

Full Length Research Paper

# The genetic diversity and population structure of common bean (*Phaseolus vulgaris* L) germplasm in Uganda

Dennis Okii<sup>1,2\*</sup>, Phinehas Tukamuhabwa<sup>1</sup>, James Kami<sup>3</sup>, Annet Namayanja<sup>2</sup>, Pamela Paparu<sup>2</sup>, Michael Ugen<sup>2</sup> and Paul Gepts<sup>3</sup>

<sup>1</sup>Makerere University, College of Agricultural and Environmental Sciences, Department of Crop production, P.O Box 7062, Kampala, Uganda.

<sup>2</sup>National Crops Resources Research Institute-Namulonge, P.O. Box 7084, Kampala, Uganda.

<sup>3</sup>University of California, Department of Plant Sciences/MS1, Section of Crop and Ecosystem Sciences, 1 Shields Avenue, Davis, CA 95616-8780, USA.

Received 15 May, 2014; Accepted 20 June, 2014

The knowledge and understanding of the genetic variability of common bean (*Phaseolus vulgaris* L.) germplasm is important for the implementation of measures addressed to their utilizations and conservation. The objective of this study was to characterize common bean in Uganda using polymorphic molecular markers for use in hybridization and variety development. Genomic DNA was extracted from plants at the first trifoliolate leaf stage growing in pots using the modified cetyltrimethylammonium bromide (CTAB) method. The gene pool membership (Andean vs. Mesoamerican) for each accession was established with the phaseolin marker. Simple sequence repeat (SSR) alleles were separated by capillary electrophoresis that provided further information on the organization of genetic diversity. The Andean and Mesoamerican genotypes were present in similar frequencies (51 vs. 49%, respectively). All SSR markers tested were polymorphic with mean polymorphism information content (PIC) of 0.8. The model-based cluster analysis of SSR diversity in the STRUCTURE software found three sub populations (K3.1, K3.2 and K3.3) genetically differentiated with moderate Wrights fixation indices ( $F_{ST}$ ) values 0.14, 0.12 and 0.09, respectively and many cases of admixture. The STRUCTURE result was confirmed by principal coordinate analysis (PCoA) which also clustered beans in three groups. Most Andean genotypes were included in K3.1 and Mesoamerican genotypes belonged to the K3.2 and K3.3 subgroups. This study sets the stage for further analyses for agronomic traits such as yield, resistance to biotic and abiotic stresses and the need for germplasm conservation.

**Key words:** Phaseolin, Simple sequence repeat (SSR), hybridization, wrights fixation index ( $F_{ST}$ ), structure.

## INTRODUCTION

The domesticated *Phaseolus vulgaris* L. ( $2n = 2x = 22$ ) consists of two major gene pools, one originating among wild beans ranging from northern Mexico to Colombia (Mesoamerican gene pool) and the other descending

from wild beans distributed from southern Peru to north-western Argentina (Andean gene pool) Freyre et al., 1996). The common bean has a widespread distribution on many continents such as Mesoamerica, South

America, Europe and Africa. It reached Uganda presumably in the 18th century via the East African coast (Gepts and Bliss, 1988). Currently, the gene pool of the domesticated species is organized into four Mesoamerican and three Andean eco-geographical races based on morphological, agronomic and ecological grounds (Singh et al., 1991a; Beebe et al., 2001). In the Andean gene pool, the races are Nueva Granada, Peru, and Chile, while in the Mesoamerican gene pool, they are Durango, Guatemala, Jalisco, and Mesoamerica (Blair et al., 2006; Diaz and Blair, 2006). The distinction the Andean and Mesoamerican gene pools is achievable, with the *Phaseolin* - marker (Kami and Gepts, 1994; Burle, 2008).

Microsatellite markers (SSRs) have also been used in common bean to construct genetic reference maps (Yu et al., 2000; Blair et al., 2003), and evaluate intra-specific diversity (Gaitán-Solís et al., 2002). These markers were also employed to study the genetic structure in Andean and Mesoamerican races in common bean (Blair et al., 2006; Diaz and Blair, 2006; Kwak and Gepts, 2009) and population structure in 192 landraces from Ethiopia and Kenya (Asfaw et al., 2009). The use of SSRs to characterize common bean germplasm in this study is justified by their high informativeness, co-dominance, and wide distribution in the genome.

The National Crops Resources Research Institute (NaCRRI) at Namulonge holds 320 distinct accessions of common bean that are morphologically distinct but have little information documented on their genetic diversity and genetic potential. This hinders the utilization of these materials as parental sources in the different breeding programs and slows the process of designing appropriate conservation strategies. The continued adoption of elite bean varieties by farmers over time has replaced some landraces that are not continuously planted in the farmers' fields (Sekabembe, 2010) leading to genetic erosion. This erosion reduces alleles and genotype frequencies from the breeders' gene pool, narrowing variation and, hence, restricting the amount of adapted genetic diversity for future breeding. For example, a previous survey conducted in South-western Uganda, a region popular for large-seeded bean production (CIAT, 2005), found that 40% of farmers had stopped growing medium and large-seeded landraces in favor of newly introduced, small-seeded varieties that are tolerant to root-rot disease.

Pest and diseases also cause loss of bean cultivars in farmer's fields (Mukankusi, 2008). Genetic diversity is necessary for the rapid genetic improvement of crop species (Trethowan and Kazi, 2008) and its studies provide a major step towards enhancing the genetic potential of the bean germplasm. Thus, the objectives of

the study was to genotype and determine the level of population structuring of common bean in Uganda using microsatellite markers for use in present and future bean breeding schemes and conservation.

## MATERIALS AND METHODS

### Plant material

The study included 100 accessions (Appendix 1) representative of the phenotypic diversity of common bean in Uganda. The selection of the bean sample specifically relied on two traits: (i) plant type/growth habit and (ii) weights of 100 seeds of each accession. These traits have been mapped as part of the crop's domestication syndrome (Koinange et al., 1996). The sample included the released bean varieties in Uganda and the ones frequently used in breeding activities. The place of collection/origin was also considered in order to have representatives from the different agro-ecological zones in Uganda. Three bean seeds per variety were planted in plastic pots and monitored until the first trifoliolate leaf stage in the greenhouse facilities of the Department of Plant Sciences, University of California Davis (UC Davis), U.S.A.

### DNA extraction, gene pool identification and genotyping

Genomic DNA was extracted from each bean accession following procedures described in Doyle and Doyle (1990). The diluted DNA samples (30 to 40 ng/μl) were subjected to Polymerase Chain Reaction (PCR) amplification. The phaseolin marker and 22 fluorescent labeled microsatellites markers labeled according to Schuelke (2000) was used in succession to determine gene pools and genotyping accessions as follows. The gene pool identification involved PCRs set up to amplify specifically a region surrounding the 15-bp tandem direct repeat of the phaseolin gene family as in Kami et al. (1995) and Burle et al. (2010). The PCR was conducted in a MJ thermocycler in a total reaction volume of 25.2 μl (containing: 18 μl double distilled water, 10X Thermopol Reaction buffer (2.5 μl), 250 μM dNTPs (2.5 μl), 1U Taq DNA polymerase (0.2 μL), 10 μM phaseolin primer (1 μl), and 50 ng/μl DNA (1 μl)). The PCR products with loading dye were loaded in a 10% vertical polyacrylamide gel electrophoresis system using 0.5X TBE buffer, and run at 130 V for 2 h. Gene pool control genotypes (BAT93 for Mesoamerican and Jalo EEP558 for Andean) were always loaded with samples. Gels were stained in 10 μl in 10 mg/ml of ethidium bromide solution and photographed for future reference. The genotypes were quantitatively assigned to either Andean or Mesoamerican gene pool based on similarity of their band patterns to the gene pool controls (Plate 1).

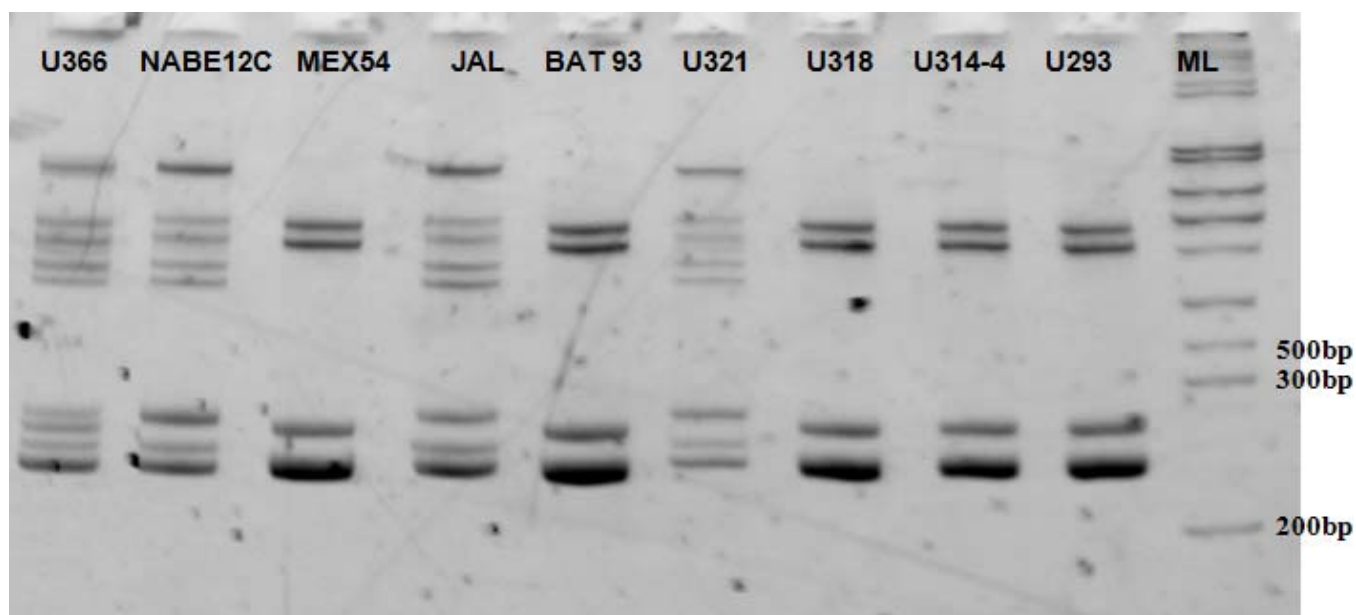
For microsatellite genotyping, two markers representing each linkage group were selected (Table 1) based on high polymorphism exhibited in common bean (Burle et al., 2010; Kwak and Gepts, 2009; Asfaw et al., 2009). Three fluorescent dyes: 6-FAM (Blue), VIC (Green), PET (Red) (Applied Biosystems Inc., USA) were attached to the 5' end of the M13 universal primer sequence (5'-TGT AAA ACG ACG GCC AGT-3'). The DNA was amplified separately for each primer pair in 96-well plates in a gradient thermal cycler (BioRad).

The PCR mix, contained in a total reaction volume of 20 μl, consisted of 1 μl of 30 ng DNA, 2.5 μl 10X PCR buffer (Roche), 2.5

\*Corresponding author. E-mail: dennis.okii@gmail.com. Tel: +256782167552. Fax: +256414531641.

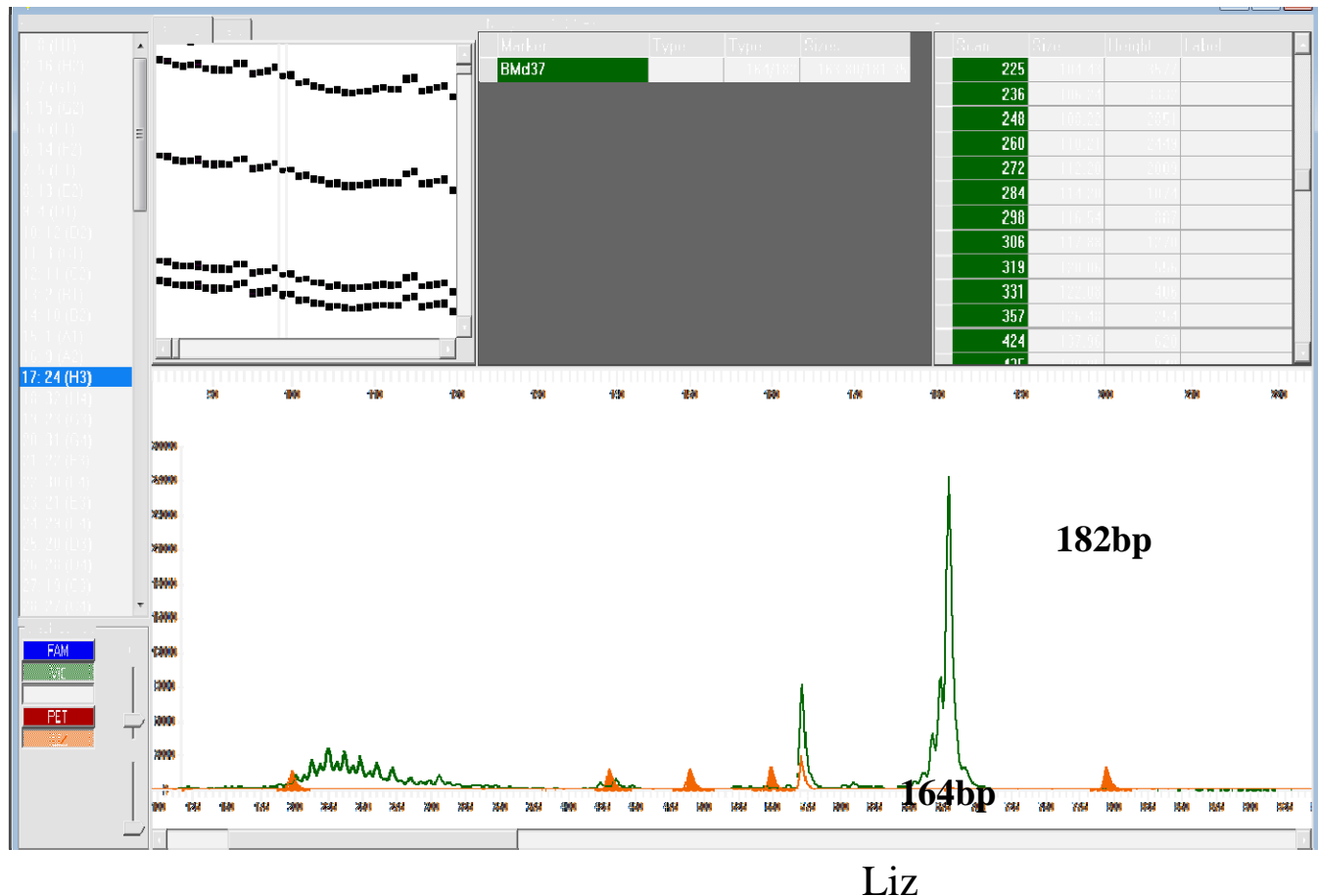
**Table 1.** Linkage group of microsatellite markers (SSRs), dye used in fluorescently labeling, SSR allele sizes and targeted sequence repeat for genotyping Common beans in Uganda.

Linkage group	Marker	Dye used	Allele size (bp)	SSR sequence
1	BM 200	PET	221	(AG)10
1	BMd10	PET	139	(GA)8
2	PVBR 107	VIC	150	(CT)16(GT)4
2	BM 156	PET	267	(CT)32
3	BM159	VIC	198	(CT)9(CA)8
3	BM197	PET	201	(GT)8
4	BM171	FAM	149	(GA)13
4	PVatgc002	FAM	144	(ATGC)4
5	BMd28	VIC	151	(GT)4
5	BM175	VIC	170	(AT)5(GA)19
6	BM187	PET	191	(CT)10(CT)14
6	BMd37	VIC	134	(AC)8
7	BM183	FAM	149	(TC)14
7	BM160	PET	211	(GA)15(GAA)5
8	BM189	FAM	114	(CT)13
8	BM211	VIC	186	(CT)16
9	PVat007	VIC	192	(AT)12
9	BM141	PET	218	(GA)29
10	GATS11B	PET	160	(CT)8
10	BMd42	VIC	149	(AT)5
11	BMd41	PET	250	(ATT)9
11	BM205b	VIC	137	(GT)11

**Plate 1.** 10% vertical polyacrylamide gel photo showing complex band patterns used to assign bean accessions to the Andean and Mesoamerican gene pool. ML is Hi-Lo DNA Marker ladder with size range 50 to 10,000 bp. The accessions JAL (JaloEEP558) and BAT93 are Andean and Mesoamerican control genotypes, respectively.

mmol each of the forward and reverse primers, 2 mmol dNTPs mix, 1 unit Taq DNA polymerase (Roche) and 2.5  $\mu$ l of double distilled water. The PCR program had an initial denaturing step of 94°C for

5 min followed by 34 cycles of 94° for 1 min, 56°C annealing for 1 min (same temperature for all markers) and 72°C for 1 min and a further primer extension at 72° for 20 min and 4°C infinite hold.



**Figure 1.** An electrophoregram genotype plot of the Strand 2.2.30 software, showing allele binning in sample 17:24(H3). In this case the most polymorphic marker from the study (BMd37) had two heterozygous alleles (peaks) with sizes 164/182 bp against the standard LIZ (orange) dye in the lower panel.

Prior to allele visualization by automated capillary electrophoresis, eight PCR products were randomly picked for each marker and run on 2% agarose gels to check for amplification.

### Capillary electrophoresis

The preparations of PCR products for capillary electrophoresis involved markers co-loaded or multiplexed in non-overlapping panels based on expected allele size (<http://phaseolusgenes.bioinformatics.ucdavis.edu/>) labeled with different dyes. Multiplex panels had two to three markers with at least 30 bp size differences to avoid background stuttering of allele peaks for the different dyes during allele binning. A master mix of 10 µl consisting of 1 µl each diluted PCR product, 5.7 µl formamide (Hi-Di) and 0.3 µl GeneScan-500 LIZ size standard (Applied Biosystems Inc., USA) was prepared in optical 96-well MicroAmp plates (Applied Biosystems Inc, USA) for electrophoresis. Fragment separation was performed using the ABI PRISM 3730 genetic analyzer instrument (Applied Biosystems, USA) at the Veterinary Genetics lab at UC Davis.

### Allele calling

The allele fragments were determined in strand 2.2.30 software

(Mellissa et al., 2000) for peak detection and fragment size matching to the reference data (Figure 1). The allele sizes were auto-calculated with reference to the internal lane size standard GeneScan-500 LIZ, ranging from 35 to 500 base pairs. The allele size bins were exported to Microsoft Office (MS) Excel (2007) program for subsequent analyses.

### Data analysis

The following genetic parameters: allele number and frequency, gene diversity, heterozygosity, and polymorphism information content (PIC) were calculated in PowerMarker software version 3.25 (Liu and Muse, 2005). The level of population structure among the bean accession was established by subjecting SSR allele sizes to a model-based program STRUCTURE version 2.3.3 (Pritchard, 2000). A Principal Coordinate Analysis (PCoA) was obtained using GenAlEx v6.4 software (Peakall and Smouse, 2006) and Un-weighted Neighbor-joining (NJ) tree using microsatellite diversity implemented in the PowerMarker program generated with the Darwin program for displaying genetic relationships among accessions and to test the results of STRUCTURE.

### Defining the population structure of common bean in Uganda

The STRUCTURE program (Pritchard and Stephens, 2000) is the

**Table 2.** Genetic diversity, observed heterozygosity (He) and number of alleles detected in 100 common bean genotypes.

Linkage group	Marker	Major allele frequency	Allele number	Gene diversity	He	PIC
1	BM 200	0.21	17	0.88	0.65	0.86
1	BMd10	0.51	15	0.69	0.72	0.67
2	PVBR 107	0.28	20	0.86	0.30	0.85
2	BM 156	0.23	26	0.89	0.51	0.88
3	BM159	0.38	13	0.69	0.06	0.64
3	BM197	0.22	14	0.86	0.96	0.85
4	BM171	0.51	11	0.68	0.10	0.65
4	PVatgc002	0.34	13	0.79	0.40	0.77
5	BMd28	0.3	15	0.83	0.66	0.81
5	BM175	0.25	21	0.87	0.45	0.86
6	BM187	0.49	12	0.71	0.40	0.68
6	BMd37	0.11	45	0.96	0.67	0.96
7	BM183	0.46	7	0.73	0.00	0.70
7	BM160	0.36	25	0.84	0.44	0.82
8	BM189	0.10	38	0.95	0.80	0.95
8	BM211	0.19	23	0.89	0.61	0.88
9	PVat007	0.17	22	0.91	0.22	0.90
9	BM141	0.29	16	0.85	0.09	0.84
10	GATS11B	0.26	12	0.83	0.33	0.80
10	BMd42	0.32	18	0.81	0.75	0.79
11	BMd41	0.36	16	0.81	0.57	0.80
11	BM205b	0.54	24	0.69	0.22	0.68
Mean		0.31	19	0.82	0.45	0.80

most widely used clustering software applied to detect population genetic structure using a defined number of pre set populations K, where each K is characterized by a set of allele frequencies at each locus. The analysis was run with 10 simulations per K value from K = 2 to 6 using 5,000 replicates for burn-in and for analysis 50,000 replicates. The “true” number of populations (K) was confirmed according to Evanno et al. (2005) using the STRUCTURE Harvester (Earl and Vonholdt, 2012), online ([http://taylor0.biology.ucla.edu/struct\\_harvest/](http://taylor0.biology.ucla.edu/struct_harvest/)) for visualizing outputs. Microsoft Excel program was used for easy conversion of estimated membership coefficient and generating the bar plot. The analysis generated the membership coefficients of each subgroup and the most correct number of subpopulations (K) using different colours according to Rosenberg et al. (2002) with each individual with a fixed length line segment partitioned into K colored components.

### K = 3 analysis

The individual membership coefficient at K = 3 from the STRUCTURE run had maximum mean probability of likelihood value of  $L(K) = -7658.3$  which led to assignment of bean accessions to K= 3 three sub populations as K3.1, K3.2 and K3.3. Accessions with a membership coefficient 0.5 (50% ancestry limits) and above were clustered in the same group at the K= 3 level. The membership coefficient data generated using the STRUCTURE software was exported to MS Excel (2007) for visualization as graphical bar plot for membership coefficient of each accession within the three subpopulations.

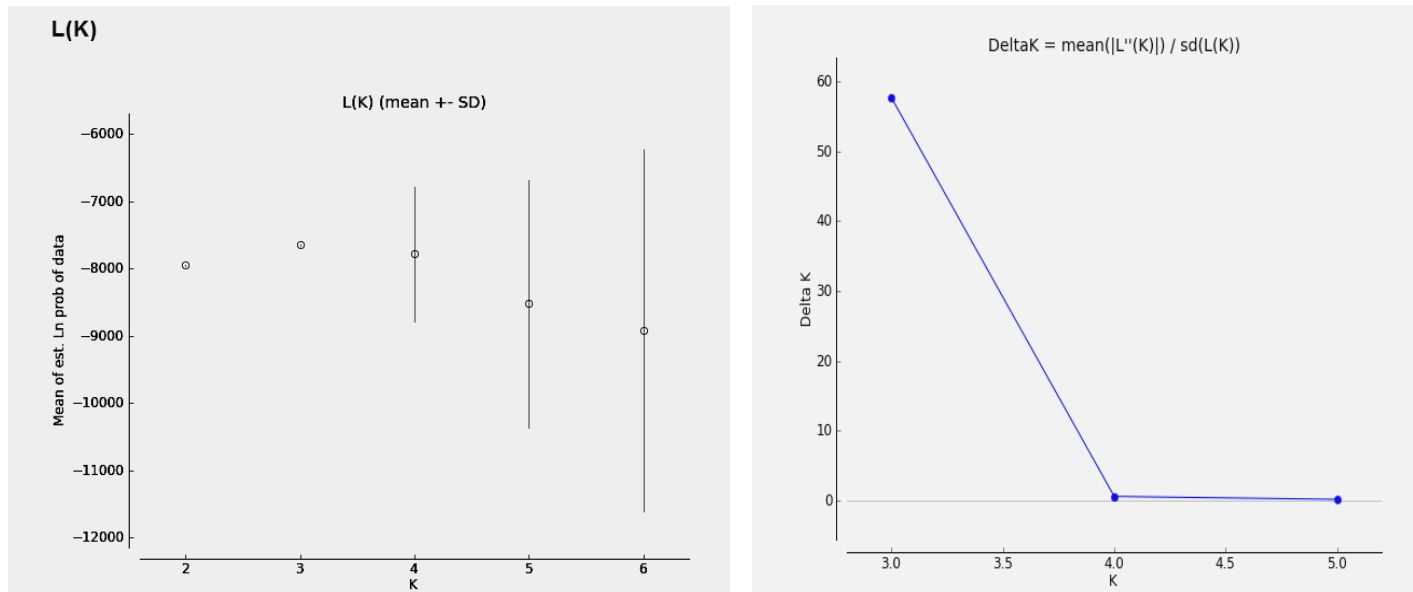
## RESULTS

### Allelic diversity of the common bean germplasm

The key parameters used to define genetic diversity among beans from Uganda are presented in Table 2. There was a high level of polymorphism with a mean of 19 alleles per locus and a range of 7 to 45 alleles in the germplasm. The frequency of the major allele ranged from 0.1 to 0.51 with mean of 0.31. In total, the 22 markers detected 423 alleles, ranging from 114 to 267 bp in size and PIC ranging from 0.64 for BM159 to 0.95 for BMd37 with mean of 0.80. The overall mean heterozygosity was 0.45 with highest heterozygosity value of 0.96 in marker BM197.

### Genetic diversity of Andean and Mesoamerican gene pools in Uganda

The common bean germplasm in Uganda comprised of 51% Andean and 49% Mesoamerican accessions based on the phaseolin analysis (Appendix 1). The mean allele number per locus was higher in the Mesoamerican group (16 alleles) than in the Andeans (13 alleles). The mean major allele frequency in the Mesoamerican gene pool



**Figure 2.** Plots of parameters L (K) and Delta K against the likely sub populations (K) generated according to Evanno et al. (2005) with sub-populations in Ugandan common bean with three sub populations (K=3) as most likely.

was 0.45 and 0.33 in the Andean gene pool (Table 3). The loci BMD37 and BM183 had the highest and lowest allele number respectively, among Ugandan common beans. The mean gene diversity detected was higher in the Mesoamerican (0.79) than in the Andean group (0.67). Mean heterozygosity was comparable between the Andean and Mesoamericans (0.46 and 0.44, respectively). The highest polymorphism was recorded in the Mesoamerican gene pool with a mean PIC value of 0.78 compared to 0.66 in the Andean. The most polymorphic locus in the two gene pools was BMD37, with PIC values of 0.94 among Andean and 0.96 among Mesoamerican genotypes. Locus BM189 also showed high polymorphism in the Andean materials. The markers BM183 and BM159 detected the highest allele frequency of 0.86 and 0.73 respectively in both gene pools with Mesoamerican group showing more genetic diversity.

### Population structure in common bean germplasm in Uganda

The STRUCTURE clustering technique identified the population membership, structuring, and admixture as shown in the K=2 and K=3 sub-populations (Figure 3). The lowest mean probability of data was recorded at L (K) = -8382.0 for K = 6 and a highest mean probability = -7658.3 for K = 3. The Evanno test found a clear maximum for Delta K at K = 3 in the plots of L (K) versus Delta (Figure 2) confirming a likely assignment of the bean germplasm to three sub-groups. The mean genetic diversity statistics (Table 4) for the three subgroups formed at K=3 were calculated as in Hunt et al. (2011)

using PowerMarker v3.25. The STRUCTURE program calculated the level of genetic differentiation or Wright fixation index of F statistics ( $F_{ST}$ ) simultaneously between the different bean sub populations according to Wright (1978). The three subpopulations had moderate differentiation, with  $F_{ST}$  values ranging from 0.05 to 0.15.

### Principal coordinate analysis and genotype associations

The principal coordinate analysis allowed the separation of the genotypes into three groups along the first coordinate with many cases of admixture (Figure 4). The smallest cluster comprised of Mesoamerican accessions U9-6 (95), U9-8 (97), UBR (92) 25 ml (98), VAX3 (99), VAX4 (100) and BAT93 (101), the Mesoamerican control genotype corresponds to the K3.3 sub-group in the STRUCTURE bar plot.

### DISCUSSION

The objective of the study was to characterize common beans in Uganda using the phaseolin marker to determine their frequencies and membership to the Andean and Mesoamerican gene pools. A subsequent analysis with fluorescently labeled SSR markers was done to assess the levels of their genetic diversity and structure. The two gene pools of domesticated common bean are present in Uganda in similar frequencies. Previous reports show a striking difference in numbers of Andean and Mesoamerican gene pools in Africa (Bellucci et al.,



**Table 3.** Genetic diversity parameters in Andean (A) and Mesoamerican (M) common beans in Uganda revealed with 22 polymorphic microsatellite markers.

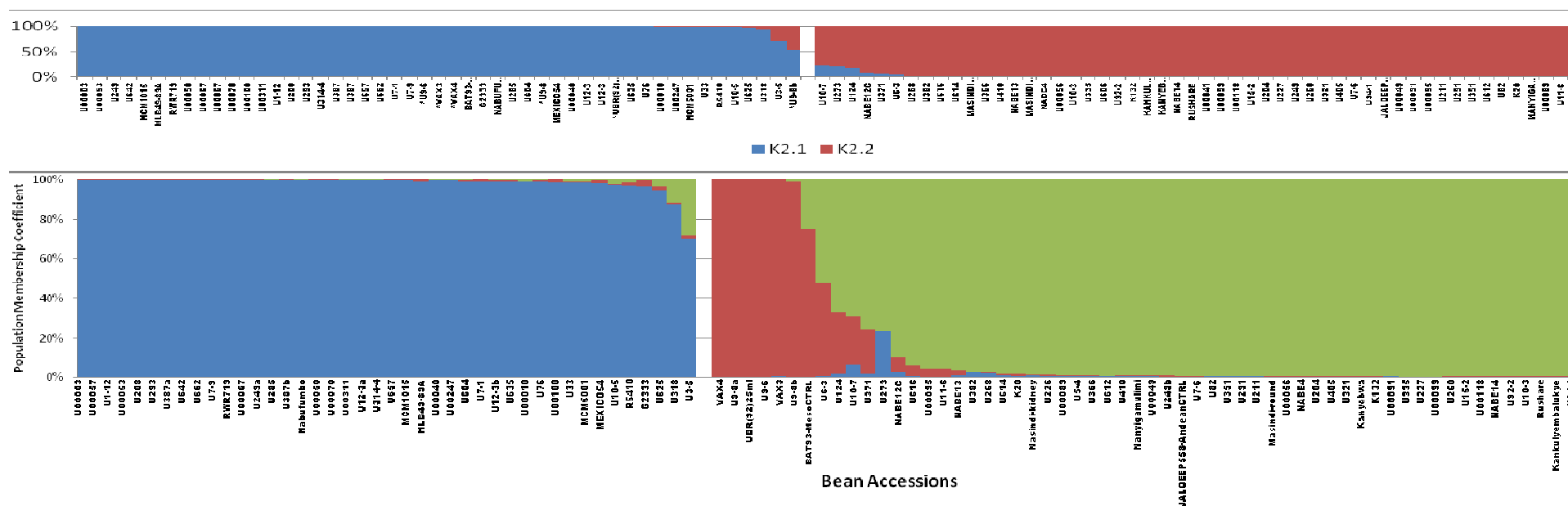
Linkage group	Marker	Major allele frequency		Allele no.		Gene diversity		Heterozygosity		PIC	
		A	M	A	M	A	M	A	M	A	M
1	BM 200	0.27	0.20	14	14	0.84	0.87	0.71	0.59	0.83	0.86
1	BMd10	0.58	0.45	11	15	0.61	0.71	0.67	0.76	0.57	0.68
2	PVBR 107	0.44	0.19	12	19	0.76	0.89	0.33	0.27	0.74	0.88
2	BM 156	0.24	0.34	15	22	0.85	0.80	0.80	0.22	0.84	0.78
3	BM159	0.71	0.73	8	10	0.46	0.46	0.06	0.06	0.41	0.44
3	BM197	0.35	0.30	8	12	0.77	0.81	0.94	0.98	0.74	0.78
4	BM171	0.53	0.49	9	9	0.6	0.69	0.10	0.10	0.53	0.65
4	PVatgc002	0.57	0.49	9	12	0.62	0.70	0.47	0.33	0.59	0.67
5	BMd28	0.44	0.21	11	12	0.74	0.86	0.61	0.71	0.71	0.84
5	BM175	0.48	0.33	8	20	0.7	0.84	0.53	0.37	0.66	0.82
6	BM187	0.84	0.27	8	11	0.28	0.82	0.14	0.67	0.28	0.80
6	BMd37	0.16	0.08	37	38	0.94	0.96	0.55	0.78	0.94	0.96
7	BM183	0.86	0.39	4	7	0.25	0.76	0.00	0.00	0.24	0.73
7	BM160	0.23	0.71	17	17	0.88	0.49	0.75	0.14	0.87	0.48
8	BM189	0.12	0.15	32	25	0.95	0.93	0.73	0.88	0.94	0.92
8	BM211	0.32	0.37	7	22	0.77	0.83	0.61	0.61	0.74	0.82
9	PVat007	0.34	0.20	14	17	0.78	0.89	0.18	0.25	0.75	0.88
9	BM141	0.46	0.25	12	15	0.73	0.87	0.08	0.10	0.71	0.86
10	GATS11B	0.25	0.27	9	12	0.83	0.81	0.22	0.45	0.81	0.79
10	BMd42	0.46	0.26	10	15	0.71	0.83	0.84	0.67	0.67	0.81
11	BMd41	0.29	0.43	13	10	0.8	0.74	0.65	0.49	0.78	0.71
11	BM205b	0.84	0.24	11	21	0.29	0.88	0.12	0.31	0.28	0.87
Mean		0.45	0.33	12.68	16.14	0.69	0.79	0.46	0.44	0.66	0.78

2014). Comparisons to previous studies in Eastern part of Africa, shows a gradient of the two gene pools, for example as you move from Malawi further North to Ethiopia, the proportion of Mesoamerican beans increases. In the southern part, that is, Malawi and Tanzania, Andean predominates (Mkandawire et al., 2004). In Kenya, Andeans dominate while in Ethiopia, Mesoamericans are more frequent (Asfaw et al., 2009). The great lakes region of Central Africa e.g Rwanda and Democratic Republic of Congo is predominated by the Mesoamerican gene pool (Blair et al., 2010). The suitability of markers for multiplexing, informativeness and efficiency in finding the levels of genetic diversity and structure of the different genetic groups was tested in this study. Data about genome-wide genetic variability was obtained quickly and accurately using a set of nine panels of multiplexed SSR markers that are well distributed throughout the genome and were scored semi-automatically. Ramachandran et al. (2003) also reported that capillary electrophoresis in combination with the use of fluorescently labeled primers provide high detection sensitivity of amplified DNA fragments. The total numbers of alleles were 423 with a high mean of 19 alleles per locus compared to other studies in common beans. Asfaw et al. (2009) found 389 alleles across 38 fluore-

scently labeled markers with a mean of 10 alleles per locus in 192 East African bean collections. Blair et al. (2010) also found 301 alleles with an average of 10 alleles per locus using 30 fluorescently labeled SSR markers from a collection of 365 common bean genotypes from Central Africa. Burle et al. (2010) reported a mean of seven alleles across 67 SSR loci in 279 common bean landraces from Brazil.

Gomez et al. (2004) found 5.7 alleles per locus in 108 small seeded individual beans from nine different sites in Nicaragua. The marked difference in frequencies of alleles recorded in this study and other studies in common bean can be attributed to differences in the number of polymorphic markers used, sample sizes, collection sites, and the geographical coverage. Common bean samples in this study included beans with different pedigrees, such as CIAT breeding lines currently used as sources of disease resistance, improved varieties such as NABE4, NABE 12C, NABE 13, NABE 14, K20, and K132, and predominantly landraces collected from farmers' fields. The high number of alleles detected in this study originates from the 22 SSRs, chosen deliberately because of their high PIC values with a high number of alleles.

Population structure refers to sub-divisions of a simple



**Figure 3.** Hierarchical organization of genetic relatedness of 102 common bean accessions based on 22 SSRs markers from STRUCTURE program analysis described in data analysis. K=2 (above) and K = 3 (below); ■ K3.1 ■ K3.2 ■ K3.3).

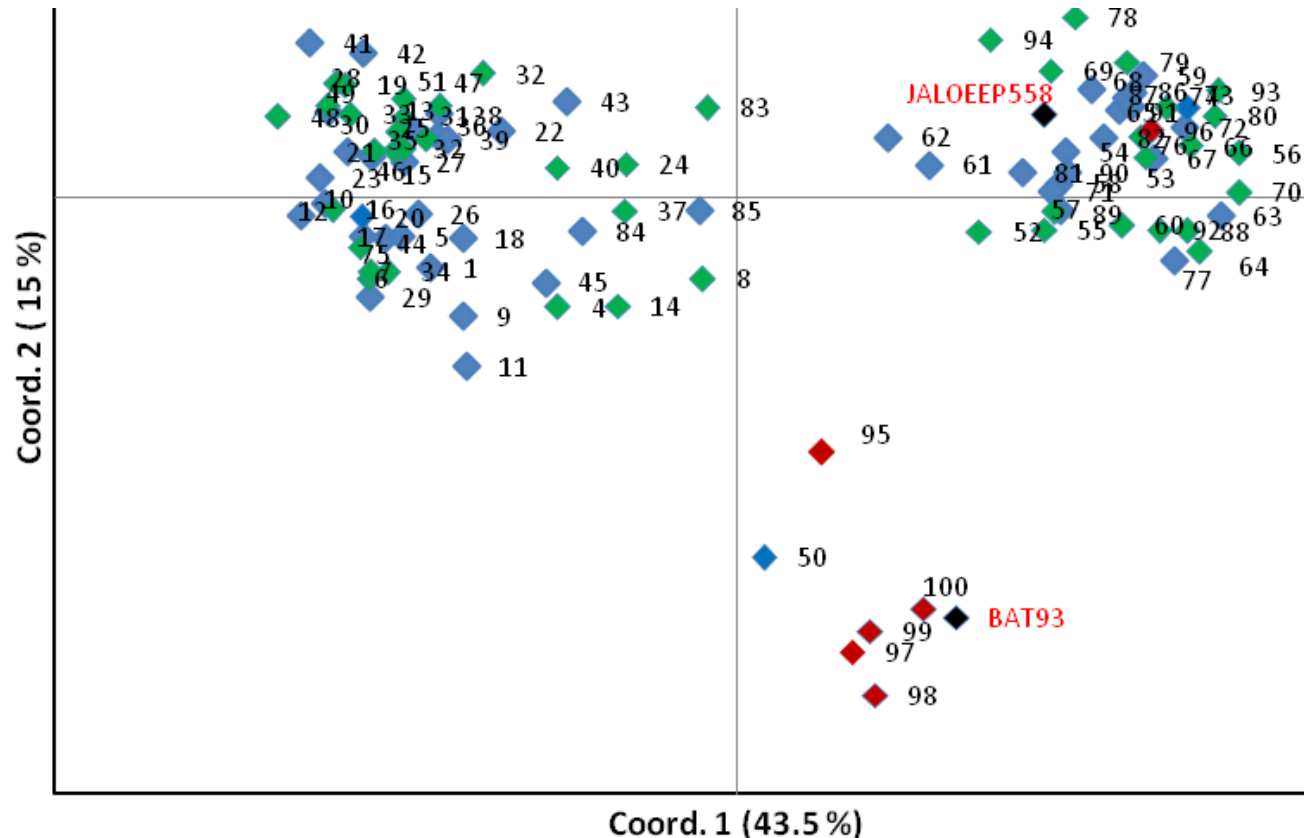
population in some way into smaller groups resulting from a single population’s deviation from Hardy-Weinberg proportions (Falush et al., 2003a) by inbreeding, selection or migration. The modeled population structure of beans in this collection was represented graphically for K= 3 and K=2 (Figure 3). Given the existence of two major gene pools in common bean, a K = 2 situation could be reasonably predicted (Gepts and Bliss, 1985; Singh et al., 1991a). The K= 2 level consisted of sub-populations K2.1 (blue) group contained the Mesoamerican control genotype BAT93, hence, this group is included in the Mesoamerican gene pool. The second group (maroon K2.2) included control genotype JaloEEP558, a typical Andean line; hence, this

group corresponds to the Andean gene pool. The STRUCTURE analysis at K=3 identified group K3.1 (blue) with (100%) Mesoamerican landraces (identified with phaseolin analysis), some of which are sources of disease resistance for bean breeding, such as RWR719 (*Pythium* root rot; Mukankusi, 2008), MEX54 (Angular leaf spot; Namayanja et al., 2006), and G2333 (Anthracnose; Nkalubo et al., 2009). Group K3.2 (maroon) identified consisted of 67% breeding lines of Mesoamerican origin such as VAX3, VAX4, and BAT93. Some accessions showed shared population membership (among K3.1, K3.2 and K3.3) and reflect the effect of

hybridization or gene flow. For example, in the K3.1, accessions U33, MCM5001, MEXICO54, U10-5, R5410, G2333, U625, U318 and U3-5 were hybrids. Among the sub group K3.3 accessions, U268, U382, NABE12C, U273, U371, U10-7 and U124, also consisted of contributions from the three sub-groups identified at K=3. Some accessions derived ancestry from two subgroups, for example, K3.2 and K3.3 with accessions U616, U6-3 and BAT93 (Mesoamerican gene pool control genotype) as shown in Figure 3.

Attempts to infer gene pool of different genotypes involved relating the sub-populations members to the known gene pool controls JaloEEP558-andeanCTRL (for Andeans) and BAT93-mesoCTRL (for Mesoamericans). The ancestry of



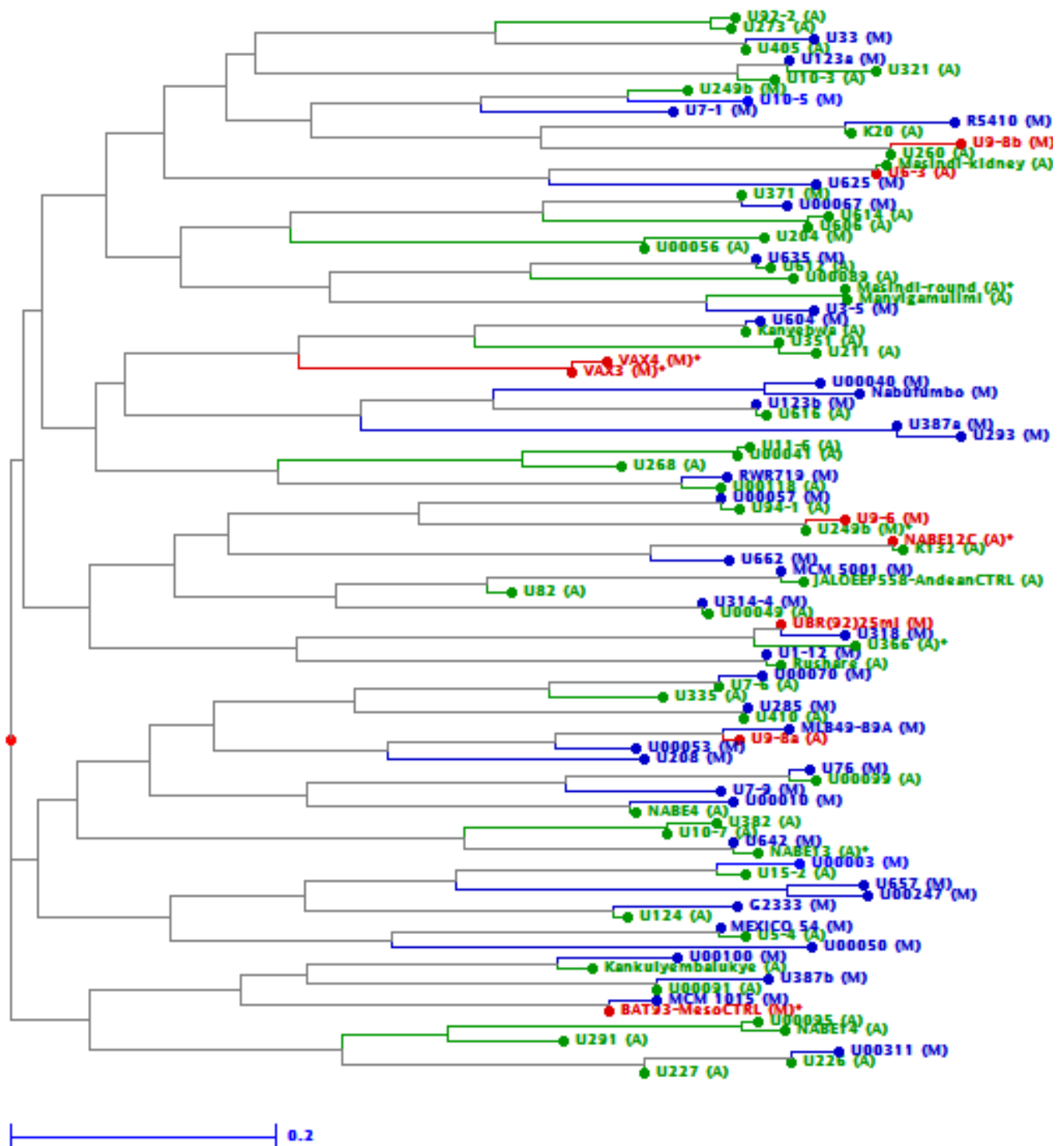


**Figure 4.** Principal coordinates analysis (PCoA) of accessions from microsatellites diversity based on the presence and absence of alleles. The three subpopulations are represented by diamond symbols, whose colour's reconcile with the three subgroups in STRUCTURE (Figure 3): K3.1 (blue), K3.2 (green), and K3.3 (maroon). The two black diamonds are controls, that is, genotypes JaloEEP558 and BAT93 for the Andean and Mesoamerican gene pools, respectively.

the Mesoamerican control genotype (BAT93-mesoCTRL) was shared between K3.1 and K3.2 (26% - green and 74% - maroon). Burle et al. (2010) used the same accession as a control for the Mesoamerican gene pool and reported this accession as a breeding genotype, which originates from a four-way cross, with all parents having Mesoamerican origin, hence explains its shared membership between two sub-population at  $K=3$ . The sharing of ancestry between genotypes is generally explained by recombination in some parts of the genome due to inter and intra gene pool crossings in breeding or natural hybridization. Membership switching is demonstrated by the PCoA (Figure 4) and NJ tree (Figure 5) with no correspondence between the SSR and phaseolin marker groupings, with a lot of gene pool membership switching. Membership switching among presumed subpopulations in common beans occurred in previous studies, involving allozymes and RAPD markers (Freyre et al., 1996) for presumed ancestral group of the Andean and Mesoamerican gene pools. According to Kwak and Gepts (2009), the lack of phaseolin polymorphism in domesticated gene pools prevents the detection of more subtle genetic differentiation between closely related

accessions at this single albeit complex locus. The same reason can be extended to this study, in addition to admixture, to explain the lack of concordance between phaseolin and the model-based approaches used to group the bean germplasm. The long cohabitation of the two cultivated gene pools possibly led to the introgression of alleles between cultivars creating hybrids with shared phenotypic traits. The identification of hybrids in similar studies was based on the intermediate position between gene pools in the NJ trees (Bellucci et al., 2014). The high levels of admixture observed within the subpopulations in the STRUCTURE, PCoA and NJ tree analyses, clearly shows that the common bean germplasm in Uganda have considerable variations for utilizations in breeding. Model-based analyses of population divisions can be performed separately in Andean and Mesoamerican, in addition to analysing the entire sample, to detect accessions membership switching, since a marked reduction in genetic differentiation is observed by analyzing separate gene pools (Kwak and Gepts, 2009).

The STRUCTURE analysis from Burle et al. (2010) found five groups in common beans landraces from Brazil



**Figure 5.** Un-weighted Neighbor-joining tree generated in Darwin program using microsatellite diversity of common bean accessions in Uganda, based on the Chord distance implemented in the PowerMarker program. Each branch is color-coded according to membership into the K=3 groups identified by STRUCTURE (same colors as in Figure 3). The names of accessions in brackets show their gene pool; (A) Andean and (M) Mesoamerican as determined with the Phaseolin marker, in methods and Plate 1. All the accessions in the blue branch are Mesoamerican, 95% of the green branch is Andean and 67% of maroon branch contains Mesoamerican accessions.

with strong genetic differentiation for the Mesoamerican gene pool and limited admixture between gene pools.

They used 67 microsatellite markers spread over the 11 linkage groups of crop's genome and Mesoamericans

four times more frequent than the Andean gene pool. In other crops, for example, African rice (*Oryza glaberrima*), Khady et al. (2011) identified three distinct populations in 74 rice varieties collected from Benin through population structure analysis. The high diversity in the Mesoamerican gene pool (both maroon and green components in STRUCTURE bar plot), compared to Andean in this study is due to farmers' preference for small seeded beans. In practice, preference for Mesoamerican bean types by some farmers (Blair et al., 2010) results in planting of many smaller weight seeds than larger seed weights. CIAT (2005) reported farmer's preference for smaller-seeded Mesoamerican genotypes to manage root rot disease as is the case in South-western Uganda, the leading common bean producing area in Uganda. These bean materials eventually find their ways to other parts of the country through various routes and activities.

The PCoA analysis graphical display showed that the Mesoamerican group was the most diverse and included many presumed hybrids. This observation shows that large variations arose from gene introgressions in breeding material and out-crosses in landraces that occur in farmers' fields. In other studies, Blair et al. (2010) conducted a PCoA and found diversity within and between gene pools in a larger collection of 365 common bean genotype from Central Africa. Introgression between gene pools was observed for 32 intermediate genotypes. Maciel et al. (2003) however, found no clear distinction between domesticated common bean samples from Brazil using AFLP markers and suggested that admixture was the possible cause. The common bean germplasm in Uganda has a moderate level of population structuring with  $F_{ST}$  values ranging from 0.09 to 0.14 for the three clusters generated from STRUCTURE analysis. The K3.3 group was the most differentiated with  $F_{ST}$  values of 0.14. The K3.1 and K3.2 groups followed with  $F_{ST}$  values of 0.12 and 0.09, respectively.

In other studies, Asfaw et al. (2009) found diversity in East African bean landraces and cultivars of Andean origin to be more differentiated ( $F_{ST}=0.331$ ) than ones in the Mesoamerican gene pool ( $F_{ST}=0.04$ ) with mean  $F_{ST}$  of 0.27 among pairs of populations analyzed. They related genetic divergence in East African bean landraces to the original differences in introduced germplasm from the primary centre of origin combined with spontaneous out-crossing in farmers' fields and further farmer selection for adaptation and production uses. Further subdivision of the two gene pool into eco-geographic races (Singh et al., 1991a) was not carried out in this study and thus recommended in future to facilitate the use of the germplasm in breeding.

## Conclusions and recommendations

The fluorescently labeled SSR markers revealed genetic

diversity and levels of population structure within the Andean and Mesoamerican common bean in Uganda. The Mesoamerican gene pool was both more structured and genetically diverse. Gene introgressions between the two gene pools have occurred in Uganda with some accessions existing as hybrid genotypes of Andean and Mesoamerican gene pools. The 22 SSR markers used in this study showed high polymorphism among common bean samples in Uganda. Recombining Andean and Mesoamerican genes in same background could generate hybrid genotypes with broader and more durable diseases resistance for the current and future breeding programs. The common bean samples studied appears to be of great importance for breeding in Uganda and regionally as it may overcome some of the difficulties in transferring traits between the two gene pools. The bean genotypes could further be evaluated in multilocation trials to screen for agronomic traits, adaptability and biotic constraints. Regular follow up of these germplasm in the farmer's fields is vital to broaden the common bean genetic base and guard against possible losses of useful segregating phenotypes in farmer's fields. The conservation of the superior landraces can be done at NaCRRI through field regenerations and duplicated at the Plant Genetic Resources Centre (PGRC) at Entebbe in Uganda, for long-term storage.

## Conflict of Interests

The author(s) declare no conflict of interests.

## ACKNOWLEDGEMENTS

The authors were supported by a fellowship of the Kirkhouse Trust (United Kingdom). We are grateful to the Bean Research and Development Programs at NaCRRI and CIAT - Kawanda for providing breeding lines. We thank Annet Namusoke, Dan Mawejeje and Jane Mukabaranga for the technical assistance at NaCRRI and helpful comments of Claire Mukankusi, Buah Stephen, Allan Male and Stephen Musoke at CIAT- Kawanda.

## REFERENCES

- Asfaw A, Blair MW, Almekinders C (2009). Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from the East African Highlands. *Theor Appl Genet.* 120:1-12.
- Beebe S, Rengifo J, Gaitán-Solis E, Duque MC, Tohme J (2001). Diversity and origin of Andean landraces of common bean. *Crop Sci.* 41:854-862.
- Bellucci E, Bitocchi E, Rau D, Rodriguez M, Biagetti E, Giardini A, Attene G, Nanni L, Papa R (2014) Genomics of Origin, Domestication and Evolution of *Phaseolus vulgaris*. In *Genomics of Plant Genetic Resources*; Tuberosa R, Graner A, Frison, E. Eds. Springer: Dordrecht, The Netherlands. 1:483-507.

- Appl Genet. 121:237-248.
- Blair MW, Giraldo MC, Buendia HF, Tovar E, Duque MC, Beebe S (2006). Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet.* 113:100-109.
- Blair MW, Gonzalez Laura F, Paul M Kimani, Louis Butare (2010). Genetic diversity, inter-gene pool introgression and nutritional quality of common beans (*Phaseolus vulgaris* L.) from Central Africa. *Theor Appl Genet.* 113:100-109.
- Blair MW, Pedraza F, Buendia H, Gaitan E, Beebe S, Gepts P, Tohme J (2003). Development of a genome wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet.* 107:1362-1374.
- Burle ML (2008). Assessing the genetic diversity of common bean (*Phaseolus vulgaris* L.) landraces from Brazil: from genetic structure to landscape distribution PhD thesis. University of California, Davis, U.S.A.
- Burle ML, Fonseca R, Kami J, Gepts P, (2010). Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a secondary center of diversity. *Theor. Appl. Genet.* 121:801-813
- CIAT (2005). Utilization of bean genetic diversity in Africa. Highlights of CIAT in Africa, No. 21. Cali.Colombia.
- Diaz LM, Blair MW (2006). Race structure within the Mesoamerican gene pool of common bean (*Phaseolus vulgaris* L.) as determined by microsatellite markers. *Theor Appl Genet.* 114:143-154.
- Doyle JJ, Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* 12:13-15.
- Earl DE, Vonholdt BM (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour.* 4(2):359-361. <http://taylor0.biology.ucla.edu/structureHarvester>
- Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol Ecol.* 14:2611-2620.
- Falush D, Stephens M, Pritchard JK (2003a). Inference of population structure: Extensions to linked loci and correlated allele frequencies. *Genetics.* 164:1567-1587.
- Freyre R, Rios R, Guzmán L, Debouck D, Gepts P (1996). Ecogeographic distribution of *Phaseolus* species (Fabaceae) in Bolivia. *Econ. Bot.* 50:195-215.
- Gaitan-Solis E, Duque MC, Edwards K, Tohme J (2002). Microsatellite repeats in common bean (*Phaseolus vulgaris* L.): isolation, characterization, and cross-species amplification in *Phaseolus* spp. *Crop Sci.* 42:2128-2136.
- Gepts P, Bliss FA (1985) F1 hybrid weakness in the common bean: differential geographic origin suggests two gene pools in cultivated bean germplasm. *J. Hered.* 76:447-450.
- Gepts P, Bliss FA (1988) Dissemination pathways of common bean (*Phaseolus vulgaris*, Fabaceae) deduced from phaseolin electrophoretic variability. II. Europe and Africa. *Econ Bot.* 42:86-104.
- Gomez OJ, Blair MW, Frankw-Lindberg BE, Gullberg U (2004) Molecular and Phenotypic diversity of common bean landraces from Nicaragua. *Crop Sci.* 44:1412-1418.
- Hunt HV, Denyer K, Packman LC, Jones MK, Howe CJ (2011). Molecular basis of the waxy phenotype in broomcorn millet (*Panicum miliaceum* L.). *Mol Biol and Evol.* 27:1478-1494.
- Kami J, Becerra-Velasquez V, Debouck DG, Gepts P (1995). Identification of presumed ancestral DNA sequences of phaseolin in *Phaseolus vulgaris*. *Evolution* 92:1101-1104.
- Kami JA, Gepts P (1994). Phaseolin nucleotide sequence diversity in *Phaseolus*. I. Intraspecific diversity in *Phaseolus vulgaris*. *Genome* 37:751-757.
- Khady ND, Ines S, Glenn G, Marie N, Ndjondjop (2011). Suitability of a selected set of simple sequence repeats (SSR) markers for multiplexing and rapid molecular characterization of African rice (*Oryza glaberrima* Steud.). *Afr. J. Biotechnol.* 10:6675-6685.
- Koinange EMK, Singh SP, Gepts P (1996) Genetic control of the domestication syndrome in common-bean. *Crop Sci.* 36:1037-1045.
- Kwak M, Gepts P (2009). Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae). *Theor Appl Genet.* 118: 979-992.
- Liu Kejun, Muse SV (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21:2128-2129.
- Maciel FL, Echeverrigaray S, Gerald LTS, Grazziotin FB (2003). Genetic relationships and diversity among Brazilian cultivars and landraces of common beans (*Phaseolus vulgaris* L.) revealed by AFLP markers. *Genet Resour. Crop Evol.* 50:887-893.
- Melissa L, Baack Eric, Toone R (2000). STRand v2.2.30 software. Veterinary Genetics Lab. Update by Toone, Rob. 2007. University of California, Davis. U.S.A.
- Mkandawire ABC, Mabagala RB, Guzman P, Gepts P, Gilbertson R L (2004). Genetic diversity and pathogenic variation of common blight bacteria (*Xanthomonas campestris* pv. *phaseoli* and *X. campestris* pv. *phaseoli* var. *fuscans*) suggests pathogen coevolution with the common bean. *Phytopathology.* 94:593-603.
- Mukankusi MC (2008). Improving resistance to Fusarium root rot in Common bean (*Phaseolus vulgaris* L.) PhD thesis. University of KwaZulu-Natal, Pietermaritzburg, South Africa.
- Namayanja A, Buruchara R, Mahuku G, Rubaihayo P, Kimani P, Mayanja S, Eyedu H (2006). Inheritance of resistance to angular leaf spot in common bean and validation of the utility of resistance linked markers for marker assisted selection outside the mapping population Euphytica 151:361-369.
- Nkalubo ST, Melis R, Derera J, Laing MD, Opio F (2009). Genetic analysis of anthracnose resistance in common bean breeding source germplasm. *Euphytica.* 167:303-312.
- Peakall R, Smouse PE, (2006). GENEALEX version 6: Genetic Analysis in Excel. Population genetic software for teaching and research. *Mol Ecol.* 6: 288-295.
- Pritchard J, Stephens DMP (2000). Inference of population structure using multilocus genotype data. *Genetics.* 155: 945-959.
- Ramachandran A, Zhang M, Goad D, Olah G, Malayer JR, Rassi ZE (2003). Capillary electrophoresis and fluorescence studies on molecular beacon-based variable length oligonucleotide target discrimination. *Electrophoresis.* 24:70-77.
- Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, Feldman MW (2002). Genetic structure of human populations. *Science.* 298:2381-2385
- Schuelke M (2000). An economic method for the Fluorescent labelling of PCR fragments. *Nat Biotechnol.* 18:233-234.
- Sekabembe CK (2010). Introduction to Agronomy, principles and practices. Fountain Publisher, Kampala, Uganda. pp. 44-49.
- Singh SP, Gepts P, Debouck DG (1991a). Races of common bean (*Phaseolus vulgaris* Fabaceae). *Econ Bot.* 45:379-396.
- Trethowan RM, Kazi AM (2008). Novel germplasm resources for improving environmental stress tolerance of hexaploid wheat. *Crop Sci.* 48:1255-1265.
- Wright S (1978). Evolution and the Genetics of Populations. Variability within and among natural populations. University of Chicago Press.
- Yu K, Park J, Poysa V, Gepts P (2000). Integration of simple sequence repeats markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). *Hereditary* 91:429-43.

**Appendix 1.** Gene pool and morphology of 100 common bean accessions from Uganda studied at 22 fluorescently labeled microsatellite loci.

Accession and status in breeding	Plant type	Gene pool	100 Seeds weight (g)	Seed colour	Place of collection
K132 <sup>I</sup>	I	A	50.1	Maroon mottled	C
K20 <sup>I</sup>	I	A	33.8	Red mottled	C
Kankulyembalukye <sup>L</sup>	II	A	37.2	Pink mottled	W
KANYEBWA <sup>L</sup>	I	A	33	Red mottled	C
MANYIGAMULIMI <sup>L</sup>	II	A	39.3	Red mottled	C
MASINDI ROUND <sup>L</sup>	I	A	35.6	Yellow	C
MASIND <sup>I</sup> KIDNEY <sup>L</sup>	I	A	31.9	Yellow	C
NABE12C <sup>L</sup>	IV	A	38.8	Cream specked	CIAT
NABE13 <sup>L</sup>	I	A	41.5	Red mottled	C
NABE14 <sup>I</sup>	I	A	42.5	Red mottled	C
NABE4 <sup>I</sup>	III	A	40	Red mottled	C
RUSHARE <sup>L</sup>	II	A	42.9	Purple	SW
U00041 <sup>L</sup>	I	A	46.8	Maroon	WN
U00049 <sup>L</sup>	IV	A	30.8	Maroon	WN
U00056 <sup>L</sup>	II	A	36.8	Red	WN
U00089 <sup>L</sup>	II	A	35.3	Purple	W (Bushenyi)
U00091 <sup>L</sup>	I	A	39.9	Red	E (Iganga)
U00095 <sup>L</sup>	I	A	41.2	Red	E (Iganga)
U00099 <sup>L</sup>	I	A	37.3	Red	E (Iganga)
U00118 <sup>L</sup>	IV	A	28.8	Yellow	W (Mbarara)
U10-3 <sup>L</sup>	II	A	50.9	Greenish-yellow	E (Mbale)
U10-7 <sup>L</sup>	III	A	35.9	Red	E (Mbale)
U11-6 <sup>L</sup>	II	A	43.8	Purple	WN (Nebbi)
U124 <sup>L</sup>	III	A	32.4	Greenish-yellow	WN
U15-2 <sup>L</sup>	II	A	33.9	Mixed	WN (Arua)
U211 <sup>L</sup>	II	A	32.9	Mixed	WN (Arua)
U226 <sup>L</sup>	II	A	41.9	Purple	WN (Arua)
U227 <sup>L</sup>	I	A	44.5	Red	WN (Arua)
U249a <sup>L</sup>	I	A	21.7	White	SW (Kabale)
U260 <sup>L</sup>	II	A	36.4	Red	SW (Kabale)
U268 <sup>L</sup>	II	A	36.3	Brown	SW (Kabale)
U273 <sup>L</sup>	II	A	43.2	Maroon	SW (Kabale)
U291 <sup>L</sup>	II	A	44	Black	WN (Arua)
U321 <sup>L</sup>	III	A	33	Black	WN (Arua)
U335 <sup>L</sup>	II	A	44.4	Purple	WN (Arua)
U351 <sup>L</sup>	I	A	40.3	Maroon	WN (Arua)
U382 <sup>L</sup>	IV	A	40	Yellow	W (Mbarara)
U405 <sup>L</sup>	II	A	40.8	Maroon	E (Iganga)
U410 <sup>L</sup>	I	A	32.2	Yellow	C (Mubende)
U5-4 <sup>L</sup>	II	A	33.5	Brown	W (Masindi)
U606 <sup>L</sup>	I	A	42.2	Black	N
U612 <sup>L</sup>	II	A	34.7	Red	WN
U614 <sup>L</sup>	II	A	40.7	Red	CT (Mukono)
U616 <sup>L</sup>	I	A	49.9	Black	NE (Kapchorwa)
U6-3 <sup>L</sup>	IV	A	43	Red	E (Iganga)
U7-6 <sup>L</sup>	I	A	42	Purple	WN Nebbi)
U82 <sup>L</sup>	I	A	43.6	Maroon	CT(Mukono)
U92-2 <sup>L</sup>	II	A	42.9	Purple	W (Mbarara)
U94-1 <sup>L</sup>	I	A	35.2	Maroon	N

## Appendix 1. Contd.

Accession and status in breeding	Plant type	Gene pool	100 Seeds weight (g)	Seed colour	Place of collection
U9-8a <sup>L</sup>	I	A	38.9	Maroon	W (Mbarara)
U366 <sup>L</sup>	IV	A	36.6	Maroon	SW (Kabale)
Jalo EEP558 <sup>C</sup>		A	35.6	Brown	Brazil
G2333 <sup>C</sup>	IV	M	24.4	Red	CIAT
MCM 1015 <sup>C</sup>	II	M	18.5	Cream	CIAT
MCM 5001 <sup>C</sup>	II	M	19.5	Cream	CIAT
MEXICO 54 <sup>C</sup>	III	M	32.4	Purple	CIAT
MLB49-89A <sup>C</sup>	II	M	31.2	Black	CIAT
NABUFUMBO <sup>L</sup>	III	M	24.8	Pink	C
R5U410 <sup>L</sup>	II	M	18.8	Black	Rwanda
RWR719 <sup>L</sup>	II	M	18.1	Red	CIAT
U00003 <sup>L</sup>	III	M	15.2	White	WN (Nebbi)
U00010 <sup>L</sup>	IV	M	21.8	White	WN (Nebbi)
U00040 <sup>L</sup>	I	M	16.7	White	WN (Nebbi)
U00050 <sup>L</sup>	IV	M	22	Red	WN (Nebbi)
U00053 <sup>L</sup>	IV	M	25.8	Maroon	WN (Nebbi)
U00057 <sup>L</sup>	III	M	22.7	Pink	W (kabale)
U00067 <sup>L</sup>	III	M	37.2	Mixed	E (Kamuli)
U00070 <sup>L</sup>	I	M	18	Cream	W (Bushenyi)
U00100 <sup>L</sup>	IV	M	20.5	White	N (Lira)
U00247 <sup>L</sup>	II	M	16.6	Mixed	E (Kapchorwa)
U00311 <sup>L</sup>	II	M	17.9	Cream	E (Kapchorwa)
U10-5 <sup>L</sup>	III	M	31.2	Black	E (Mbale)
U1-12 <sup>L</sup>	IV	M	27.4	Red	W (Kisoro)
U123a <sup>L</sup>	III	M	21.8	Mixed	N
U123b <sup>L</sup>	III	M	21.9	Mixed	WN(Arua)
U204 <sup>L</sup>	III	M	44.4	Black	E(Iganga)
U208 <sup>L</sup>	III	M	37.7	Cream	WN(Nebbi)
U249b <sup>L</sup>	III	M	32.9	Red	W(kabale)
U285 <sup>L</sup>	II	M	35.2	Mixed	W(kabale)
U293 <sup>L</sup>	III	M	33.3	Cream	C(Mubende)
U314-4 <sup>L</sup>	III	M	31.5	Cream	W(Mbarara)
U318 <sup>L</sup>	II	M	16.9	Mixed	WN(Nebbi)
U33 <sup>L</sup>	I	M	18.5	Mixed	WN
U3-5 <sup>L</sup>	III	M	15	Mixed	W(Masindi)
U371 <sup>L</sup>	III	M	34	Maroon	CT(Mubende)
U387a <sup>L</sup>	III	M	26.7	Mixture	WN(Nebbi)
U387b <sup>L</sup>	IV	M	29.1	Red	W(kabale)
U604 <sup>L</sup>	III	M	24.5	Red	W(kabale)
U625 <sup>L</sup>	I	M	18.3	White	E(Kapchorwa)
U635 <sup>L</sup>	II	M	18	Cream	WN (Nebbi)
U642 <sup>L</sup>	II	M	16.1	Mixed	WN (Nebbi)
U657 <sup>L</sup>	III	M	20.1	White	W(Mbarara)
U662 <sup>L</sup>	III	M	32.9	Mixed	E(Kapchorwa)
U7-1 <sup>L</sup>	I	M	18.9	Red mottled	WN (Nebbi)
U76 <sup>L</sup>	I	M	19.2	Mixed	WN(Nebbi)
U7-9 <sup>L</sup>	I	M	30.9	Cream striped	WN (Nebbi)
U9-6 <sup>L</sup>	II	M	18.8	Cream	WN (Nebbi)
U9-8b <sup>L</sup>	III	M	18.2	Maroon	WN (Arua)
UBR9225ML <sup>C</sup>	II	M	18.4	White	CIAT



## Appendix 1. Contd.

Accession and status in breeding	Plant type	Gene pool	100 Seeds weight (g)	Seed colour	Place of collection
VAX3 <sup>C</sup>	II	M	21.9	Red	CIAT
VAX4 <sup>C</sup>	II	M	20.8	Cream	CIAT
BAT93 <sup>C</sup>	II	M	20.5	Cream	Brazil

**Accession name:** superscript is status in breeding: I=improved, L=landrace, C=CIAT line. Gene pool: A= Andean and M = Mesoamerican determined as described in the methods with the corresponding weights of 100 seeds. **Gene pool controls** are accessions: No.52 (for Andean) and 102 (for Mesoamerican). **Plant types:** I = Determinate growth habit, II = Indeterminate growth habit, III= Indeterminate prostrate growth with some climbing ability and IV= Indeterminate growth habit with strong climbing ability. **Place of collection:** Region with district name in parentheses (C = Central, W = Western, WN = West Nile, E = Eastern and N = Northern) and CIAT (regional headquarters in Uganda). **Seed coat colour:** mixed are accessions that segregated with seeds of multiple colours.