## A. Maquet · I. Zoro Bi · M. Delvaux B. Wathelet · J.-P. Baudoin

# Genetic structure of a Lima bean base collection using allozyme markers

Received: 3 June 1997 / Accepted: 17 June 1997

Abstract Genetic diversity and structure within a Lima bean (*Phaseolus lunatus* L.) base collection have been evaluated using allozyme markers. The results obtained from the analysis of wild and cultivated accessions confirm the existence of Andean and Mesoamerican gene pools characterised by specific alleles. Wild and cultivated accessions of the same gene pool are grouped. The Andean natural populations have a very limited geographic distribution between Ecuador and northern Peru. The Mesoamerican wild form extends from Mexico up to Argentina through the eastern side of the Andes. Andean and Mesoamerican cultivated accessions of pantropical distribution contribute substantially to the genetic diversity of the Lima bean base collection. Population genetic parameters, estimated from allozymes, confirmed the predominant selfing mating system of the Lima bean. The selfing mating system, the occurrence of small populations, and low gene flow lead to an interpopulation gene diversity  $(D_{ST} = 0.235)$  higher than the intrapopulation gene diversity  $(H_s = 0.032)$ . On the basis of the results, guidelines are given to preserve and exploit the genetic diversity of this threatened species. The results also confirm the independent domestication of the Lima bean in at least two centres, one of which is located at medium elevation in the western valleys of Ecuador and northern Peru.

Communicated by P. M. A. Tigerstedt

M. Delvaux · B. Wathelet

Key words *Phaseolus lunatus* L. · Genetic diversity · Plant conservation · Domestication · Gene pool

#### Introduction

The Lima bean (*Phaseolus lunatus* L.) is the second most important food legume in the genus *Phaseolus* after the common bean (*P*. *vulgaris* L.) but before the runner bean (*P*. *coccineus* L.), the tepary bean (*P*. *acutifolius* Gray) and *P*. *polyanthus* Greenm. (National Research Council 1989). The Lima bean is grown in most of the low-elevation tropical areas of the world where its landraces are progressively being abandoned by farmers and replaced by Vigna unguiculata (L.) Walp., *Lablab purpureus* (L.) Sweet and *Cajanus cajan* Millsp. (Maquet and Baudoin 1997). Improved varieties of these species are either more resistant to dry conditions or to bruchids, or else are earlier with more stable yields. At the same time, wild populations of *P*. *lunatus*, often found on borders of deciduous forests, are disappearing throughout Latin America due to de-forestation or urbanisation (Maquet and Baudoin 1997).

To preserve the genetic diversity of the Lima bean, recent collecting missions were conducted in Latin America by the International Plant Genetic Resources Institute (IPGRI, Rome, Italy), the ''Centro Internacional de Agricultura Tropical'' (CIAT, Cali, Colombia), and national institutes (Debouck 1987 a, b, 1988 b, 1989, 1990 a, b, 1995). In 1991, the Lima bean world collection held at CIAT contained 1496 accessions comprising 90% cultivars, 9% wild accessions and 1% weedy accessions (Maquet 1991). Evaluation of the genetic diversity of a base collection is helpful for an effective management, the establishment of a core collection, the guidance of further prospections, as well as the preservation of genetic integrity during seed rejuvenation. Using seed-protein markers, Maquet

A. Maquet  $(\boxtimes) \cdot$  I. Zoro Bi  $\cdot$  J.-P. Baudoin Faculté Universitaire des Sciences Agronomiques, Unité de Phytotechnie des Régions Intertropicales, 2 Passage des Déportés, B-5030 Gembloux, Belgium E-mail: phytotrop@fsagx.ac.be

Faculté Universitaire des Sciences Agronomiques, Unité de Chimie Organique et Biologique, 2 Passage des Déportés, B-5030 Gembloux, Belgium

et al. (1990), Gutiérrez et al. (1995), and Maquet (1995) suggested the existence of two gene pools. A Mesoamerican gene pool distributed from Mexico up to Argentina, including Brazil, showed a basic seedprotein electrophoretic banding pattern (''M'') with several variants. In South America, the Mesoamerican wild form is found along the eastern side of the Andes. An Andean gene pool circumscribed in South America, was characterised by a different basic seed-protein pattern (''A'') with other variants. A hybrid pattern (''H'') was also observed in cultivated accessions from Peru and Colombia. The small-seeded Mesoamerican landraces (''Sieva'' and ''Potato'' cultigroups) showed reduced genetic diversity; for example, one of the pattern variants (M1) was present at a high frequency along the whole geographic area of distribution from Mexico to Argentina. Such a founder effect was not observed among large-seeded Andean landraces (Big-Lima cultigroup), a possible consequence of insufficient sampling (Gutiérrez et al. 1995).

Isozymes are useful markers to study plant systematics (Crawford 1990; Murphy et al. 1990), crop evolution (Doebley 1990), the genetics of populations (May 1992), and plant conservation (Schaal et al. 1991; Gepts 1995). Studies on the genetic diversity of the Lima bean remain limited. Allozyme or DNA markers were used for taxonomic studies in the genus *Phaseolus* (West and Garber 1967; Hamann et al. 1995; Jaaska 1996) but only a few Lima bean accessions were sampled. Other studies were carried out to assess the relationships among Lima bean landraces (Nienhuis et al. 1995; Lioi and Lotti 1996). In order to help guiding ex situ and in situ conservation, a base collection involving several widely distributed wild accessions and landraces was studied using allozymes.

### Materials and methods

#### Plant materials

From the CIAT base collection, we selected 235 Lima bean accessions originating from all over Latin America and the Carib zone (Maquet 1991). The material included 49 wild accessions (14 from South America, and 35 from Mesoamerica), seven weedy accessions or hybrids between wild and cultivated forms (three from southern Andes and four from Mesoamerica), and 179 landraces with a wide range of seed sizes, shapes, and colours and belonging to the three cultigroups or to intermediate cultivated forms as described by Debouck (1990 b). For the selected accessions, either original or rejuvenated seeds were sampled.

#### Electrophoretic analysis

For each accession, 1*—*5 seeds were analysed. To evaluate the genetic diversity in the wild Lima bean accessions, we selected ten enzymes obtained from the cotyledon tissues: alcohol dehydrogenase (ADH, E.C. 1.1.1.1), cytosol amino-peptidase (CAP, E.C. 3.4.11.1), dihydrolipoamide dehydrogenase (DDH, E.C. 1.8.1.4), en-

dopeptidase (END, E.C. 3.4.23.6), general protein (GP, non-specific), glucose-6-phosphate isomerase (GPI, E.C. 5.3.1.9), isocitrate dehydrogenase (IDH, E.C. 1.1.1.42), malate dehydrogenase (MDH, E.C. 1.1.1.37), phosphogluconate dehydrogenase (PGDH, E.C. 1.1.1.44), and phosphoglucomutase (PGM, E.C. 5.4.2.2). Electrophoretic analyses were conducted for ADH, CAP, END, GPI, IDH, MDH, PGDH and PGM on horizontal 10% starch (Sigma  $#$  S-4501) gels containing  $3\%$  sucrose (Sigma  $\#$  S-8501) and according to the method of Maquet et al. (1994). DDH and GP enzymes were separated on PAGE according to the method used by Delvaux (1994). Five-day old imbibed cotyledons were ground in a potassium phosphate 0.1 M pH 7.0 buffer as described by Hussain et al. (1988). The crude homogenate absorbed onto paper-wicks was used for starch-gel electrophoresis following the method described by Murphy et al. (1990) and using the histidine-citrate pH 5.7 buffer system. Polyacrylamide gels were prepared according to Laemmli (1970). The loading gel had 3.1% T and 2.6% C in a 0.5 M Tris-HCl pH 6.8 buffer. The separator gel was prepared to obtain 8.2% T and 2.6% C in a Tris-HCl 1.5 M pH 8.8 buffer. The electrode buffer was a 0.025 M Tris, 0.019 M glycine buffer at pH 8.3. Homogenates (6  $\mu$ l) were electrophoresed for 5 h (DDH) or 7 h (GP) at 80 mA and 500 V in a Biorad Protean II apparatus. The histochemical staining procedures used were those of Hussain et al. (1988) for CAP, DDH, END, GP, GPI and MDH, and those of Murphy et al. (1990) for ADH, IDH, PGDH and PGM.

Loci were labelled sequentially with those migrating closest to the anodal end designated as number 1 (Koenig and Gepts 1989). The first and the last stacks correspond to accession G25221, a Mexican wild form, considered as a standard for our analysis. The allele from this genotype is designated as 100 and all other allozymes are assessed according to their relative distance from the standard. The genetic control and the quaternary structure of the analysed enzyme systems were discussed previously by Maquet et al. (1994) and Zoro Bi et al. (1997 b).

#### Genetic-variation analysis

Allozyme diversity was estimated within all sampled accessions and within groups of accessions using the programme *Gen*-*Survey* (Vekemans and Lefebvre 1997). Genotype and allele frequencies, the percentage of polymorphic loci  $(P)$ , the mean number  $(n_a)$  and the effective number  $(n_e)$  of alleles per locus, and gene diversity were calculated both at the species level  $(H_{\rm es})$  and the within-accession level  $(H_e)$ . These parameters were compared with those obtained by Hamrick et al. (1991) who classified 473 taxa using eight species traits: major phyletic group, life form, geographic range, regional distribution, breeding system, seed dispersal mechanism, mode of reproduction, and successional status. The parameters were also estimated after grouping Lima bean accessions according to their geographic origin and their electrophoretic seed-protein patterns obtained from previous studies (Maquet 1991, 1995; Gutiérrez et al. 1995). As suggested by Weir (1990) in the case of non-panmictic populations, we used genotypic frequencies rather than allelic frequencies to test for differentiation among groups of accessions with contingency tables. After the alleles codification, the genotypes were tentatively grouped into different classes to obtain an expected frequency at least equal to five. In this case, a significance test was performed to evaluate if the genotype frequencies in the Andean and Mesoamerican groups were equal. For an expected frequency lower than five, an angular transformation was employed (Dagnelie 1978).

Out of the 235 accessions, 72 were selected and used to calculate Nei's gene-diversity statistics (Nei 1973). Three to five seeds were analysed for each accession. Total gene diversity  $(H_T)$  and mean diversity within accessions  $(H<sub>s</sub>)$  were calculated for each polymorphic locus. Gene diversity due to variation among accessions  $(D_{ST})$ was related to the total diversity to determine the proportion residing among accessions  $(G_{ST})$ . When the comparison involved two closely related groups, we used the Student *t*-test for paired

observations, locus by locus, to determine their difference in gene diversity (Nei 1987). This test was also performed to evaluate the difference between the fixation index (f) in original and rejuvenated seeds of *P*. *lunatus*. Nei's corrected genetic distances (Nei 1987) were calculated using the *Phylip* software (Felsenstein 1993), and the distance matrix was used to elaborate a cluster based on the "weighted pair group arithmetic averaging" link (WPGMA) with the *Systat* software (Wilkinson 1989).

## Results and discussion

The ten enzyme systems show 13 loci expressing 32 alleles. As shown in Table 1, the number of alleles per locus varies between one (e.g. *Pgdh*-*1* locus) and four (*Cap*, *End*, and *Pgdh*-*2*). The *Pgm*-*2* locus expresses a null allele and only the homozygous genotype *Pgm*- $2^{n/n}$  specific for the Andean region is noted. The independence tests for the genotype frequencies between the two seed-protein families are significant for all loci (Table 1). Allelic frequencies at the species level or according to the geographic region and to the seedprotein family are given in Table 2. Lima bean accessions classified according to geographic regions do not show a clear picture of the specific alleles. Indeed, some accessions originating in the Andes or in Brazil have a Mesoamerican seed-protein pattern and alleles (e.g. allele *Adh*-*261*). In South America, Mesoamerican and Andean Lima bean accessions can be in contact. In Brazil, small-seeded accessions are mainly distributed in the Northeast and along the Atlantic coast, while large-seeded landraces are localised in the southern and central parts (Maquet 1991). On account of the presence of small-seeded accessions in South America, we have grouped Lima bean accessions according to the seed-protein families. Alleles *Adh*-*261*, *Cap<sup>94</sup>*, *End<sup>88</sup>*,  $End^{139}$ , *Ddh-2<sup>87</sup>*, and  $Gp^{111}$  are found only among accessions showing a Mesoamerican seed-protein pattern. These alleles are observed in both wild and cultivated accessions except for allele *End<sup>139</sup>*, identifying a wild and a weedy accession, and for allele  $Gp^{III}$ , specific to the cultivated accessions. Three alleles (*Cap<sup>118</sup>*, *Idh<sup>120</sup>*, and *Pgdh*-*2135*) are endemic in the Andean gene pool. The former is found only among cultivated accessions while the two others are observed among wild accessions. The remaining alleles are distributed in both gene pools. Such shared alleles limit the risk of loss in a gene pool especially in the narrowly distributed Andean one.

Among the weedy form, only the alleles *Adh*-*2100*, *Gp<sup>100</sup>*, *Idh<sup>100</sup>*, *Mdh*-*3100*, *Pgdh*-*1100*, *Pgdh*-*2100* and *Pgm*-*1100* are noted for these loci. The other loci are polymorphic:  $Cap^{100}$  (allelic frequency = 0.2),  $Cap^{108}$ (0.8), *End<sup>100</sup>* (0.8), *End<sup>139</sup>* (0.2), *Gpi*-*196* (0.5), *Gpi*-*1100* (0.5), *Mdh*-*2100* (0.2), *Mdh*-*2140* (0.8), *Ddh*-*2<sup>87</sup>* (0.4) and *Ddh*-*2100* (0.6). These results are explained by the genetic nature of these materials arising from hybridisations between a wild form and a sympatric landrace.

The dendrogram in Fig. 1 depicting the genetic relationship among the Lima bean accessions clearly shows two main clusters. The first includes mainly accessions originating from the Andes and having an Andean seed-protein pattern. The second cluster contains Mesoamerican and Andean accessions but are characterised by a Mesoamerican seed-protein pattern. Both Andean and Mesoamerican landraces are grouped with their respective wild relatives. Such a cluster can be interpreted as the result of two independent domestications, as found in *P*. *vulgaris* by Singh et al. (1991). Considering only the landraces, the dendrogram (Fig. 2) does not show a separation between the two Mesoamerican cultigroups ("Sieva" and "Potato"). Gutiérrez et al. (1995) and Maquet (1995) showed that a Mesoamerican seed-protein variant (M1) is found in approximately 95% of the small-seeded landraces. Such a genetic uniformity based on seed protein is not reflected by the study on allozymes: landraces are clustered in different groups suggesting various allelic combinations (Fig. 2).

Some accessions exhibit characteristics from both gene pools: for example G25844 and G25290 with a Mesoamerican seed-protein pattern and Andean alleles, and G25417 with an Andean seed-protein pattern and Mesoamerican alleles (Fig. 1 and 2). These accessions and the two Peruvian entries showing a hybrid seed-protein pattern (G25941 and G25577) can be considered as products of hybridisation between genotypes from both gene pools. Indeed, accessions with a hybrid seed-protein pattern are characterised by alleles belonging to both gene pools (Table 2).

Table 3 indicates the genetic diversity of *P*. *lunatus* in relation to different categories. At the level of the base collection, 75% of the loci are polymorphic with an average of 2.6 alleles per locus. However, the alleles are not evenly distributed for each locus  $(n_e = 1.35)$ . The mean genetic diversity for the 473 taxa reviewed by Hamrick et al. (1991) was 0.149. In comparison, the genetic diversity of *P*. *lunatus* at the species level  $(H_{\rm es} = 0.26)$  is relatively important and higher than species with mixed-mating  $(H_{es} = 0.12)$  or short-lived perennial species ( $H_{es} = 0.12$ ) but closer to those of widely distributed species ( $H_{es} = 0.20$ ). This similarity is in relation to the wide distribution of the Lima bean in Latin America. According to Hamrick et al. (1991), geographic range accounts for the largest proportion of the explained variation in genetic diversity at the species level.

Within accessions, individuals are polymorphic, on average, at only 4% of their loci (Table 3). This very low value is probably due to our small sample size (1*—*5 seeds analysed per accession) but also to the mating system of *P*. *lunatus*. Indeed, the mating system accounts for the largest proportion of the variation in plant genetic diversity at the population level (Hamrick et al. 1991): for selfing species, 20% of the loci were polymorphic per population, whereas for outcrossed wind-pollinated species, nearly 50% of the loci were polymorphic. According to our results, the genetic



Table 1 Distribution of absolute genotypic frequencies in *P*. *lunatus* L. according to the seed-protein family. NB: this table contains  $n$ tains  $t_0 h$ la  $thio$  $\frac{1}{2}$  $\frac{1}{2}$  $\ddot{r}$  $\overline{\phantom{a}}$  $\ddot{\phantom{a}}$ É,  $\ddot{\phantom{1}}$  $\frac{1}{2}$  $\overline{a}$  $\overline{a}$  $\mathbf{a}$  $\ddot{\phantom{0}}$  $\ddot{\cdot}$  $\ddot{a}$  $\cdot$  $\ddot{\tau}$  $dim A$ میں<br>میاہ ا.<br>ا Table 1 Distribu

983

Homozygous genotype with the null allele



984



Fig. 1 Dendrogram depicting the genetic relationship of *P*. *lunatus* L. according to biological status. Biological status (*S*: wild form; *W*: weedy form; L: landraces), seed-protein pattern (A: Andean; M: Mesoamerican; *H*: Hybrid), country of origin, and accession number in the CIAT collection are presented on the left side. *ARG*: Argentina; *BOL*: Bolivia; *BRA*: Brazil; *COL*: Colombia; *CRI*: Costa Rica; *EC*º: Ecuador; *G*¹*M*: Guatemala; *HND*: Honduras; *JAM*: Jamaica; MEX: Mexico; PAN: Panama; PER: Peru; SLV: Salvador

diversity of *P*. *lunatus* within accessions ( $H_e = 0.03$ ) is half the average genetic diversity of selfing species  $(H_e = 0.074)$ . Andean and Brazilian accessions seem to have a higher genetic diversity than Mesoamerican accessions. But this trend is reversed when Lima bean accessions are grouped according to their seed-protein pattern (Table 3). In fact, the genetic diversity between



wild and cultivated accessions, as well as between the accessions originating from the two geographic regions or between both gene pools, is not significantly different (Table 4).

On average, Lima bean shows 76% and 24% of the total diversity, respectively, among and within accessions (Table 5). The mating system and life form of the species are both usually highly associated with differences in  $G_{ST}$  values (Hamrick et al. 1991). This behaviour is even more important with Lima bean than with other selfing species ( $G_{ST} = 0.51$ ) or outcrossed species  $(G_{ST} = 0.10)$ .

These results indicate that species with a limited potential for gene flow show more differentiation among populations than do species with a greater potential for gene flow. Gene-flow potential is correlated with pollination and seed-dispersal mechanisms, as well as with longevity and plant size. In a study of flower and pollen dispersal in wild populations of *P*. *lunatus*from the Central Valley of Costa Rica, Hardy et al. (1997) showed that the gene dispersal is sufficiently restricted to expect the occurrence of local genetic differentiation among populations.

A hierarchical analysis of gene diversity (Nei 1987) was performed to estimate the differentiation among groups of accessions based on biological status, geographic origin and seed-protein pattern (Table 6). The total genetic diversity  $(H_T)$  is divided into diversity within accessions  $(H_C)$ , among accessions within groups  $(D_A)$ , and among accessions within differential  $(D_{CS})$ , and among groups  $(D_{ST})$ . A substantial differentiation among groups of accessions is only observed for groups based on seed-protein pattern: specific alleles are identified in each gene pool.

The low proportion of variation within accessions can be explained by various factors, such as: the seedrejuvenation phase, the few seeds tested per accession, the sample size of the original collected population, the mating system, and the original variation within the populations at their natural sites. In our study, 76% of Lima bean accessions were rejuvenated. This phase would probably decrease the genetic diversity within accessions as, generally, only a few seeds were used to multiply each accession. Nonetheless, the fixation index (f) is not significantly different between the rejuvenated and original seeds (Table 7). In fact, Maquet et al. (1996) showed that the wild Lima bean in Costa Rica maintained, on average, 80% and 20% of its total diversity respectively among and within populations. In addition, all the wild populations analysed were significantly inbred as shown by Weir and Cockerham's fixation indices estimated by the jackknife

Fig. 2 Dendrogram depicting the genetic relationship among the landraces of *P*. *lunatus* L. Cultigroup (*S*: Sieva, *P*: Potato, *B*: Big Lima, *IS*: intermediate Sieva, *IP*: intermediate Potato, *IB*: intermediate Big Lima), seed protein pattern (*A*: Andean; *M*: Mesoamerican; *H*: Hybrid), country of origin, and accession number in the CIAT

collection are presented on the left side. *ARG*: Argentina; *BOL*: Bolivia; *BRA*: Brazil; *COL*: Colombia; *CRI*: Costa Rica; *ECU*: Ecuador; GTM: Guatemala; HND: Honduras; JAM: Jamaica; *MEX*: Mexico; *PAN*: Panama; *PER*: Peru; *SLV*: Salvador

Table 3 Genetic diversities of *P*. *lunatus* L. according to different categories



<sup>a</sup> Total number of accessions analysed

 $b$  Percentage of polymorphic locus at 5% level

<sup>c</sup> Average number of alleles per locus

 $\frac{d}{dt}$  Effective number of alleles

<sup>e</sup> Nei's gene diversity

<sup>f</sup> Not calculated for accessions with one seed

Table 4 Student *t*-test of differences for gene diversity on *P*. *lunatus* L. according to biological status, geographic origin, and seed protein family



<sup>a</sup> Total number of accessions analysed

<sup>b</sup>Total gene diversity averaged over all loci

<sup>c</sup> Student *t*-test for paired observations

<sup>d</sup> Degrees of freedom

<sup>e</sup> Probability value

Table 5 Analysis of genetic diversity within and among accessions of *P*. *lunatus* L

Locus	$H_T^a$	$H_s^b$	$D_{ST}^{\quad c}$	$G_{ST}^{\dagger d}$
$Adh-2$	0.181	0.029	0.152	0.842
Cap	0.512	0.029	0.483	0.944
$Ddh-2$	0.527	0.024	0.502	0.954
End	0.507	0.097	0.410	0.809
Gp				
$Gpi-I$	0.415	0.026	0.389	0.937
<b>Idh</b>	0.008	0.008	0.001	0.078
$Mdh-2$	0.472	0.045	0.427	0.905
$Mdh-3$	0.031	0.015	0.016	0.524
$Pqdh-1$				
$Pqdh-2$	0.052	0.019	0.033	0.628
$Pqm-1$	0.467	0.030	0.437	0.936
Mean	0.317	0.032	0.285	0.755
SE <sup>e</sup>	0.070	0.008	0.066	0.088

<sup>a</sup> Gene diversity in the base collection

<sup>b</sup> Gene diversity within accessions

<sup>c</sup> Gene diversity between accessions

<sup>d</sup> Coefficient of gene differentiation

<sup>e</sup> Standard error

method (Maquet 1995). The high level of inbreeding was caused in part by local inbreeding within populations  $(f = 0.891)$  and substantial differentiation among populations ( $\theta = 0.814$ ). On the basis of the relation Table 6 Hierarchical gene diversity of *P*. *lunatus* L. according to biological status, geographic origin, and seed protein family



<sup>a</sup> Gene diversity in the base collection

<sup>b</sup> Gene diversity within accessions

<sup>e</sup> Gene diversity among accessions within groups

<sup>d</sup> Gene diversity among groups

between the fixation index (f) and the outcrossing rate (t) (Nei 1987), Maquet et al. (1996) estimated an outcrossing rate of 5.8% confirming the facultative autogamy system of the Lima bean. According to Hartl and Clark (1989), the drift resulting from a selfing breeding system is also more pronounced in small populations, a situation commonly found in the Lima bean: landraces are cultivated in small backyards as part of a subsistence agriculture (Debouck 1987 c; 1988 a, b) and natural populations are small sparse demes belonging to a meta-population (Rocha 1997).

Table 7 Fixation indices (f) estimated in original and rejuvenated seeds of *P*. *lunatus* L

Locus	Original seeds f <sup>a</sup>	Regenerated seeds f
$Adh-2$	1.000	1.000
Cap	1.000	0.768
$Ddh-2$	0.000	0.286
End	0.610	0.955
$Gpi-I$	1.000	0.657
Idh	0.614	0.000
$Mdh-2$	1.000	0.821
$Mdh-3$	1.000	1.000
$Pgdh-2$	0.293	0.063
$Pgm-1$	0.000	0.841
Mean	0.724	0.710
	$t^{\rm b}$	P <sup>c</sup>
	0.096	0.925

<sup>a</sup> Fixation index

<sup>b</sup> Student *t*-test for paired observations ( $df = 9$ )

<sup>e</sup> Probability value

#### **Conclusions**

The use of allozymes complements the data obtained from seed-protein electrophoresis (Maquet et al. 1990; Gutiérrez et al. 1995; Maquet 1995) and preliminary allozyme analysis (Maquet et al. 1993), providing a clearer picture of the organisation of genetic diversity in *P*. *lunatus*.

The total genetic variation maintained within *P*. *lunatus* can be partitioned in a hierarchical manner on the basis of its distribution among regions, botanical forms, accessions, and individuals within accessions. The total gene diversity of the Lima bean at the species level  $(H_{es} = 0.26)$  is similar to that of *P. acutifolius* (Schinkel and Gepts 1989) or *P*. *coccineus* (Escalante et al. 1994) but higher than that of *P*. *vulgaris* wild forms ( $H_{es} = 0.13$ ) as indicated by Koenig and Gepts (1989). According to our results, the genetic diversity of the Lima bean is evenly distributed among the Mesoamerican and the Andean gene pools. This confirms the results obtained by Nienhuis et al. (1995) using RAPD markers. Our study also shows that wild and cultivated forms of the Lima bean are characterised by similar levels of gene diversity and are not differentiated within each gene pool. In contrast, using seed storage-protein markers, Gutiérrez et al. (1995) and Maquet (1995) found a less variable genetic diversity in the Mesoamerican landraces compared to their wild ancestors. They attributed this difference to a founder effect. By using RAPD markers, Fofana et al. (1997) showed that the wild form and landraces of each gene pool were genetically differentiated. A higher genetic diversity was observed among Lima bean landraces than among the wild accessions. Such discrepancies between studies could result either from the different nature of the markers or from the samples used.

Within the base collection of *P*. *lunatus*, gene diversity is mainly distributed among accessions ( $G_{ST}$  = 0.755). Gene flow, through pollen and seed dispersal, should however favour the homogenisation of population structure, acting against the effect of drift and diversifying selection. Maquet et al. (1996) reported a high differentiation ( $G_{ST} = 0.803$ ) among wild Lima bean populations. With such a value, the average number of successful immigrants per generation, Nm (Slatkin and Barton 1989), is estimated to be 0.056, which corresponds to one migrant per 18 generations. This low rate of gene flow is not sufficient to introduce genetic variation and to counterbalance genetic drift. Most populations are genetically depauperate. In addition, in a population dynamic study conducted in the Central Valley of Costa Rica, Degreef et al. (1997) pointed to a common demographic instability of the wild populations of Lima bean at least in human disturbed habitats. This instability, expressed by periodic high reductions in population size (population bottlenecks) or by the establishment of new populations from one or a few individuals (founding events), increases the effects of genetic drift and inbreeding.

Lima bean germplasm is characterised by a high inbreeding coefficient  $(f = 0.891)$  which influences fitness through inbreeding depression. However, populations with a long history of inbreeding should be less vulnerable to inbreeding depression than typically outbreeding populations (Charlesworth and Charlesworth 1979). This is the case with the Lima bean, considered as a predominantly selfing species (Allard and Workman 1963; Free 1993; Maquet et al. 1996). In spite of inbreeding and a low gene flow rate, the intrapopulation gene diversity estimated either from the base collection  $(H_s = 0.032)$  or from the wild populations of Costa Rica  $(H_s = 0.058)$  in the study of Maquet et al. (1996) is significantly different from zero. The intrapopulation gene diversity of the Lima bean is higher than those of other selfing species such as *P*. *acutifolius*  $(H<sub>S</sub> = 0.0004)$  (Schinkel and Gepts 1989) and *P. vul*garis ( $H_s = 0.006$ ) (Koenig and Gepts 1989).

Estimating levels of genetic variation at the species and accession levels is helpful for managers. A conservation programme of *P*. *lunatus* must include wild and cultivated forms from both gene pools. As the genetic diversity is distributed mainly among accessions, more populations or accessions should be preserved to ensure the retention of allelic and genotypic diversity for both gene pools and botanical forms. In addition, the low gene-flow rate and the predominantly selfing mating system result in populations being spatially structured, as shown by Zoro Bi et al. (1997 a). Such a structural organisation implies that a population must be sampled all over its area during the collecting process.

Where habitats are in immediate danger of destruction (Maquet and Baudoin 1997), the collection of germplasm and its maintenance ex situ become necessary in addition to in situ conservation. Developing procedures for reducing the size of a collection to a manageable and accessible level (a core collection) is becoming one of the more important issues in the management and utilisation of plant germplasm collections (Brown 1989). A fraction ranging from 20 to 30% of the total plant collection was estimated to be the best or nearly the best (Yonezawa et al. 1995). A strategy for maintaining more entries, with fewer individuals per entry, would be efficient for species with higher selfing rates. When the collection is subdivided into a number of groups of accessions, like those obtained in our study, Yonezawa et al. (1995) suggest that one should sample each group in proportion to the range of its genetic diversity.

Lima bean landraces could also be maintained in situ as part of a farmer conservation system (Altieri and Merrick 1987; Brush 1991). In fact, *P*. *lunatus* is progressively being replaced in some regions by other food legumes. As it is a long process to re-introduce a crop plant, in a study conducted in Cuba, Esquivel and Hammer (1988) proposed to maintain Lima bean landraces as part of the traditional horticultural system.

Genetic diversity and its organisation are not only the result of natural evolution and plant breeding but also of domestication. The results from allozymes confirm that the Lima bean was domesticated at least twice. A first domestication centre may be localised somewhere in Ecuador and the northern region of Peru where the Andean landraces originate. It is more difficult to define the second centre of domestication since landraces expressing a Mesoamerican type are distributed from Mexico to Argentina. Geographic isolation and the predominant selfing mating system are probably two major causes of the genetic differentiation between the two gene pools. No founder effect was observed with allozyme data as the genetic diversities between wild and cultivated forms are not significantly different.

No specific alleles distinguishing the Sieva and Potato cultivars within the Mesoamerican gene pool were identified using our enzyme systems. This result is in accordance with those found by Gutiér rez et al. (1995) and Maquet (1995) based on seed-storage proteins, and with those observed by Lioi and Lotti (1996) using allozymes. Fofana et al. (1997), using RAPD markers on 17 Mesoamerican landraces, noted a low but significant differentiation between the Sieva and Potato cultigroups. No clear divergence occurred between the two Mesoamerican cultigroups. Selection in these landraces, in particular for their seed type, may have taken place in genetically related ancestral materials which have diverged more recently.

Maximising the numbers of distinct alleles (allelic richness) is the most appropriate theoretical objective in constructing core collections (Brown and Schoen 1994). From our analysis, it appears that Lima bean has a low average number of alleles per locus both at the species level  $(n_a = 2.58)$  and the within-population level  $(n_a = 1.14, \text{ Maquet et al. } 1996)$ . Other types of genetic markers like microsatellites (Schaal et al. 1991) could be helpful to better evaluate the allelic richness and the genetic structure of Lima bean.

Acknowledgements This study was financed by the Belgian Administration for Co-operation to Development (A.G.C.D., Bruxelles, Belgique) and supervised by the International Plant Genetic Resources Institute (I.P.G.R.I., Roma, Italy). We thank the Centro Internacional de Agricultura Tropical (C.I.A.T., Cali Colombia) for providing seeds of Lima bean accessions and Dr. X. Vekemans (Laboratoire de Génétique et d'Écologie Végétales, Université Libre de Bruxelles, Belgium) for his useful comments.

## References

- Allard RW, Workman PL (1963) Population studies in predominantly self-pollinated species. IV. Seasonal fluctuations in estimated values of genetic parameters in lima bean populations. Evolution 17 :470*—*480
- Altieri MA, Merrick LC (1987) In situ conservation of crop genetic resources through maintenance of traditional farming systems. Econ Bot 41:86*—*96
- Brown AHD (1989) The case for core collections. In:Brown AHD, Frankel OH, Marshall DR, Williams JT (eds) The use of plant genetic resources. Cambridge University Press, Cambridge, UK, pp 136*—*156
- Brown AHD, Schoen DJ (1994) Optimal sampling strategies for core collections of plant genetic resources. In:Loeschcke V, Tomiuk J, Jain SK (eds) Conservation genetics. Birkhaüser Verlag, Basel, Switzerland, pp 357*—*370
- Brush SB (1991) A farmer-based approach to conserving crop germplasm. Econ Bot 45 :153*—*165
- Charlesworth D, Charlesworth B (1979) The evolutionary genetics of sexual systems in flowering plants. Proc R Soc Lond B 205 : 513*—*530
- Crawford DJ (1990) Enzyme electrophoresis and plant systematics. In: Soltis DE, Soltis PS (eds) Isozymes in plant biology. Chapman and Hall, London, England, pp 146*—*164
- Dagnelie P (1978) Théorie et méthode statistiques. II. Les méthodes de l'inférence statistique. Les Presses Agronomiques de Gembloux A.S.B.L., Gembloux, Belgique
- Debouck DG (1987 a) *Phaseolus* germplasm collection in Central and Southern Mexico. Trip Report, Centro Internacional de Agricultura Tropical (C.I.A.T.), Cali, Colombia
- Debouck DG (1987 b) *Phaseolus* germplasm collection in Central Costa Rica. January 4*—*January 16, 1987. Trip Report, AGPG/IBPGR: 87/52, International Board for Plant Genetic Resources (I.B.P.G.R.), Roma, Italia
- Debouck DG (1987 c) Recolección de germoplasma de *Phaseolus* en el centro y centro-sur del Perú; Abril 27-Mayo 27, 1987. Trip report, AGPG/IBPGR: 87/112, International Board for Plant Genetic Resources (I.B.P.G.R.), Roma, Italia
- Debouck DG (1988 a) *Phaseolus* germplasm collection in central and eastern Guatemala. November 29*—*December 22, 1987. Trip Report, Centro Internacional de Agricultura Tropical (C.I.A.T.), Cali, Colombia
- Debouck DG (1988 b) Recolección de germoplasma de *Phaseolus* en Bolivia. Abril 23*—*Mayo 14, 1988. Trip Report, Centro Internacional de Agricultura Tropical (C.I.A.T.), Cali, Colombia
- Debouck DG (1989) Recolección de germoplasma de *Phaseolus* en el norte del Peru´ . Julio 19*—*Agosto 11, 1989. Trip Report, Centro Internacional de Agricultura Tropical (C.I.A.T.), Cali, Colombia
- Debouck DG (1990 a) Collecting *Phaseolus* germplasm in Colombia. August*—*September 1990. Final report for USDA/C.I.A.T./ ICA supported mission, IBPGR, Roma, Italia
- Debouck DG (1990 b) Collecting *Phaseolus* germplasm in Ecuador. Trip Report, Centro Internacional de Agricultura Tropical (C.I.A.T.), Cali, Colombia
- Debouck DG (1995) Germplasm exploration for the genera *Manihot* and *Phaseolus*in western and central Guatemala. Final report for C.I.A.T./USAC supported mission, International Plant Genetic Resources Institute (I.P.G.R.I.), Rome, Italy
- Degreef J, Baudoin J-P, Rocha OJ (1997) Case studies on breeding systems and its consequences for germplasm conservation. 2. Demography of wild Lima bean populations in the Central Valley of Costa Rica. Genet Resour Crop Evol 44:1*—*10
- Delvaux M (1994) Contribution à l'étude de la variabilité génétique du haricot de lima (*Phaseolus lunatus* L.) par le développement de méthodes électrophorétiques mono et bidimensionnelles sur gels de polyacrylamide. Travail de fin d'études, Faculté des Sciences Agronomiques, U.E.R. Chimie organique et biologique, Gembloux, Belgique
- Doebley J (1990) Isozymic evidence and the evolution of crop plants. In: Soltis DE, Soltis PS (eds) Isozymes in plant biology. Chapman and Hall, London, England, pp 165*—*191
- Escalante AM, Coello G, Eguiarte LE, Piñero D (1994) Genetic structure and mating systems in wild and cultivated populations of *Phaseolus coccineus* and *P*. *vulgaris* (Fabaceae). Am J Bot 81:1096*—*1103
- Esquivel H, Hammer K (1988) The "conuco"—an important refuge of Cuban plant genetic resources. Kulturpflanze 36:451*—*463
- Felsenstein J (1993) PHYLIP (Phylogeny Inference Package), Distributed by the author, Department of Genetics, University of Washington, Seattle, USA
- Fofana B, Vekemans X, du Jardin P, Baudoin J-P (1997) Genetic diversity in Lima bean (*Phaseolus lunatus* L.) as revealed by RAPD markers. Euphytica:157*—*165
- Free JB (1993) Insect pollination of crops, Academic Press, London, England
- Gepts P (1995) Genetic markers and core collections. In: Hodgkin T, Brown AHD, van Hintum TJL, Morales EAV (eds) Core collections of plant genetic resources. John Wiley and Sons, Chichester, UK, pp 127*—*146
- Gutiérrez S.A, Gepts P, Debouck DG (1995) Evidence for two gene pools of the lima bean, *Phaseolus lunatus* L., in the Americas. Genet Resour Crop Evol 42 :15*—*28
- Hamann A, Zink D, Nagl W (1995) Microsatellite fingerprinting in the genus *Phaseolus*. Genome  $38:507-515$
- Hamrick JL, Godt MJW, Murawski DA, Loveless MD (1991) Correlations between species traits and allozyme diversity: implications for conservation biology. In: Falk DA, Holsinger KE (eds) Genetics and conservation of rare plants. Oxford University Press, New York, USA, pp 75*—*86
- Hardy O, Dubois S, Zoro Bi I, Baudoin J-P (1997) Gene dispersal and its consequences on the genetic structure of wild lima bean (*Phaseolus lunatus* L.) populations in Costa Rica. Plant Genet Resour Newslett (FAO/IPGRI) 109 :1*—*6
- Hartl DL, Clark AG (1989) Principles of population genetics. Sinauer Associates, Sunderland, Massachusetts, USA
- Hussain A, Bushuk W, Ramı´rez H, Roca W (1988) A practical guide for electrophoretic analysis of isoenzymes and proteins in cassava, field beans and forage legumes. Centro Internacional de Agricultura Tropical (C.I.A.T.), Cali, Colombia
- Jaaska V (1996) Isoenzyme diversity and phylogenetic affinities among the *Phaseolus* beans (*Fabaceae*). Pl Syst Evol 200: 233*—*252
- Koenig R, Gepts P (1989) Allozyme diversity in wild *Phaseolus vulgaris*: further evidence for two major centers of genetic diversity. Theor Appl Genet 78:809*—*817
- Laemmli IK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227 :680*—*685
- Lioi L, Lotti C (1996) Allozyme variability in cultivated Lima bean (*Phaseolus lunatus* L.). Bean Improv Coop Annu Rep 39: 249*—*250
- Maquet A (1991) Lima bean (*Phaseolus lunatus* L.) catalogue. Centro Internacional de Agricultura Tropical (C.I.A.T.), Cali, Colombie
- Maquet A (1995) Étude de la diversité génétique de la légumineuse alimentaire *Phaseolus lunatus* L. par l'analyse de caractères morphophysiologiques et de marqueurs protéiques. Thèse de Doctorat, Faculté Universitaire des Sciences Agronomiques, Département "Agronomie, Economie et Développement", U.E.R. Phytotechnie des Régions Intertropicales, Gembloux, Belgique
- Maquet A, Baudoin J-P (1997) Apercu de la distribution néotropicale de *Phaseolus lunatus*. Belg J Bot: 93*—*116
- Maquet A, Gutierrez A, Debouck DG (1990) Further biochemical evidence for the existence of two gene pools in lima beans. Bean Improv Coop Annu Rep 33:128*—*129
- Maquet A, Wathelet B, Baudoin J-P (1993) A study on isoenzyme polymorphism in lima bean (*Phaseolus lunatus* L.). In: Roca WM, Maver JE, Pastor-Corrales MA, Tohme J (eds) *Phaseolus* Beans Advanced Biotechnology Research Network, BARN. Centro Internacional de Agricultura Tropical (C.I.A.T.), Cali, Colombia, pp 87*—*96
- Maquet A, Wathelet B, Baudoin J-P (1994) Étude du réservoir génétique de la légumineuse alimentaire *Phaseolus lunatus* L. par l'analyse électrophorétique d'isozymes. Bull Rech Agron Gembloux 29 :369*—*381
- Maquet A, Zoro Bi I, Rocha OJ, Baudoin J-P (1996) Case studies on breeding systems and its consequences for germplasm conservation. 1. Isoenzyme diversity in wild Lima bean populations in central Costa Rica. Genet Resour Crop Evol 43: 309*—*318
- May B (1992) Starch-gel electrophoresis of allozymes. In: Hoelzel AR (ed) Molecular genetic analysis of populations *—* a practical approach. Oxford University Press, Oxford, England pp 1*—*27
- Murphy RW, Sites Jr. JW, Buth DG, Haufler CH (1990) Proteins I. Isozyme electrophoresis. In: Hillis DM, Moritz C (eds) Molecular systematics. Sinauer Associates, Inc., Sunderland, Massachusetts, USA, pp 45*—*126
- National Research Council (1989) Lost crops of the Incas: littleknown plants of the Andes with promise for worldwide cultivation. National Academy Press, Washington, D.C., USA
- Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70: 3321*—*3323
- Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, USA
- Nienhuis J, Tivang J, Skroch P, dos Santos B (1995) Genetic relationships among cultivars and landraces of lima bean (*Phaseolus lunatus* L.) as measured by RAPD markers. J Am Soc Hort Sci 120:300*—*306
- Rocha OJ, Macaya G, Baudoin J-P (1977) Three-years monitoring wild populations of *Phaseolus lunatus* L. in the central valley of Costa Rica: causes of local extinctions and re-colonizations. Plant Genet Resour Newslett (in press)
- Schaal BA, Leverich WJ, Rogstad SH (1991) A comparison of methods for assessing genetic variation in plant conservation biology. In: Falk DA, Holsinger KE (eds) Genetics and conservation of rare plants. Oxford University Press, New York, USA, pp 123*—*134
- Schinkel C, Gepts P (1989) Allozyme variability in the tepary bean, *Phaseolus acutifolius* A. Gray. Plant Breed 102: 182*—*195
- Singh SP, Nodari R, Gepts P (1991) Genetic diversity in cultivated common bean. I. Allozymes. Crop Sci 31 :19*—*23
- Slatkin M, Barton NH (1989) A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43:1349*—*1368
- Vekemans X, Lefebvre C (1997) On the evolution of heavy metaltolerant populations in *Armeria maritima*: evidence from allozyme variation and reproductive barriers. J Evol Biol 10:175*—*191
- Weir BS (1990) Intraspecific differentiation. In: Hillis DM, Moritz C (eds) Molecular systematics. Sinauer Associates, Inc., Sunderland, Massachusetts, USA, pp 373*—*410
- West NB, Garber ED (1967) Genetic studies of variant enzymes. I. An electrophoretic survey of esterases and leucine aminopeptidases in the genus *Phaseolus*. Can J Genet Cytol 9 :640*—*645
- Wilkinson L (1989) SYSTAT: The system for statistics. SYSTAT, Inc., Evanston, Illinois, USA
- Yonezawa K, Nomura T, Morishima H (1995) Sampling strategies for use in stratified germplasm collections. In: Hodgkin T, Brown AHD, van Hintum TJL, Morales EAV (eds) Core collections of plant genetic resources. John Wiley and Sons, Chichester, UK, pp 35*—*53
- Zoro Bi I, Maquet A, Baudoin J-P (1997 a) Spatial patterns of allozyme variants within three wild populations of *Phaseolus lunatus* L. from the central valley of Costa Rica. Belg J Bot 129:149*—*155
- Zoro Bi I, Maquet A, Wathelet B, Baudoin J-P (1997 b) Genetic control of alcohol dehydrogenase, malate dehydrogenase, and phosphoglucomutase isozymes in Lima bean (*Phaseolus lunatus* L.). Plant Breed 116:181*—*185

.