

## High diversity of cultivated lima beans (*Phaseolus lunatus*) in Brazil consisting of one Andean and two Mesoamerican groups with strong introgression between the gene pools

R.N.O. Silva<sup>1</sup>, A.C.A. Lopes<sup>2</sup>, R.L.F. Gomes<sup>2</sup>, J.G. Pádua<sup>3</sup> and M.L. Burle<sup>3</sup>

<sup>1</sup> Campus Universitário Amílcar Ferreira Sobral, Universidade Federal do Piauí, Floriano, PI, Brasil.

<sup>2</sup> Centro de Ciências Agrárias, Universidade Federal do Piauí, Teresina, PI, Brasil.

<sup>3</sup> Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF, Brasil.

Corresponding author: R.N.O. Silva

E-mail: jraio@ufpi.edu.br

Genet. Mol. Res. 18 (4): gmr18441

Received July 18, 2019

Accepted December 07, 2019

Published December 31, 2019

DOI <http://dx.doi.org/10.4238/gmr18441>

**ABSTRACT.** The lima bean (*Phaseolus lunatus*) has been cultivated in Brazil since pre-colonization times and remains an important source of food and income for small farmers. Nevertheless, the species has not been extensively studied in this country. We assessed the genetic diversity of 183 lima bean landraces collected from different regions of Brazil, maintained by Embrapa, the Brazilian Government Agricultural Research Corporation. Twelve microsatellite markers were used, and seven morphological descriptors were applied. The genetic parameters suggested high diversity of the Brazilian collection of lima beans, with a mean gene diversity of 0.68 and number of alleles varying from 5 to 15 among sites. Based on a Bayesian model using molecular data, three sub-populations were identified in the sample: one predominantly from the Andean gene pool of the species (large seeds, mean 100-seed weight of 80g), and two predominantly from the Mesoamerican pool (both groups with a mean 100-seed weight of 34g). Another large group was composed of accessions classified as potential hybrids among the different sub-populations. All the accessions collected in the Krahô indigenous reserve were allocated in the Andean sub-population, and these indigenous accessions represented most of this Andean group. All

the three sub-populations identified included accessions collected from far-apart sites in different geographic regions of Brazil. There was considerable introgression between the Andean and the Mesoamerican gene pools of cultivated *P. lunatus*.

**Key words:** Genetic diversity; Lima bean; *Phaseolus lunatus*; Germplasm; Microsatellite; Structure

## INTRODUCTION

The lima bean (*Phaseolus lunatus*) is an important grain legume and is a source of protein and other nutrients in many developing countries, contributing all essential amino acids required for human nutrition (Chel-Guerrero et al., 2012). The species is one of the main traditional crops in Mexico (Martínez-Castilho et al., 2008), and it is part of the regular diet in Africa, where it is cultivated in association with maize, sorghum, sweet potato, coffee, cotton and yam (Asante et al., 2008). The crop has considerable importance in Brazil, where it is mostly cultivated on small farms. According to Vieira (1992), the lima bean used to be cultivated in most of the states of Brazil, in small to large areas and, occasionally, as a cash crop. Nowadays, the main growing region in the country is the northeast, where the lima bean constitutes an important alternative as a source of food and income (Lopes et al., 2010; Oliveira et al., 2004). The lima bean crop is also very important in the context of food security for different indigenous Brazilian tribes, such as the Krahôns (Terezinha Dias, Embrapa Recursos Genéticos e Biotecnologia, personal communication). For some of these indigenous communities, the grains of this crop represent the main source of protein in periods of scarcity.

Different studies raise the hypothesis that the species presents two major gene pools: one group with small seeds and wild types distributed in Mexico, Central America and the eastern part of the Andes (the Mesoamerican gene pool), and the other group with large seeds and wild types distributed predominantly in the western part of the Andes, in Ecuador and northern Peru (Debouck et al., 1987; Gutiérrez Salgado et al., 1995; Fofana et al., 2001; Motta-Aldana et al., 2010; Serrano-Serrano et al., 2010). The study by Serrano-Serrano et al. (2010) supports the Andean origin of wild lima beans, followed by an early divergence due to geographic isolation related to major geographic events, such as the uplift of the Andes and the closure of the Isthmus of Panama. The Andean and the Mesoamerican gene pools of the lima bean are usually identified based on molecular analysis, both in the wild and in the cultivated pools of the species (Gutiérrez Salgado et al., 1995; Lioi et al., 1998; Fofana et al., 2001). This division of the species into these two gene pools may present practical impacts for the use of the respective genetic resources. In the case of the common bean (*P. vulgaris*), which presents a pattern of evolution and genetic structure similar to the lima bean, the genetic divergence between the Andean and Mesoamerican groups has led to reproductive isolation between these groups (Gepts and Bliss, 1985). Baudoin (1988) also reports some degree of reproductive isolation between these two gene pools of the lima bean. However, the lima bean most probably presents a weaker level of differentiation between its Andean and Mesoamerican pools, if compared to the common bean, for instance. Both the studies by Lioi et al (1998) and Fofana et al. (2001) with *P. lunatus* reported an intermediary group, which most probably resulted from the introgression between the Andean and Mesoamerican lima bean gene pools. In our opinion, however, the degree of genetic structure and the levels of introgression between these two gene pools in the lima bean crop remain to be further investigated.

*Phaseolus lunatus* diversity has been assessed based on different types of characters, such as morphological, biochemical, and molecular (Gutiérrez-Salgado et al., 1995; Zoro Bi et al., 2003; Martínez-Castillo et al., 2004; Martínez-Castillo et al., 2006; Martínez-Castillo et al., 2008; Silva et al., 2017). The species presents high levels of diversity, if compared to other predominantly selfing species, such as *P. vulgaris*, for example (Zoro Bi et al., 2003). Most studies of lima bean crop diversity used accessions that were collected from Central America and Mexico. To the best of our knowledge, few studies have assessed the diversity of the cultivated lima bean using a sample that is broadly geographically represented, and numerous accessions from South America (Lioi et al., 1998; Fofana et al., 2001; Silva et al., 2017).

Although Brazil is not a primary center of diversity for the lima bean, large diversity of the species' cultivated pool is expected in the country, since various types have been cultivated in the different regions of the country (Lopes et al., 2010; Silva et al., 2017). Despite its importance for Brazil, the species has received attention from the Brazilian scientific community in just the last 16 years. Since 2003, the Plant Breeding Program at the Federal University of Piauí has carried out multi-disciplinary studies on the lima bean (Santos et al., 2009; Lopes et al., 2010); this program aims to obtain short cycle genotypes with a determinate growth habit. A few studies have accessed the characterization and use of the species' genetic resources in Brazil (Santos et al., 2002; Guimarães et al., 2007; Alves et al., 2008; Silva et al., 2017). Based on a multivariate analysis using both quantitative and qualitative morphological traits, Silva et al. (2017) concluded that both the Andean and the Mesoamerican lima bean gene pools were cultivated in Brazil, with high levels of genetic diversity.

The Brazilian Agricultural Research Corporation, Embrapa, under the aegis of the Brazilian Ministry of Agriculture, Livestock, and Food Supply, holds and conserves a large germplasm collection of the species in its cultivated form, in the long term collection (Colbase) at Cenargen, the Embrapa Genetic Resources and Biotechnology Center. The accessions stored in Colbase were obtained predominantly from the extensive collection expeditions performed in different regions of Brazil, since the 60's and, therefore, represent the different regions where the crop was cultivated in the country. Some accessions of this lima bean collection were collected by other Brazilian institutions in the 1960s and were deposited in Colbase-Cenargen for security purposes, as was the case of the accessions deposited by the Federal University of Viçosa. In 2009, Embrapa formalized an Active Genebank of *P. lunatus*; since then, accessions of this important collection have been multiplied, characterized, and made available for institutions that request its germplasm.

We examined the genetic diversity and the structure of the *P. lunatus* conserved in the Embrapa collection. In a previous publication (Silva et al., 2017), we investigated the genetic diversity of a lima bean sample from this germplasm collection; however, this was based only on morphological characters. In this study we used microsatellite markers and some morphological descriptors of the species to access the diversity of the lima bean cultivated in Brazil.

## MATERIAL AND METHODS

### Sampling

We used 192 accessions of *P. lunatus* in the study ([Supplementary 1](#)). Most of these accessions are maintained in the Embrapa Active Genebank of *P. lunatus*, which in Cenargen, the Embrapa Genetic Resources and Biotechnology Center, located in Brasília, DF. This active gene bank is composed of accessions collected from all the major regions where the crop has been cultivated, as a result of the extensive collection expeditions performed by Embrapa and

other institutions, such as the Federal University of Viçosa, in the last 60 years. The accessions in this study sample were chosen to represent the maximum geographic diversity inside the country. Nine other accessions that were collected in different countries were included in the study sample. The Federal University of Piauí imported these accessions from the Centro Internacional de Agricultura Tropical (CIAT) as promising genotypes for the characteristics of determinate growth habit and short cycle. For the germplasm collected in Brazil, in this study we used samples that have not been regenerated yet.

## Marker analysis

Genomic DNA was extracted from young leaves, following the cetyltrimethylammonium bromide (CTAB) method described by Doyle and Doyle (1987), with the adaptation of triturating the plant material with Fatstrep (Bio 101 Savant) equipment. Twenty-four microsatellite markers were tested, and 12 were used in the genetic analysis (Table 1).

**Table 1.** Microsatellite markers tested and the performance in the *Phaseolus lunatus* sample.

Marker	Motif Type	Reference	Performance <sup>†</sup>
AG1	(GA) <sub>8</sub> GGTA(GA) <sub>5</sub> GGGGACG(AG) <sub>4</sub>	(Gaitán-Solis et al., 2002)	C
BM114	(TA) <sub>8</sub> (GT) <sub>10</sub>	(Gaitán-Solis et al., 2002)	+M
BM140	(GA) <sub>30</sub>	(Gaitán-Solis et al., 2002)	+P
BM141	(GA) <sub>29</sub>	(Gaitán-Solis et al., 2002)	+P
BM146	(CTGTTG) <sub>4</sub> -(CTG) <sub>4</sub> -(TTG) <sub>3</sub> (CTG) <sub>3</sub> -(CTG) <sub>4</sub>	(Gaitán-Solis et al., 2002)	+P
BM154	(CT) <sub>17</sub>	(Gaitán-Solis et al., 2002)	+P
BM155	(CA) <sub>8</sub>	(Gaitán-Solis et al., 2002)	+P
BM156	(CT) <sub>32</sub>	(Gaitán-Solis et al., 2002)	+P
BM160	(GA) <sub>15</sub> (GAA) <sub>5</sub>	(Gaitán-Solis et al., 2002)	+P
BM164	(GT) <sub>9</sub> (GA) <sub>21</sub>	(Gaitán-Solis et al., 2002)	+P
BM170	(CT) <sub>5</sub> CCTT(CT) <sub>12</sub>	(Gaitán-Solis et al., 2002)	+P
BM172	(GA) <sub>23</sub>	(Gaitán-Solis et al., 2002)	NA
BM183	(TC) <sub>14</sub>	(Gaitán-Solis et al., 2002)	+P
BM189	(CT) <sub>13</sub>	(Gaitán-Solis et al., 2002)	+M
BM211	(CT) <sub>16</sub>	(Gaitán-Solis et al., 2002)	C
BM212	(CA) <sub>13</sub>	(Gaitán-Solis et al., 2002)	+P
BMd-12	(AGC) <sub>7</sub>	(Blair et al., 2003)	+M
BMd41	(ATT) <sub>9</sub>	(Blair et al., 2003)	NA
BMd42	(AT) <sub>5</sub>	(Blair et al., 2003)	NA
BMd53	(GTA) <sub>5</sub>	(Blair et al., 2003)	+M
GATS 91	(GA) <sub>17</sub>	(Gaitán-Solis et al. 2002)	+P
Pvag003	(AG) <sub>8</sub>	(Yu et al., 2000)	+M
Pvat001	(TA) <sub>22</sub>	(Yu et al., 2000)	NA
Pvatcc001	(ATCC) <sub>3</sub> (AG) <sub>2</sub> (TAC) <sub>3</sub>	(Yu et al., 2000)	+M

<sup>†</sup>+ = successful amplification; NA = no amplification; C = 50% amplification; P = polymorphic primer; M = monomorphic primer.

The markers used in the genetic analysis were distributed in seven linkage groups of the common bean (Table 2). All of these markers were developed for the common bean (*P. vulgaris*) (Gaitán-Solis et al., 2002; Yu et al., 2000; Blair et al., 2003), and 15 of those were tested previously on the lima bean (Gaitán-Solis, et al., 2002). We used the method of fluorescent labeling of microsatellite fragments amplified by PCR described by Schuelke (2000). PCR reactions contained approximately 20 ng of genomic DNA, reaction buffer (10mMTris-HCL, pH 8.4, 50 mMCLCl), 250µM dNTP, 3 mM of MgCl<sub>2</sub>, 0.04 µM forward primer (with the M13 tail), 0.16 µM reverse primer, 0.16 µM fluorescent dye (with M13 tail) and 0.5 U of TaqDNA polymerase (Invitrogen®). PCR cycles were as follows: 5 min at 94°C; 30 cycles of 45 s at

94°C, 45 s at 56°C and 45 s at 72°C; 8 cycles of 45 s at 94°C, 45 s at 53°C and 45 s at 72°C; a final extension of 10 min at 72°C. Each PCR reaction contained just one locus of microsatellite, and the amplified products were detected in an automatic analyzer of fragments (ABI 377, Perkin Almer). The detection and the estimation of fragments sizes were performed with GeneScan3.1.2 software (Applied Biosystems). Genotyper 2.5.2 software (Applied Biosystems) was used to filter the peaks and to detect the genotypes. Allelic values were adjusted with Flexibin software (Amos et al., 2007).

**Table 2.** Genetic parameters estimated for the *Phaseolus lunatus* sample based on microsatellite markers.

Marker <sup>†</sup>	Linkage group	SSR sequence	Mallelef	No. of alleles	Gene diversity	H <sub>o</sub>	PIC
Bm140	4	(GA) <sub>30</sub>	0.41	12	0.73	0.000	0.69
Bm141	9	(GA) <sub>29</sub>	0.39	10	0.74	0.037	0.70
Bm146	1	(CTGTTG) <sub>4</sub> -(CTG) <sub>4</sub> - (TTG) <sub>3</sub> (CTG) <sub>3</sub> -(CTG) <sub>4</sub>	0.37	5	0.74	0.000	0.69
Bm154	9	(CT) <sub>17</sub>	0.45	13	0.72	0.054	0.69
Bm155	5	(CA) <sub>8</sub>	0.73	5	0.43	0.000	0.39
Bm156	2	(CT) <sub>32</sub>	0.44	7	0.72	0.048	0.68
Bm160	7	(GA) <sub>15</sub> (GAA) <sub>5</sub>	0.47	7	0.60	0.058	0.52
Bm164	2	(GT) <sub>9</sub> (GA) <sub>21</sub>	0.50	5	0.63	0.000	0.56
Bm170	6	(CT) <sub>5</sub> CCTT(CT) <sub>12</sub>	0.27	15	0.85	0.024	0.84
Bm183	7	(TC) <sub>14</sub>	0.44	5	0.66	0.000	0.59
Bm212	9	(CA) <sub>13</sub>	0.53	6	0.58	0.011	0.50
GATS91	2	(GA) <sub>17</sub>	0.35	10	0.80	0.065	0.78
Mean	-	-	0.45	8.33	0.68	0.025	0.64

<sup>†</sup>All the markers were developed by Gaitan-Solis et al. (2002) and were genomic; Mallelef= major allele frequency; H<sub>o</sub>= observed heterozygosity; PIC= polymorphism information content.

## Analysis of genetic data

Power Marker 3.25 software (Liu and Muse, 2005) was used to calculate the following genetic parameters. Major allele frequency, allele number, gene diversity (or expected heterozygosity) and observed heterozygosity were calculated according to Weir (1996). The polymorphism information content (PIC) calculation followed Botstein et al. (1980). Because all the markers used in this study were developed, originally, for another species (*P. vulgaris*), we did not include the monomorphic markers in the statistical analysis for diversity. The Bayesian model of Pritchard et al. (2000), implemented in Structure 2.1 software, was used to define the population structure and to assign individuals to sub-populations. The program was run with a preset number of populations (K) ranging from 1 to 10, with a running length of 5000 burn-in and 50,000 Markov Chain Monte Carlo (MCMC) repetitions. Twenty independent simulations were performed for each K. The admixture model and the correlated allele frequencies were used. Accessions with a sub-population membership coefficient of less than 0.8 were identified as potential hybrids. The following parameters were calculated to identify the number of sub-populations that best reflects the structure in the study sample: the likelihoods (posterior probabilities) of simulations for each preset K, the standard deviations of likelihoods, and Delta K (Evanno et al., 2005). Delta K was calculated using the application Structure Harvester (Earl and vonHoldt, 2012), available at the following web page: <http://taylor0.biology.ucla.edu/structureHarvester/>. A neighbor joining tree was reconstructed using the Shared-allele distance (Chakraborty and Jin, 1993) and Mega software (Tamura et al., 2007).

## Morphological characterization

The following morphological descriptors were applied to the sample: background color of seed, seed shape, seed width, seed thickness, seed weight, number of days from emergence to flowering and growth habit. Except for pod width and seed thickness (Vilhordo et al., 1996), all the above-mentioned morphological descriptors followed IBPGR (1982). For the seed descriptors, 10 seeds were evaluated per accession, and means and modes were calculated. Growth habit and number of days to flowering were measured in two replicates (two plants) of the accessions sown in pots in a completely randomized design in a greenhouse located in Embrapa Recursos Genéticos e Biotecnologia, in Brasília (15°43' S and 47°54' W). For the classification of seeds' shape and seeds' profiles, the J and H indices were used, respectively, following Puerta Romero (Vilhordo et al., 1996). The J index is given by the ratio seed length/seed width (L/W) and was used to classify the seeds into the following types of shape: spherical (J ranging from 1.16 to 1.42), elliptic (J 1.43 to 1.65) and oblong/short reniform (J 1.66 to 1.85). The H index is given by the ratio seed thickness/seed width (T/W) and was used to classify the seeds into the following types of profiles: flattened (H < 0.69), semi-flattened (H 0.70 to 0.79), and full (H > 0.80). A pachymeter was used to obtain the linear measures of the seeds.

## RESULTS AND DISCUSSION

### Polymorphism and microsatellite diversity

From the 24 microsatellite markers tested in our *P. lunatus* sample, 18 produced peaks of expected size and with consistent shapes over the complete evaluated sample (Table 2). Four microsatellite loci did not amplify in our lima bean study (BMd41, Pvat001, BMd42 and BM172), and two microsatellite loci (G1 and BM211) amplified in just some of the accessions. From the 18 microsatellite markers that presented amplification in the complete sample, six were monomorphic: BM114, BM189, BMd12, BMd53, PVag003 and PVatcc001. Most of the microsatellites tested in our study (17 out of 24) were developed by Gaitán-Solis et al. (2002), and the other seven markers were developed by Yu et al. (2000) and Blair et al. (2003). Only the markers developed by Gaitán-Solis et al. (2002) worked for our lima bean sample and were polymorphic. Gaitán-Solis et al. (2002) also reported no amplification for locus BM172 on *P. lunatus* and monomorphism for loci BM114 and BM189 when tested in two accessions of the lima bean from different origins, Peru and Guatemala.

The polymorphism information content ranged from 0.39 (BM155) to 0.84 (BM170) among these markers (Table 3). The number of alleles ranged from 5 (BM155) to 15 (BM170). Major allele frequency was lower than 50% for all but three of the polymorphic markers (Table 3). The observed heterozygosity in the study sample ranged from 0 to 0.065, with mean 0.025. Gene diversity ranged from 0.43 to 0.85 in the different loci evaluated. The mean gene diversity in the Brazilian lima bean sample was 0.68.

### Sub-populations (groups) and introgression – “Structure” modeling

Mean likelihoods of the Structure models increased with a higher number of preset K (Table 3). However, according to Pritchard et al. (2000) and Evanno et al. (2005), this increase in the likelihoods for a higher number of preset populations simulated with Structure modeling may constitute a model bias. According to Evanno et al. (2005), the parameters of standard deviation of the models and Delta K are more appropriate for investigating the best number of

populations that express the structure of the study sample. From all the preset  $K$  tested, the standard deviations of likelihoods were the lowest at  $K = 3$ . The standard deviations for a putative division of the studied sample in just an Andean and a Mesoamerican genepool ( $K = 2$ ; just two sub-populations) was higher than for other preset  $K$ s ( $K = 3, K = 4, K = 5, K = 6$  and  $K = 7$ ). Delta  $K$  was the highest at  $K=3$ , and the difference from the  $s$  highest preset  $K$  simulation was significant, as it was more than 10 times higher in magnitude. Therefore, our results for both Delta  $K$  and for standard deviations of likelihoods supported, very strongly, three sub-populations as the best representation of the structure in the Brazilian sample of lima bean.

The accessions collected from the Krahô indigenous reserve, most of them with large seeds, were allocated on the right-hand side of the Structure bar plot of membership coefficient (Figure 1): the groups marked in yellow in the different  $K$  simulations. Most of the Krahô accessions were allocated in the same group for the different preset  $K$ s ( $K = 2, 3, 4$  and  $5$ ), even though they presented high diversity in their seed morphology (Moraes et al., 2017). The accessions of determinate growth habit, introduced from the CIAT (Centro Internacional de Agricultura Tropical), were allocated in the red group for  $K = 3$  and in the blue group for  $K = 2$ .

**Table 3.** Mean likelihoods of Pritchard et al.'s (2000) Bayesian models, their standard deviations and Delta  $K$  for simulations for different Structure simulations within each preset  $K$ .

$K$	Ln $P(D)$ †		Delta $K$
	Means	Standard deviations	
1	-619.310	1.462.621	-
2	-5.419.730	21.335.296	10.666.831
3	-4.875.730	1.427.217	239.749.865
4	-4.673.905	2.668.229	15.819.480
5	-4.514.290	7.849.029	5.819.191
6	-4.400.350	9.456.744	3.470.539
7	-4.319.230	17.904.986	0.050265
8	-4.237.210	40.528.936	0.244640
9	-4.145.275	29.972.441	2.893.491
10	-4.140.065	204.390.638	-

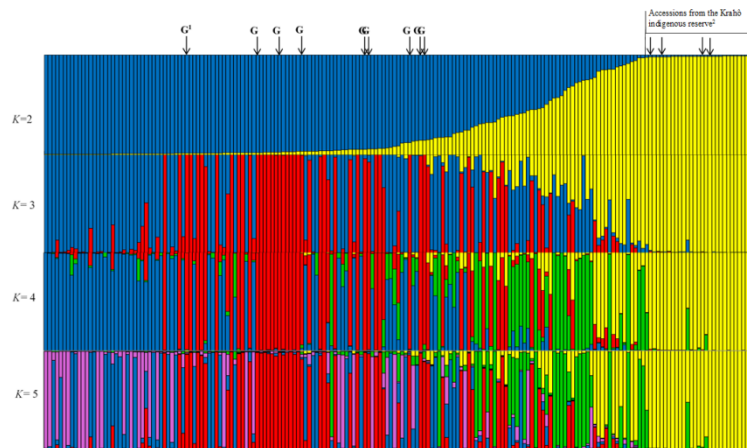
†20 iterations for each  $K$  level

When *Structure* was simulated for  $K = 2$  (a putative subdivision of the sample into an Andean and a Mesoamerican gene pool), a large number of accessions presented a hybrid background (Figure 1), suggesting a high degree of introgression between these two gene pools, if these two pools indeed exist in the studied sample. Most of the accessions that presented a hybrid background for  $K = 2$  also presented a hybrid background when the program modeled three sub-populations in the studied sample. When *Structure* was run for  $K = 3$ , each one of the three sub-populations identified presented some degree of introgression with the other two sub-populations. From the potential hybrids identified at  $K=3$ , some presented backgrounds estimated from two sub-populations, while other accessions presented backgrounds estimated from the three sub-populations.

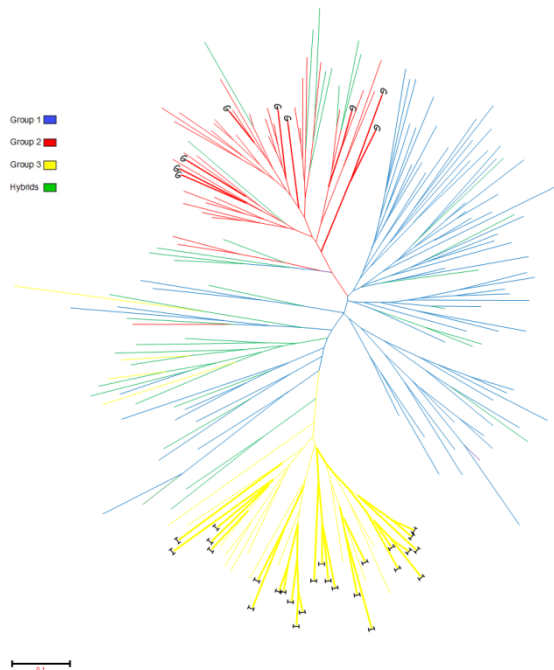
### Neighbor joining tree

Most of the accessions classified as group “3” for the *Structure* modeling considering  $K = 3$  (Figure 1), represented in yellow, clustered together in the reconstructed tree (Figure 2). A few accessions from this group clustered close to accessions classified as potential hybrids (in green). Except for one accession that clustered close to the potential hybrids identified for  $K = 3$ , all the other accessions of group “2” (group with accessions of determinate growth habit; in red)

clustered together in the reconstructed tree. The accessions of group “1” (in blue) formed different clusters that were located between groups “2” and “3” in the reconstructed tree. The accessions classified as potential hybrids for  $K = 3$  clustered predominantly closer to the accessions classified in group “1”. A few accessions classified as potential hybrids clustered together with group “2”.



**Figure 1.** Structure bar plot of membership coefficient for all the accessions of the lima bean sample, listed in the same order, with  $K$  ranging from 2 to 5. “G” are the accessions introduced from CIAT (Centro Internacional de Agricultura Tropical). All the accessions inside the rectangular area marked in Group 3 (yellow bars for  $K = 3$ ) were collected in the Kraho indigenous reserve, except for the accessions marked with arrows.



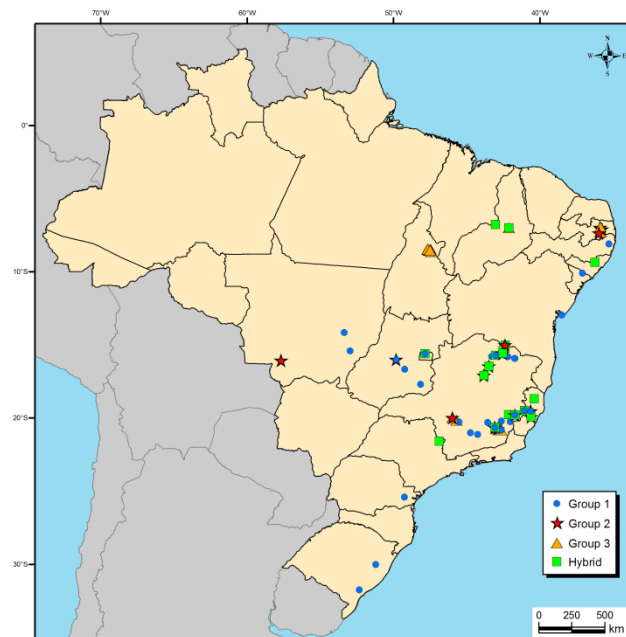
**Figure 2.** Neighbor-joining tree reconstructed for the *Phaseolus lunatus* sample based on Shared-allele distance and 12 polymorphic molecular markers, and colored according to sub-populations for *Structure*  $K = 3$ . “G” are the accessions introduced from CIAT (Centro Internacional de Agricultura Tropical) “I” are the accessions collected in the Kraho indigenous reserve.



### Morphology and geography of sub-populations (groups) & major gene pools

Because *Structure* simulations suggested three sub-populations as the best representation of the structure and the neighbor joining analysis also supported this result, we presented the morphological characteristics for each of the sub-populations estimated for *Structure*  $K = 3$  (Table 4).

Group (sub-population) “3”: represented in yellow in Figures 1 and 2; formed by 39 accessions. Most of these accessions (26) were collected in the Krahô indigenous Reserve, located in Tocantins State, in the Northern region of Brazil (Supplementary 1; Figure 3). The other accessions were collected in three other States located many miles apart (Minas Gerais, 4 accessions; Paraíba, 2; Piauí, 1), or did not present informed collecting sites (six accessions). This group presented the genotypes with the largest seed sizes. Although the majority of the accessions were of large seeds (61%), small seeded genotypes were also allocated in this group in a frequency of 13% (Table 4). Different from the other sub-populations, seed shapes in this group were predominantly elliptical or spherical, and the large majority of the accessions presented a flattened seed profile. In a similar way as in the other groups, the seed colors were diverse. This group presented the longest mean flowering cycles, although the diversity was also large for this trait. Growth habit was predominantly of type III (indeterminate climber). The longer cycles observed suggested that the respective accessions are, most probably, more primitive landraces (Baudoin, 1988). Based on the predominance of large seeded accessions with flattened profiles, we concluded that the group is composed, predominantly, of the Andean gene pool of the species (Debouck et al., 1987; Gutierrez-Salgado et al., 1995). The ranges of seed sizes are compatible with the ranges reported by Gutierrez-Salgado et al. (1995) for the Andean group of a broad sample of the lima bean collected in Latin America, with the difference of smaller seed sizes for the largest seeds in the Brazilian collection.



**Figure 3.** Collecting sites of the *Phaseolus lunatus* accessions in Brazil, colored according to sub-populations for *Structure*  $K = 3$ .

**Table 4.** Morphological characteristics and geographic distribution of the three sub-populations of the *Phaseolus lunatus* sample identified with *Structure* simulation ( $K = 3$ ).

Group	100 seed weight	Seed size class	Seed color	Seed shape	Seed profile	Days to flowering	Growth habit	Collecting State/Country
1 Blue color N = 68	17.89 - 72.54g 33.69 g	Small (84%), large (12%), medium (4%)	Brown (43%), cream, white, black, red, purple red	SPE (90%) ELP (10%)	FLA (59%) SFLA(29%) FUL(12%)	35 – 153; 103.61	II (9%) III (91%)	Minas Gerais (63%); Rio Grande do Sul; Espírito Santo; Distrito Federal; Non identified State in Brazil; Bahia; Goias; Mato Grosso; Pernambuco; Alagoas; Sergipe; Paraná;
2 Red color N = 44	17.52 - 72.00 g 33.92 g	Small (89%), large (7%), medium (4%)	Cream (32%), brown (25%), pink, white, red, black, yellow	SPE (86%) ELP (14%)	FLA(64%) SFLA (34%) FUL(2%)	26 – 130; 78.22	I (23%) II (16%) III (61%)	Minas Gerais (34%); Non identified State in Brazil; Espírito Santo; Mato Grosso; Paraíba; Bahia; Goias; Philippines; Ghana; USA; Trinidad and Tobago;
3 Yellow color N = 39	31.45 - 143.19 g 80.16 g	Large (61%), medium (26%), small (13%)	White (44%), brown, cream, black, red, grey, green	ELP (46%) SPE (41%) OSR(13%)	FLA(95%) SFLA(5%)	54 – 150; 119.21	II (15 %) III (85%)	Tocantins (67%); Non identified State in Brazil; Minas Gerais; Piauí; Paraíba;
4 Hybrid N = 41	27.74 - 147.03 g 53.17 g	Small (68%), large (20%), medium (12%)	White (34%), brown, cream, black, red	ELP (37%) SPE(58%) OSR (5%)	FLA (61%) SFLA (32%) FUL (7%)	37 – 151; 99.85	II (10%) III (90%)	Minas Gerais (41%); Distrito Federal; Non identified State in Brazil; São Paulo; Piauí; Bahia; Alagoas; Espírito Santo;

Group (sub-population) “1”: represented in blue; it was the largest sub-population identified and the one with the broadest geographic distribution in Brazil (Figure 3), in 11 different States (Table 4). Most of the accessions were collected from Minas Gerais State (Table 4). The group presented large ranges of seed weight (the three classes of seed size in the species), and most (84%) were of small sizes according to Castiñeiras et al.’s (1991) classification (<46g). The shapes of the seeds were predominantly spherical. Seed profiles were predominantly flattened, but the group also presented a large frequency of semi-flattened seed profiles. The accessions presented different seed colors and were also very diverse for their flowering cycle. All the accessions in this group were indeterminate, and most of them presented type III (indeterminate climber) growth habit. Based on the higher frequency of small seed sizes, we concluded that this group presents, most probably, a predominance of the Mesoamerican gene pool of the species (Gutierrez-Salgado et al., 1995).

Group (sub-population) “2”: represented in red; it was formed by 44 accessions and included accessions that were collected in six States of Brazil (Table 4) from different regions (Figure 3), as well as all the imported accessions, introduced from the Philippines, the United States of America and Trinidad and Tobago. Although this group also presented accessions with the different classes of seed size in the species (Debouck et al., 1989; Castiñeiras et al., 1991), the small seeded accessions predominated significantly (89%;

Table 4). The majority of the accessions presented spherical seed shapes and the seed profiles were predominantly flattened and semi-flattened. Seed colors were also very diverse. All the accessions of determinate growth habit (type I) and short cycles to flowering were allocated in this group. However, most of the accessions were also of type III growth habit. Although this group included the accessions with the shortest cycle to flowering, a large diversity for flowering cycle was also identified. Based on seed sizes, we concluded that this sub-population also presents a predominance of the Mesoamerican genepool (Debouck et al., 1987; Guitierrez-Salgado et al., 1995).

The accessions classified as potential hybrids according to the *Structure* simulation ( $K = 3$ ; threshold of 0.80 for membership coefficient) formed a group with 41 accessions (group “4” in Table 4). Accessions in this group were collected predominantly in the Southeast and Northeast regions of Brazil (Figure 3; Table 4), in seven different States. This group also presented a large variation for seed sizes and flowering cycles, and presented an intermediary distribution of seed sizes, if compared to the putative Andean group (group “3”; predominance of large seeds) and the putative Mesoamerican groups (groups “1” and “2”; predominance of small seeds). In a similar way to the putative Andean group (“3”), this group also presented some accessions of large seed sizes (e.g., 147 g/100 seeds) and a predominance of accessions with flattened seed profile.

## DISCUSSION

The microsatellite markers developed by Gaitán-Solís et al. (2002) for *P. vulgaris* were successfully used in *P. lunatus*. To the best of our knowledge, no microsatellite markers have so far been developed specifically for lima beans. Considering the importance of the lima bean crop for food security in different regions of the world, one can presume that there is demand for research in the field of molecular markers for the *P. lunatus* species.

The estimated genetic parameters in the present study suggested the high diversity of the Brazilian collection of lima beans, when compared to other collections of the species. Our estimations of mean gene diversity (0.68) and the range for number of alleles among sites (5 to 15) were similar to the ones obtained by Martínez-Castillo et al. (2006) for wild populations of the species (0.69 of mean gene diversity; 4 to 16 alleles; 8 microsatellite loci). Gene diversity of the studied sample was also high if compared to other studies that accessed the domesticated pool of the lima bean: ranges from 0.119 to 0.28 for Maquet et al. (1997), Castiñeiras et al. (2007), and Martínez-Castillo et al. (2008), using different types of markers, though. Our results for observed heterozygosity in the studied sample - ranges of 0 to 0.065, mean of 0.025 - were consistent with the predominantly selfing behavior of the species and with previous studies (Ouédraogo and Baudoin, 2002; Zoro Bi et al., 2003).

Based on the *Structure* modeling parameters (Delta K and standard deviations of likelihoods for different simulations) we concluded that the Brazilian collection of lima bean was composed of three sub-populations (or groups). This conclusion was also supported by the neighbor joining clustering analysis.

Integrating the morphological data in our analysis, in the present study we concluded that one of the sub-populations identified was predominantly composed of the Andean genepool of the species, while the two other sub-populations presented a

predominance of the Mesoamerican pool. These results were in accordance with Serrano-Serrano et al. (2010), Andueza-Noh et al. (2013) and Ruiz-Gil et al. (2016), who also reported one Andean and two Mesoamerican groups in broad lima bean samples (wild and domesticated accessions), based on chloroplast DNA sequences. In a previous study that we carried out with part of the same sample (without the 26 accessions collected in the Krahô indigenous reserve) and based just on morphological traits (Silva et al., 2017), we also identified three groups in the lima bean cultivated in Brazil. However, we suspected that one group presented the Mesoamerican genepool, another group presented the Andean pool and the third group presented intermediate characteristics between these two gene pools (Silva et al., 2017). With the inclusion of the molecular analysis, we hope that the present study has brought a more reliable view of the organization of the diversity in the Brazilian lima bean collection.

We confirmed that both the Andean and the Mesoamerican gene pools of the lima bean are cultivated in Brazil. Most of the Brazilian lima bean samples grouped in sub-populations predominantly composed of the Mesoamerican gene pool (58%); around one fifth of the sample grouped in the sub-population with a predominance of the Andean gene pool; and the other fifth was classified as potential hybrids among the different sub-populations. As most of the accessions allocated in the Andean group had only recently been introduced in the germplasm collection of Embrapa (the 26 collected in the Krahô indigenous reserve, in 2010), we concluded that the Andean background of the lima bean was even less frequent in the previous germplasm collection of Embrapa.

All the three sub-populations identified in the Brazilian sample (one Andean and two Mesoamerican), as well as the hybrid group, presented accessions collected in far-apart sites in Brazil, in the different geographic regions (Southeast, Northeast and Midwest regions). One of the Mesoamerican sub-populations, however, presented a wider geographic distribution in the country, with accessions also collected in the South region. All the lima bean accessions collected in the Krahô indigenous reserve were allocated in the Andean sub-population, and always together and in the same sub-populations in the different *Structure* simulations. An Andean predominance in a Brazilian sample could be expected in pre-colonization cultures due to geography, as Brazil is closer to the region where the Andean lima bean was domesticated than it is to the region where the Mesoamerican pool was domesticated. The first records of contacts between the Krahô indigenous people and the European colonizers dates from the beginning of the nineteenth century, in Maranhão State (Melatti, 1972; 1978), located north-east of the region where they live nowadays (Tocantins State). According to Melatti (1972; 1978) this indigenous tribe migrated from Maranhão to Tocantins State due to conflicts and pressures for land. Most probably, these lima bean types reached this indigenous people through the trade routes and migrations during pre-colonization times.

An important aspect observed in our study was the strong introgression between the Mesoamerican and the Andean gene pools in lima beans cultivated in Brazil. Because one of us studied the diversity of *P. vulgaris* in Brazil using a similar methodology (Burle et al., 2010), and considering that both *Phaseolus* species (*P. lunatus* and *P. vulgaris*) present Andean and Mesoamerican gene pools, we were able to compare aspects of diversity and introgression between the Brazilian domesticated pools of these *Phaseolus* species. We concluded that the introgression between the Andean and the Mesoamerican gene pools cultivated in Brazil was larger in the *P. lunatus* species (present study) than in *P.s vulgaris*

(Burle et al. 2010). In our study the simulations of the *Structure* modeling for two sub-populations (putatively, Andean and Mesoamerican pools) resulted in the largest standard deviations and in low values of Delta K, and a large number of potential hybrids between these groups were identified. However, in the case of *P. vulgaris*, the common bean (Burle et al. 2010), the best performance of the *Structure* modeling was obtained when two sub-populations were simulated and only a single potential hybrid between these two sub-populations (gene pools) was identified. In addition, while in the present study some introgression was observed among sub-populations of different gene pools (Andean and Mesoamerican) for the different *Structure* simulations, in the case of the Brazilian common bean, the introgression was observed only among the different Mesoamerican groups, and the single Andean group was always maintained separate in the different *Structure* simulations (K ranging from 2 to 10) (Burle et al., 2010). The strong introgression among the Andean and the Mesoamerican gene pools observed in our study was also in accordance with the high morphological variability observed within groups in our previous study, conducted with part of the same Brazilian lima bean sample and based on a multivariate analysis of both quantitative and qualitative morphological descriptors (Silva et al., 2017). As we stated previously (Silva et al 2017), and according to Fofana et al. (1997), lima bean landraces may be more prone to exchanging genes between the Andean and Mesoamerican gene pools through cross-hybridization than the wild accessions.

This study provided a first glance at the organization of the diversity of the lima bean cultivated in Brazil, based on molecular data. The large genetic diversity observed in this study (based on the estimated genetic parameters), and the presence of an Andean and two Mesoamerican gene pools, indicates that Brazil is a sary center of diversity of the lima bean. We expect that subsequent studies can improve the knowledge related to this understudied species. The germplasm collection of lima bean conserved in Embrapa is available for exchange upon request (<http://alelo.cenargen.embrapa.br/>), and we expect that this study will facilitate the management and use of this important germplasm collection.

## ACKNOWLEDGMENTS

A CAPES (Brazil) fellowship to RNOS is gratefully acknowledged. Funds for laboratory and greenhouse evaluations were provided by Embrapa. We are grateful to Dr. Vânia C. R. Azevedo and Sérgio Noronha for assistance in DNA extraction and map elaboration, respectively. We acknowledge the assistance of Luciane G. Silva and Malu D. de Melo in the greenhouse experiment. We are also grateful to Dr. Terezinha Dias, as the person responsible for the incorporation of the accessions from the Krahô indigenous reserve in the Embrapa germplasm collection of *P. lunatus*.

## CONFLICTS OF INTEREST

The authors declare no conflict of interest.

## REFERENCES

- Alves AU, Oliveira AP, Alves AU, Dornelas CSM, et al. (2008). Lima beans production and economic revenue as function of organic and mineral fertilization. *Hortic. Bras.* 26: 251-254. <http://dx.doi.org/10.1590/S0102-05362008000200024>

- Amos W, Hoffmann JI, Frodsham A, Zhang L, et al. (2007). Automated binning of microsatellite alleles: problems and solutions. *Mol. Ecol. Notes* 7: 10-14. <http://dx.doi.org/10.1111/j.1471-8286.2006.01560.x>
- Andueza-Noh RH, Serrano-Serrano ML, Chacón Sánchez MI, Sánchez del Pino I, et al. (2013). Multiple domestications of the Mesoamerican gene pool of lima bean (*Phaseolus lunatus* L.): evidence from chloroplast DNA sequences. *Genet. Resour. Crop Ev.* 60: 1069-1086. <http://dx.doi.org/10.1007/s10722-012-9904-9>.
- Asante IK, Offei SK, Addy R and Carson AG (2008). Phenotypic and seed protein analysis in 31 lima bean (*Phaseolus lunatus*) accessions in Ghana. *West Afr. J. Appl. Ecol.* 12: 98-106. <http://dx.doi.org/10.4314/wajae.v12i1.45775>
- Baudoin JP (1988). Genetic resources, domestication and evolution of lima bean, *Phaseolus lunatus*. In: Genetic resources of Phaseolus bean. (Gepts P, ed.). Kluwer Academic Publishers, Holland.
- Blair MW, Pedraza F, Buendia HF, Gaitán-Solís E, et al. (2003). Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 107: 1362-1374. <http://dx.doi.org/10.1007/s00122-003-1398-6>.
- Botstein D, White RL, Skolnick M and Davis RW (1980). Construction of a genetic map in man using restriction fragment length polymorphism. *Am. J. Hum. Genet.* 32: 314-331.
- Burle ML, Fonseca JR, Kami JA and Gepts P (2010). Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a sary center of diversity. *Theor. Appl. Genet.* 121: 801-813. <http://dx.doi.org/10.1007/s00122-010-1350-5>.
- Castiñeiras L, Esquivel MA, Rivero N and Mariño A (1991). Variabilidad de la semilla de *Phaseolus lunatus* L. en Cuba. *Rev. Jard. Bot. Nac.* 12: 109-114.
- Castiñeiras L, Guzmán FA, Duque MC, Shagarodsky T, et al. (2007). AFLPs and morphological diversity of *Phaseolus lunatus* L. in Cuban home gardens: approaches to recovering the lost ex situ collection. *Biodivers. Conserv.* 16: 2847-2865. <http://dx.doi.org/10.1007/s10531-006-9025-x>
- Chakraborty R and Jin L (1993). A unified approach to study hypervariable polymorphisms: Statistical considerations of determining relatedness and population distances. In: DNA fingerprinting, state of the science. (Pena SDJ, Chakraborty R, Epplen JT, Jeffreys AJ, eds.). Birkhauser Verlag, Basel.
- Chel-Guerrero L, Domínguez-Magaña M, Martínez-Ayala A, Dávila-Ortiz G, et al. (2012). Lima bean (*Phaseolus lunatus* L) protein hydrolysates with ACE-I inhibitory activity. *Food Nutr. Sci.* 3: 511-521. <http://dx.doi.org/10.4236/fns.2012.34072>.
- Debouck DG, Maquet A and Posso CE (1989). Biochemical evidence for two different gene pools in lima beans, *Phaseolus lunatus* L. *Annual report of the Bean Improvement Cooperative* 32: 58-59.
- Debouck DG, Liñan Jara JH, Campana Sierra A and De La Cruz Rojas JH (1987). Observations on the domestication of *Phaseolus lunatus* L. *Plant Genet. Resour. Newsl.* 70: 26-32.
- Doyle JJ and Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19: 11-15.
- Earl DA and vonHoldt BM (2012). Structure Harvester: a website and program for visualizing Structure output and implementing the Evanno method. *Conserv. Genet. Resour.* 4 (2): 359-361. doi: 10.1007/s12686-011-9548-7. Accessed 3 December 2019.
- Evanno G, Regnaut S and Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14: 2611-2620. <https://dx.doi.org/10.1111/j.1365-294X.2005.02553.x>.
- Fofana B, du Jardin P and Baudoin JP (2001). Genetic diversity in the Lima bean (*Phaseolus lunatus* L.) as revealed by chloroplast DNA (cpDNA) variations. *Genet. Resour. Crop Ev.* 48: 437-445. <https://doi.org/10.1023/A:1012016030318>.
- Fofana B, Vekemans X, Jardin P and Baudoin JP (1997). Genetic diversity in lima bean (*Phaseolus lunatus* L.) as revealed by RAPD markers. *Euphytica* 95: 157-165. Doi: 10.1023/A:1002910114869.
- Gaitán-Solís E, Duque MC, Edwards KJ and 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. <https://doi.org/10.2135/cropsci2002.2128>.
- Gepts P and Bliss FA (1985). F1 hybrid weakness in the common bean: differential geographic origin suggests two gene pools in cultivated bean germplasm. *J. Hered.* 76: 447-450.
- Guimarães WNR, Martins LSS, Silva EF, Ferraz GMG, et al. (2007). Caracterização morfológica e molecular de acessos de feijão-fava (*Phaseolus lunatus* L.). *Rev. Bras. Eng. Agr. Amb.* 11: 37-45. <http://dx.doi.org/10.1590/S1415-43662007000100005>.
- Gutiérrez-Salgado A, Gepts P and Debouck DG (1995). Evidence for two gene pools of the lima bean, *Phaseolus lunatus* L., in the Americas. *Genet. Resour. Crop Ev.* 42: 15-28. <http://dx.doi.org/10.1007/BF02310680>.
- IBPGR (1982). Descriptors for Lima bean. International Board for Plant Genetic Resources, n.36, Rome.
- Lioi L, Lotti C and Galasso I (1998). Isozyme diversity, RFLP of the rDNA and phylogenetic affinities among cultivated Lima beans, *Phaseolus lunatus* L. (Fabaceae). *Plant Syst. Evol.* 213: 153-164. <http://dx.doi.org/10.1007/BF00985196>.
- Liu K and Muse SV (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21: 2128-2129. <http://dx.doi.org/10.1093/bioinformatics/bti282>.

- Lopes ACA, Gomes RLF, Silva RNO, Costa EMR, et al. (2010). Diversidade Genética. In: A cultura do feijão-fava na Região Meio-Norte do Brasil. (Araújo ASF, Lopes ACA and Gomes RLF, eds.). Edufpi, Teresina.
- Maquet A, Zoro Bi I, Delvaux M, Wathelet B, et al. (1997). Genetic structure of a Lima bean base collection using allozyme markers. *Theor. Appl. Genet.* 95: 980-991. <http://dx.doi.org/10.1007/s001220050651>.
- Martínez-Castillo J, Colunga-Garciamarín P and Zizumbo Villarreal D (2008). Genetic erosion and in situ conservation of Lima bean (*Phaseolus lunatus* L.) landraces in its Mesoamerican diversity center. *Genet. Resour. Crop Ev.* 55: 1065-1077. <http://dx.doi.org/10.1007/s10722-008-9314-1>.
- Martínez-Castillo J, Zizumbo-Villarreal D, Gepts P, Delgado-Valerio P, et al. (2006). Structure and genetic diversity of wild population of lima bean (*Phaseolus lunatus* L.) from the Yucatan Peninsula, Mexico. *Crop Sci.* 46: 1071-1080. <http://dx.doi.org/10.2135/cropsci2005.05-0081>.
- Martínez-Castillo J, Zizumbo-Villarreal D, Perales-Rivera H and Colunga-Garciamarín P (2004). Intraspecific diversity and morpho-phenological variation in *Phaseolus lunatus* L. from the Yucatan Peninsula Mexico. *Econ. Bot.* 5: 354-380. [https://doi.org/10.1663/0013-0001\(2004\)058\[0354:IDAMVI\]2.0.CO;2](https://doi.org/10.1663/0013-0001(2004)058[0354:IDAMVI]2.0.CO;2)
- Melatti JC (1972). O messianismo Krahô. Herder/Edusp, São Paulo.
- Melatti JC (1978). Ritos de uma tribo Timbira. vol.1. Ática, São Paulo.
- Moraes CS, Dias TAB, Costa SPP, Vieira RC, et al. (2017). Catálogo de fava (*Phaseolus lunatus* L.) conservada na Embrapa. Documentos, 351. Embrapa Recursos Genéticos e Biotecnologia, Brasília.
- Motta-Aldana JR, Serrano-Serrano ML, Hernández-Torres J, Castillo-Villamizar G, et al. (2010). Multiple origins of Lima bean landraces in the Americas: Evidence from chloroplast and nuclear DNA polymorphisms. *Crop Sci.* 50: 1773-1787. <https://dx.doi.org/10.2135/cropsci2009.12.0706>
- Oliveira AP, Alves EU, Alves AU, Dornelas CSM, et al. (2004). Produção de feijão-fava em função do uso de doses de fósforo em um Neossolo Regolítico. *Hortic. Bras.* 22: 543-546. <http://dx.doi.org/10.1590/S0102-05362004000300008>.
- Ouédraogo M and Baudoin JP (2002). Comparative analysis of genetic structure and diversity in wild Lima bean populations from the Central Valley of Costa Rica, using microsatellite and isozyme markers. *Annual Report of the Bean Improvement Cooperative* 45:240-241.
- Pritchard JK, Stephens M and Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics.* 155: 945-959.
- Ruiz-Gil PJ, Chepe-Cruz G, Andueza-Noh RH, Ortiz-García MM, et al. (2016). Differentiation and genetic diversity of *Phaseolus lunatus* wild populations from Chiapas, Mexico, and their genetic relationships with MI and MII groups. *Bot. Sci.* 94: 701-712. <http://dx.doi.org/10.17129/botsci.699>.
- Santos JO, Araújo ASF, Gomes RLF, Lopes ACA, et al. (2009). Ontogenia da nodulação em feijão-fava. *Rev. Bras. Ciênc. Agrar.* 4: 426-429. <http://dx.doi.org/10.5039/agraria.v4i4a9>.
- Santos D, Corlett FMF, Mendes JEMF and Wanderley Júnior JSA (2002). Produtividade e morfologia de vagens e sementes de variedades de fava no Estado da Paraíba. *Pesqui. Agropecu. Bras.* 37: 1407-1412.
- Schuelke M (2000). An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol.* 18: 233-234. <http://dx.doi.org/10.1038/72708>.
- Serrano-Serrano ML, R. Andueza-Noh H, Martínez-Castillo J, Debouck DG, et al. (2012). Evolution and domestication of lima bean in Mexico: Evidence from ribosomal DNA. *Crop Sci.* 52: 1698-1712. <http://dx.doi.org/10.2135/cropsci2011.12.0642>.
- Serrano-Serrano ML, Hernández-Torres J, Castillo-Villamizar G, Debouck DG, et al. (2010). Gene pools in wild Lima bean (*Phaseolus lunatus* L.) from the Americas: evidences for an Andean origin and past migrations. *Mol. Phylogenet. Evol.* 54: 76-87. <http://dx.doi.org/10.1016/j.ympev.2009.08.028>.
- Silva RNO, Burle ML, Padua JG, Lopes ACA, et al. (2017). Phenotypic diversity in lima bean landraces cultivated in Brazil, using the Ward-MLM strategy. *Chil. J. Agric. Res.* 77(1): 35-40.
- Tamura K, Dudley J, Nei M and Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4. *Mol. Biol. Evol.* 24: 1596-1599. <http://dx.doi.org/10.1093/molbev/msm092>.
- Vilhordo BW, Araújo RS, Rava CA, Stone LF, et al. (1996). Morfologia. In: Cultura do feijoeiro comum no Brasil. (Araújo RS, Rava CA, Stone LF and Zimmerman MJO, eds.). Potafos, Piracicaba.
- Yu K, Park SJ, Poysa V and Gepts P (2000). Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). *J. Hered.* 91: 429-434. <https://doi.org/10.1093/jhered/91.6.429>.
- Weir BS (1996). Genetic data analysis II. Sinauer Associates Inc., Sunderland.
- Zoro Bi I, Maquet A and Baudoin JP (2003). Population genetic structure of wild *Phaseolus lunatus* (Fabaceae), with special reference to population sizes. *Am. J. Bot.* 90: 897-904. <https://doi.org/10.3732/ajb.90.6.897>.