Genet Resour Crop Evol (2017) 64:1355–1364
DOI 10.1007/s10722-016-0441-9

RESEARCH ARTICLE

Estimation of natural outcrossing rate and genetic diversity in Lima bean (*Phaseolus lunatus* L. var. *lunatus*) from Brazil using SSR markers: implications for conservation and breeding

J. S. Penha · A. C. A. Lopes · R. L. F. Gomes · J. B. Pinheiro · J. R. Assunção Filho · E. A. Silvestre · J. P. G. Viana · J. Martínez-Castillo

Received: 15 August 2015 / Accepted: 23 August 2016 / Published online: 6 September 2016

Abstract Lima bean (*Phaseolus lunatus* L.) is an important food source in Brazil, especially in the northeast region, where its production and consumption are high. The goals of the present study were to estimate natural outcrossing rates and genetic diversity levels of Lima bean from Brazil, using ten microsatellite loci to obtain information for their conservation and breeding. Fourteen accessions were selected from an experiment in field with open-pollinated and with the presence of pollinating insects. Twelve seeds of each of the 14 selected accessions were grown in screenhouse for tissue harvest and DNA extraction. The multilocus model was used to determine the reproductive system. The outcrossing rate was 38.1 % ($t_m = 0.381$; $t_s = 0.078$), and the results indicated a mixed mating system with a predominance of selfing ($1 - t_m = 61.9$ %). The biparental inbreeding rate was high ($t_m - t_s = 0.303$) and the multilocus correlated paternity was quite high ($r_{p(m)} = 0.889$), indicating that the progeny was mostly composed of full sibs. The average effective number of pollen donors per maternal plant (N_{ep}) was low (1.12), and the fixation index for maternal genotypes (F_m) was 0.945, indicating that most genitors resulted from inbreeding. The studied families presented considerable genetic variability: $A = 6.10$; $\%P = 30$; $H_e = 0.60$ and $H_o = 0.077$. Total diversity was high ($H_T = 0.596$), and a portion was distributed within families ($H_S = 0.058$). In addition, diversity was higher between families ($D_{ST} = 0.538$), and genetic differentiation was high ($G_{ST} = 0.902$). The results presented here can be used in the implementation of Lima bean conservation and breeding programs in Brazil.

Keywords Genetic variability · Landraces · Microsatellite markers · *Phaseolus lunatus* · Reproduction system

J. S. Penha (✉) · J. B. Pinheiro · J. R. Assunção Filho
Genetics Department, Luiz de Queiroz College of Agriculture, University of São Paulo, Avenue Pádua Dias, N. 11, Piracicaba, São Paulo C.E.P 13418-900, Brazil
e-mail: josilane_penha@hotmail.com

A. C. A. Lopes · R. L. F. Gomes
Plant Technology Department, Campus Ministro Petrônio Portella, Federal University of Piauí, Avenue Universitária, Teresina, Piauí C.E.P 64.049-550, Brazil

E. A. Silvestre · J. P. G. Viana
Biology Institute, University of Campinas, Avenue Bertrand Russel, University City Zeferino Vaz, Campinas, São Paulo C.E.P 13083-970, Brazil

J. Martínez-Castillo
Centro de Investigación Científica de Yucatán (CICY), Calle 43 N. 130, Colonia Chuburná de Hidalgo, C. P. 97200 Mérida, Yucatán, Mexico

Introduction

Developing breeding programs and propose strategies for conservation of crops, it is necessary to know at

least two, interrelated, aspects: (1) the natural outcrossing rates of the species of interest, and (2) the levels of genetic diversity present in the target gene pool (natural populations, landraces, gene banks). The study of plant reproduction systems can be used to estimate the outcrossing rate between individuals and determine the mode of gene transmission between generations (Wright 1921; Hamrick and Loveless 1986). The determination of the outcrossing rate is among the first steps to know in order to establish the genetic parameters of a species towards its conservation or its improvement (Sebbenn et al. 2000). Natural outcrossing rates should be determined using multilocus models, which produce more accurate estimates. The mixed-mating model proposed by Ritland and Jain (1981) is frequently used; it states that populations reproduce through a combination of selfing at rate *s* and outcrossing at rate *t*. The basic assumptions for its application are: (1) homogeneous pollen pools for all maternal genotypes; (2) alleles at different loci segregate independently; and (3) no selection or mutation between fertilisation and the time of analysis (Ritland and Jain 1981; Ritland 1990). However, violations of the assumption of homogeneity of allele frequencies of ovules and pollen have little effect on the estimation of multilocus population outcrossing rates when the number of loci used is higher than four or five (Ritland and Jain 1981).

The other fundamental component for the development of conservation and breeding programs is the characterization of genetic diversity. Such characterization results in improved knowledge of the germplasm and is necessary for subsequent stages to develop improved material and species conservation strategies (Valls 2007). Genetic diversity has been studied through phenotypic analyses, agro-morphological characteristic measurements, and molecular-level variation studies. Several molecular methods of germplasm characterization have been developed; those that utilize molecular markers to reveal DNA sequence polymorphisms have been of great importance in this area (Ferreira et al. 2007). Molecular markers are DNA fragments that can genetically differentiate individuals, they are available in a large number (larger than phenotypic markers) and do not exert epistatic or pleiotropic (Jiang 2013). Molecular markers have numerous applications, including the study of genetic diversity, estimations of outbreeding and inbreeding levels, construction of genetic maps, characterization of germplasm, and marker-assisted selection (Jiang 2013; Borém and Miranda 2009).

Microsatellites, also known as simple sequence repeats (SSR), are a popular type of molecular markers for the study of DNA polymorphisms and consist of DNA segments in a sequence of one to six nucleotides repeated in tandem. These markers are frequently found in eukaryotes and used to differentiate among individuals within a species (Buso et al. 2003; Borém and Miranda 2009). The analysis of microsatellites consists of using a pair of primers that are specific to conserved DNA sequences and flank a DNA repeat. Fragments that may occur randomly along the genome or specifically within genes are amplified through PCR (Ramalho et al. 2012). Loci polymorphism is shown by differences in the number of repeats (Ferreira et al. 2007). Microsatellites are important, particularly because of the significant amount of genetic information generated as a consequence of their high mutation rate and, hence, polymorphism (Buso et al. 2003).

Lima bean (*Phaseolus lunatus* L.) is one of the five domesticated species of *Phaseolus* (Fabaceae) and the second most cultivated species of this genus after the common bean (*P. vulgaris* L.) (Baudoin et al. 2004). This crop originated in the Neotropics, where it evolved for at least 6000 years (Kaplan and Lynch 1999). There are two Lima bean botanical varieties: a wild variety, *P. lunatus* var. *silvester* Baudet; and a domesticated variety, *P. lunatus* var. *lunatus* (Baudet 1977). Lima bean is an herbaceous species with a short annual or short-lived perennial life cycle. It is self-compatible and has a mixed reproduction system where selfing is favoured by the synchronised maturation of pollen grains and stigma as well as their proximity within the floral bud (Webster et al. 1979). However, outcrossing rates between 0.02 and 48 % have been reported (Baudoin 1988; Zoro Bi 1999); differences in outcrossing rates are attributed to variations in population size, genotype, population density, flower morphology and phenology, and pollinator availability (Fausto et al. 2001). Bees (*Apis mellifera*) are reported to be the main pollinators of Lima bean (Hardy et al. 1997).

Although outcrossing rate is an important factor influencing the distribution of genetic diversity in plants (Hamrick 1982), in Lima bean the outcrossing rate has seldom been determined. Most studies have been conducted mostly using wild populations from

Costa Rica (Baudoin 1988; Zoro Bi 1999). Lima bean studies have so far focused on the determination of genetic diversity. Using allozymes, Maquet et al. (1997) found a value of genetic diversity (h) of 0.26 for the *P. lunatus* base collection held at the germplasm bank of the International Center Tropical Agriculture (CIAT, Colombia). These authors observed significant and higher diversity than that reported for other species with mixed reproduction systems and short-lived perennials ($h = 0.12$) (Hamrick and Murawski 1991). Castiñeiras et al. (2007) observed lower levels of diversity ($h = 0.11$) for Lima bean grown in Cuba using amplified fragment length polymorphism (AFLP) markers. Martinez-Castillo et al. (2008) observed $h = 0.29$ for Lima bean grown by Mayan farmers in the Yucatan peninsula, Mexico using Inter-Simple Sequence Repeat (ISSR) markers. More recently, Montero-Rojas et al. (2013) reported $h = 0.34$ using SSR markers for Lima bean grown in three Caribbean countries.

Lima bean is an important crop in Brazil, especially in the northeast region where it is grown and used as a food and alternative source of income for small farmers (Azevedo et al. 2003; Oliveira et al. 2004). The price of Lima bean can reach 10–15 USD per kg, which is three times higher than the price of cowpea (*Vigna unguiculata* (L.) Walp.). However, Lima bean is still a subsistence crop, which is farmed without the use of production technologies. In addition, the occurrence of pests and diseases, such as anthracnose, affects the production quality and yield, resulting in low crop productivity (Lopes et al. 2010). Explorations of Lima bean from several regions of Brazil has revealed a high number of landraces and morphological variation of seeds, suggesting that the genetic diversity present in Brazil may be as high as the one reported for the Yucatan Peninsula, considered as a center of genetic diversity for this crop (Martínez-Castillo et al. 2004, 2008, 2012). However, few studies have been published on the genetic diversity of Lima bean from Brazil (Carmo et al. 2013; Guimarães et al. 2007; Santos et al. 2002) and none have reported on their outcrossing rates. The goals of the present study were to determine the natural outcrossing rates and to estimate the genetic diversity levels of Lima bean landraces from Brazil using microsatellite markers.

Materials and methods

Plant material

The plants used in this study were selected from a previous experiment with 156 accessions obtained of Germplasm Bank of the Federal University of Piauí and developed at Experimental Field of Luiz de Queiroz College of Agriculture - ESALQ/USP during July/2012 to January/2013. Each access was represented by a single plant, with open-pollinated and with the presence of pollinating insects. It was selected 14 accessions dispersed throughout the field (Table 1). Twelve seeds of each accession (mother plants) were taken and planted in screenhouse. DNA was extracted from each descent, totaling 168 plants genotyped. The plant tissue was harvested approximately 25 days following sowing when the first trifoliate leaves were fully expanded. The leaves were immediately frozen in liquid nitrogen, freeze-dried for 48 h (−50 °C, 0.04 mbar), homogenised using a knife mill (MA048, MARCONI), and stored in falcon tubes at room temperature.

Table 1 Data of 14 Lima bean accessions from Brazil used to determine the outcrossing rate and genetic diversity, using 10 microsatellite loci

Families	Accessions	Origin	Seed classification[a]
1	UFPI-252	Minas Gerais	Small seeds
2	UFPI-267	Piauí	Large seeds
3	UFPI-464	Piauí	Large seeds
4	UFPI-503	Maranhão	Small seeds
5	UFPI-537	Unknown	Small seeds
6	UFPI-587	Paraíba	Medium seeds
7	UFPI-596	Paraíba	Small seeds
8	UFPI-613	Paraíba	Large seeds
9	UFPI-627	Paraíba	Small seeds
10	UFPI-713	Piauí	Medium seeds
11	UFPI-731	Ceará	Medium seeds
12	UFPI-740	Ceará	Small seeds
13	UFPI-743	Ceará	Medium seeds
14	UFPI-755	Ceará	Large seeds

[a] Classification based on the seed shape and weight, with the identification of three groups: Small seeds, with 35–50 g for 100 seeds), Medium seeds, with 50–70 g for 100 seeds) and Large seeds, over 70 g for 100 seeds)

DNA extraction and amplification

The study was performed in the Laboratory of Genetic Diversity and Breeding of the Escola Superior de Agricultura Luiz de Queiroz–ESALQ, University of São Paulo–USP in 2013. DNA was extracted from frozen-dried homogenised material according to the cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle 1987). DNA was quantified through electrophoresis in 0.8 % agarose gel with *Syber safe* comparison with the 25, 50 and 100 ng phage lambda DNA standards. DNA samples were diluted with ultra pure water to a final concentration of 10 ng/μL. Ten pairs of primers of SSR markers designed and optimized for common bean by Gaitán-Solís et al. (2002) were used in the present study (Table 2). The forward primer was synthesized with a M13 bacterial plasmid tail at the 5′ end. PCR reactions were performed in a 20 μL final volume containing 10 ng DNA, 1X buffer, 2.0 mM $MgCl_2$, 250 μg/mL bovine serum albumin (BSA), 2.0 μM deoxynucleotide (dNTP), 1 U Taq, 0.2 μM forward primer, 0.4 μM reverse primer, 0.2 μM M13-IrDye 700 or 800, and MilliQ water using a LifePro thermal cycler (BIOER). The PCR program consisted of an initial denaturation at 94 °C for 2 min, followed by 45 cycles at 94 °C for 30 s, 49 °C for 30 s and 72 °C for 30 s and a final extension at 72 °C for 10 min. This program was used for all primers pairs. Genotyping was performed with a LI-COR 4300 DNA Analyzer automated sequencer (Li-Cor, Biosciences) using infrared fluorescence technology. The alleles in each sample were identified using SAGA GT software (Li-Cor, Biosciences).

Data analysis

The reproductive system was analyzed according to the mixed-mating model (Ritland and Jain 1981) using the software Multilocus MLTR 3.2 (Ritland 2008). The estimated parameters included multilocus population outcrossing rate (t_m), single-locus population outcrossing rate (t_s), biparental inbreeding rate ($t_m - t_s$), correlation of selfing (r_s), multilocus correlated paternity [$r_{p(m)}$], selfing ($1 - t_m$), average effective number of pollen donors per maternal plant (N_{ep}), and fixation index for maternal genotypes [$F_{(m)}$].

Table 2 Data of 10 microsatellite loci used for genotyping of 14 Lima bean maternal families

Name	Sequences	Motif	Range of allele size (bp)
AG1	F: CATGCAGAGGAAGCAGAGTG	$(GA)_8GGTA(GA)_5GGGGACG(AG)_4$	164–173
	R: GAGCGTCGTCGTTTCGAT		
BM140	F: TGCACAACACACATTTAGTGAC	$(GA)_{30}$	177–193
	R: CCTACCAAGATTGATTTATGGG		
BM141	F: TGAGGAGGAACAATGGTGGC	$(GA)_{29}$	191–204
	R: CTCACAAACCACAACGCACC		
BM146	F: GAGATGAGTCCTTTCCCTACCC	$(CTGTTG)_4(CTG)_4$ $(TTG)_3(CTG)_3(CTG)_4$	291–299
	R: TCGAGACACAATTTATGAAGGC		
BM154	F: TCTTGCGACCGAGCTTCTCC	$(CT)_{17}$	214–220
	R: CTGAATCTGAGGAACGATGACCAG		
BM156	F: CTTGTTCCACCTCCCATCATAGC	$(CT)_{32}$	232–238
	R: TGCTTGCATCTCAGCCAGAATC		
BM160	F: CGTGCTTGGCGAATAGCTTTG	$(GA)_{15}(GAA)_5$	178–210
	R: CGCGGTTCTGATCGTGACTTC		
BM211	F: ATACCCACATGCACAAGTTTGG	$(CT)_{16}$	213–229
	R: CCACCATGTGCTCATGAAGAT		
BM212	F: AGGAAGGGATCCAAAGTCACTC	$(CA)_{13}$	190–230
	R: GAACTTTCAGGTATTGATGATGAAG		
GATS91	F: GAGTGCGGAAGCGAGTAGAG	$(GA)_{17}$	234–267
	R: TCCGTGTTCCTCTGTCTGTG		

Genetic diversity was analyzed at two levels: (1) at each locus for the entire dataset; and, (2) in each family. The estimators were: number of alleles per locus, percentage of polymorphic loci (%P), expected (H_E) and observed (H_O) heterozygosity, and fixation index (f). All the analyses were performed using the software GDA (Genetic Data Analysis) (Lewis and Zaykin 2000). Allelic frequencies and genetic diversity distribution within and between families were analyzed using Nei statistics (Nei 1973) with the software FSTAT (Goudet 2002).

Results and discussion

Natural outcrossing rate in Lima bean

The estimated multilocus (t_m) and single-locus (t_s) outcrossing rates were 0.381 and 0.078, respectively, indicating a mixed mating system ($t < 0.95$) with a predominance of selfing (Table 3). The percentage of selfing ($1 - t_m$) was therefore 61.9 %. Lima bean is a hermaphroditic species with reproductive structures that allow selfing. The potential factors responsible for the high rate of natural outcrossing in Lima bean are stigma and style protruding to the outside of the keel and stigma receptivity lasting for many hours (Baudoin 1988). In addition, larger flower production (relative to the common bean), the presence of floral and extra-floral nectaries, the presence of pollinators and environmental conditions also contribute to increased outcrossing rates. High gene flow increases genetic diversity of the species and high outcrossing rates used in breeding result in the segregation of different traits and may increase the difficulty of obtaining pure lines, thus requiring the adoption of isolation measures to prevent gene flow (Ortega 1974).

Zoro Bi et al. (2005) studied ten wild populations of Lima bean using isozyme markers and found variations in the multilocus (0.027–0.268) and single-locus (0.024–0.246) outcrossing rates, with natural outcrossing rates between 2.7 and 26.8 %. Similar results were reported by Harding and Tucker (1964), who studied Lima bean cultivars using morphological markers and observed rates between 3.2 and 24.2 %. The high outcrossing rates observed in the present study may be due to favorable environmental conditions such as mild temperatures, presence of pollinators and the short distance (1 m) between plants and efficiency of microsatellite markers, which have been shown to be more informative than enzymatic and morphological markers. Kageyama et al. (2003) studied diversity in tree species using different genetic markers and observed higher discriminatory power in the detection of alleles per locus using microsatellite markers as compared to enzymatic markers.

The difference between multilocus and single-locus ($t_m - t_s$) outcrossing rates has been used to calculate the occurrence of biparental inbreeding, i.e., mating between related individuals. The difference between multilocus and single-locus outcrossing rates found in this study was positive (0.303), indicating the occurrence of 30.3 % inbreeding. This high rate likely resulted from the occurrence of preferential mating caused by proximity with related genotypes because the mother plants were not randomly distributed at the experimental site and flowering was synchronized, with plants that flowered within the same time period only exchanging pollen among each other. Zoro Bi et al. (2005) found no significant values of biparental inbreeding (0.002–0.022) in wild Lima bean, which

Table 3 Estimates of mating system parameters of 14 Lima bean maternal families, using 10 microsatellite loci

Parameters	Estimates (standard error)[a]
Multilocus outcrossing rate (t_m)	0.381 (0.026)
Single-locus outcrossing rate (t_s)	0.078 (0.013)
Biparental inbreeding (t_m–t_s)	0.303 (0.057)
Correlation of selfing (r_s)	0.174 (0.111)
Multilocus correlated paternity [$r_{p(m)}$]	0.889 (0.221)
Average effective number of pollen donors per maternal plant (N_{ep})	1.120
Fixation index for maternal genotypes [$F_{(m)}$]	0.945 (0.026)

[a] Estimated with 1000 bootstraps

indicated the absence of biparental inbreeding and mating between unrelated individuals. The correlation of selfing (r_s), which measures the probability of finding two individuals generated by selfing, was low but significant (0.174). This indicates that individuals generated by selfing were randomly distributed among the families, resulting in wide variations in outcrossing rates between mother plants.

Biparental crossing was evaluated by calculating the multilocus correlated paternity [$r_{p(m)}$], which measures the probability of two random individuals with the same pollen donor, and a $r_{p(m)}$ value of 0.889 was obtained. This high value indicates significant relationships among descendants (full-sibs). High and significant [$r_{p(m)}$] values indicate that the effective number of pollen donors is small (Ritland 1989). The correlated paternity can be used to estimate the average effective number of pollen donors per maternal plant (N_{ep}), i.e., the mean number of individuals that contributed pollen to a mother plant. The number of pollen donors was very low (1.12), and the high correlated paternity [$r_{p(m)}$] and low number of pollen donors per maternal plant (N_{ep}) may have resulted from the different flowering periods because of genetic differences between mother plants, which resulted in preferential mating between few individuals, with plants that flowered together exchanging more pollen than plants flowering at other times.

The fixation index for maternal genotypes [$F_{(m)}$] was 0.945; this value indicates that a significant portion of the parental generation resulted from inbreeding. This may be because these are seeds of local varieties homogeneous and do not mix with other varieties. A lower fixation index ($F_{(m)} = 0.504$) was reported by Zoro Bi et al. (2005) for wild Lima bean, indicating that the tested genitors resulted from mating between related and unrelated individuals in equal percentage.

Genetic diversity in Lima bean from Brazil

For the entire dataset, all SSR loci tested were polymorphic. Sixty-one alleles were detected, with the number of alleles per locus varying from 4 to 10, with a mean of 6.10 (Table 4). The loci with a higher number of alleles were BM212 and GATS19 (ten alleles), and those with a lower number were AG1, BM146 and BM154 (four alleles). Observed heterozygosity (H_O) varied from 0.024 (locus BM211) to 0.292 (locus BM160), with a mean of 0.077; and the expected heterozygosity (H_E) was from 0.167 (locus AG1) to 0.784 (locus GATS91), with a mean of 0.60. We can compare our results with studies made on Lima bean landraces from several regions from America. Martinez-Castillo et al. (2008) analyzed the genetic diversity of 21 landraces planted by the Mayan farmers from the Yucatán Peninsula, Mexico, reporting an $H_E = 0.290$ using 91 ISSRs markers. For the same region, Martínez-Castillo et al. (2012) analyzed landraces collected in 1979 and in 2007 using eight SSR loci, reporting a $H_E = 0.18$ and 0.05, respectively. Montero-Rojas et al. (2013) characterized 50 landraces from the Caribbean region using 24 SSR loci. They found 64 alleles, with a mean of 2.67. Also, these authors reported an $H_O = 0.403$ and an $H_E = 0.337$. Compared with these studies, our results could indicate the existence of high genetic diversity levels in the accessions of Lima bean of the Germplasm Bank of the UFPI. However, caution is recommended considering the differences among these studies (sampling strategies, sample size, number and type of markers, origin of the samples).

Martinez-Castillo et al. (2006, 2014) reported a $H_E = 0.69$ and a 0.61, respectively, for wild Lima bean populations from Mexico using SSR markers. Interestingly a similar value was found in the present study ($H_E = 0.60$). High values such as these have been reported just for wild germplasm but not for domesticated one. One possible explanation for the H_E

Table 4 Parameters of genetic diversity per locus, number of alleles per locus (A), expected heterozygosity (H_E) and observed heterozygosity (H_O) of 14 Lima bean families, using 10 microsatellite loci

Loci	A	H_E	H_O
AG1	4	0.167	0.071
BM140	5	0.713	0.065
BM141	6	0.745	0.048
BM146	4	0.686	0.065
BM154	4	0.524	0.065
BM156	7	0.619	0.042
BM160	5	0.417	0.292
BM211	6	0.680	0.024
BM212	10	0.642	0.053
GATS19	10	0.784	0.042
Mean	6.1	0.600	0.077

found in the present study is that the material used consisted of local unimproved cultivars without great loss of genetic diversity. Thus, the Lima bean landraces from Brazil can be considered unimproved germplasm and these could be conserving high genetic diversity levels that can be used for crop breeding.

High total diversity was observed for all dataset ($H_T = 0.596$), with a small portion attributed to intrapopulation diversity ($H_S = 0.058$). Diversity between families (D_{ST}) was 0.538. Of the total genetic diversity observed, 90.2 % was caused by genetic differences between families ($G_{ST} = 0.902$) (Table 5). This diversity distribution pattern is expected for species with selfing or mixed mating systems, with a predominance of the former where the gene flow is low or absent, resulting in low genetic diversity within populations (H_S low) and high genetic diversity between populations (D_{ST} and G_{ST}) as well as differentiation between families. This indicates the need for a large number of genotype collections in germplasm banks so that the maximum diversity possible can be used in the implementation of breeding programs. Montero-Rojas et al. (2013) found $H_T = 0.411$, $H_S = 0.396$, $G_{ST} = 0.036$. Castiñeiras et al. (2007) using AFLP molecular markers in landraces planted in the Cuban home gardens reported $H_T = 0.119$, $H_S = 0.119$. The high value observed in the present study ($G_{ST} = 0.902$) may have been a result of the obtained material from a germplasm bank collection, which due to several generations of predominantly selfing, resulted in the homogenization of the crop and absence of gene flow between different varieties. This is supported by the high fixation index for maternal genotypes observed (Fm = 0.945).

At the family level, the mean number of alleles varied from 1.1 (families 4, 13 and 14) to 1.8 (family 9) (Tables 6); the percentage of polymorphic loci (%P) was from 10 (families 4, 13 and 14) to 60 % (family 9), with a mean of 30 %. The H_E varied from 0.015 (family 13) to 0.120 (family 8), with a mean of 0.060; and the H_O varied between 0.016 (family 13) and 0.20 (family 8). The high values observed for families 1, 5, 7 and 8 may have been because of the occurrence of higher outcrossing percentages within the families. H_O was higher than H_E within families, showing an excess of heterozygotes relative to the Hardy–Weinberg Equilibrium (HWE) this generation; this observation can be attributed to a high natural outcrossing rate. The number of alleles found indicated considerable allelic diversity in the studied accessions even when taking into account the predominance of selfing in the germplasm analyzed. The low H_O and H_E values found within families was expected because of the predominance of selfing and the high fixation index for maternal genotypes (Fm = 0.945). These low values are expected in

Table 5 Nei's parameters of genetic diversity (1973): intrapopulation diversity (*Hs*), total diversity (H_T), diversity between families (D_{ST}), and percentage genetic diversity between families (G_{ST}) of 14 families of Lima bean, using 10 microsatellite loci

Loci	*HS*	H_T	D_{ST} (H_T-H_S)	G_{ST} (D_{ST}/H_T)
AG1	0.065	0.165	0.100	0.606
BM140	0.063	0.712	0.649	0.911
BM141	0.043	0.743	0.700	0.942
BM146	0.042	0.684	0.642	0.939
BM154	0.060	0.524	0.463	0.885
BM156	0.040	0.618	0.577	0.935
BM160	0.155	0.416	0.261	0.628
BM211	0.023	0.679	0.655	0.966
BM212	0.052	0.641	0.588	0.918
GATS19	0.040	0.782	0.742	0.948
Mean	0.058	0.596	0.538	0.902

Table 6 Parameters of genetic diversity per family, average number of alleles per locus (A), percentage of polymorphic loci (% P), expected heterozygosity (H_E), observed heterozygosity (H_O), and fixation index (f) of 14 families of Lima bean using 10 microsatellite loci

Maternal families	A	% P	H_E	H_O	F
1	1.40	30	0.095	0.141	−0.527
2	1.50	30	0.048	0.050	−0.312
3	1.30	30	0.025	0.025	0.000
4	1.10	10	0.029	0.033	−0.158
5	1.60	50	0.110	0.158	−0.467
6	1.30	30	0.032	0.033	−0.023
7	1.40	40	0.077	0.125	−0.667
8	1.40	30	0.120	0.200	−0.714
9	1.80	60	0.088	0.092	−0.043
10	1.40	20	0.053	0.058	−0.092
11	1.20	20	0.031	0.033	−0.047
12	1.50	50	0.078	0.083	−0.068
13	1.10	10	0.015	0.016	−0.048
14	1.10	10	0.022	0.025	−0.100
Mean	1.36	30	0.060	0.077	−0.316

species with a predominance of selfing because most of their loci are homozygous. Even with a significant outcrossing rate was not sufficient for a considerable increase in diversity. In selfing species such as common bean, H_O may be very low and may reach zero (Burle et al. 2010).

The fixation index (f) is an important parameter in population genetics because it allows the measurement of the ratio between homozygotes and heterozygotes in different populations (Kageyama et al. 2003). Interestingly, in the present study the fixation index was negative (mean = −0.316) for the different families (except for family 3, which exhibited f = 0). This result indicates a higher percentage of heterozygotes than expected under HWE, which could be explained because of the considerable outcrossing rate observed. Deviations from HWE were possibly caused by the reproductive subdivision into groups because of preferential mating (Sebbenn et al. 1998). Only family 3 may have been under HWE. Excess of heterozygotes had been reported for Lima bean landraces. Montero-Rojas et al. (2013) found an excess of heterozygotes (−0.1618) in landraces from Caribbean region using SSR markers.

Implications for conservation and breeding

Understanding the genetic potential of the Lima bean accessions stored in germplasm banks of Brazil is extremely important for the conservation of this crop and its use in breeding programs. Despite its importance as a crop, improved Lima bean varieties have not been produced for Brazil and there is little information on its system of reproduction and genetic diversity. The present study showed that the Lima bean from Brazil has a 38.1 % natural outcrossing rate, indicating a mixed mating system with a predominance of selfing. Most crossings occurred between related individuals, thus resulting in progenies of full-sibs as well as half-sibs and selfing-sibs. Thus, multiplication and regeneration of the germplasm bank material should therefore be performed under controlled conditions, such as screenhouse planting or field planting with significant distances between different genotypes to prevent pollen exchange between plants. Such processes increase the speed at which different plant lines are generated, which is one advantage for breeding programs that are limited by time and have a high demand for improved varieties. The significant selfing rate, high correlated paternity, and low diversity within families have implications for the conservation and breeding of *P. lunatus* because they indicated that a large sample size is necessary to retain maximum diversity for ex situ conservation. Also, high genetic diversity was observed among Lima bean accessions from Brazil, higher to that reported for other regions from America and similar to the reported for wild populations from the Yucatan Peninsula in Mexico, which is considered a center of genetic diversity for *P. lunatus*. It could be explained because of the existence of an excess of heterozygotes because of the high outcrossing rate, which led to increased diversity. Genetic diversity was low within families, with 90.2 % of the total diversity occurring between families. This result is consistent with the mating system of this species, which has a predominance of selfing and results in homogenization within families and increased differentiation between families. The observed genetic diversity shows that several families have the potential for use in breeding programs. Data obtained from the present study must be combined with the results of other studies, including morphological, biochemical and agronomical characterisations, with the goal of improving the knowledge of *P. lunatus* and thus promote conservation and breeding programs better suited to the needs of the Brazilian farmers.

Acknowledgments The authors thank the Laboratory of Genetic Diversity and Breeding of the Luiz de Queiroz College of Agriculture (ESALQ), where the present study was performed. Thanks are also due to Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq for the financing and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior- CAPES for the scholarship granted to the first author.

References

Azevedo JN, Franco LJD, Araújo ROC (2003) Composição química de sete variedades de feijão-fava. Embrapa Meio-Norte, Teresina

Baudet JC (1977) The taxonomic status of the cultivated types of Lima bean (*Phaseolus lunatus* L.). Trop Grain Legume 7:29–30

Baudoin JP (1988) Genetic resources, domestication and evolution of Lima bean, *Phaseolus lunatus*. In: Gepts P (ed) Genetic resources of *Phaseolus* bean. Kluwer Academic Publishers, Amsterdam, pp 393–407

Baudoin JP, Rocha O, Degreef J, Maquet A, Guarino L (2004) Ecogeography, demography, diversity and conservation of *Phaseolus lunatus* L. in the central valley of Costa Rica. Systematic and ecogeographic studies on crop genepools. Int Plant Genet Resourc Inst, Rome

Borém A, Miranda GV (2009) Melhoramento de plantas, 5th edn. UFV, Viçosa

Burle ML, Fonseca JR, Kami JA, Gepts P (2010) Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a secondary center of diversity. Theor Appl Genet 121:801–813

Buso GSC, Ciampi AY, Moretzhon MC, Amaral ZPS, Brondani RV (2003) Marcadores microssatélites em espécies vegetais—desenvolvimento e caracterização de marcadores microsatélites em espécies vegetais tropicais. Biotecnol Ciênc Desenvolv 30:46–50

Carmo MDS, Gomes RLF, Lopes ACA, Penha JS, Gomes SO, Assunção Filho JR (2013) Genetic variability in subsamples of determinate growth Lima bean. Crop Breed Appl Biotechnol 13:158–164

Castiñeiras L, Guzmán FA, Duque MC, Shagarodsky T, Cristóbal R, De Vicente MC (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

Doyle JJ, Doyle JL (1987) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Fausto JA, Eckhart VM, Geber MA (2001) Reproductive assurance and the evolutionary ecology of self-pollination in *Clarkia xantiana*. Am J Bot 88:1794–1800

Ferreira ME, Moretzsohn MC, Buso GSC (2007) Fundamentos de caracterização molecular de germoplasma vegetal. In: Nass LL (ed) Recursos genéticos vegetais. Embrapa Recursos genéticos e Biotecnologia, Brasília, pp 379–420

Gaitán-Solís E, Duque MC, Edwards KJ, Tohme J (2002) Microsatellite repeats in common bean (*Phaseolus vulgaris* L.): isolation, characterization, and cross-species amplification in *Phaseolus* ssp. Crop Sci 42:2128–2136

Goudet J (2002) FSTAT, a program to estimate and test gene diversities and fixation indices. Version 2.9.3.2. Université de Lausanne, Lausanne. http://www.unil.ch/popgen/softwares/fstat.htm. Acesso 22 dezembro 2013

Guimarães WNR, Martins LSS, Silva EF, Ferraz GMG, Oliveira FJ (2007) Caracterização morfológica e molecular de acessos de feijão-fava (*Phaseolus lunatus* L.). Rev Bras de Eng Agrícola e Ambient 11:37–45

Hamrick JL (1982) Plant population genetics and evolution. Am J Bot 69:1685–1693

Hamrick JL, Loveless MD (1986) Isozyme variation in tropical trees; procedures and preliminary results. Biotropica 8:201–207

Hamrick JL, Murawski DA (1991) Levels of allozyme diversity in populations of uncommon neotropical tree species. J Trop Ecol 7:395–399

Harding J, Tucker CL (1964) Quantitative studies on mating systems. I. evidence for the non-randomness of outcrossing in *Phaseolus lunatus*. Heredity 19:369–381

Hardy O, Dubois S, Zoro BII, Baudoin JP (1997) Gene dispersal and its consequences on the genetic structure of wild populations of Lima bean (*Phaseolus lunatus*) in Costa Rica. Plant Genet Resour Newsl 109:1–6

Jiang GL (2013) Molecular markers and marker-assisted breeding in plants. In: Andersen SB (Ed) Plant breeding from laboratories to fields. InTech, University of Copenhagen. doi:10.5772/3362

Kageyama PY, Sebbenn AM, Ribas LA, Gandara FB, Castellen M, Perecim MB, Vencovsky R (2003) Diversidade genética em espécies arbóreas tropicais de diferentes estágios sucessionais por marcadores genéticos. Sci For 64:93–107

Kaplan L, Lynch TF (1999) *Phaseolus* (Fabaceae) in archaeology: AMS radiocarbon dates and their significance for pre-Colombian agriculture. Econ Bot 53:261–272

Lewis P, Zaykin D (2000) Genetic data analysis: computer program for the analysis of allelic data version 1.0. http://www.eeb.uconn.edu/people/plewis/software.php. Accessed 27 Dec 2013

Lopes ACA, Gomes RLF, Araújo ASF (2010) A cultura do feijão-fava no Meio-Norte do Brasil. EDUFPI, Teresina

Maquet A, Zoro Bi I, Delvaux M, Wathelet B, Baudoin JP (1997) Genetic structure of a Lima bean base collection using allozyme markers. Theor Appl Genet 95:980–991

Martinez-Castillo J, Zizumbo-Villarreal D, Gepts P, Delgado-Valerio P, Colunga-Garcíamarín P (2006) Structure and genetic diversity of wild population of Lima bean (*Phaseolus lunatus* L.) from the Yucatan Peninsula. Mex Crop Sci 46:1071–1080

Martinez-Castillo J, Colunga-Garcíamarín P, 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 Evol 55:1065–1077

Martínez-Castillo J, Zizumbo-Villarreal D, Perales-Rivera H, Colunga-GarcíaMarín P (2004) Intraspecific diversity and morpho-phenological variation in *Phaseolus lunatus* L. from the Yucatan Peninsula, Mexico. Econ Bot 58(3):354–380

Martínez-Castillo J, Camacho-Pérez L, Coello-Coello J, Andueza-Noh R (2012) Wholesale replacement of Lima bean (*Phaseolus lunatus* L.) landraces over the last 30 years in northeastern Campeche, Mexico. Genet Resour Crop Evol 59:191–204

Martínez-Castillo J, Camacho-Pérez L, Villanueva-Viramontes S, Andueza-Noh RH, Chacón-Sánchez MI (2014) Genetic structure withinthe Mesoamerican gene pool of wild *Phaseolus lunatus* (Fabaceae) from Mexico as revealed by microsatellite markers: implication for conservation and the domestication of the species. Am J Bot 101(5):1–14

Montero-Rojas M, Ortiz M, Beaver JS, Siritunga D (2013) Genetic, morphological and cyanogen content evaluation of a new collection of Caribbean Lima bean (*Phaseolus lunatus* L.) landraces. Genet Resour Crop Evol 60:2241–2252

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Oliveira AP, Alves EU, Alves AU, Dornelas CSM, Silva JA, Porto ML, Alves AV (2004) Produção de feijão-fava em função do uso de doses de fósforo em um neossolo regolítico. Hortic Bras 22:543–546

Ortega SV (1974) Polinización cruzada natural de la caraota (*P. vulgaris* L.) en Venezuela. Agron Trop 24:27–32

Ramalho MAP, Santos JB, Abreu AFB, Nunes JAR (2012) Aplicações de genética quantitativa no melhoramento de plantas autógamas, 1st edn. UFLA, Lavras

Ritland K (1989) Correlated matings in the partial selfer *Mimulus guttatus*. Evolution 43:848–859

Ritland K (1990) A series of fortran computed programs for estimating plant mating systems. J Hered 81:235–237

Ritland K (2008) Multilocus mating system program MLTR (Software) version 3.2. http://genetics.forestry.ubc.ca/ritland/programs.html. Accessed 27 Dec 2013

Ritland K, Jain SK (1981) A model for the estimation of outcrossing rate and gene frequencies using independent loci. Heredity 47:35–52

Santos D, Corlett FMF, Mendes JEMF, 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

Sebbenn AM, Siqueira ACMF, Kageyama PY, Machado JAR (1998) Parâmetros genéticos na conservação de cabreúva—*Myroxylon peruiferum* LF Allemão. Sci For 53:31–38

Sebbenn AM, Kageyama PY, Siqueira ACMF, Zanatto ACE (2000) Sistema de cruzamento em populações de *Cariniana legalis* (Mart.) O. Ktze.: impli-cações para a conservação e o melhoramento genético. Sci For 58:25–40

Valls JFM (2007) Caracterização de recursos genéticos vegetais. In: Nass LL (ed) recursos genéticos vegetais. Embrapa Recursos Genéticos e Biotecnologia, Brasília, pp 281–342

Webster BD, Lynch SP, Tucker CL (1979) A morphological study of the development of reproductive structures of *Phaseolus lunatus* L. J Am Soc Hort Sci 104:240–243

Wright S (1921) Systems of mating. Genetics 6:111–178

Zoro Bi I (1999) Variabilité génétique des populations sauvages de *Phaseolus lunatus* L. dans la vallée centrale du Costa Rica et ses implications dans la mise ou point d' une stratégíe de conservation in situ. Ph. D. thesis. Faculté Universitaire des Sciences Agronomiques, Gembloux, Belgium

Zoro Bi I, Maquet A, Baudoin JP (2005) Mating system of wild *Phaseolus lunatus* L. and its relationships with population size. Heredity 94:153–158