Identification of the variability of a common bean collection through morphological, physiological, biochemical, and molecular relationships

Identificación de la variabilidad de una colección de fríjol común mediante la relación de parametros morfológicos, fisiológicos, bioquímicos y moleculares

Gustavo A. Ligarreto M.[1] and Orlando Martínez W.[2]

ABSTRACT

The present study was carried out in Corpoica, C.I. Tibaitata (Mosquera, Colombia) at an altitude of 2,540 m a.s.l. in six environments composed of different seasons and at an altitude of 1,485 m a.s.l. in the International Center of Tropical Agriculture (CIAT), Calima, Colombia. Morphological, physiological, biochemical, and molecular descriptors were used to estimate the genetic variability between 36 Colombian bean accessions, of which four were wild and the others cultivated. Diacol Calima (Nueva Granada from the Andean gene pool) and ICA Pijao (Mesoamerican from the mesoamerican gene pool) were used as controls. The combined analysis of the qualitative and quantitative variables was carried out with Gower distance and an unified data matrix with 315 descriptors. The relationships between the genetic distances differentiated the bean collection into two genetic groups: Andean and Mesoamerican. The following groups of characteristics presented high association: total morphology with qualitative morphology (P=0.91), physiological evaluation with grouped evaluation of morphological, quantitative, and physiological characteristics (P=0.91), characterization of isoenzymatic and molecular markers with respect to just molecular markers (P=0.99) and the characterization of all the studied markers in relation to the molecular and isoenzymatic markers (P=0.88).

Key words: grain legumes, germplasm, genetic variation, isoenzymatic, molecular markers.

RESUMEN

El trabajo se realizó en Corpoica, C.I. Tibaitatá (Mosquera, Colombia), a 2.540 msnm, de altitud, en seis ambientes constituidos por diferentes localidades y semestres y a una altidud de 1.485 msnm en el Centro Internacional de Agricultura Tropical (CIAT) en Calima, Colombia. Se usaron descriptores morfológicos, fisiológicos, bioquímicos y moleculares para estimar la variabilidad genética entre 36 accesiones colombianas de fríjol, de las cuales cuatro son silvestres y las otras cultivadas. Se incluyen los testigos Diacol Calima (raza Nueva Granada) e ICA Pijao (raza Mesoamerica), uno de cada acervo genético. El análisis mixto de variables cualitativas y cuantitativas se efectuó con la distancia Gower, para el efecto, en una misma matriz de datos se unieron 315 descriptores. Las relaciones entre las distancias genéticas discriminaron la colección de fríjol en dos grupos genéticos andinos y mesoamericanos. Se presentó alta asociación entre las caracterizaciones siguientes: morfológica total con morfológica cualitativa (P=0,91), evaluación fisiológica con la evaluación conjunta de caracteres morfológicos cuantitativos y fisiológicos (P=0,91), caracterización de marcadores isoenzimáticos más moleculares con respecto a sólo marcadores moleculares (P=0,99) y en la caracterización de todos los marcadores estudiados en relación con los moleculares y moleculares más isoenzimas (P=0,88).

Palabras clave: leguminosas de grano, germoplasma, variación genética, isoenzimas, marcadores moleculares.

Introduction

The beans of the *Phaseolus* genus are considered a priority for conservation because they are a source of protein, vitamins, dietary fiber, minerals, and calories at a low cost for consumers in the diets of the populations of Latin America and Africa (Pérez *et al*., 2011), where almost 75% of the global production is found. The bean is produced by poor farmers and almost always in crops that are associated with corn in some areas in Antioquia, Nariño, Cundinamarca and Boyaca (Vélez-Vega *et al*., 2007). On average, production is low, around 500 kg ha^{-1} in Latin America with sales in the local markets and 700 kg ha^{-1} in Africa with subsistence systems (Pachón *et al*., 2009).

For the evaluation of genetic diversity, rapid scientific advances, especially in molecular biology and informative technologies, have introduced new techniques for the conservation and use of phytogenetic resources, as well as their broad application, which offers new opportunities that increase the efficiency of food production (Franco and Hidalgo, 2003).

Studies with different types of morphological, biochemical, and molecular markers on the bean have confirmed

the presence of two centers of domestication in Central America and in the southern Andes (Delgado-Salinas, 2014). The variability of phaseolins seems to indicate a third domestication center in the northern Andes, particularly in Colombia (Hernández-López *et al.*, 2013). Recent reports suggest the Valle de Oaxaca as a Mesoamerican domestication area and area for diversity of the *Phaseolus* genus with a great amount of diversity in the bean plants on the farms of the producers (Worthington *et al.*, 2012). According to Kwak *et al.* (2012), the pattern of domestication in *Phaseolus* is one of multiple occurrences, involving several species (five domesticated species), different gene pools within species (Andean and Mesoamerican gene pools) and repeated selection of the same trait within gene pools. This situation may be due to the relatively young age of the genus amounting to some 4 million years.

Throughout the world, the availability of germplasm is seriously limited due to the lack of information relative to the genetic potential, passport data and inventory. In addition, the evaluation data of the collections is not completely analyzed and the accessions are almost always classified by similar characteristics within homogenous categories (Cole-Rodgers *et al.*, 1997). The methods of statistical analysis of diversity provide vital information on the aggregated value in germplasms and their use. However, when grouping analysis is going to be carried out, inadequacies are generated, such as the type of variable to be used: quantitative, discrete, ordinal, and/or multi-state, due to the fact that each of these has a significant difference within the analysis (Franco and Hidalgo, 2003). The measurement of the distance used in the combined analysis depends on the types of variables and the measurement scale (Franco *et al.*, 1997).

In the present study, a Colombian bean collection was used as a reference for processes of genetic improvement for adaptation, absent the knowledge on its genotypical and phenotypical characteristics. The objective was to integrate different types of descriptors, morphological, physiological and genetic, that allow for the enrichment of the knowledge of the germplasm bank of the country's species, promoting their conservation and use. Isoenzymatic markers and RAPDs were used because they have proved useful in studies on the diversity and the classification of bean types for genetic archives.

Materials and methods

Germplasm

This study included 29 cultivated common bean accessions, of which four corresponded to improved varieties such as Diacol Andino, ICA Tundama, ICA Cerinza and L-34400; two controls: Diacol Calima (Nueva Granada race) and ICA Pijao (Mesoamerican race), four wild ones and the uncultivated Argentina 2 (Tab. 1). Different numbers of accession were used due to the need to include different accessions as reference patterns per characterization. The accessions form part of the germplasm that is administered by the Corporación Colombiana de Investigación Agropecuaria-Corpoica, located at the Tibaitata Research Center (Mosquera, Colombia) and were randomly selected from this collection of 165 shrub-type and 110 semi-climbing type accessions.

TABLE 1. Bean accessions from the Colombian collection.

CIAT number identification	Growth habit	Pool
G4698 Tolima 16	3	M
G4674 Nariño 7 A	3	M
-o- * Tolima 17-2	1	A
G4544 Antioquia 20	2	A
G11576 Peru 5	3	A
G4644 Cundinamarca 148	1	A
-o- * Mexico 497	1	A
G4543 Antioquia 19	1	A
G4700 Tolima 16 B	3	M
G4551 Antioquia 27 A	1	A
G7895 Peru 40	1	A
G4610 Cauca 34	3	M
G4545 Antioquia 21	3	M
G4648 Huila 5	3	M
-o- * Tolima 17-1	2	M
-o- * Mexico 496	1	A
-o- * Mexico 171	1	A
G4706 Boyaca 18	1	A
G11525 Peru 154	1	A
G4706 Tolima 43	1	A
G5772 Diacol Andino	1	A
G14016 ICA Tundama	2	A
-o- Line 34400	1	A
-o- ICA Cerinza	1	A
G11292 Argentina 1	2	M
G11293 Bolivia 7	1	A
-o- * Peru 224	3	A
G4679 Nariño 12	1	A
G4534 Antioquia 10	1	A
G3729 Argentina 2	1	A
G5773 ICA Pijao	2	M
G4435 Diacol Calima	1	A
G11034 Silvestre Mexicano (DGD 414)	4	M
G19890 Silvestre Argentino (DGD 626)	4	A
G21244 Silvestre Peruano (DGD 1956)	4	A
G24405 Silvestre Colombiano (OT 114)	4	A

Growth habit: 1, determinate shrubby; 2, indeterminate shrubby e; 3, indeterminate prostrated; 4, indeterminate climbing.
Pool according to morphological description, A, Andean; M, Meso-American (Ligarreto and Ocampo, 2012).
* Accessions not found in CIAT.

Morphological and physiological descriptors

The seeds were sown in Corpoica, C.I. Tibaitata (Mosquera, Colombia), located at 4°42' N and 72°12' W, at an altitude of 2,540 m a.s.l., in six environments composed of different seasons, with precipitation per crop cycle of 149, 242, 267, 292, 396 and 473 mm and an average daily temperature between 12.8 and 13.5°C; and in Calima Darien of the CIAT (Calima, Colombia), located at 3°56' N and 76°31' W, at an altitude of 1,485 m a.s.l., with a precipitation of 495 mm per crop cycle and an average daily temperature of 20°C. A random complete block design was used with 29 accessions and four repetitions. The experimental unit had an area of 9.6 m^2 and four 4 m-long furrows at a distance of 0.6 m between them, with a distance of 10 cm between the plants.

The data were collected for each environment with 22 quantitative characteristics and 42 qualitative characteristics (Tab. 2) (Muñoz *et al.*, 1993). The six continuous variables

TABLE 2. Morphological and physiological variables evaluated in the Colombian bean collection.

Qualitative morphology	Quantitative morphology	Physiological
Cotyledon colors	Emergence days	Total leaf area
Hypocotyl color	Hypocotyl length	Root dry weight
Rib color of the primary leaves	Epycotyllength	Stem dry weight
Branch color	Primary leaf width	Leaf dry weight
Standard limb color	Primary leaf length	Floral dry weight
Standard limb color distribution	Length of central leaflet (4° node)	Pod dry weight
Flower veining	Central foilage width (4° node)	Seed dry weight
Standard collar color	Principal stem length	Total dry weight
Standard collar color distribution	Cover height	Chlorophyll a
Calyx color	Number of nodes	Proportion of clorophyll a
Bracteole color	Vein length	Chlorophyll b
Bracteole formation	Vein width	Proportion of clorophyll b
Bracteole size	Lenght of the vein tips	Total chlorophyll
Growth habit	Veins per plant	Proportion of dry root
Stem color	Seeds per vein	Proportion of dry stem
Stem pubescence	Foliar area of central foilage	Proportion of dry leaf
Leaf color	Yield per plant	Proportion of dry flowers
Central foilage formation	Yield per lot	Proportion of dry pod
Ramification type	Days to flowering	Proportion of dry seeds
Turnover	Days to physiological maturity	Net assimilation rate (NAR)
Reaction to wear	Days to harvest	Relative growth rate (RGR)
Reaction to mildew	Weight of 100 seeds	Leaf area ratio (LAR)
Reaction to Chrysomelidae		Specific leaf area (SLA)
Vein color		Leaf weight ratio (LWR)
Vein color distribution		Leaf area duration (LAD)
Transversal vein cut formation		Stomatic resistence
Plant vein distribution		Radiation
Vein color		Leaf temperature
Vein color distribution		Photosynthesis
Vein profiles		Conductance of CO_2
Vein tip type		Internal conductance to CO_2
Degree of curvature in the tip vein		Transpiration
Curvature direction of the tip		Absorbance at 649
Vein consistency		Absorbance at 665
Primary seed color		
Secondary seed color		
Secondary color distribution		
Head aspect		
Seed veining		
Color surrounding the hilum		
Seed formation		
Genetic archive		

reported by Ligarreto (2003) were considered, such as high heritability, for the description of the morpho-agronomical variability of the bean to know: the weight of 100 seeds, the length of the pods, the number of pods and nodes, the length of the pod apexes, and the days needed for physiological maturity. As a result, a matrix of misbalanced data was generated because the characteristics could not always be measured in the seven environments. The measurements were taken individually in 10 plants that were in competition in each parcel, except for the yield calculation, which was obtained for the entire plot. The valuation of the different states of the qualitative variables was carried out based on the descriptors developed by CIAT (Muñoz *et al.,* 1993). The states of the descriptors in the evaluation of pests and diseases were calculated according to the scales of the standardized systems for the evaluation of bean germplasm, described by CIAT (Van Schoonhoven and Pastor-Corrales, 1991).

The physiological variables dry weight (DW) and foliar area were evaluated in the environments for one plant per plot, starting at 17 d after emergence of the crop, when the plants entered the primary leaves stage, until the physiological maturity stage. Measurements of 34 variables were taken every 8 d for a total of 22 readings. Based on the DW and foliar area data, the NAR (net assimilation rate), RGR (relative growth rate), LAR (leaf area ratio), LAD (leaf area duration), SLA (specific leaf area) and LWR (leaf weight ratio) growth indices were calculated following the traditional methods proposed by Radford (1967). The dynamic of the dry material expressed as a percentage was calculated based on the DW accumulated over time per organ for each one of the 29 evaluated bean accessions (Tab. 2).

The leaf area was measured in fresh material in all the leaves of the plants using a Foliar-3000 (LICOR, Lincoln, NE) device; afterwards, the material was dried in a Mettler PE 600 (EquipNet, San Francisco, CA) oven at 70°C for 48 h to attain constant weight and determine total DW and the organ specific DW.

Measurements of net photosynthetic gas exchange and stomatal conductance were taken on a fully expanded leaf with a portable photosynthesis system (LI-6250; LICOR, Lincoln, NE) device between the hours of 9:00 and 12:00 HR using a trifoliate from the upper portion of one plant per treatment in the reproductive phase, the phenological flowering stages (R6) and the formation of veins phase (R7). The a, b and total chlorophyll contents were determined by the spectrophotometry of V4, R6 and R7 stage plant extracts at an absorbance of 649 and 665 nm.

Isoenzyme tests

The isoenzymatic analysis was carried out with the tissue extracts of the primary leaves and roots obtained from 10-d-old plantlets, taking into account the modifications established by CIAT (1996). Twelve isoenzymatic systems were tested to know: αβ-EST, GOT, αβ-ACP, DIA, PRX, ASD, 6-PGDH, MDH, IDH, ME, PGI and PGM. The isoenzymatic loci, with respect to their alleles, were identified as suggested by Koenig and Gepts (1989); the most common allele was identified with a value of 100 and the others were measured by the difference in millimeters from the four wild bean cultivars used as the standard (Ligarreto and Ocampo, 2012).

RAPD molecular markers

The method described by Dellaporta *et al.* (1983) was used to extract the DNA: 5 g young foliar tissue samples were collected with liquid nitrogen, the DNA concentration was qualified with a TKO 100 fluorometer (Hoefer, Scientific Instruments, San Francisco, CA), and then the samples were suspended at a volume of 500 μL, with a final concentration of 10 ng μL^{-1}. 13 primers or markers were tested: OV10, O20, P7, W6, X11, G10, K12, L4, AA19, AD12, AK6, H19 and W12 from Operon Technologies, previously selected in the CIAT laboratory due to their high polymorphism in bean species. The PCR reaction was carried out in a PTC-100 thermocycler (MJ Research, Waltham, Massachusetts). The amplification of the products was done by electrophoresis in 1.5% p/v agarose gel, with a 280 V current and a resistance of between 80-90 mA; ethidium bromide dying was used for the visualization. Low UV light was used for the photography.

Statistical methods

The quantitative agronomic and physiological variables were processed with the program: SAS®, version 9.4 (SAS Institute, Cary NC). The quantitative variables were processed with the Princomp procedure in order to select a suitable number of components ($\lambda \geq 1$) that estimated the variance and the contribution to the total variability suggested by the collected data.

In order to establish the similitude between the accessions due to the quantitative variables, Euclidean distance was used, defined as: $D^2_{ij} = \sum^p_{k=1}(x_{ik} - x_{jk})^2$, which represented a geometric measurement between the evaluated characteristics. The combined analysis of all the qualitative and quantitative variables was carried out with Gower distance (d) (Gower, 1971), which provided values between 0 (maximum similitude s=1) and 1 (minimum similitude s=0) with d=1-s (Franco *et al.,* 1997). For the effect, the descriptors

were combined in a data matrix as follows: 22 quantitative morphological variables, 42 qualitative morphological variables, of which 13 were binary and 29 were nominal, 34 quantitative physiological variables, 43 biochemical binary markers and 174 binary molecular variables.

The statistical packages SYN-TAX®, version 5.0 (Podani, 1993), and NTSYS®, version 2.0 (Exeter Software, Setauket-East Setauket, NY) (Rohlf, 1998) were used to construct dendrograms that provided information on the nearness and dissimilitude between the accessions through the use of combined analysis with Euclidean and Gower distances. To determine the degree of association between the different evaluation and characterization types, a correlation analysis was carried out using the distances between the bean accessions.

Results

Variability based on the morphological and physiological characteristics

The principal components analysis for the 22 morphological characteristics and the 34 physiological quantitative characteristics for the 29 bean accessions in seven environments indicated that the first 12 components, with a characteristic value of $\lambda \geq 1$, represented 88.80% of the variation, explained for each component, and the sum was the accumulated proportion of the total variance.

The variables with a higher absolute value within each vector were those that contributed more to the total variation, as a consequence, the more important characteristics in the first vector dimension were the number of nodes, the days for floration and for harvest maturity, and the number of pods per plant as morphological characteristics and the DW of the pods (5) and foliar area ratio as physiological characteristics. The second component was of a physiological order, in which the foliar area of the plantlets and plants and total DW of the plants were important. The third and fourth components were composed of chlorophyll variables, the fifth component was expressed as a high stomatic resistance magnitude, and the yield per parcel, the specific foliar area, the radiation, and the leaf temperature were the sixth component.

The seventh component was composed of the bean DW and the foliar weight ratio; the eighth component corresponded to the pod width; the ninth component was photosynthesis; the tenth component expressed the pod length; the eleventh component corresponded to the CO_2 conductance; and the twelfth component was related to the number of seeds per pod.

The above description of the components suggests that the yield, the yield components, the growth habit represented by the number of nodes, the precocity, the leaf area, the total DW, and the indices of growth: leaf area ratio and specific leaf area, photosynthesis, and the chlorophyll variables were the descriptors that contributed the most to the discrimination of the bean genotypes; the numerical data of these components are not shown in this article but their direct results are in the dendrogram of Fig. 1.

There was a notable, large group of 17 cultivars of an Andean origin in the products of the morphological and physiological ratios, coming from Colombia, Peru, and Bolivia; along with a group of seven materials that represented the Mesoamerican group that came from Colombia and Argentina; and a third group of the northern Andes with an origin center in Peru, integrated by Peru 224, Peru 5 and Tundama with one of their progenitors: Peru 5; with Cauca 34 and L-34400, with growth habits III and I, respectively, far from the other cultivar groups due to their physiological behavior.

### Classification based on the isoenzymatic and RAPD markers

The cultivated and wild bean accessions demonstrated variation in 10 tested enzymatic systems: αβ-EST, GOT, αβ-ACP, DIA, PRX, ASD, 6-PGDH, MDH, IDH and ME, and monomorphism in the systems: PGI and PGM. The 12 isoenzymatic systems presented 19 bands of activity, of which 14 (74%) were polymorphic loci with 43 combinations of different bands.

In the RAPD molecular markers, the 13 primers that were used produced 174 visualized polymorphic bands in the samples of the studied bean accessions, with an average of 13.38 bands per primer and a polymorphism percentage of 97%. The 43 binary variables, products of the isoenzymatic markers, and the 174 RAPD polymorphic variables were used to create a Jaccard distance matrix that related accession pairs with a dendrogram (Fig. 2).

In plotting the relationships between the accessions in Fig. 2, two dissimilar groups were visualized due to their genetic grouping: the first group contained 10 Mesoamerican accessions, including two controls: Pijao, which is a cultivated bean typical of this group with S phaseolin and the M type of Central America; and one was the wild accession of Mexico (DGD-414) with phaseolin M10, characteristic of some
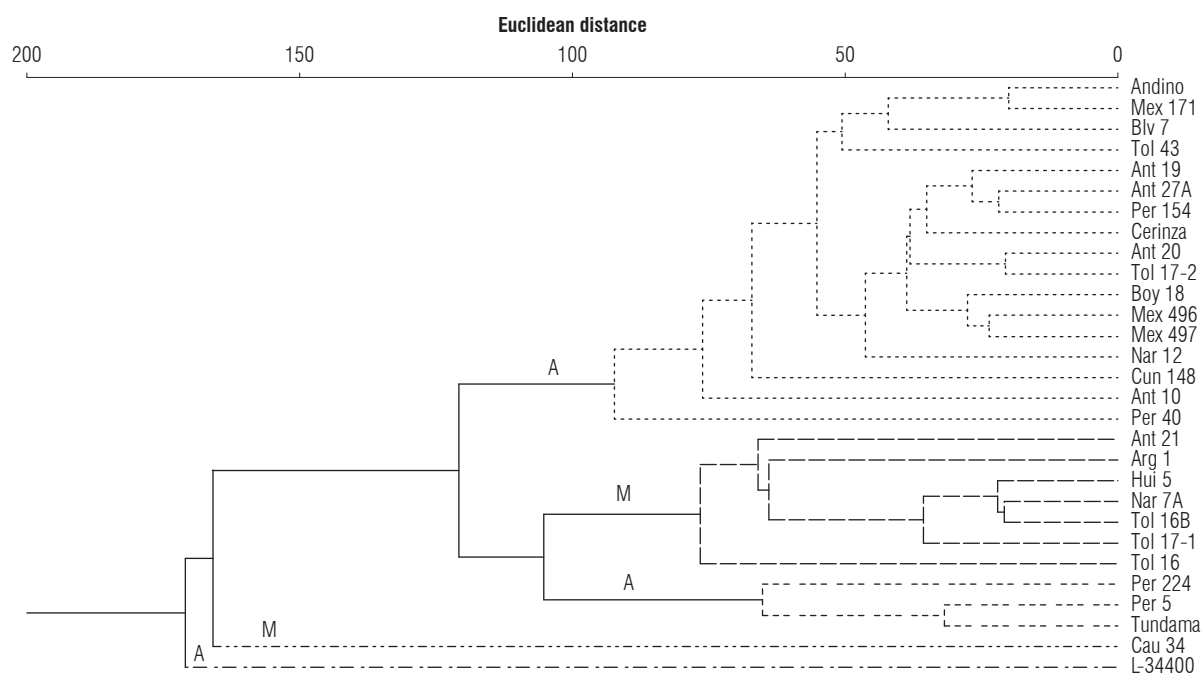
FIGURE 1. Morpho-physiological relationships of 29 bean accessions. Groups: A, Andean; M, Mesoamerican.
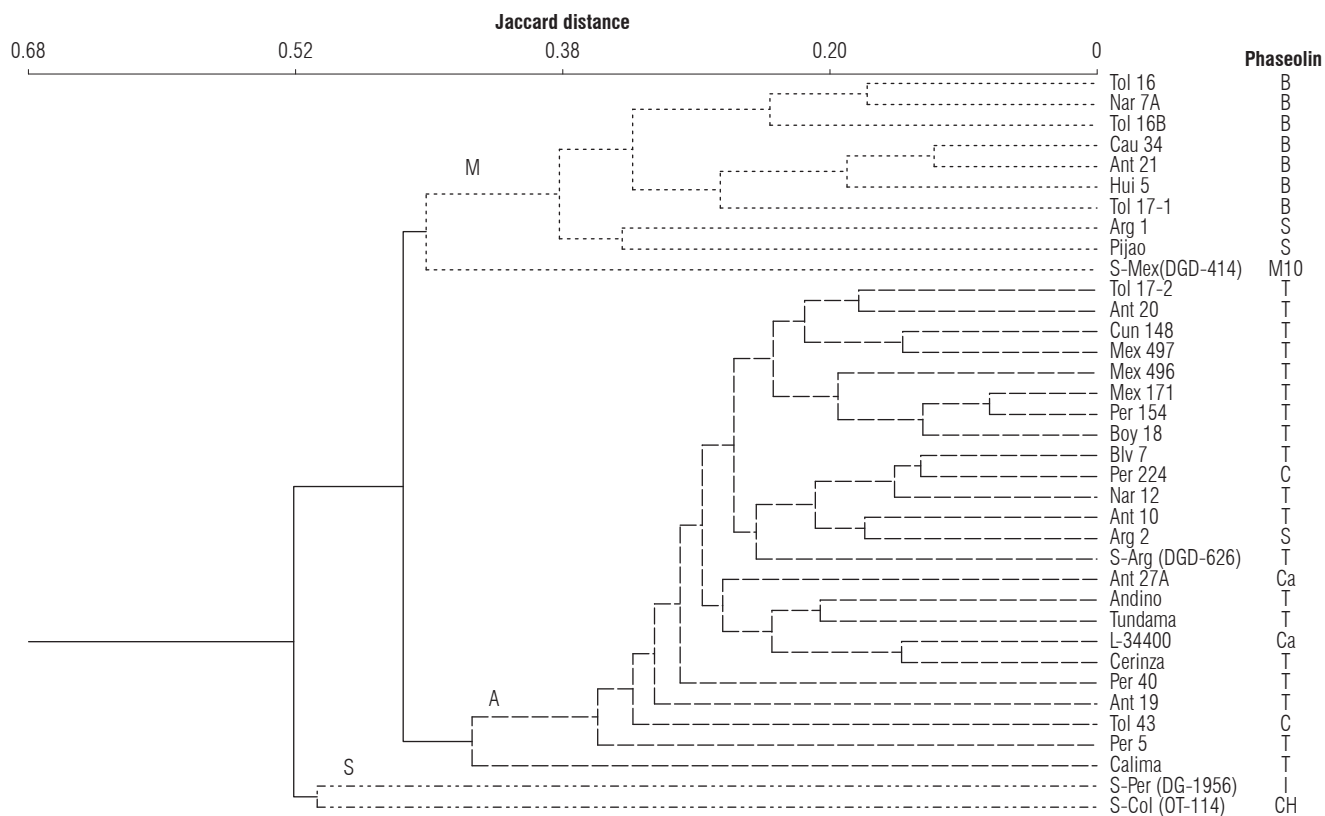
FIGURE 2. Genetic analysis of 36 bean accessions based on RAPD and isoenzymatic patterns. Groups: A, Andean; M, Mesoamerican, and S, wild. Phaseolin: B, Boyaca 22; S, Sanilac; T, Tendergreen; C, Contender; Ca, Durango race *"ojo de cabra"*; I and CH, wild.

wild Mesoamerican accessions. The cultivated accessions of this group came from Colombia and Argentina and are characterized by their small, red and cream colored seeds.

The second group was composed of 24 cultivars, two of which were from the southern Andes of Argentina (one is cultivated and presented S type phaseolin and Andean

morphological patterns; the other one was wild with T phaseolin). The other materials are cultivated in the northern Andes, include the Calima control, representative of the Nueva Granada race, and are characterized by large, red, cream, and cream-speckled red colored seeds and T, C, and Ca type phaseolins.

The polymorphic genetic markers also differentiated a group composed of two wild accessions of the Andean group from Colombia and Peru that possessed CH and I type phaseolins, respectively, and that were found close to the second group of Andean group cultivars.

**Variability based on different types of markers**

The relationships between the genetic distances for the bean collection based on the 315 descriptors are seen in the dendrogram of Fig. 3, where it is clearly seen that the two groups are differentiated both from the morphophysiological view and the biochemical and molecular view.

A genetic group combined of all the quantitative and qualitative variables of the different markers was composed of 22 Andean accessions that included genotypes of the southern and northern Andes with a predominance of seed protein pattern types T and C, and by eight cultivars of the Mesoamerican group, subgrouped by the B type phaseolin for the materials from Colombia and one accession of Argentina with S phaseolin (Fig. 3).

In the analysis of the association between the analog matrices of the distances of the different characterization modulations of the bean collection, the correlation coefficient (ρ) started to be significant when it was equal to or greater than 0.34 and highly significant at ρ=0.46 (Tab. 3). There was high association between the total morphological characterization with the qualitative morphological characterization ρ=0.91, measured with Gower distance; the physiological evaluation with the evaluation together with the physiological and quantitative morphological variables (ρ=0.91), in units of Euclidean distance; the characterization of isoenzymatic markers together with RAPD, with respect to just RAPD markers (ρ=0.99), compared by Jaccard distance; and the characterization of all the studied markers with respect to the RAPD markers and the RAPD and the isoenzymatic markers ρ=0.88, quantified by Gower distance.

Lower degrees of association and high significance were seen, among others, in the following morphological and physiological descriptors: quantitative morphology against morphology for six quantitative variables with ρ=0.55, total morphology with quantitative morphology ρ=0.78, and quantitative morphology associated with quantitative morphology and physiology with ρ=0.60.

The following relationships presented significant degrees of association ($P \leq 0.05$): total characterization with total morphology ρ=0.42, total characterization with physiological
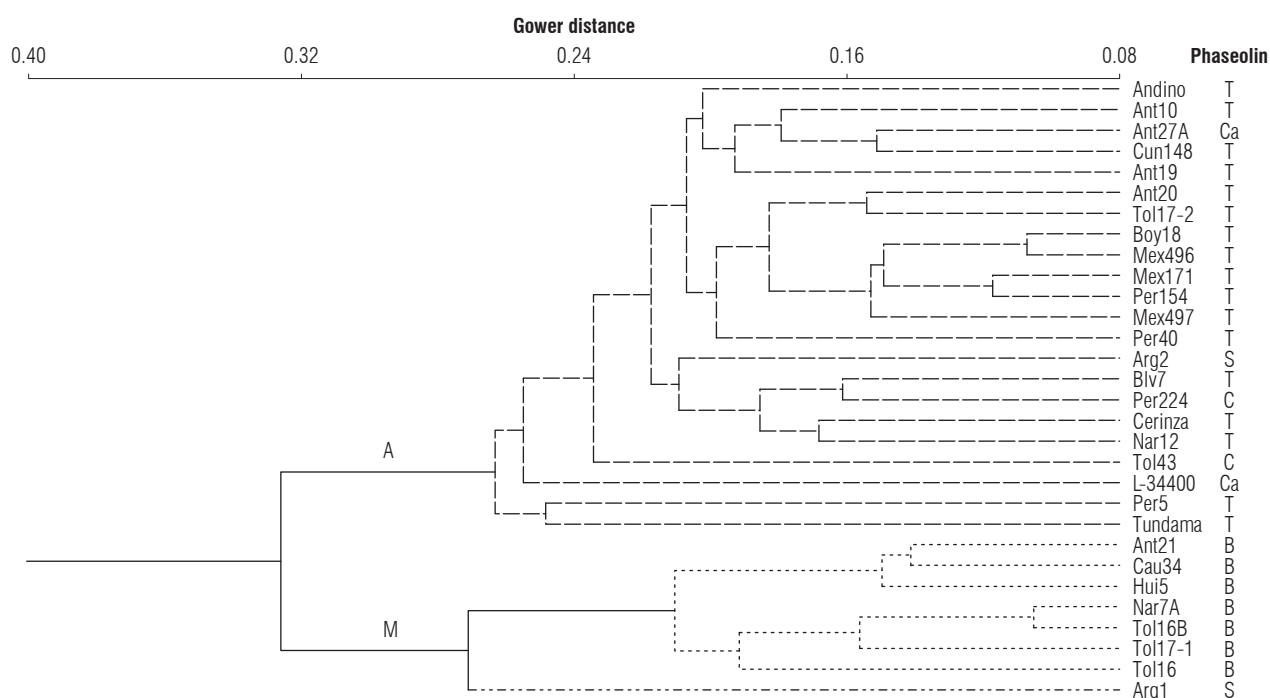
FIGURE 3. Genetic relationship of bean accessions as generated by a combined characterization with morphological, physiological, biochemical, and molecular markers. Groups: A = Andean, M = Mesoamerican. Phaseolin: B, Boyaca 22; S, Sanilac; T, Tendergreen; C, Contender; Ca, Durango race *"ojo de cabra"*.

TABLE 3. Degrees of association between the different types of characterization and measurements of dissimilitude in the bean collection.

Characterization types	Quantitative morphology	Morphology (six quantitative variables)	Qualitative morphology	Total morphology	Physiology	Quantitative morphology and physiology	Isoenzymatic	RAPDs	Isoenzymatic and RAPDs	Total characterization
Distances	Euclidean	Euclidean	Gower	Gower	Euclidean	Euclidean	Jaccard	Jaccard	Jaccard	Jaccard
Quantitative morphology	1									
Morphlology (six quantitative variables)	0.55**	1								
Qualitative morphology	0.48**	0.11	1							
Total morphology	0.78**	0.34*	0.91**	1						
Physiology	0.33	0.19	0.00	0.2	1					
Quantitative morphology and physiology	0.69**	0.47**	0.15	0.46*	0.91**	1				
Isoenzymatic	0.19	0.22	-0.11	0.03	0.02	0.12	1			
RAPDs	0.25	0.54**	-0.16	0.06	0.13	0.27	0.45*	1		
Isoenzymatic and RAPDs	0.27	0.53**	-0.15	0.07	0.11	0.26	0.59**	0.99**	1	
Total characterization	0.56**	0.60**	0.15	0.42*	0.39*	0.56**	0.50**	0.88**	0.88**	1

* Significant ($P \leq 0.05$). ** Highly significant ($P \leq 0.01$). Number of observations=30.

evaluation ρ=0.39, and molecular markers verses isoenzymatic ones with ρ=0.60; and a low degree of correlation ρ=0.15 between total characterization and quantitative morphological characterization, without a statistical significance; and an equally low and negative correlation for the latter with the isoenzymatic and molecular patterns, which, at the same time, presented an association of ρ=-0.16 and ρ=-0.11, respectively. On the other hand, the high correspondence values between the six quantitative variables with the total characterization led to the idea that the selection for these variables is being carried out indirectly by some genetic constitution.

Discussion

For the bean, the first studies on genetic diversity defined two genetic groups in terms of a simple locus that corresponded to phaseolin (Gepts *et al.*, 1986). Later, the isoenzymatic data demonstrated that the separation between the groups is not limited to a simple locus or region of chromosomes, but that there are differences between those of Central America and those of the Andes with respect to morpho-agronomic, adaptation, and disease resistance characteristics (Singh *et al.*, 1991a, Ligarreto and Ocampo, 2012). However, as physiological and morphological characteristics are expressed phenotyipcally, the accessions can be similar in these characteristics but different in their genomes and can also present dissimilarity in the phenotypic characteristics and have genetic nearness, as in the cultivars with growth habit I that can be seen in the two genetic groups (Cole-Rodgers *et al.,* 1997).

The molecular and isoenzymatic markers allow for the characterization of the accessions by their genotype, the establishment of relationships of evolution and the revealing of genetic diversity levels. In this sense, integrating the different types of morphological, isoenzymatic, and molecular markers as well as physiological variables generates complementary information and provides aggregated value for the genetic resources (Singh *et al.*, 1991b; Cole-Rodgers *et al.*, 1997).

The results of the multivariate analysis of the quantitative physiological and morphological characteristics indicated that the number of nodes, the time until floration, the DW of the grains, of the pods, and of the total plant, the foliar area, and the foliar area ratio contributed the most to the degree of discrimination of the cultivars into the Andean and Mesoamerican genetic groups, corroborating the findings of Singh *et al.* (1991a,b). According to Gepts *et al.* (1986) and Singh *et al.* (1991c), the separation between the two genetic groups results from the divergence of domestication and wild populations with specific geographic and ecological adaptations for the subgroups within each group formed by the types. While, Worthington *et al.* (2012), recently suggested that the Valle de Oaxaca in Mexico is an area of domestication and diversity for the *Phaseolus* genus.

The large amount of morphological and physiological variability represented by the dendrogram of Fig. 1 is the product of the evolution of the accessions of each group on the farms of the producers and the obtainment of varieties

generated from the use of a relatively broad genetic base of cultivated varieties. The fact that various accessions of Colombia are grouped as Mesoamerican represents the contacts that were made between the two large centers of genetic diversity of the bean in the geographic zone of Colombia during the different domestication processes, as well as the domestication of the bean from the two pools in independent events (Kwak and Gepts, 2009), providing repeated selections for the generation of changes in the morphological and physiological markers in the plants within the same genetic pool, as in the selection towards plants that are more compact in growth pattern through the use of generally herbaceous wild relatives.

The nearness of the Peruvian bean groups, Peru 5, Peru 224 and Tundama, derived from Peru 5, with the Meso-americans is due to common characteristics, indeterminate growth habit, high total dry weight and foliar area with the determined Andean accessions that, at the same time, originated variations for other derived physiological parameters, such as the growth indices. Also, the nearness between these accessions can be understood because *P. vulgaris* of northern Peru and Ecuador represents a fraction of the genetic diversity that migrated from central Mexico in ancient times, which at the same time, shared physiological and morphological markers (Bitocchi *et al.*, 2012).

The separation of the Cauca 34 (Mesoamerican group) and the L-34400 (Andean group) accessions by the two genetic groups in Fig. 1 is due to the fact that L-34400, a typical Andean accession, demonstrated a large increase in the proportion of leaf area ratio for each unit of weight (LAR), expressed as the ratio between the total leaf area and the total plant dry weight, at the start of the R7 phenological phase, as the Cauca 34 cultivar of the Mesoamerican group did between stages R7 and R8 of the formation and filling of the pods, a behavior that coincided with the rapid decrease in the tendency of their corresponding Andean and Mesoamerican groups, which presented maximum LAR peaks in the floration stage (R6 stage), a behavior described by Ligarreto (2013).

The dynamic of the introduction of wild *P. vulgaris* into cultivated accessions has allowed for the generation of genetic variation through hybridization, which is important in the selection of varieties for agriculture, which have been used for self-consumption. It is common in some regions of Colombia to find low subsistence systems that sow with the use of "jumbled" or mixed beans with different sizes, colors, and life cycles that provide various harvests in different periods, an aspect that has resulted in crossover between the levels of complexity and the subsequent maintenance of crossed populations on farms.

The improved cultivars: Diacol Andino (Sánchez x Estrada Rosado), ICA Tundama, ICA Cerinza (Ant. 10 Algarrobo x L 3043 and Ant. 8 Uribe Redondo x Ant. 26 Sanchez) and L-34400 (L 34104 x ICA Cerinza) present morpho-physiological divergences that allow for their placement in different genetic combinations (Fig. 1) and for the indication that there are no losses in diversity through the generation of new varieties, which can be increased by the combination of incorporated favorable genes of adapted progenitors that are traditionally cultivated in the production zones (Voysest *et al.*, 1994). Certain grain types have predominated in the generation of new varieties that have a demand in local markets, principally large ones, as well as in the low selection of maize-bean associated crop systems.

The contrasts in the levels of variability between the morphological and biochemical and molecular markers is due to the characteristics that are expressed in many genes. The morphological variables of high heritability constitute an atypical sample of the majority of the genes of the genome and their genetic action has contributed to a high response to the selection during and after domestication, while the isoenzymatic, the phaseolin, and the molecular markers did not have a strong effect on the phenotype that determined the morphology of the cultivars, but offered additional advantages that can reveal genetic relationships between the accessions through a higher representation of the genome (Singh *et al.*, 1991b).

The biochemical and molecular markers confirmed the separation of the collection into two large genetic groups: Central America and the Andes, as well as into cultivated and wild ones. These results coincide with other studies that suggest molecular and biochemical markers as important instruments in the classification of beans by groups or degree of domestication (Gepts and Debouck, 1991; Singh *et al.* 1991b).

The I phaseolin pattern of the DGD 1956 accession of northern Peru (Cajamarca province, San Pablo), lacking a high molecular weight band of 52 kD, is more similar to the phaseolin type of the wild Colombian accessions than the one from the southern Andes or the wild one from Mexico; an equal behavior is seen at the molecular level that can be partly explained by genetic flow because the accession has remained and evolved in a zone of transition between two Mesoamerican and Andean populations (Ligarreto and Ocampo, 2012).

The S type phaseolin Argentina-2 accession may be grouped into the Andean group, containing some accessions of beans of Colombia and the southern Andes, in which the contrasting molecular and morphological characterizations could be due to the frequent hybridization and recombinations of the indirect mutations that have occurred in the evolutionary processes of the distinct lines that originated the two large classification groups, followed by processes of selection during domestication in different places and times (Ligarreto and Ocampo, 2012).

Similarly, when the cultivated accessions were integrally valued with qualitative and quantitative morphological, physiological, biochemical, and molecular RAPD markers, they were classified by groups that correspond to the Mesoamerican and Andean genetic groups (Fig. 3), with association between the different descriptor types and between these and the total characterization. The high correlation of the molecular and biochemical markers with the total characterization of all the studied marker types is explained by the majority of the combined matrix variables corresponding to these markers (Hernández-López *et al.*, 2013).

The high degree of association between the six morphological variables of high heritability could be due to the morphological characteristics that have been conserved for adaptive reasons after random mutations have occurred. On the other hand, it was found that typical morphological descriptors, central leaflet and bract forms, and phaseolins of the Mesoamerican group appeared in some Andean cultivars, reflecting the existence of an intermediate morphology between the Central American and southern Andes groups (Mitrick *et al.*, 1997; Pérez-Vega *et al.*, 2011).

Conclusions

The analysis of the genetic variability of the bean collection with different types of continuous combined and categorical binary descriptors in the same algorithm generated consistent information between the different types of characteristics and markers for the simultaneous evaluation of the same, allowing for the analysis, the presentation and the interpretation of the results, providing what could be a useful methodology in studies on the characterization of any species.

The implications of these results for the improvement of the species is based on taking advantage of the genetic variability that exists in the Colombian bean collection through crosses within each genetic group or between the

groups to take advantage of the "heterosis" of the cultivars that for years have been adapting and have been sharing similar habits, in this way, the Andean group could be used as a carrier of quality characteristics of large grains and altitude adaption and the Mesoamerican group as a carrier of good yield and adaptation to low zones.

Literature cited

Bitocchi, E., L. Nanni, E. Bellucci, M. Rossi, A. Giardini, P. Spagnoletti Zeuli, G. Logozzo, J. Stougaard, P. McClean, G. Attene, and R. Papa. 2012. Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. Proc. Natl. Acad. Sci. E788-E796.

CIAT, Centro Internacional de Agricultura Tropical. 1996. Protocolos de laboratorio de electroforesis. Cali, Colombia.

Cole-Rodgers, P., D.W. Smith, and P.W. Bosland. 1997. A novel statistical approach to analyze genetic resource evaluations using *Capsicum* as an example. Crop Sci. 37, 1000-1002.

Delgado-Salinas, A. 2014. Biodiversity and systematics of *Phaseolus* L. (Leguminosae). Legume Perspectives 2, 5-7.

Dellaporta, S.L., J. Wood, and J.B. Hicks. 1983. A plant DNA minipreparation: version II. Plant Mol. Biol. Rep. 1, 19-21.

Franco, J., J. Crossa, J. Villaseñor, S. Taba, and S.A. Eberhart. 1997. Classifying mexican maize accesions using hierarchical and density search methods. Crop Sci. 37, 972-980.

Franco, T.L. and R. Hidalgo. 2003. Análisis estadístico de datos de caracterización morfológica de recursos fitogenéticos. Technical Bulletin No. 8. Instituto Internacional de Recursos Fitogenéticos (IPGRI), Cali, Colombia.

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

Gepts, P. and D. Debouck. 1991. Origin, domestication, and evolution of common bean (*Phaseolus vulgaris* L.). pp. 7-53. In: Van Schoonhoven, A. and O. Voysest (eds.). Common beans: research for crop improvement. CAB, Wallingford, UK; CIAT, Cali, Colombia.

Gower, J.C. 1971. A general coefficient of similarity and some of its properties. Biometrics 27, 857-874.

Hernández-López, V.M., M.L.P. Vargas-Vázquez, J.S. Muruaga-Martínez, S. Hernández-Delgado, and N. Mayek-Pérez. 2013. Origen, domesticación y diversificación del fríjol común. Avances y perspectivas. Rev. Fitotec. Mex. 36, 95-104.

Koenig, R. and P. Gepts. 1989. Segregation and linkage of genes for seed proteins, isozymes, and morphological traits in common bean (*Phaseoulus vulgaris*). J. Hered. 80, 455-459.

Kwak, M. and P. Gepts. 2009. Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae). Theor. Appl. Genet. 118, 979-992.

Kwak, M., O. Toro, D.G. Debouck, and P. Gepts. 2012. Multiple origins of the determinate growth habit in domesticated common bean (*Phaseolus vulgaris*). Ann. Bot. 110, 1573-1580.

Ligarreto, G.A. 2003. Análisis de la variabilidad genética en fríjol. pp. 40-49. In: Franco, T.L. and R. Hidalgo (eds.). Análisis estadístico de datos de caracterización morfológica de recursos fitogenéticos. Technical Bulletin No. 8. Instituto Internacional de Recursos Fitogenéticos (IPGRI), Cali, Colombia.

Ligarreto, G.A. 2013. Componentes de variancia en variables de crecimiento y fotosíntesis en fríjol común (*Phaseolus vulgaris* L). Rev. UDCA Act. & Div. Cient. 16, 87-96.

Ligarreto, G. A. and C. Ocampo. 2012. Genetic diversity in a Colombian bean (*Phaseolus vulgaris* L.) collection as assessed by phaseolin patterns and isoenzymatic markers. Agron. Colomb. 30, 179-187.

Mitrick, A.J., P.W. Skroch, J. Nienhuis, P. Hinrichsen, G. Bascur, and C. Muñoz-Schick. 1997. Gene pool classification of common bean landraces from Chile based on RAPD and morphological data. Crop Sci. 37, 605-613.

Muñoz, G., G. Giraldo, and J. Fernández de Soto. 1993. Descriptores varietales: arroz, fríjol, maíz, sorgo. Publication No. 177. CIAT, Cali, Colombia.

Pachón, N.A., D.F. Gracia, and G.A. Ligarreto. 2009. Yield evaluation of fourteen populations of climbing bean (*Phaseolus vulgaris* L.) segregating lines with anthracnose (*Colletotrichum lindemuthianum*) resistance genes. Agron. Colomb. 27, 7-13.

Pérez-Vega, J.C., M.W. Blair, F. Monserrate, and G. Ligarreto M. 2011. Evaluation of an Andean common bean reference collection under drought stress. Agron. Colomb. 29, 17-26.

Podani, J. 1993. Syntax. Computer programs for multivariate data analysis in ecology systematic. Version 5.0. Scientia Publishing. Budapest, Hungary.

Radford, P.J. 1967. Growth analysis formulae-their use and abuse. Crop Sci. 7, 171-174.

Rohlf, F.J. 1998. NTSYS-pc. Numerical taxonomy and multivariate analysis system. Version 2.0, Exeter Publishing, Setauket, NY.

Singh, S.P., J.A. Gutiérrez, A. Molina, C. Urrea, and P. Gepts. 1991a. Genetic diversity in cultivated common beans: II Marker-based analysis morphological and agronomic traits. Crop Sci. 31, 23-49.

Singh, S.P., R. Nodari, and P. Gepts. 1991b. Genetic diversity in cultivated common bean: I. Allozymes. Crop Sci. 31, 19-23.

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

Van Schoonhoven, A. and M.A. Pastor-Corrales. 1991. Sistema estándar para la evaluación de germoplasma de fríjol. CIAT, Cali, Colombia.

Vélez V., L.D., J. Clavijo P., and G.A. Ligarreto M. 2007. Análisis ecofisiológico del cultivo asociado maíz (*Zea mays* L.) - fríjol voluble (*Phaseolus vulgaris* L.). Rev. Fac. Nal. Agr. Medellín 60, 3965-3984.

Voysest, O., M.C. Valencia, and M.C. Amezquita. 1994. Genetic diversity among Latin American Andean and Mesoamerican common bean cultivars. Crop Sci. 34, 1100-1110.

Worthington, M., D. Soleri, F. Aragón-Cuevas, and P. Gepts. 2012. Genetic composition and spatial distribution of farmer-managed *Phaseolus* bean plantings: an example from a Village in Oaxaca, Mexico. Crop Sci. 52, 1722-1735.