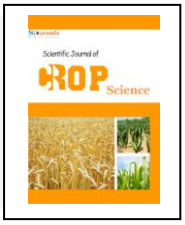
Sjournals
Scientific Journals
Scientific Journal of Crop Science (2012) 1(4) 63--75
ISSN 2322-1690

Contents lists available at Sjournals
Scientific Journal of CROP Science
Journal homepage: www.Sjournals.com



Original article

Comparing genetic diversity and population structure of common beans grown in Kyrgyzstan using microsatellites

S. Hegay[a,*], M. Geleta[b], T. Bryngelsson[b], L. Gustavsson[b], H. P. Hovmalm[b], R. Ortiz[b]
[a]*Department of Agronomy, Kyrgyz National Agrarian University, Kyrgyzstan.*
[b]*Department of Plant Breeding and Biotechnology, Swedish University of Agricultural Sciences, Sweden.*

[*]Corresponding author; Department of Agronomy, Kyrgyz National Agrarian University, Kyrgyzstan.

ARTICLE INFO

Article history:
Received 25 October 2012
Accepted 25 November 2012
Available online 25 Nov 2012

Keywords:
Common bean
Genetic diversity
Simple sequence repeats
Structure
Phaseolus vulgaris

ABSTRACT

Common bean (*Phaseolus vulgaris* L.) is an important export crop in Kyrgyzstan. The aim of this study was to assess the extent of genetic diversity, determine the population structure, and relate to the main gene pools grown in Kyrgyzstan. Twenty-eight common bean accessions (including five Kyrgyz cultivars, and main references from the Mesoamerica and South America) were evaluated with microsatellites. Nine polymorphic microsatellites were used to estimate genetic diversity and heterozygosity. The number of alleles per microsatellite locus ranged from 2 to 4 and there were a total of 24 alleles. The observed heterozygosity of each accession over all loci ranged from 0 to 0.11 (with an average of 0.01), while the expected average heterozygosity was 0.05, which could reflect the self-pollinating breeding behavior of common beans. The analysis of molecular variance further revealed that 94.71% of the total variation was accounted by differences among accessions (F_{st} =0.947; $p<0.001$). Cluster analysis grouped accessions in two gene pools: 16 belong to the Andean and 12 to the Mesoamerican gene pool. The microsatellites separated accessions in Mesoamerican gene pool from Durango and Jalisco races, which were grouped together. We also observed that the most divergent accessions were the Kyrgyz cultivars, which may be related to the Mesoamerican races. Andean accessions were less diverse than Mesoamerican accessions in this study. This research confirms the ability of microsatellites to differentiate common bean accessions, even using a small sample

size, and to be able to assign modern cultivars to their gene pools or races.

1. Introduction

Common beans (*Phaseolus vulgaris* L.) are among the most important food legumes in human diets. Dry beans production was estimated as 23 million t in 2010; i.e., an increase of 0.5 million in the last decade (FAO, 2010). This legume crop is an annual, self-pollinated diploid (2n = 2x = 22) plant with cleistogamy. Cross-pollination may happen by honey bees as vectors (Ferreira *et al*., 2000).

Common beans originated in Meso and South America (Vavilov, 1926), which are regarded as the major centers of diversity for this crop, which relate to the two major gene pools: Mesoamerican (MA) and Andean (A). The two gene pools are further divided into races. The Mesoamerican gene pool, which spreads from Mexico throughout Central America, includes Mesoamerica (M), Durango (D), and Jalisco (J) races, whereas the Andean gene pool consists of Nueva Granada (NG), Chile (C) and Peru (P) races (Singh, 1988, Singh *et al*., 1991b). The environments where Andean races are grown in a slightly cooler environment than the Mesoamerican races (Debouck, 1993). The common bean races in the two distinct gene pools are defined by their morphology, growth habit, color, shape and size of seeds, leaf shapes, bracteoles and their eco-geographical distribution (Singh and Urrea, 1990, Urrea and Singh, 1991). Small to medium size beans are a feature of Mesoamerican genepool whereas medium to large beans characterized the Andean gene pool.

There have been various diversity assessments using morphological descriptors (Evans, 1976, Kaplan, 1981), the seed protein phaseolin (Gepts and Bliss, 1985), allozymes (Gepts *et al*., 1986, Koenig and Gepts, 1989, Singh *et al*., 1991a), and DNA markers such as amplified fragment length polymorphism or AFLP (Beebe *et al*., 2001, Tohme *et al*., 1996), random amplified polymorphic DNA or RAPD (Beebe *et al*., 2000), microsatellites –also known as simple sequence repeats or SSR (Blair *et al*., 2006, Blair *et al*., 2011b, Zhang *et al*., 2008), and single nucleotide polymorphisms or SNP (Blair *et al.,* 2011a). Microsatellite markers (SSR) have been very useful for studying genetic diversity of common beans and to distinguish between the Mesoamerica and Andean gene pools and their respective races.

Common beans were likely introduced to Central Asia by the Soviets in the last century. The annual production of beans (67,000-70,000 t) has become an important activity for the country because 90% of it generates about US$ 43 million year^{-1} through export trade mainly with Bulgaria, Russia and Turkey. Smallholders are their main growers in Kyrgyzstan. After harvest, Kyrgyz farmers select beans according to their shape and color, and used them as seeds for the next growing season. These seeds are also exchanged among neighbors. Increasing our knowledge on the diversity and population structure of the beans grown in Kyrgyzstan will assist on conserving this genetic endowment and its appropriate use in bean breeding. The objective was to assess the extent of genetic diversity, determine population structure, and relate to the main gene pools of common beans grown in Kyrgyzstan with the aid of microsatellites.

2. Materials and methods

2.1. Plant materials and DNA extraction

We characterized with microsatellites 28 common bean accessions: five Kyrgyz cultivars, one wild type from Armenia, and the remaining, which were kindly provided by Michigan State University (East Lansing), and United States Department of Agriculture (Pullman), were used as reference sets of the two gene pools. The first true leaf (7-9 days old) of 10 seedlings of each accession were taken randomly for DNA extraction (Warwick and Gugel, 2003). The quality of DNA was checked by electrophoresis in a 1.5% agarose gel containing ethidium bromide. The DNA concentration was adjusted using a Nanodrop® ND-1000 spectrophotometer (Saveen Werner, Sweden).

2.2. PCR pre-amplification and electrophoresis

After optimizing the protocol, we found that of the 11 microsatellite provided by Invitrogen Life Technologies (USA), only 9 were polymorphic primers (Table 1). Their linkage groups follows Blair *et al*. (2003) and Hanai *et al*. (2010). The PCR reaction was performed in a total volume of 25 µl containing 1× PCR buffer with 1.5 mM $MgCl_2$, 0.2 mM of each dNTPs, 7.5 pmol μl^{-1} each of the forward and reverse primers (Sigma-Aldrich AB, Sweden), 0.5 Units of Tag polymerase (Saveen Werner AB, Sweden) and 10 ng μl^{-1} of DNA. The negative control was included (without DNA in a reaction) to prevent contaminant DNA. A 50 bp DNA ladder (GeneRulerTM, Fermentas Life Sciences) was used as molecular size range.

Pre-amplification was performed in 96-well plates on a Gene Amp PCR system 9700 (Applied Biosystems Inc, USA) for optimized each primer. The touchdown PCR program consists of following steps: denaturation for 3 min at 95°C followed by 10 cycles of 30 s denaturing at 94°C, 30 sec annealing at 70°C reducing by 1°C every cycle, and 45 s extension at 72°C. This step was followed by 30 cycles of 30 s at 94°C, 30 s at 60°C and the last cycle was followed 20 min of product extension at 72°C. Amplified products were stored in a refrigerator until carrying the electrophoresis.

Table 1
Microsatellite provided by invitrogen life technologies (USA).

Primer name	Linkage group	Repeat motif	Allele number	Fluorescent label	Primer sequence
BMd9[a]	b 04	CATG	3	VIC	F: TATGACACCACTGGCCATACA
					R: CACTGCGACATGAGAGAAAGA
BMd16[a]	b 04	CATG	1	VIC	F: ATGACACCACTGGCCATACA
					R: GCACTGCGACATGAGAGAAA
BMd17[a]	b 02	CGCCAC	3	6FAM	F: GTTAGATCCCGCCCAATAGTC
					R: AGATAGGAAGGGCGTGGTTT
BMd18[a]	b 02	TGAA	2	NED	F: AAAGTTGGACGCACTGTGATT
					R: TCGTGAGGTAGGAGTTTGGTG
BMd33[a]	b11	ATT	4	6FAM	F: TACGCTGTGATGCATGGTTT
					R: CCTGAAAGTGCAGAGTGGTG
BMd53[a]	b 05	GTA	2	NED	F: TGCTGACCAAGGAAATTCAG
					R: GGAGGAGGCTTAAGCACAAA
BMd54[a]		CCT	3	6FAM	F: GGCTCCACCATCGACTACTG
					R: GAATGAGGGCGCTAAGATCA
PVM075[b]	b 09	GAT	2	6FAM	F: ATTGGAAGGGGGATGAACCT
					R: TAGGAGAGTGCCCAGTGCTT
PVM145[b]		TCC	1	VIC	F: TTTCAGTTCGGGATTGTTCC
	b 05				R: ATTGGTGGAGGTGGGAGAG
PVM148[b]	b 03	CCA	3	NED	F: ACCTCAAAACCCACCACAAA
					R: GAAGTGCTCCCAGATGAAGG
PVM152[b]		TTG	2	HEX	F: ATTTTGGAGCGAAACAGCAT
					R: GAGAACCTCGTCGTCGTCTT

Primers-pairs used to amplify the microsatellite loci and their diversity. The names given are after [a]Blair *et al.* (2003), [b]Hanai *et al.* (2010)

About 5µl of each PCR product was separated by electrophoresis using 1.5% agarose gel and visualized using ethidium bromide. The confirmed amplified PCR products were further run on polyacrylamide gels (CleanGel 10% 52S; ETC Electrophorase-technic, Germany) supplied with rehydration buffer (Tris-phosphate buffer pH 8.4 and Bromophenol Red) and electrode (Tris-Borate buffer pH 8.6). Five µl of sample loading buffer (20% sucrose, 10% ficoll, 0.05% bromophenol blue, 5 M urea and 1 mM EDTA) were mixed with 6 µl of each PCR product, and the mixture was loaded for running in a polyacrylamide gel. The procedure for rehydrating the gel was that used for horizontal Multiphor II Electrophoresis Unit (GE Healthcare Bio-Sciences AB, Sweden). A 50 bp DNA ladder (GeneRulerTM, Fermentas Life Sciences) was used to estimate the molecular size of the bands. To visualize the

bands the gel was silver-stained using the Hoefer Automated Gel stain (Pharmacia Biotech, USA) using the protocol recommended by the manufacturers. The selected forward primers were fluorescently 5´ labeled with 6FAM™, VIC™, HEX™, NED™ fluorescent dyes. The reverse primer were PIG-tailed with "*GCTTCT*" to prevent mismatch on the template strand of a single nucleotide by *Taq* polymerase to the PCR product, as reported Ballard *et al.* (2002). The PCR amplification with the labeled primers was done as described above for the pre-amplification stage. The PCR amplified products were multiplexed into panels as indicated by Geleta *et al.* (2012). The PCR products were analyzed using ABI Prism 3730 DNA Analyzer (Applied Biosystems) at the Genomics Core Facility of the University of Gothenburg in Sweden.

2.3. Data sequence analysis

Peak identification and fragment allele sizing were analyzed using GeneMarker® V2.2.0 software (SoftGenetics, LLS, State College, Pennsylvania) based on the internal Genescan-500 LIZ size standard. The peak scores were based on single or numerous ratios and peak shapes for a co-dominant locus for each individual from 28 populations. Observed results were manually recorded.

Number and percentage of polymorphic loci, expected and observed heterozygosities were calculated using POPGENE version 1.31 (Yeh and Boyle, 1997). The average genetic diversity of common beans samples was estimated based on Nei's gene diversity (Nei, 1978). Roger's standard genetic and cluster analysis and bootstrapping were conducted using Free Tree-Freeware program (Pavlicek *et al.*, 1999). TreeView (32) 1.6.6 program (Page, 1996) was used to display the trees. Software's STRUCTURE version.2.3.4, DISTRACT version1.1, STRUCTURE HARVESTER (Earl and vonHoldt, 2012, Pritchard *et al.,* 2000, Rosenberg, 2004) were used to denote population structure and to visualize with distinct colors groups and subgroups. The NTSYSpc program (Rohlf, 2000) was used to perform the Jaccard´s similarity matrix and principal co-ordinate analysis. The overall genetic diversity of common bean accessions was estimated through the analysis of molecular variance (AMOVA) using Arlequin 3.5 (Excoffier and Lischer, 2010).

3. Results

3.1. Microsatellite polymorphism

Nine out of 11 microsatellites were polymorphic with a varying degree. *BMd16* and *PVM145* were monomorphic loci. There were 24 distinct alleles across the polymorphic loci among the 28 common bean accessions (Table 2). The maximum number of alleles (NA) per locus was 4 (for *BMd33* locus). There were on average 2.67 alleles per polymorphic locus. The fragment size of the alleles ranged from 99 bp (*BMd17 and BMd33*) to 230 bp (*PVM075*). The observed heterozygosity at each polymorphic locus (Ho) ranged from 0 (*BMd18, PVM075, PVM152*) to 0.0087 (*BMd9*), which may ensued from both purifying selection done by common bean breeders and self-pollination (as corroborated by the mean Ho, which was 0.0042). The Ho of *BMd9*, *BMd33*, *BMd54* and *PVM148* was higher than the average across microsatellite loci.

3.2. Genetic diversity

The polymorphic information content (PIC) was 0.427 (Table 3). The percentage of polymorphic loci (P) ranged from 0 to 66.7%. PI 527537 from Burundi showed 66.7% P because it appears to be a seed mixture derived from both gene pools. Accessions PI 337090, PI 527537, PI 208776, PI 543043 and PI 416043 accounted for 44% or higher polymorphic loci. The overall mean estimates of percent polymorphic loci (%P) were 14.7, while Shannon´s diversity index (I) and Nei's gene diversity were 0.0685 and 0.0435, respectively.

The mean number of alleles (NA) per accession over all loci ranged from 1 to 1.7 and the overall mean number of alleles per accession was 1.1 (Table 3). PI 527537 had the highest number of alleles across loci. The observed heterozygosity of each accession over all loci (Ho) ranged from 0 to 0.11 and averaged 0.01. There were 22 (out of 28) accessions whose microsatellites were homozygous across all loci. The expected heterozygosity of each accession over all loci (He) ranged from 0 to 0.30 with an average of 0.05.

3.3. Genetic variation and genetic distances

Analysis of molecular variance (AMOVA) was used to estimate genetic variation among populations, races, and gene pools (Table 4). Variation among and within accessions was highly significant ($P < 0.001$). There was a

significant genetic variation between the races (F_{ST} = 0.9388). Variation between the two gene pools was 76.71% ($P < 0.001$).

The Nei's standard genetic distance between pairs ranged from 0 to 2.349. There was low genetic similarity for 25 accessions pairs whose Nei´s standard genetic distances vary between 2.092 and 2.349 (data not shown). The most divergent accessions were the Kyrgyz cultivars Kytayanka and Lopatka. A high genetic similarity (i.e., Nei´s standard genetic distances > 0.005) was noted among 47 accessions pairs, although they differ in their country of origin.

3.4. Cluster analysis, principal coordinate analysis and population structure

The unweighted pair group method with arithmetic mean (UPGMA, Fig. 1) using Roger´s genetic distance, and the neighbor-joining analysis based on the Jaccard´s similarity coefficient (Fig. 2) with 100% bootstrap support gives two main clusters, each belonging to the main common bean gene pools. The accessions in the UPGMA dendrogram were further separated into two Mesoamerican sub-clusters (Ia and Ib) and Andean sub-clusters (IIa and IIb).

The principal coordinate analysis (PCoA) split the accessions in two groups: 16 belonging to the Andean gene pool (Group II) and 12 to the Mesoamerican gene pool (Group I). The first two PCoA explain 79.7% of total variation with the first principal co-ordinate accounting for 67.9% of the total variation (Fig. 3). The Mesoamerica race belongs to Group 1b and was separated from the Durango and Jalisco races, which were together in Group 1a.

Population structure of accessions was characterized with STRUCTURE version 2.3.4 (Fig. 4). The highest peak was K=2 clearly separated the Andean and Mesoamerican gene pools. The next peak K=3 divided the Mesoamerican gene pool into two subgroups: (1) Mesoamerica race and (2) Durango and Jalisco races that were grouped together (Fig. 4).

Table 2
Microsatellite loci and their diversity among 28 common bean accessions.

Primer name	Observed allele size (bp)	Shannon index	Observed hetero-zygosity	Expected hetero-zygosity	Nei´s distance	F_{st}
BMd9	136, 139, 151	0.2458	0.0087	0.1155	0.1153	0.5398
BMd16	138					
BMd17	99, 105, 117	1.0020	0.0036	0.6104	0.6093	0.9234
BMd18	160, 161	0.6908	0.0000	0.4985	0.4976	0.9395
BMd33	99, 102, 108, 111	1.0814	0.0073	0.6100	0.6088	0.8393
BMd53	109, 112	0.6800	0.0036	0.4878	0.4869	0.9372
BMd54	151, 159, 162	0.4686	0.0076	0.2795	0.2790	0.8770
PVM075	227, 230	0.6761	0.0000	0.4839	0.4831	0.9435
PVM145	209					
PVM148	188, 194, 197	0.7595	0.0073	0.5069	0.5060	0.8847
PVM152	203, 206	0.4255	0.0000	0.2577	0.2573	0.9472
Mean		0.6700	0.0042	0.4278	0.4270	0.8981
Standard deviation		0.2665	0.0036	0.1706	0.1703	

Table 3
Percentage of polymorphism (%P), heterozygosity and diversity measurements of common bean accessions and Kyrgyz cultivars

Accession	Country	%P	Mean number of allele
W6 23905	Armenia	0.0	1.0
PI 638895	Argentina	11.1	1.1
W6 9748	Australia	11.1	1.1
PI 337090	Brazil	44.4	1.4
W6 9655	Bulgaria	0.0	1.0
PI 527537	Burundi	66.7	1.7
PI 557467	Chile	0.0	1.0
PI 640943	China	11.1	1.1
PI 207262	Colombia	0.0	1.0
PI 415928	Ecuador	11.1	1.1
PI 415994	Former Soviet Union	0.0	1.0
TO	France	22.2	1.2
PI 451886	Guatemala	11.1	1.1
PI 361240	India	11.1	1.1
PI 289531	Italy	11.1	1.1
PI 577694	Uzbekistan	22.2	1.2
Bokser	Kyrgyzstan	11.1	1.1
Ryabaya	Kyrgyzstan	0.0	1.0
Kytayanka	Kyrgyzstan	0.0	1.0
Lopatka	Kyrgyzstan	22.2	1.2
Yubka	Kyrgyzstan	0.0	1.0
AB 136	Mexico	11.1	1.1
PI 208776	Nicaragua	44.4	1.4
PI 549795	S.Africa	0.0	1.0
PI 543043	Spain	44.4	1.4
PI 181954	Syria	0.0	1.0
PI 618815	USA	0.0	1.0
PI 416043	Iran	44.4	1.4
Mean		14.7	1.1
Standard deviation		17.5	0.1

Table 3
(continue)

Accession	Observed hetero-zygosity	Expected hetero-zygosity	Shannon index	Nei´s gene index	Gene pool	Race
W6 23905	0.00	0.00	0.0000	0.0000		
PI 638895	0.02	0.03	0.0556	0.0356		
W6 9748	0.00	0.05	0.0679	0.0467	A[d]	NG[d]
PI 337090	0.03	0.07	0.1304	0.0706		
W6 9655	0.00	0.00	0.0000	0.0000		
PI 527537	0.00	0.30	0.4477	0.3015		
PI 557467	0.00	0.00	0.0000	0.0000	A[a]	C[a]
PI 640943	0.00	0.02	0.0361	0.0200		
PI 207262	0.00	0.00	0.0000	0.0000		
PI 415928	0.00	0.05	0.0770	0.0556		
PI 415994	0.00	0.00	0.0000	0.0000		
TO	0.02	0.06	0.0986	0.0617	MA[b]	D[b]
PI 451886	0.00	0.03	0.0556	0.0356		
PI 361240	0.00	0.02	0.0361	0.0200		
PI 289531	1.11	1.11	0.0221	0.0106		
PI 577694	0.00	0.09	0.1357	0.0933		
Bokser	0.10	0.10	0.0221	0.0106		
Ryabaya	0.00	0.00	0.0000	0.0000		
Kytayanka	0.00	0.00	0.0000	0.0000		
Lopatka	0.00	0.06	0.0917	0.0556		
Yubka	0.00	0.00	0.0000	0.0000		
AB 136	0.02	0.04	0.0577	0.0374	MA[b]	Me[b]
PI 208776	0.01	0.14	0.2039	0.1295	MA[f]	Me[f]
PI 549795	0.00	0.00	0.0000	0.0000		
PI 543043	0.00	0.08	0.1445	0.0800	A[e]	
PI 181954	0.00	0.00	0.0000	0.0000	MA[c]	
PI 618815	0.00	0.00	0.0000	0.0000		
PI 416043	0.00	0.16	0.2379	0.1562		
Mean	0.01	0.05	0.0685	0.0435		
Standard deviation	0.02	0.06	0.095	0.062		

Gene pool and race are given to some accessions according to [a] Singh and Teran, (1995), [b] Pastor Corrales *et al.* (1995), [c] Johnson and Gepts (1994), [d] Strausbaugh *et al.* (1999), [e] Alves-Santos *et al.* (2002), [f] Kwak *et al.* (2009).

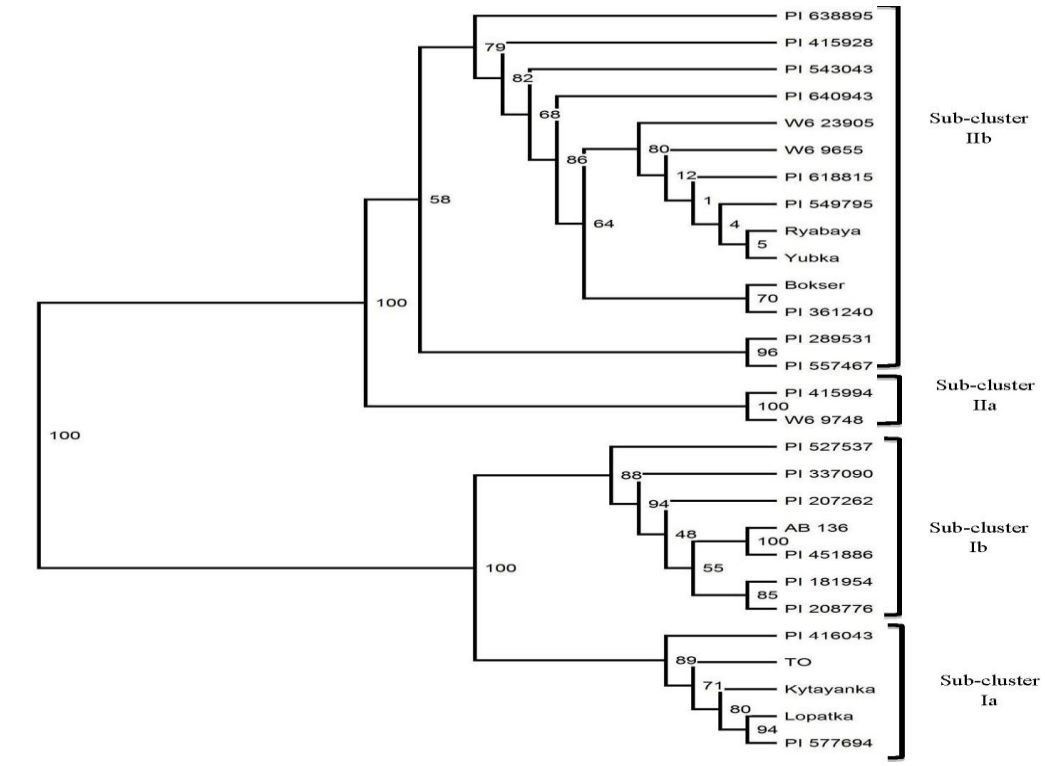


Fig. 1. Unweighted pair group method with arithmetic mean dendrogram based on Rogers' genetic distance for 28 common beans accessions. The value branches are the bootstrap value generated by 1000 resampling using the FreeTree software.

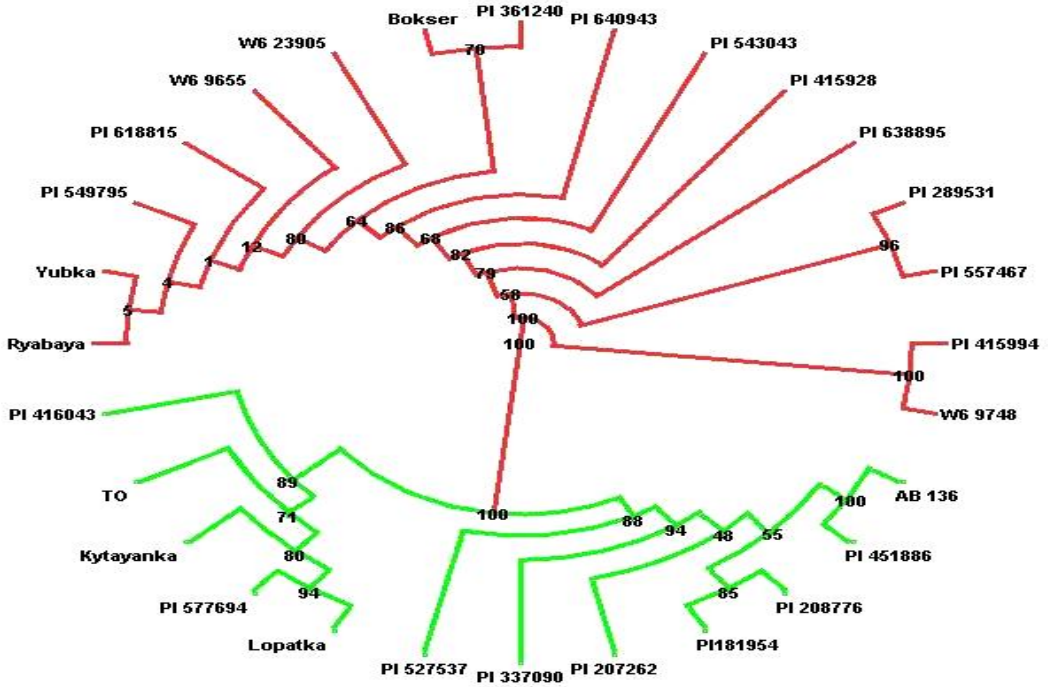


Fig. 2. Neighbor-joining tree based on microsatellite data using Rogers' genetic distance for 28 common beans population. The value branches are the bootstrap value generated by 1000 resampling in the FreeTree program. Accessions in red branching are from the Andean gene pool while Mesoamerican accessions are in the green branching.

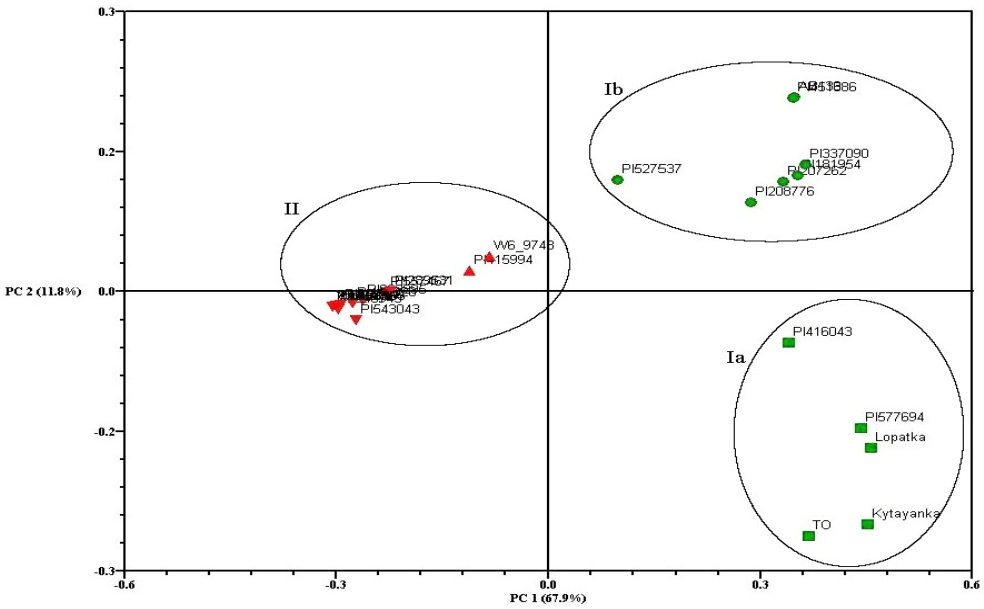


Fig. 3. Two principal coordinates of the principal co-ordinate analysis based on a microsatellite genetic similarity matrix for 28 common beans accessions. The plot was generated from Nei's similarity matrix using NTSYSpc software. Group Ia and Ib include the Mesoamerican gene pool while the Group II includes the Andean gene pool. Durango and Jalisco race accessions are in Group Ia while Mesoamerica race accessions are in Group Ib.

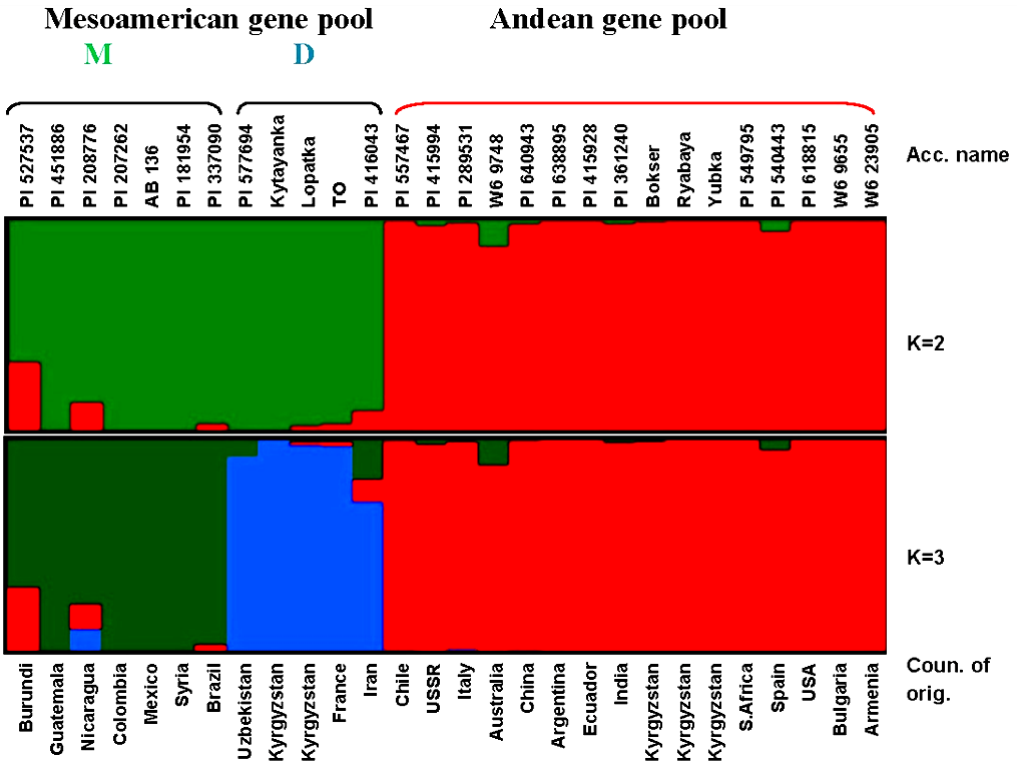


Fig. 4. Results of STUCTURE analysis at K = 2 to K = 3 for 28 common bean accessions. Their country of origin indicated at the bottom of the chart. The colors indicate the sub-groups (races) based on Andean and Mesoamerican gene pools at structure analysis K = 3. The letters M and D are for the Mesoamerica and Durango-Jalisco race grouping, respectively.

4. Discussion

4.1. Allelic variation

This genetic diversity study provides for the first time insights on common bean cultivars grown by farmers in Kyrgyzstan. The bean cultivars grown in Kyrgyzstan appear to be closer to the Nueva Granada race from the Andean gene pool and to the Durango-Jalisco grouping from the Mesoamerica gene pool as revealed by the clustering using microsatellite data. Two of these cultivars (Kytayanka and Lopatka) were the most distinct among the accessions included in this study. Nonetheless, the accessions included in our research had a narrow allele range and low allelic diversity.

The total allelic diversity and allelic richness observed in the 28 accessions included in our study was smaller that the diversity noted by Diaz *et al*. (2011) among 92 landraces from Colombia when using 45 microsatellites. They found a total of 436 alleles. Similarly, Blair *et al.* (2009) observed a total 679 alleles in 604 common bean accessions from Africa, America and Europe using 36 microsatellites. Blair *et al*. (2011b) also detected a total 204 allele based on 32 microsatellites in 101 accessions from Mexico, while Yu *et al*. (1999) noticed seven microsatellites (with 2-10 alleles per locus, and an average of 4.4 per locus) in 12 common bean breeding lines.

The observed heterozygosity in our research was overall low (0.05), which could reflect the self-pollinating breeding behavior of common beans, as well as selection of homozygous lines for cultivar release, and the nature of the sample (pure lines) made available from the US gene banks. The observed heterozygosity in our study was higher in the Andean gene pool (0.076) than in the Mesoamerican gene pool (0.006), but the variance component of the AMOVA (Table 4) for Mesoamerican gene pool (0.747) was higher than the Andean gene pool (0.363). This result was not surprising since Duarte *et al*. (1999) and Blair *et al*. (2010) also noted higher polymorphism for Mesoamerican gene pool vis-à-vis the Andean gene pool using fluorescent microsatellites.

4.2. Genetic relationships

Accessions were grouped according to Andean and Mesoamerican gene pools and gene diversity were estimated among groups. The longer repeat of nucleotides often contains more polymorphic alleles than shorter repeats. In our study, the observed genetic variation (as revealed by the AMOVA) was higher among accessions than within accessions. This result could ensue from the inbred nature of common beans, effective barriers for gene flow among populations (in spite of seed exchange between farmers), and human selection of pure lines for use as cultivars. The genetic variation observed in the American centers of diversity was higher than elsewhere, which was also noticed by Blair *et al*. (2009).

The structure, cluster analyses, and PCoA defined two main groups, which correspond to Mesoamerican and Andean accessions origins (Figs. 1, 2, 3 and 4). The grouping of common beans at STRUCTURE K=3 further divided the Mesoamerican gene pool in two sub-groups (as also noted in the PCoA): Mesoamerica race and together the Durango plus Jalisco races, which are from Mexico. Diaz and Blair (2006) also found that accessions from races Durango and Jalisco were grouped together at K=3 because of their geographic origin. Singh *et al.* (1991a) and Beebe *et al*. (2000) were able to differentiate Durango and Jalisco races using morphology and RAPD markers, respectively.

The diversity indicators used in our research depend on many factors including the method of sampling used, number and size ranges of loci characterized, and marker distribution on the genome from gene coding or non-gene coding regions. Nonetheless, our study shows the ability of microsatellites to discriminate among common bean accessions, even using a small sample size, and to putatively assign modern cultivars to their gene pools or races.

Acknowledgements

This work is part of the PhD thesis of the first author. The study was financing by the Swedish International Development Agency (Sida). We thank the Michigan State University (East Lansing) and United States Department of Agriculture (Pullman) for providing common bean germplasm used in this study. We are grateful to Annia Zborowska at Swedish University of Sciences (SLU), Alnarp for her laboratory assistance.

Table 4

Analysis of molecular variance (AMOVA) for common bean accessions based on microsatellite polymorphism: (A) for all 28 accessions, (B) only for the 12 Mesoamerican accessions, (C) only for the 16 Andean accessions, (D) grouping the accessions according to two known gene pools, (E) grouping the accessions according to known races, (F) grouping the accessions according to primary and secondary centers of diversity as per Blair *et al.* (2009).

Groups	Sources of variation	Degrees of freedom	Variance components	Variation (%)	Fixation index (F_{ST})	Probability
(A) All accessions	Among accessions	27	Va=1.62092	94.71	0.94706	Va and F_{ST} =0.00000
	Within accessions	532	Vb=0.09060	5.29		
	Total	559	1.71152			
(B) Mesoamerican accessions	Among accessions	11	Va=0.74701	78.97	0.78972	Va and F_{ST}=0.0000
	Within accessions	228	Vb=0.198	21.03		
	Total	239	0.94591			
(C) Andean accessions	Among accessions	15	Va=0.36275	83.15	0.83148	Va and F_{ST}=0.0000
	Within accessions	304	Vb=0.07352	16.85		
	Total	319	0.43627			
(D) Accessions as per gene pools	Among gene pools	1	Va= 2.14964	76.71	0.95459	Vc and F_{ST}=0.0000
	Among accessions within gene pools	26	Vb=0.52529	18.76	0.80498	Vb and F_{SC}=0.0000
	Within accessions	532	Vc= 0.12726	4.54	0.76713	Va and F_{CT}=0.0000
	Total	559	2.80219			
(E) Accessions as per races	Among races	4	Va =1.52836	73.48	0.93882	Vc and F_{ST}=0.0000
	Among accessions within races	23	Vb=0.4244	20.40	0.76932	Vb and F_{SC}=0.0000
	Within accessions	532	Vc=0.12726	6.12	0.73478	Vc and F_{CT}=0.0000
	Total	559	2.08002			
(F) Primary center versus secondary center of diversity	Among groups	1	Va=0.08460	4.72	0.92894	Vc and F_{ST}=0.0000
	Among accessions within groups	26	Vb=1.57890	88.17	0.92541	Vb and F_{SC}=0.0000
	Within accessions	532	Vc=0.12726	7.11	0.04724	Va and F_{CT}=0.19062
	Total	559	1.79075			

References

Ballard, L., Adams, P., Bao, Y., Bartley, D., Bintzler, D., Kasch, L., Petukhova, L., Rosato, C., 2002. Strategies for genotyping: Effectiveness of tailing primers to increase accuracy in short tandem repeat determinations. Biomol. Tech. 13, 20-29.

Beebe, S., Rengifo, J., Gaitan, E., Duque, M.C., Tohme, J., 2001. Diversity and origin of Andean landraces of common bean. Crop Sci. 41, 854-862.

Beebe, S., Skroch, P.W., Tohme, J., Duque, M.C., Pedraza, F., Nienhuis, J., 2000. Structure of genetic diversity among common bean landraces of middle American origin based on correspondence analysis of RAPD. Crop Sci. 40, 264-273.

Blair, M.W., Cortes, A.J., Chavarro, M.C., 2011a. SNP marker diversity in common bean (Phaseolus vulgaris L.). Theo. App. Gen. 123, 827-845.

Blair, M.W., Diaz, L.M., Buendia, H.F., Duque, M.C., 2009. Genetic diversity, seed size associations and population structure of a core collection of common beans (Phaseolus vulgaris L.). Theo. App. Genet. 119, 955-972.

Blair, M.W., Chaves, A., Tofino, A., Calderon, J.F., Palacio, J.D., 2010. Extensive diversity and inter-genepool introgression in a worldwide collection of indeterminate snap bean accessions. Theor. App. Genet. 120, 1381-1391.

Blair, M.W., Giraldo, M.C., Buendia, H.F., Tovar, E., Duque, M.C., Beebe, S.E., 2006. Microsatellite marker diversity in common bean (Phaseolus vulgaris L.). Theo. App. Genet. 113, 100-109.

Blair, M.W., Diaz, L.M., Gill-Langarica, H.R., Rosales-Serna, R., Mayek-Perez, N., Acosta-Gallegos, J.A., 2011b. Genetic relatedness of Mexican common bean cultivars revealed by microsatellite markers. Crop Sci. 51, 2655-2667.

Blair, M.W., Pedraza, F., Buendia, H.F., Gaitan-Solis, E., Beebe, S.E., Gepts, P., Tohme, J., 2003. Development of a genome-wide anchored microsatellite map for common bean (Phaseolus vulgaris L.). Theo. App. Genet. 107, 1362-1374.

Debouck, D.G., Toro, O., Paredes, O.M., Johnson, W.C., Gepts, P., 1993. Genetic diversity and ecological distribution of Phaseolus vulgaris (Fabaceae) in northwestern South America. Econ. Bot. 47, 408-423.

Diaz, L.M., Blair, M.W., 2006. Race structure within the Mesoamerican gene pool of common bean (Phaseolus vulgaris L.) as determined by microsatellite markers. Theo. App. Genet. 114, 143-154.

Diaz, L.M., Buendia, H.F., Duque, M.C., Blair, M.W., 2011. Genetic diversity of Colombian landraces of common bean as detected through the use of silver-stained and fluorescently labelled microsatellites. Plant Genet. Res. 9, 86-96.

Duarte, J.M., dos Santos, J.B., Melo, L.C., 1999. Genetic divergence among common bean cultivars from different races based on RAPD markers. Genet. Mol. Bio. 22, 419-426.

Earl, D.A., vonHoldt, B.M., 2012. Structure harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conser. Genet. Res. 4, 359-361.

Evans, A.M., 1976. Beans. Phaseolus spp. In N.W. Simmonds (ed.) Evol. of crop plants. Longman, London, England. pp. 168-172.

Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Eco. Res.10, 564-567.

FAO, 2010. FAOSTAT. Food and Agriculture Organization of the United Nations, Rome, Italy. Accessed at http://www.fao.org/ on May 2012.

Ferreira, J.J., Alvarez, E., Fueyo, M.A., Roca, A., Giraldez, R., 2000. Determination of the outcrossing rate of Phaseolus vulgaris L. using seed protein markers. Euph.113, 259-263.

Geleta, M., Herrera, I., Monzón, A., Bryngelsson, T., 2012. Genetic diversity of Arabica coffee (Coffea arabica L.) in Nicaragua as estimated by simple sequence repeat markers. Sci. World J.pp. 1-11.

Gepts, P., Bliss, F.A., 1985. F1-hybrid weakness in the common bean - Differential geographic origin suggests two gene pools in cultivated bean germplasm. J.of Hered. 76, 447-450.

Gepts, P., Osborn, T.C., Rashka, K., Bliss, F.A., 1986. Phaseolin-protein variability in wild forms and landraces of the common bean (Phaseolus vulgaris). Evidence for multiple centers of domestication. Econ. Bot. 40, 451-468.

Hanai, L.R., Santini, L., Camargo, L.E.A., Fungaro, M.H.P., Gepts, P., Tsai, S.M., Vieira, M.L.C., 2010. Extension of the core map of common bean with EST-SSR, RGA, AFLP, and putative functional markers. Mol. Breed. 25, 25-45.

Kaplan, L.,1981. What is the origin of the common bean? Econ. Bot. 35, 240-254.

Koenig, R., Gepts, P., 1989. Allozyme diversity in wild Phaseolus vulgaris: further evidence for two major centers of genetic diversity. Theo. App. Gen. 78, 809-817.

Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genet. 89, 583-590.

Page, R.D.M., 1996. TREEVIEW. An application to display phylogenetic trees on personal computer. Comp. App. in the Biosci. 12, 357-358.

Pavlicek, A., Hrda, S., Flegr, J., 1999. FreeTree-freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap jackknife analysis of the tree robustness. Application in the RAPD analysis of genus Frenkelia. Folia Bio. 45, 97-99.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genet. 155, 945-959.

Rohlf, F.J., 2000. NTSYS-pc: Numerical taxonomy and multivariate analysis system. Setanket, New York.

Rosenberg, N.A., 2004. DISTRUCT: a program for the graphical display of population structure. Mol. Eco. Notes. 4, 137-138.

Singh, S.P., 1988. Gene pools in cultivated dry bean. Ann. Rep. Bean Imp. Coop. 31, 180-182.

Singh, S.P., Urrea, C.A., 1990. Variation for bracteoles and its association with races of common bean. Ann. Rep. Bean Imp. Coop. 33, 112.

Singh, S.P., Nodari, R., Gepts, P., 1991a. Genetic diversity in cultivated common bean. 1. Allozymes. Crop Sci. 31, 19-23.

Singh, S.P., Gepts, P., Debouck, D.G., 1991b. Races of common bean (Phaseolus vulgaris, Fabaceae). Econ. Bot. 45, 379-396.

Tohme, J., Gonzalez, D.O., Beebe, S., Duque, M.C., 1996. AFLP analysis of gene pools of a wild bean core collection. Crop Sci. 36, 1375-1384.

Urrea, C.A., Singh, S.P., 1991. Variation for leaflet shape in wild and cultivated landraces of common bean. Ann. Rep. Bean Imp. Coop. 34, 133.

Vavilov, N.I., 1926. Studies on the origin of cultivated plants. Bull. of Appl. Bot. and Plant Breed. (Leningrad). 16, 1-248.

Warwick, S.I., Gugel, R.K., 2003. Genetic variation in the Crambe abyssinica - C. hispanica - C. glabrata complex. Genet. Res. and Crop Evol. 50, 291-305.

Yeh, F.C., Boyle, T.J.B., 1997. Population genetic analysis of codominant and dominant markers and quantitative traits. Belg. J. of Bot. 129, 157.

Yu, K.F., Park, S.J., Poysa, V., 1999. Abundance and variation of microsatellite DNA sequences in beans (Phaseolus and Vigna). Genome. 42, 27-34.

Zhang, X.Y., Blair, M.W., Wang, S.M., 2008. Genetic diversity of Chinese common bean (Phaseolus vulgaris L.) landraces assessed with simple sequence repeat markers. Theo. App. Genet. 117, 629-640.