ORIGINAL PAPER

Genetic diversity, inter-gene pool introgression and nutritional quality of common beans (*Phaseolus vulgaris* L.) from Central Africa

Matthew W. Blair · Laura F. González ·
Paul M. Kimani · Louis Butare

Abstract The Great Lakes region of Central Africa is a major producer of common beans in Africa. The region is known for high population density and small average farm size. The common bean represents the most important legume crop of the region, grown on over a third of the cultivated land area, and the per capita consumption is among the highest in the world for the food crop. The objective of this study was to evaluate the genetic diversity in a collection of 365 genotypes from the Great Lakes region of Central Africa, including a large group of landraces from Rwanda as well as varieties from primary centers of diversity and from neighboring countries of Central Africa, such as the Democratic Republic of Congo and Uganda, using 30 fluorescently labeled microsatellite markers and automated allele detection. In addition, the landraces were evaluated for their seed iron and zinc concentration to determine if genetic diversity influenced nutritional quality. Principal coordinate and neighbor-joining analyses allowed the separation of the landraces into 132 Andean and 195 Mesoamerican (or Middle American) genotypes with 32 landraces and 6 varieties

intermediate between the gene pools and representing inter-gene pool introgression in terms of seed characteristics and alleles. Genetic diversity and the number of alleles were high for the collection, reflecting the preference for a wide range of seed types in the region and no strong commercial class preference, although red, red mottled and brown seeded beans were common. Observed heterozygosity was also high and may be explained by the common practice of maintaining seed and plant mixtures, a coping strategy practiced by Central African farmers to reduce the effects of abiotic and biotic stresses. Finally, nutritional quality differed between the gene pools with respect to seed iron and zinc concentration, while genotypes from the intermediate group were notably high in both minerals. In conclusion, this study has shown that Central African varieties of common bean are a source of wide genetic diversity with variable nutritional quality that can be used in crop improvement programs for the region.

M. W. Blair (✉) · L. F. González
Bean Project, International Center for Tropical Agriculture/Centro Internacional de Agricultura Tropical (CIAT), A.A. 6713, Cali, Colombia
e-mail: m.blair@cgiar.org

P. M. Kimani
Department of Crop Science, University of Nairobi,
P.O. Box 29053, Nairobi, Kenya

L. Butare
Institut des Sciences Agronomiques du Rwanda (ISAR),
Rubona Station, Butare, Rwanda

Introduction

Highland Africa is a secondary center of diversity for common bean (*Phaseolus vulgaris* L.) and the crop is an important contributor to food security in the region (Broughton et al. 2003; Asfaw et al. 2009). While the largest producers of Eastern and Central Africa are the Democratic Republic of Congo (DRC), Kenya, Tanzania and Uganda, consumption is highest around the Great Lakes, especially in the adjacent areas of Burundi, Rwanda, southwest Uganda and eastern DRC (Allen et al. 1989; Wortmann et al. 1998). Trade flow between regions and countries is notable with smaller towns serving as collection centers and small markets in their own right. Across all of Africa, a total of 4 million hectares are planted annually

with a total production of around 2.8 million tons (Broughton et al. 2003).

Within East Africa, the areas bordering the Great Lakes have particularly high per capita common bean consumption rates (above 40 kg/year). The region is intensively farmed due to the high populations with common bean as one of the major crops of each country. The region benefits from bimodal rainfall and therefore common bean can often be planted twice a year and the total production is high. For example Rwanda and Burundi, two of the countries in the region each produce between 200,000 and 250,000 tons yearly with an additional 102,000 and 65,000 hectares in neighboring DRC and Uganda (Wortmann et al. 1998). Rwandese production has been stable except during the civil war of the mid 1990s when production fell to 150,000 tons (Sperling 2001). It is estimated that up to 95% of Rwandan and Burundian farmers produce beans mostly for subsistence, accounting for 22–30% of cultivated land area with an average productivity of 800–1,000 kg/ha. Some imports of common beans occur coming by land from neighboring countries. Common beans are thought to provide up to 65% of protein consumed in the region and to be the second most important source of calories providing 32% of energy supply.

The major limitations to bean production in the Great Lakes region are: biotic constraints, especially fungal and viral diseases such as angular leaf spot, anthracnose, ascochyta blight, bean common mosaic virus and root rots or insect pests such as bean stem maggot; as well as abiotic stresses, such as acid soil, low soil phosphorus or nitrogen and intermittent or seasonal drought (Allen et al. 1989). Low soil fertility is addressed by micro-fertilization especially for climbing beans while liming and irrigation are not common, therefore aluminum toxicity and water limitations are serious (Kelly et al. 2001). During the dry season, some beans are grown in low-lying areas with high water tables in the valleys, but most of the production is on steep to gently sloped hillsides where erosion risks are high. Some artificial terracing or contour lines made up of vitiver grass help to reduce erosion; while agroforestry species especially *Eucalyptus* provide staking material for climbing beans, which are often grown in higher fertility areas or in backyards (Graf et al. 1991). Chemical control of insects or diseases is uncommon, but some biocontrol methods and crop rotations have been recommended (Allen et al. 1989).

Phenotypic diversity of common beans is very high in the Great Lakes region due to the fact that much of the common bean crop, especially in Rwanda and neighboring areas, is grown as varietal mixtures with consumers accepting a wide range of seed colors (Lamb and Hardman 1985; Sperling 2001). In addition, agro-ecologies and environments in the region range from 800 to 2,400 meters above sea level and from semi-arid savannah to humid montane forests. Varietal mixture has been postulated to be a mechanism of coping with environmental variability and with diversifying production in small plots, with early and late maturing components together providing harvestable products over a long period. This is especially important where the bean seed is moved from one agro-ecology to another (for example up and down the hillsides or from valley bottom to steep slopes) and where common bean leaves, pods, green seed and dry grain are all consumed at different times of the crop cycle.

Common bean diversity in Central Africa has been under threat due to a variety of circumstances. First among these, economic and agronomic developments have led to some emphasis on single-component varieties over multi-component mixtures. This is driven by more urban consumers demanding pure lines rather than mixtures, since they are easier to prepare due to uniformity. Secondly, the introduction of new varieties has displaced some traditional varieties. In this regard, new climbing beans brought into the region since the 1980s have been widely adopted due to their high yield potential and this has led to farm intensification. Root rots, which have increased as land pressure has led to reduced crop rotation, have also resulted in the disappearance of susceptible genotypes. In the case of Rwanda in the early 1990s, and more recently in neighboring areas of Burundi and DRC, the social upheavals of genocide, civil war and displacement have accelerated many of these changes even if the seed supply system has stayed intact in some cases (Sperling 2001).

In this study, our goal was to evaluate the genetic diversity found in Central African germplasm including the "seeds of hope" collection of Rwandan landraces collected in the early 1990s by Institut des Sciences Agronomiques du Rwanda (ISAR) and International Center for Tropical Agriculture (CIAT) as well as control genotypes from other Great Lakes region countries and from the Andean and Mesoamerican or Middle American gene pools. In total, we analyzed a set of 365 genotypes with a set of 30 fluorescent microsatellite markers that have proved valuable for fingerprinting common bean gene pools and races. Characterization of this germplasm is important since landraces may have disease or insect resistance and specific adaptation or consumer traits. In this regard, we also measured nutritional quality of the accessions.

Materials and methods

Plant materials

A total of 365 genotypes were evaluated in this study, including 355 landraces from the "Seeds of Hope"

collection maintained as part of the complete gene bank at ISAR in Rwanda, as well as six varieties from DRC, Rwanda and Uganda and four control genotypes from CIAT that represented the large-seeded types from the Andean gene pool (G19833, G4494) and the small-seeded types from the Mesoamerican or Middle American (MA) gene pool (DOR364, ICA Pijao), where these latter terms are used interchangeably given their similar use in previous literature (Blair et al.2009b; Díaz and Blair 2006). The Central African checks included the varieties Dore de Kirundo (medium-sized yellow seed), Namulenga (medium-sized cream-striped seed) and two versions of Ituri Matata (two varieties with medium-sized white seed) from DRC, Urugezi (medium-sized red mottled seed) from Rwanda and Kanyebwa (medium-sized pink mottled seed) from Uganda. Each of the Central African varieties were uniform in color and similarly each landrace in the collection was of a single seed type as any off types were removed. The "Seeds of Hope" collection can be considered a baseline of common bean diversity that was present in Rwanda in the 1990s and stored in the ISAR gene bank, while the check genotypes and varieties are stored at CIAT.

DNA extraction

DNA extraction involved washing three seeds in distilled water, scarifying to insure uniform germination and pre-germinating in darkness on germination paper in the laboratory using 0.5 mM calcium sulfate (Ca_2SO_4) to avoid pathogen contamination. The first trifoliate leaves of 8-day-old seedlings were collected, frozen at −80 C and lyophilized overnight in a MODULYoD-115 lyophilizer (Thermo Corp). We then proceeded to the DNA extraction step using the method of Mahuku (2004), based on a potassium acetate precipitation. DNA was re-suspended in TE buffer and DNA quality was evaluated on 0.8% agarose gels followed by quantification with Hoescht H 33258 dye on a Hoefer DyNA fluorometer (DNA Quant^{TM} 200). DNA was diluted to 10 ng/μl for genotyping.

Phenotypic characterization

Data were taken on seed color and size according to CIAT scales for commercial classes (CIAT 1987) where 1–9 represent colors from light to dark and S, M and L represent seed sizes. Seed pattern was also recorded according to CIAT pattern classifications. In addition, nutritional quality was assessed as described in Blair et al. (2009a) where seed iron and zinc concentrations were measured by atomic absorption spectrophotometry on a ground sample of seeds of each genotype in the CIAT analytical services laboratory.

Microsatellite amplification

Microsatellite amplification used 30 simple sequence repeat (SSR) markers from a four-color, fluorescent common bean marker kit that we developed in Blair et al. (2009b). Markers were PCR amplified with 6-FAM, NED, PET or TET 5′labeled forward primers and unlabeled reverse primers on PTC-200 thermocyclers (MJ Research, Watertown, MA, USA) with standard reaction mix and cycling conditions as described for the marker kit. Dilution and evaluation of the fluorescent PCR products in multi-mixed panels were also as described in Blair et al. (2009b). PCR products were run on an ABI PRISM 3730xl fragment analyzer (Applied Biosystems, Foster City, CA, USA) at the Cornell University (Biotechnology Resource Center) and allele sizes were determined by comparing with Genescan LIZ500 size standard using GeneMapper v. 3.7 software and the second-order least squares size calling method.

Data analysis

Alleles were binned according to Blair et al. (2009b) and then used to construct a genetic similarity matrix based on the proportion of shared alleles ($PS = \sum PS_i/n$, where i = each locus, n = total number of loci and PS_i= proportion of shared alleles at locus i) calculated in the SAS software (SAS Institute, Cary, NC, USA). This matrix was the basis for a principal coordinate analysis (PCoA) with the goal of determining gene pool assignments, which were represented in a 3D graph using the GRAPH/G3D module from SAS. A dendrogram was then built using UPGMA clustering and a Euclidean distance coefficient from the PCoA in the TREEPLOT subprogram of NTSYS v. 2.10 (Rohlf 2002). Genotype associations were also analyzed with Darwin v. 5.0 software (Perrier et al. 2003) using the simple matching coefficient and neighbor-joining algorithm for confirmation of major nodes and to determine which genotypes were introgressed between gene pools. Finally, PowerMarker v. 3.25 (Liu and Muse 2005) was used to evaluate the number of alleles, Nei's genetic diversity/expected heterozygosity (Nei 87/H_e) and observed heterozygosity (H_o) for each marker across all genotypes and then across the genotypes within each gene pool or introgressed between gene pools.

Results

Allelic diversity

A total of 301 alleles were identified in the entire study giving an average of 10.0 alleles per locus for the

30 microsatellites evaluated, of which 16 were genomic SSRs and 14 were gene-based SSRs (Table 1). The range in allele number was 2–40 with the marker BM187 showing the highest number of alleles followed by BM143, GATs91 and BM157 with 31, 19 and 18 alleles, respectively. All of these markers were genomic. The highest number of alleles found for a gene-based microsatellite was for BMd1 with 16 alleles, followed by PV-gaat001 with 9 alleles. All other gene-based microsatellites had fewer than six alleles. The mean number of alleles for genomic microsatellites was more than twice that of gene-based

microsatellites, with significant difference shown with a t test ($P = 0.0022$). Nei's heterozygosity values were correlated with allele number ($r = 0.671$, $P < 0.01$) and again genomic microsatellites had a significantly higher value than gene-based microsatellites (t test, $P = 0.0067$).

The observed heterozygosity on average was 0.19 across all 30 markers evaluated. All the markers detected a single locus, except for one (BM188) which detected two loci (A and B) distinguished by the pattern of stutter band amplification and size range as described in Blair et al. (2009b). The markers with the highest levels of observed

Table 1 Genetic diversity, observed heterozygosity and number of alleles in 355 Rwandan accessions of common bean, analyzed with 30 microsatellite markers, with number of alleles also shown for Andean and Mesoamerican (MA) group common bean genotypes

| Locus	Genetic diversity, overall	Observed heterozygosity, overall	No. of alleles, overall	No. of alleles, Andean	No. of alleles, MA
Genomic					
AG01	0.32	0.14	6	3	6
BM139	0.52	0.13	11	8	7
BM143	0.89	0.44	31	24	20
BM156	0.88	0.31	18	13	10
BM160	0.66	0.13	16	15	29
BM165	0.85	0.18	9	7	28
BM172	0.51	0.13	9	8	6
BM175	0.78	0.14	12	7	10
BM183	0.81	0.17	14	11	11
BM187	0.92	0.41	40	35	17
BM188-a	0.51	0.01	4	2	4
BM188-b	0.85	0.65	16	12	15
BM201	0.72	0.18	9	6	9
BM205	0.66	0.02	7	5	6
GATs54	0.52	0.00	3	3	3
GATs91	0.91	0.20	19	12	10
Average	0.71	0.20	14.0	10.7	9.8
Gene-based					
BMd01	0.81	0.61	16	13	15
BMd02	0.50	0.12	2	2	2
BMd08	0.53	0.00	5	2	5
BMd16	0.57	0.11	6	5	6
BMd17	0.67	0.30	5	4	5
BMd18	0.32	0.12	5	3	5
BMd20	0.66	0.22	5	3	5
BMd45	0.49	0.09	4	4	2
BMd46	0.50	0.00	4	2	3
BMd47	0.51	0.01	4	3	3
Pv-ctt001	0.72	0.39	5	4	5
Pv-ag003	0.51	0.00	3	1	3
Pv-gaat001	0.56	0.43	9	7	7
Pv-cct001	0.02	0.01	4	2	4
Average	0.53	0.17	5.5	3.9	5.0
Overall average	0.62	0.19	10.2	7.5	7.6

heterozygosity were BMd01 and BM188B, which were used despite being multi-banding in previous studies since they were expected to be single copy in the fluorescent amplifications. The majority of the other microsatellites presented observed heterozygosity values below 0.2, with BM143, BM156, BM187, BMd17, BMd20, Pv-ctt001 and Pv-gaat001 being those with higher levels.

Principal coordinate analysis and landrace associations

The principal coordinate analysis allowed the separation of the genotypes into two gene pool groups with 130 Andean genotypes and 193 MA genotypes. The identity of each group was confirmed by the gene pool controls, which were grouped as expected: G4494 and G19833 with large-seeded genotypes representing the Andean gene pool and DOR364 and ICA Pijao with small-seeded genotypes representing the Mesoamerican or Middle American gene pool (Fig. 1a). Substantial diversity was found in the genotypes in the PCoA analysis and this was distributed both within and between gene pools. Dimension 1 represented within-gene pool diversity in both gene pools, while dimension 2 represented between-gene pool divergence. In this latter dimension, introgression between the gene pools was found

Fig. 1 Principal coordinate analysis of Central African genotypes and their division into gene pools and subgroups. Andean group genotypes are represented by hearts, while Mesoamerican (MA) and inter-gene pool introgression group genotypes are represented by other symbols
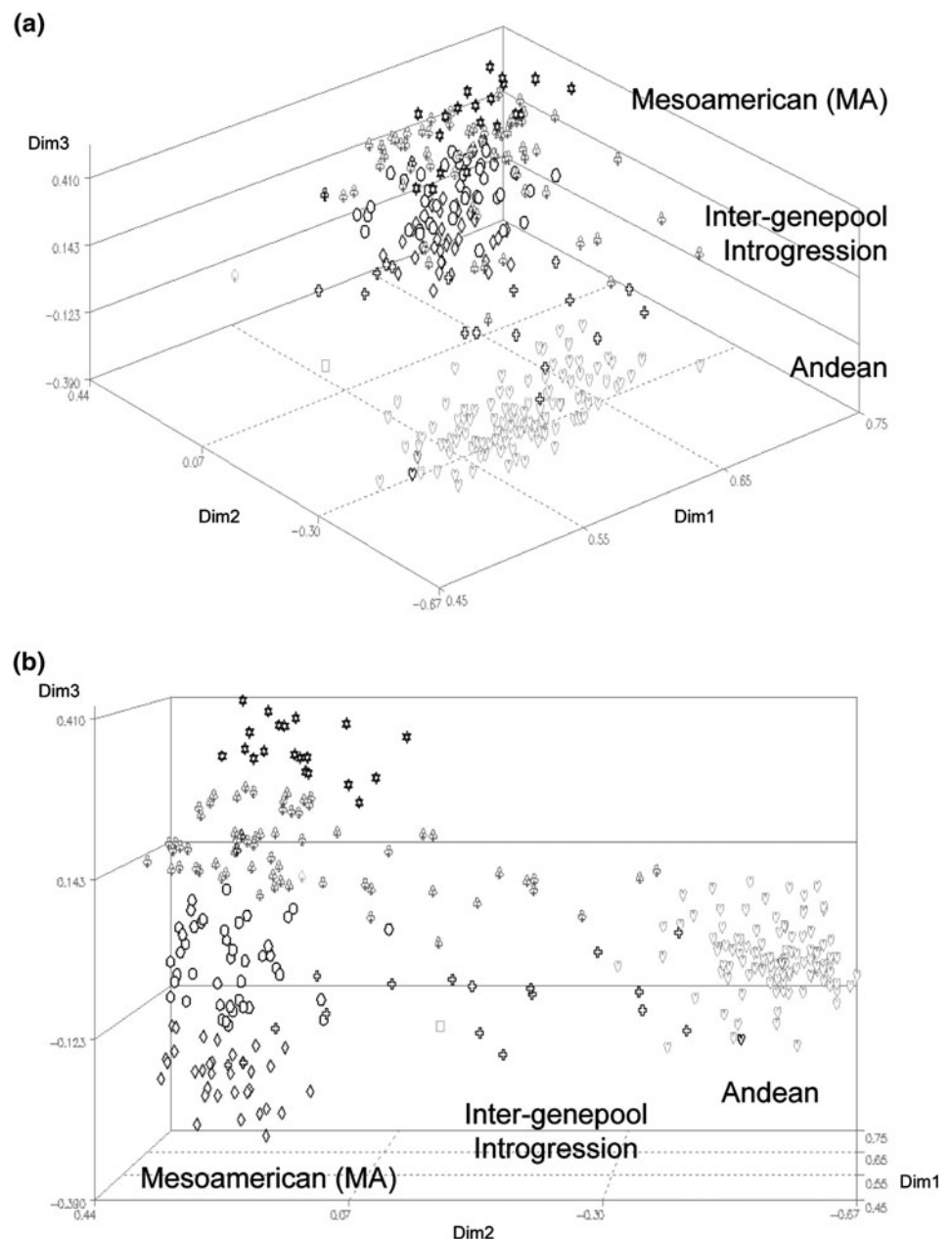
for 32 genotypes that were intermediate between the two groups (Fig. 1b). Finally, dimension 3 represented within-gene pool diversity mainly in the Mesoamerican or Middle American gene pool.

Alternate clustering with Darwin software and neighbor-joining analysis also showed that the Andean and Meso-american/Middle American groups had equivalent amounts of diversity detected with the Dice coefficient (Fig. 2). Interestingly, in this analysis the MA control genotypes, both from race Mesoamerica, were closely associated, while the Andean control genotypes were on different branches of the dendogram, possibly representing differences between race Nueva Granada and race Peru individuals. This analysis also showed the introgressed or intermediate genotypes, some clustering with the Andean group genotypes and some with the MA group genotypes. Finally the varieties from DRC and Uganda were found together with the introgressed Andean (Dore de Kirundo,

Kanyebwa, both Ituri Matata and Urugezi) and introgressed MA groups (Namulenga).

In the dendogram generated with NTSYS based on shared alleles and UPGMA analysis, we also observed the separation of Andean and MA groups and the presence of the control genotypes in their respective clusters (data not shown). Euclidean distance between the 132 genotypes clustering in the Andean group did not surpass 0.27, while for the 208 genotypes clustering in the MA group values went up to 0.40. Introgressed genotypes were less evident as the separation between gene pool groups was at Euclidean distance of 0.85 and all genotypes were assigned to one or the other of the two major clusters. Ten of the genotypes that were intermediate between the gene pools in the PCA or in the neighbor-joining tree described above, and that had mixed seed characteristics, grouped with the Andeans in the UPGMA dendogram, while 22 grouped with the Mesoamerican or Middle American accessions.

Fig. 2 Dendrogram showing the relationships between 365 Central African landraces and varieties as well as primary center controls based on neighbor-joining analysis and simple matching coefficient. Mesoamerican (*MA*) group genotypes are identified with *red* lines (*lower half of the tree*), Andean group genotypes in *green* (*upper half of the tree*) and inter-gene pool introgression groups (*intermediate*) in *blue*. Control genotypes, varieties and race identity are shown in text adjacent to appropriate, *lighter colored lines* for each gene pool
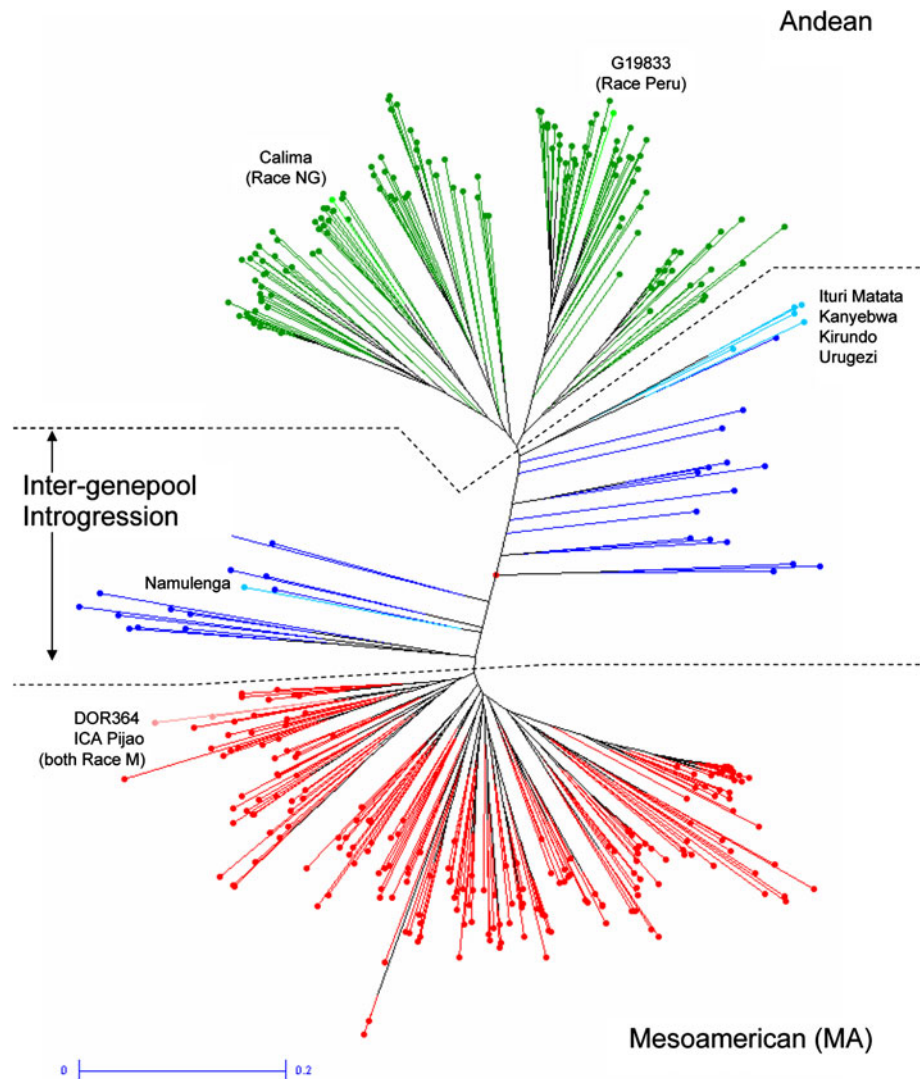
Seed color and pattern traits

Seed color was highly variable in the collection of 355 Rwandan landraces with most landraces being of a single commercial type, but with some within-landrace variability observed. Between-landrace variability was extreme on the other hand, with all seed size and color classes and combinations represented. Red and brown seed colors were the most common especially compared to white, pink, purple or gray, and many landraces had mottled seed rather than other seed coat patterns such as striped, stippled or pinto (Table 2). Cream, yellow and black-seeded landraces were also common as were some pinto types and seed with uniform background, but evident venation. A larger number of Rwandan landraces had seed sizes that were small (<25 g/100 seed) or medium (25 to 40 g/100 seed) compared to large-seeded (>40 g/100 seed) landraces, which only represented 8% of the total. The proportion of medium and small-seeded types was even with 165 and 164 genotypes, respectively, falling into these categories. Only 26 genotypes were large seeded.

Phenotypic diversity in each gene pool

Contingency tests were used to evaluate if the frequency of landraces from each category of seed size or color was associated with gene pool identity or independent of this classification. Purple color was found to be exclusively Andean, while there were no white beans from this gene pool. Pink, brown and black colors were found to be predominantly in the MA group, while cream, yellow and red colors were found in both gene pool groups. In terms of seed pattern, mottled, stippled and pinto seed was found in Rwandan landraces of both gene pools, but, surprisingly, striped seed was more common in the Andean group, although in both gene pool groups most genotypes had non-patterned seed coats. Venation of the seed coat was found for some MA group genotypes, but not in Andean group genotypes. As expected, seed size was correlated with gene pool identity; Andean group landraces were predominantly medium or large seeded and all MA group landraces were either small or medium seeded. In other words, none of the MA group genotypes had large seeds.

The two gene pool groups also differed in terms of nutritional quality traits, with Andean beans having a tendency for higher average seed iron concentration, but significantly lower seed zinc concentration than MA and introgressed-type beans (Table 3). This was also reflected in the difference between the minimum and maximum values for seed mineral concentration, which varied according to gene pools, especially in the case of seed zinc concentration. Meanwhile, the genotypes from the

Table 2 Frequency distribution of seed characteristics found in the Andean and Mesoamerican (MA) gene pool groups and inter-gene pool introgressed genotypes of common bean

Characteristic	Andean	Inter-gene pool introgressed	Mesoamerican (MA)	Total
Seed size				
Large	26	0	0	26
Medium	100	17	48	165
Small	4	15	145	164
Seed color				
1 (white)	0	3	3	6
2 (cream)	31	3	18	52
3 (yellow)	22	9	30	61
4 (brown)	17	4	73	94
5 (pink)	1	1	3	5
6 (red)	54	6	27	87
7 (purple)	2	0	0	2
8 (black)	2	5	35	42
9 (other-gray)	1	1	4	6
Seed coat pattern				
NA (no pattern)	48	21	134	203
J (stippled)	6	3	23	32
M (mottled)	34	4	16	47
P (pinto)	4	3	7	21
R (striped)	38	0	1	39
V (venation)	0	1	12	13
Total	130	32	193	355

Table 3 Average (±SD) and range of seed iron and zinc concentration found in the Andean and Mesoamerican (MA) gene pool groups and inter-gene pool introgressed genotypes

Gene pool	Fe (ppm)				Zn (ppm)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Overall	68.2	8.4	45.3	97.6	34.7	3.4	25.1	49.1
Andean	69.2[A]	8.4	45.8	89.5	33.0[A]	3.1	25.1	44.9
Mesoamerican (MA)	67.4[A]	8.4	45.4	97.5	35.4[B]	3.2	29.2	49.1
Inter-gene pool introgressed	71.0[A]	7.6	59.0	84.4	35.1[B]	3.5	27.9	39.6

Ryan-Einot-Gabriel-Welsch multiple range tests with significance at $P < 0.05$ probability threshold for average seed iron and zinc concentrations in the Andean, Mesoamerican and introgressed groups indicated by superscript letters after the mean values

inter-gene pool introgression group showed the highest average seed iron concentration and similar seed zinc concentration as in MA group beans.

Discussion

Genetic diversity in Central African genotypes

All of the fluorescent markers used to evaluate the Central African varieties were found to be polymorphic and in general the markers detected a large number of alleles compared to previous studies of germplasm from other regions. This can be explained by the large number of landraces evaluated and genetic diversity of the germplasm (a topic to which we will return) or by the high sensitivity of the detection system and the specific markers selected for evaluation. Some of the same markers found to be very polymorphic in the present study, such as BM143, BM187 and GATS91, were also found to be the most variable in previous studies (Blair et al. 2006, 2007, 2009b; Díaz and Blair 2006; Gaitán et al. 2002) and our study used a pre-selected set of 30 markers for panel design from the most polymorphic markers of this previous research.

Compared to other regions, the allelic diversity found in the Central African germplasm was high with an average of ten alleles per locus, but not as high as that found across the entire range of common beans throughout the world. For example, Blair et al. (2009b) found an average of 18.4 alleles per locus in an evaluation of 604 diverse genotypes with 36 gene-based and genomic microsatellites for a worldwide core collection of common beans, while Blair et al. (2006) found an average of only 7.8 alleles per locus in a mini-core collection of 44 diverse parental genotypes. Gomez et al. (2004), evaluating small-seeded beans from Nicaragua found an average of 5.7 alleles across seven genomic markers and 108 individuals from nine collection

sites. As in previous evaluations of diversity, we found it useful to combine genomic microsatellites with gene-based microsatellites, since these latter markers have proven their utility for stable analysis of phylogenetic relationships, especially for gene pool identity (Blair et al. 2003, 2006, 2009b).

Overall genetic diversity (0.62) was found to be especially high in the Central African landraces compared to previous studies of other germplasm sets with less polymorphic nuclear markers such as isoyzmes, RFLPs or RAPDs. For example, Becerra-Velazquez and Gepts (1994), analyzing both gene pools, found genetic diversity of 0.38 with RFLP markers, while Paredes-Lopez and Gepts (1995) found a genetic diversity of only 0.16 with isozyme markers and Beebe et al. (2001) found even lower values with AFLP markers in Andean beans. RAPD markers have generally given intermediate values of diversity (Skroch et al. 1998; Beebe et al. 2000). Meanwhile, in one study with chloroplast markers, genetic diversity was 0.37 (Chacón et al. 2005).

In comparison, microsatellite markers have generally detected the highest levels of genetic diversity of any nuclear marker type for cultivated common beans reaching as high as 0.64 depending on the germplasm analyzed, the detection system used and whether they were landraces or breeding lines from a specific region or from a wider geographical area (Blair et al. 2006; Hanai et al. 2007; Zhang et al. 2008). For example when evaluating the landraces from Nicaragua, the highest number of alleles detected was 13 (Gomez et al. 2004), while when a broader set of landraces from the Mesoamerican or Andean gene pools representing various countries was evaluated then up to 15 and 23 alleles were found for a single locus (Díaz and Blair 2006; Blair et al. 2007). When both gene pools were analyzed together with some of the same markers, up to 50 alleles were found (Blair et al. 2009b). The high diversity found in the Central African landraces is promising for the use of this germplasm in common bean improvement for the region. The high genetic diversity in the Great Lakes region of Central Africa may reflect the selection for favorable adaptation to the range of agro-ecologies found there, to the large number of diseases, insects and abiotic constraints that afflict the crop in this intensively farmed region or to the variability in consumer preferences for various grain types.

High observed heterozygosity in some Rwandan genotypes

Another aspect that was notable about the collection that was analyzed was their high observed heterozygosity (average 0.19) that reflected the conditions under which they were multiplied and their origin as landraces that were

often maintained within seed mixtures by Rwandan farmers (Sperling 2001). By comparison, accessions from the CIAT gene bank analyzed by Blair et al. (2007, 2009b) or Díaz and Blair (2006) had lower levels of observed heterozygosity. The high heterozygosity observed may be also a reflection of the high genetic diversity of the landraces and crossing between landraces, though we selected seeds with similar phenotypic appearance for DNA extraction.

Among the elements that maintain diversity in Central Africa, seed mixing appears to be very important. Seed mixtures are used for various agronomic and cultural reasons (David and Sperling 1999). For example, seed mixtures are often preferred for home cooking or for local markets. Meanwhile, unique seed colors are only selected for sale to urban populations or export to a niche market. Therefore, as a general rule, production is not limited to a single seed type. Rather, a wide range of seed colors and sizes are grown, especially for home consumption. Regular seed sale and exchange between farmers or between regions allows mixtures to maintain their diversity, especially when new seed types are added to a given farmer's seed stock. Furthermore, recombination between varietal components is likely, depending on the season and nick between genotypes. Seed selection from season to season is based mainly on seed quality and blemishes, but not on seed color or size. In summary, the lack of selection based on seed type maintains the diversity generated by mixing or inter-crossing.

Varietal mixtures have advantages in stabilizing the production system since different components have different phenologies in terms of architecture, rooting depth, flowering date and maturity period, which allow at least some plants within the mixture to thrive under different conditions of drought, poor soils or excess moisture (Sperling 1992). This is important under low fertility conditions with fertilizer application only 1.3% of the recommended amounts. Furthermore, production is more evenly distributed in time and space when determinate and indeterminate bean plants are grown together. In addition, varietal mixtures may play a role in suppressing weeds since indeterminate types can compete with other plants helping the determinate plants to grow. Alternatively, the variability in genotypes within a varietal mix prevents disease pathogens or insect pests from getting established. Finally, varietal mixtures can help in marketability among consumers who may prefer to buy mixtures or by providing a seed type for every consumer (Sperling and Loevinsohn 1993).

Gene pools, seed colors and nutritional quality associations

In terms of population structure, the Central African landraces were divided into Andean and Mesoamerican gene pool groups according to the principal coordinate analysis, but many accessions presented evidence of introgression between the gene pools. Molecular separation of common bean accession into gene pools has been observed in a large number of studies, both with genotypes from the primary centers of origin in the Americas (Singh et al. 1991b; Becerra-Velazquez and Gepts 1994; Blair et al. 2006) and in secondary centers of diversity in Europe (Rodiño et al. 2003; Sicard et al. 2005), Asia (Zhang et al. 2008) and parts of Africa (Kairaillah et al. 1990). Seed morphology of the Andean and Mesoamerican group genotypes agreed with the characteristics reported for each gene pool by Singh et al. (1991a) with large or medium seeds for Andean group genotypes and mostly small seeds for Mesoamerican group genotypes.

Previous studies of southern and East African bean germplasm has found that most genotypes tend to be from the Andean gene pool (Martin and Adams 1987; Wortmann et al. 1998) and this was the case for varieties from DRC and Uganda. However, it was surprising that among the Rwandan landraces, more genotypes were from the Mesoamerican gene pool (58.3%) than from the Andean gene pool (37%) with the rest being the result of introgression. The predominance of Mesoamerican types in Rwanda may be a result of a recent increase in root rots to which Andean beans, especially determinate types, are less resistant. Another possibility is that land pressure and the need for higher productivity have led to greater use of Mesoamerican germplasm, since Mesoamerican genotypes are often higher yielding per plant than Andean genotypes. Alternatively, the input of germplasm from breeding programs established in the 1970s through the national programs at ISAR and INERA as well as the long-standing relationship with the CIAT bean program may have influenced the increase in Mesoamerican types (Voss 1992). This certainly was the case with climbing bean introductions, which were of Mesoamerican origin and may have displaced some local varieties (Graf et al. 1991; Sperling and Muyaneza 1995).

In practical terms, Mesoamerican beans may be preferred in some settings because a given weight of Mesoamerican seed plants has a larger area than the same amount of a larger seeded Andean genotype. This is important to farmers who often save a limited amount of seed between seasons. Furthermore, farmers would be interested in selling as much of their production as possible and therefore would not want to save the larger seeded types of which a greater weight would be needed to plant in the same area as a smaller weight of small-seeded types. Finally, introgressed genotypes may have been selected because they often have higher disease resistance and yield potential than pure Andean beans. They also have the medium seed size and quality characteristics of Andean beans, but are often smaller seeded than pure Andean beans resulting in saving of planting time.

The frequency of seed color differences agreed with previous gene pool descriptions (Singh et al. 1991a), but showed high diversity in both the Andean and Mesoamerican group genotypes from Central Africa. For example, black seed was found at a low frequency in the Andean group, while brown seed was also fairly common in this group, which is not typical of Andean landraces in other regions. It was also surprising to find a large number of yellow beans in the Mesoamerican group, since this color was not highly prevalent in this gene pool. Likewise, yellow or mottled beans with a yellow background were found in the inter-gene pool introgressed group. Wortmann et al. (1998) have highlighted previously the prevalence of small red, large red and red mottled seed types in eastern and southern Africa, which were also common in the germplasm analyzed here.

It seems therefore that seed color genes have been extensively selected in both gene pool groups in Central Africa with definite preferences by producers and consumers. In addition, seed colors have perhaps been moved between gene pool groups and commercial classes through hybridization and inter-gene pool introgression. Farmer experience with varietal selection and the importance of bean color and cooking characteristics have been documented in participatory plant breeding programs (Voss 1992; Sperling et al. 1993; Trutmann et al. 1996) and in the overall distribution of beans in eastern and southern Africa (Wortmann et al. 1998). The most popular and representative seed types in Central Africa appear to be red, red mottled, brown, yellow or cream beans, with white and black beans also found but not highly consumed. White and yellow beans command higher prices in urban areas where they are preferred for restaurants, but are more difficult to manage agronomically due to root rot susceptibility.

Nutritionally, the landraces were also found to vary and although this trait might not have been consciously selected, it was notable that Andean beans had higher seed iron concentration but lower seed zinc concentration compared to Mesoamerican beans. The differences between Andean and Mesoamerican or Middle American beans for seed iron and zinc concentrations have been noted before (Islam et al. 2004) and may have to do with loading and distribution of the minerals in the seed (Ariza-Nieto et al. 2007).

Introgression between the gene pools

As mentioned earlier, some genotypes showed inter-gene pool introgression both in their position in the principal coordinate analysis and in their seed characteristics: grouping with one gene pool but having characteristics of the other gene pool. Extensive evidence for introgression between the gene pools has been found by various authors

for beans from different secondary centers of diversity (Durán et al. 2005; Rodiño et al. 2003; Zhang et al. 2008) and also in two studies of the CIAT core collection (Islam et al. 2004; Blair et al. 2009b). Preservation of the products of inter-gene pool introgression is more likely among farmers who produce for home consumption and save the seed of all the segregants resulting from any natural hybridization for planting in subsequent generation. Highly commercial farmers who rouge out off types in the field or remove seed mixture components from their harvests are less likely to preserve these segregants, but subsistence farmers would be likely to save them. Subsistence farmers are also more likely to keep seed mixtures intact and to practice less strict rouging or selection in their landraces, which in turn maintains the high level of diversity and encourages additional rounds of hybridization between divergent genotypes and gene flow between the gene pools.

Inter-gene pool introgression may produce interesting combinations of traits, such as higher adaptability to environmental stresses, diseases or insects, and have surprising benefits such as higher nutritional quality. Indeed, one notable result of this study was to confirm that genotypes showing introgression tend to have higher seed iron concentration as found previously in germplasm evaluations (Islam et al. 2004) and in inter-gene pool crosses where transgressive segregants can have higher seed mineral concentrations than either parent (Blair et al. 2009a).

In summary, this study identified a high level of genetic diversity in landraces and varieties from Central Africa. As such, germplasm from this region is an integral part of the secondary diversity of common beans found in eastern and southern Africa as a whole. This diversity is especially valuable considering the high level of bean production in the region, which, compared to other regions of the world, is among the most concentrated and vitally important to its inhabitants. For example, in Rwanda and Burundi, the population densities are among Africa's highest, reaching 350 people per km^2, and common beans are the major source of protein for this population. In conclusion, common bean is an agriculturally complex crop in Central Africa, and the bean germplasm in the region appears to be a valuable source of diversity for breeding programs that should be preserved and prioritized for crop improvement.

References

Allen DJ, Dessert M, Trutmann P, Voss J (1989) Common beans in Africa and their constraints. In: Schwartz HF, Pastor-Corrales Marcial A (eds) Bean production problems in the tropics, 2nd edn. International Center for Tropical Agriculture, Cali, p 350

Ariza-Nieto M, Blair MW, Welch RM, Glahn RP (2007) Screening of bioavailability patterns in eight bean (Phaseolus vulgaris L.) genotypes using the Caco-2 cell in vitro model. J Agr Food Chem 55:7950–7956

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

Becerra-Velazquez L, Gepts P (1994) RFLP diversity of common bean (Phaseolus vulgaris L.) in its centres of origin. Genome 37:256–263

Beebe S, Skroch P, Tohme J, Duque MC, Pedraza F, Nienhhuis 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

Beebe S, Rengifo J, Gaitán-Solís E, Duque MC, Tohme J (2001) Diversity and origin of Andean landraces of common bean. Crop Sci 41:854–862

Blair MW, Pedraza F, Buendia H, Gaitán-Solis 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

Blair MW, Giraldo MC, Buendía HF, Tovar E, Duque MC, Beebe SE (2006) Microsatellite marker diversity in common bean (Phaseolus vulgaris L.). Theor Appl Genet 113:100–109

Blair MW, Diaz JM, Hidalgo R, Diaz LM, Duque MC (2007) Microsatellite characterization of Andean races of common bean (Phaseolus vulgaris L.). Theor Appl Genet 116:29–43

Blair MW, Astudillo C, Grusak M, Graham R, Beebe S (2009a) Inheritance of seed iron and zinc content in common bean (Phaseolus vulgaris L.). Mol Breed 23:197–207

Blair MW, Díaz LM, Buendia HF, Duque MC (2009b) Genetic diversity, seed size associations and population structure of a core collection of common beans (Phaseolus vulgaris L.). Theor Appl Genet 119:955–973

Broughton WJ, Hernandez G, Blair MW, Beebe SE, Gepts P, Vanderleyden J (2003) Beans (Phaseolus spp.): model food legumes. Plant Soil 252:55–128

Chacón MI, Pickersgill S, Debouck DG (2005) Domestication patterns in common bean (Phaseolus vulgaris L.) and the origin of Mesoamerican and Andean cultivated races. Theor Appl Genet 110:432–444

CIAT (Centro Internacional de Agricultura Tropical) (1987) Standard system for the evaluation of bean germplasm. Schoonhoven AV, Pastor Corrales MA (eds). Cali, Colombia

David S, Sperling L (1999) Improving technology delivery mechanisms: lessons from bean seed systems research in Eastern and Central Africa. Agric Hum Values 16:381–388

Díaz 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

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

Durán LA, Blair MW, Giraldo MC, Machiavelli R, Prophete E, Nin JC, Beaver JS (2005) Morphological and molecular characterization of common bean (Phaseolus vulgaris L.) landraces from the Caribbean. Crop Sci 45:1320–1328

Gaitán 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

Gomez O, Blair MW, Frankow-Lindberg B, Gullberg U (2004) Molecular and phenotypic diversity of common bean landraces from Nicaragua. Crop Sci 4:1412–1418

Graf W, Voss J, Nyabyenda P (1991) Climbing bean introduction in Southern Rwanda. In: Tripp R (ed) Planned Change in Farming Systems: progress in on-farm Research. Wiley, UK, p 182

Hanai LR, de Campos T, Camargo LEA, Benchimol LL, de Souza AP, Melotto M, Carbonell SAM, Chioratto AF, Consoli L, Formighieri EF, Siquiera MF, Tsai SM, Vieira MLC (2007) Development, characterization and comparative analysis of polymorphism at common bean SSR loci isolated from genic and genomic sources. Genome 50:266–277

Islam FM, Beebe S, Muñoz M, Tohme J, Redden RJ, Basford KE (2004) Using molecular markers to assess the effect of introgression on quantitative attributes of common bean in the Andean gene pool. Theor Appl Genet 108:243–252

Kairaillah MM, Adams MW, Sears BB (1990) Mitochondrial polymorphisms of Malawian bean lines: further evidence for two major gene pools. Theor Appl Genet 80:753–761

Kelly V, Mpyisi E, Shingiro E, Nyarwaya JB (2001) Agricultural Intensification in Rwanda: an elusive goal fertilizer use and conservation investments. In: Rwanda (ed.), vol 3. FSRP/DSA Publication

Lamb EM, Hardman LL (1985) Final report of: survey of bean varieties grown in Rwanda. January 1984–June 1985. AID-Rwanda Local Crop Storage Cooperative Research, Rwanda

Liu K, Muse SVB (2005) PowerMarker: integrated analysis environment for genetic marker data. Bioinformatics 21:2128–2129

Mahuku GS (2004) A simple extraction method suitable for PCR based analysis of plant, fungal and bacterial DNA. Plant Mol Biol Rep 22:71–81

Martin GB, Adams MW (1987) Landraces of Phaseolus vulgaris (Fabaceae) in Northern Malawi, I: regional variation. Eco Bot 41:190–203

Paredes-Lopez O, Gepts P (1995) Extensive introgression of middle American germplasm into Chilean common bean cultivars. Genet Resour Crop Evol 42:29–41

Perrier X, Flori A, Bonnot F (2003) Data analysis methods. In: Hamon P, Seguin M, Perrier X, Glaszmann JC (eds) Genetic diversity of cultivated tropical plants. Enfield Science Publishers, Montpellier, pp 43–76

Rodiño AP, Santalla M, De Ron AM, Singh SP (2003) A core collection of common bean from the Iberian Peninsula. Euphytica 131:165–175

Rohlf F (2002) NTSys-pc v.2.10: numerical taxonomy system. Exeter, NY

Sicard D, Nanni L, Porfiri O, Bulfon D, Papa R (2005) Genetic diversity of Phaseolus vulgaris L. and P. coccineus L. landraces in central Italy. Plant Breeding 124:464–472

Singh S, Gepts P, Debouck D (1991a) Races of common bean (Phaseolus vulgaris, Fabaceae). Econ Bot 45:379–396

Singh S, Gutierrez A, Molina A, Urrea C, Gepts P (1991b) Genetic diversity in cultivated common bean. Crop Sci 31:23–29

Skroch PW, Nienhuis J, Beebe S, Tohme J, Pedraza F (1998) Comparison of Mexican common bean (Phaseolus vulgaris L.) core and reserve germplasm collection. Crop Sci 38:488–496

Sperling L (1992) Farmer participation in the development of bean varieties in Rwanda. In: Moock J, Rhoades R (eds) Diversity, farmer knowledge and sustainability. Cornell University Press, Ithaca, p 225

Sperling L (2001) The effect of the civil war on Rwanda's bean seed systems and unusual bean diversity. Biodivers Conserv 10:989–1009

Sperling L, Loevinsohn M (1993) The dynamics of improved bean varieties among small farmers in Rwanda. Agric Syst 41:441–453

Sperling L, Muyaneza S (1995) Intensifying production among smallholder farmer: the impact of improved climbing beans in Rwanda. Afr Crop Sci J 3(1):40–51

Sperling L, Loevinsohn M, Ntabomvura B (1993) Rethinking the farmer's role in plant breeding: local bean experts and on-station selection in Rwanda. Experim Agric 29:509–519

Trutmann P, Voss J, Fairhead J (1996) Local knowledge and farmer perceptions of bean diseases in the Central Africa Highlands. Agric Hum Values 13:112–125

Voss J (1992) Conserving and increasing on-farm genetic diversity: farmer management of varietal bean mixtures in Central Africa. In: Moock JL, Rhoades RE (eds) Diversity, farmer knowledge and sustainability. Cornell University Press, Ithaca

Wortmann CS, Kirkby RA, Eledu CA, Allen DJ (1998) Atlas of common bean (*Phaseolus vulgaris* L.) production in Africa. CIAT Pan-African Bean Research Alliance

Zhang X, Blair MW, Wang S (2008) Genetic diversity of Chinese common bean (*Phaseolus vulgaris* L.) landraces assessed with simple sequence repeat markers. Theor Appl Genet 117:629–640