breeding history involving intercrossing and introgression between germplasm from diverse backgrounds, overlaid with strong selection pressure for agronomic and quality characteristics (Mather et al., 2004).

In the results of the Duncan's Multiple Range Test with mean of 16 phenotypic traits data, significant difference was observed among two genetic groups and an admixed group in 12 qualitative and quantitative traits. This result indicated that significant phenotypic variations of mungbean could be generated by a number of morphological and agronomical characters, and verified genetic diversity among genetic groups.

In conclusion, these 15 SSR markers were used successfully to detect genetic variation in mungbean germplasm for reasonable management and crossbreeding purposes. These data will also be used to develop collection strategies and to establish core collections in the RDA mungbean collection.

Table 6. Duncan's Multiple Range Test among population genetic groups using 16 phenotypic traits.

Group	No. acc	HC	SC	LS	GH	DF	DM	PL	PW
1	296	2.2B	3.3A	1.4B	1.5B	56.5C	71.4B	10.1A	5.1A
2	354	2.8A	3.3A	1.2C	1.9A	59.3B	75.0A	10.2A	5.0A
Admixed	42	2.4B	3.4A	1.5A	2.1A	61.6A	77.9A	9.5B	4.9B
F-value[a]		40.58***	2.02	12.57***	21.2***	12.86***	11.08***	3.24*	2.5
Group	No. acc	SL	SDW	PH	BN	PNP	SNP	SW	YI
1	296	5.0A	3.9A	91.8B	12.4A	86.6A	13.5B	5.1A	43.4A
2	354	4.7B	3.8B	103.5A	13.2A	91.3A	14.0A	4.7B	44.3A
Admixed	42	4.6B	3.7C	111.1A	12.9A	88.5A	13.5B	4.4B	38.3B
F-value		19.01***	9.87***	19.12***	7.01**	1.56	15.09***	12.64***	1.05

* HC: hypocotyl color, SC: seed coat color, LS: luster on seed surface, GH: growth habit, DF: days to flowering, DM: days to maturing, PL: pod length, PW: pod width, SL: seed length, SDW: seed width, PH: plant height, BN: number of first branches, PNP: number of pods per plant, SNP: number of seeds per pod, SW: 100-seed weight, YI: yield per plant.
[a] F-value of ANOVA.
*** Significant at $P < 0.0001$, ** $P < 0.001$, * $P < 0.05$.

Acknowledgements This work was supported by the Agenda Program (20090801-030-001), the Technology Development Program for Agriculture and Forestry, Ministry for Food, Agriculture, Forestry and Fisheries, Republic of Korea and National Academy of Agricultural Science, RDA, Republic of Korea.

References

Anderson JA, Churchill GA, Autrique JE, Tanksley SD and Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome 36: 181-186.

Bae KG, Nam SW, Kim KN and Hwang YH (2004) Growth of soybean sprouts and concentration of CO_2 produced in culture vessel affected by watering methods. Kor.ean J. Crop Sci. 49: 167-171.

Barkley NA, Roose ML, Krueger RR and Federici CT (2006) Assessing genetic diversity and population structure in a *citrus* germplasm collection utilizing simple sequence repeat markers (SSRs). Theor. Appl. Genet. 112: 1519-1531.

Betal S, Chowdhury PR, Kundu S and Raychaudhuri SS (2004) Estimation of genetic variability of *Vigna radiata* cultivars by RAPD analysis. Biol. Planta. 48: 205-209.

Botstein D, White RL, Skolnick M and Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Human Genet. 32: 314.

Capo-chichi LJA, Weaver DB and Morton CM (2003) The use of molecular markers to study genetic diversity in *Mucuna*. Tropic. Subtropic. Agroecosys. 1: 309-318.

Cuc LM, Mace ES, Crouch JH, Quang VD, Long TD and Varshney RK (2008) Isolation and characterization of novel microsatellites markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). BMC Plant Biol. 8: 55 doi:10.1186/1471-2229-8-5.

Evanno G, Regnaut S and Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Molec. Ecol. 14: 2611-2620.

Gupta PK, Roy JK and Prasad M (2001) Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Curr. Sci. 80: 524-535.

Gwag JG (2008) Molecular diversity assessment and population structure analysis in mung bean, *Vigna radiata* (L.) Wilczek. PhD Thesis Seoul National University pp 1-112.

Humphry M, Konduri V, Lambrides C, Magner T, McIntyre C, Aitken E and Liu C (2002) Development of a mungbean (*Vigna radiata*) RFLP linkage map and its comparison with lablab (*Lablab purpureus*) reveals a high level of colinearity between the two genomes. Theor. Appl. Genet. 105: 160-166.

Kapila RK, Yadav RS, Plaha P, Rai KN, Yadav OP, Hash CT and Howarth CJ (2008) Genetic diversity among pearl millet maintainers using microsatellite markers. Plant Breed. 127: 33-37.

Kumar SV, Tan SG, Quah SC and Yusoff K (2002a) Isolation of microsatellite markers in mungbean, *Vigna radiata*. Molec. Ecol. Notes 2: 96-98.

Kumar SV, Tan SG, Quah SC and Yusoff K (2002b) Isolation and characterization of seven tetranucleotide microsatellite loci in mungbean, *Vigna radiata*. Molec. Ecol. Notes 2: 293-295.

Lakhanpaul S, Chadha S and Bhat KV (2000) Random amplified polymorphic DNA (RAPD) analysis in Indian mung bean (*Vigna radiata* (L.) Wilczek) cultivars. Genetica 109: 227-234.

Lee JR, Hong GY, Dixit A, Chung JW, Ma KH, Lee JH, Kang HK, Cho YH, Gwag JG andPark YJ (2008). Characterization of microsatellite loci developed for *Amaranthus hypochondriacus* and their cross-amplifications in wild species. Conserv. Genet. 9: 243-246.

Lee YS and Kim YH (2004) Changes in postharvest respiration, growth, and vitamin C content of soybean sprouts under different storage temperature conditions. Kor.ean J. Crop Sci. 49: 410-414.

Litt M and Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am. J. Human Genet. 44: 397.

Liu K and Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21: 2128-2129.

Malhotra VV, Singh S and Singh KB (1974) Yield components in greengram (*Phaseolus aureus* Roxb). Indian J. Agric. Sci. 44: 136-141.

Maras M, Jelka SV, Branka J and Vladimir M (2008) The efficiency of AFLP and SSR markers in genetic diversity estimation and gene pool classification of common bean (*Phaseolus vulgaris* L.) Acta Agric. Slovenica 91: 87-96.

Mather DE, Hyes PM, Chalmers KJ, Eglinton J, Matus I, Richardson K, Von Zitzewitz J, Marquez-Cedillo L, Hearnden P and Pal N (2004) Use of SSR marker data to study linkage disequilibrium and population structure in *Hordeum vulgare*: Prospects for association mapping in barley. In: International barley genetics symposium, Brno, Czech Republic, 20-26 June 2004, pp. 302-307

Mullis K, Faloona F, Scharf S, Saiki R, Horn G and Erlich H (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb. Symp. Quant. Biol. 51: 263-273.

Newbury H and Ford-Lloyd BV (1997) Estimation of genetic diversity. In: Plant genetic conservation-the in situ approach, Maxted N, Ford-Lloyd BV, Hawkes JG. Eds., Chapman and Hall, London, UK, pp. 192-206.

Pallottini L, Garcia E, Kami J, Barcaccia G and Gepts P (2004) The genetic anatomy of a patented yellow bean. Crop Sci. 44: 968-977.

Petit E, Balloux F and Goudet J (2001) Sex-biased dispersal in a migratory bat: A characterization using sex-specific demographic parameters. Evolution 55: 635-640.

Poehlman JM (1991) The mungbean. Published in India by Mohan Primlani for Oxford & IBH Publishing Co, pp. 195-222.

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

Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Queller DC, Strassmann JE and Hughes CR (1993) Microsatellites and kinship. Trends in Ecol. Evol. 8: 285-288.

Roussel V, Koenig J, Beckert M and Balfourier F (2004) Molecular diversity in French bread wheat accessions related to temporal trends and breeding programmes. Theor. Appl. Genet. 108: 920-930.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nature Biotech. 18: 233-234.

Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucl.eic Acids Res. 17: 6463-6471.

Tomooka N, Lairungreang C, Nakeeraks P, Egawa Y and Thavarasook C (1992) Center of genetic diversity and dissemination pathways in mungbean deduced from seed protein electrophoresis. Theor. Appl. Genet. 83: 289-293.

Vos P, Hogers R, Bleeker M, Reijans M, Van De, Lee T, Hornes M, Frijters A, Pot J, Peleman J and Kuiper M (1995) AFLP: A new technique for DNA fingerprinting. Nucl.eic Acids Res. 23: 4407-4414.

Weber JL and May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Human Genet. 44: 388.

Welsh J and McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucl.eic Acids Res. 18: 7213-7218.

Westman AL and Kresovich S (1997) Use of molecular marker techniques for description of plant genetic variation. In: Callow JA, Ford-Lloyd BV, Newbury HJ (eds) Plant Genetic Resources: Conservation and Use CAB Int., Wallingford, UK, pp. 9-48.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA and Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl.eic Acids Res. 18: 6531-6535.

Yifru T, Hammer K, Huang XQ and Roder MS (2006) Regional patterns of microsatellite diversity in Ethiopian tetraploid wheat accessions. Plant Breed. 125: 125-130.