



Gene pools in wild Lima bean (*Phaseolus lunatus* L.) and their origin

Martha L. Serrano-Serrano¹, Jorge Hernández Torres¹, Genis Castillo-Villamizar¹, Daniel G. Debouck² & Maria I. Chacón S.³

¹ Centro de Innovación en Biotecnología Industrial y Biología Molecular, Escuela de Biología, Universidad Industrial de Santander, Carrera 27 Calle 9, Bucaramanga, Colombia; ² Unidad de Recursos Genéticos, CIAT, Km 17 Via Cali-Palmira, Colombia; ³ Facultad de Agronomía, Universidad Nacional de Colombia, Ciudad Universitaria, Bogotá, D.C., Colombia.

Introduction

In the Neotropics, the current distribution patterns of flora have been shaped by relatively recent events such as the closure of the Isthmus of Panama (7 million years ago, Mya), the uplift of the Andes (during the last 10 Mya) and past climatic fluctuations during the Quaternary. The genus *Phaseolus* of Neotropical origin is distributed in both North and South America and offers an opportunity to assess the influence of past historical events on the current distribution patterns of the Neotropical flora. *Phaseolus lunatus*, the Lima bean, is widely distributed in lowlands and middle elevations in seasonally dry forests in both Mesoamerica and South America, and is one of the five species domesticated within this genus. Within wild Lima beans, two gene pools, the *Mesoamerican* and the *Andean*, have been recognized on the basis of geographical distribution, morphological, biochemical and molecular traits (1,2). The *Andean* gene pool is restricted to the western slope of the Andes of Ecuador and northern Peru (3). The *Mesoamerican* gene pool is widespread and occurs in Mexico, Central America, the Caribbean and South America (Colombia and east of the Andes). It has been suggested that the species has an Andean origin (4) and very little is known on how it reached its widespread current distribution (Figure 1). The aim of this research was to study the genetic structure of wild *P. lunatus* in the Americas and to assess the processes that could have influenced the current geographical distribution of genetic diversity by a combined phylogeographical and phylogenetic approach. For this purpose, nuclear (ITS/5.8S) and chloroplast (intergenic spacers *atpB-rbcL* and *trnL-trnF*) DNA markers were analyzed in a representative collection of wild Lima bean accessions and six allied species from the sections *Paniculati*, *Coriacei* and *Falcati*. The results suggest that the genetic diversity of the species may be organized into three gene pools which would have diverged from each other very early after the formation of the species in the Andes of South America.

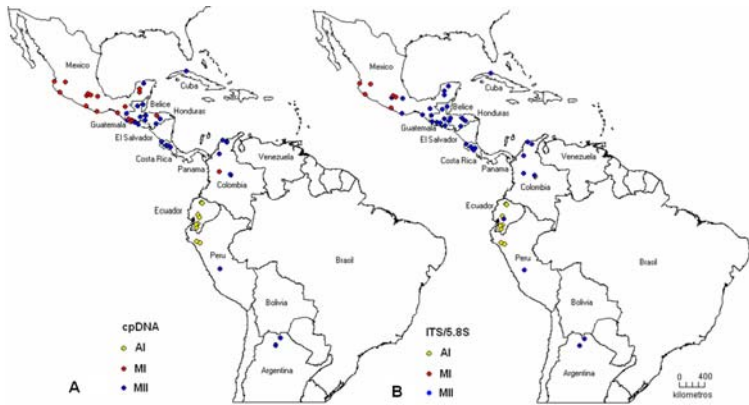


Figure 1. Geographic distribution of gene pools within wild Lima beans in the Americas. A. Distribution of cpDNA gene pools. B. Distribution of ITS/5.8S gene pools.

Plant Material

All plant material used in this study was obtained from the International Center for Tropical Agriculture in Cali, Colombia. A total of 59 wild Lima bean accessions were selected on the basis of their geographic distribution and complete passport data. A total of 12 accessions that contain Mesoamerican (*P. leptostachyus*, *P. marshallii* and *P. novoleonensis*) and Andean (*P. augusti*, *P. pachyrhizoides* and *P. bolivianus*) allied species, were also included.

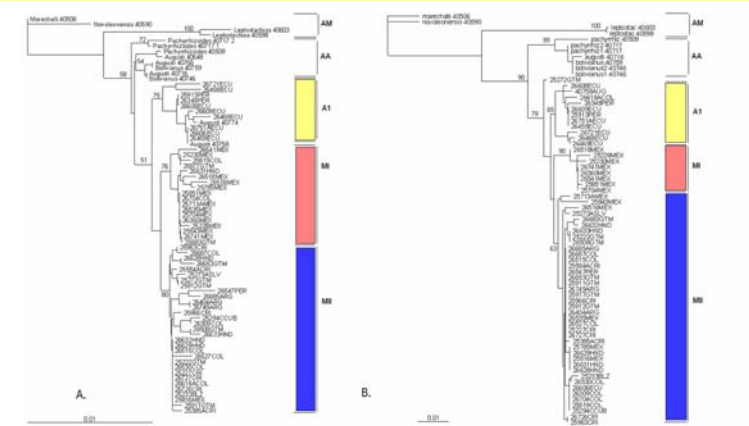


Figure 2. NJ topology based on cpDNA (A) and ITS/5.8S data (B) for wild *P. lunatus* and allied Mesoamerican (AM) and Andean (AA) species. MEX=Mexico, GTM= Guatemala, BLZ= Belize, HND= Honduras, SLV= El Salvador, CRI=Costa Rica, COL= Colombia, CUB= Cuba, ECU= Ecuador, PER= Peru, BOL= Bolivia, ARG= Argentina. AI, MI and MII indicate gene pools within wild Lima bean. Numbers on branches indicate bootstrap values (%).

Organization of genetic structure of wild Lima beans into three gene pools

The genetic structure of wild Lima beans was investigated by means of sequencing of two cpDNA intergenic spacers (*atpB-rbcL* and *trnL-trnF*) and the ITS/5.8S region. Two genetic distance matrixes were obtained under the JC model and two dendrograms were built by applying the NJ algorithm (Figure 2). Genetic divergence among NJ clusters was estimated by means of the statistics G_{ST} , N_{ST} , Nm and by means of AMOVA analyses (Table 1). The present results support the organization of the genetic diversity within wild Lima beans into three gene pools: AI, MI and MII. These three gene pools occur in mostly non-overlapping geographic areas (Figure 1) and genetic differentiation among each other is very high (Table 1). Gene pool AI is restricted to the Andes of Ecuador and northern Peru, gene pool MI is mostly found in Mexico in the area to the north and northwest of the Isthmus of Tehuantepec, and gene pool MII is widespread and found in Mexico (to the east and southeast of the Isthmus of Tehuantepec), Central America, Colombia, southern Peru and Argentina (Figure 1). The fact that the three gene pools in wild Lima bean occur in mostly non-overlapping areas suggests they may be adapted to different ecological conditions. Also, the high genetic differentiation among gene pools indicate that they have accumulated unique genetic variation and therefore they should be taken into account in current conservation programs for this species.

The genetic structure of wild Lima beans into three gene pools sharply contrasts with previous studies where mainly two gene pools have been proposed (the *Andean* and the *Mesoamerican* ones) (1,4). Part of these discrepancies may be in part explained by the analysis of a different set of samples and by the molecular techniques used. It is worth to highlight that in the present study a larger number of samples from Mexico (where gene pool MI is mainly distributed) have been analyzed, which contrasts with previous studies where the sampling in Mexico has been very limited. Apart from this, in the present study sequencing of DNA was performed on all samples while in previous studies indirect methods to study the variation at the nucleotide level were used.

Group configuration		cpDNA	ITS/5.8S
AI vs. (MI+MII)			
Source of variation %	Within groups	43.22	48.43
	Between groups	56.78	51.57
Fixation indexes	Φ_{ST}	0.568**	0.516**
	G_{ST}	0.141	0.090
	N_{ST}	0.643	0.549
Gene flow	Nm	0.14	0.21
AI vs. MI vs. MII			
Source of variation %	Within groups	35.37	29.70
	Between groups	64.63	70.30
Fixation indexes	Φ_{ST}	0.646**	0.73**
	G_{ST}	0.291	0.127
	N_{ST}	0.734	0.745
Gene flow	Nm	0.09	0.09
MI vs. MII			
Source of variation %	Within groups	40.39	23.72
	Between groups	59.61	76.28
Fixation indexes	Φ_{ST}	0.596**	0.763**
	G_{ST}	0.199	0.062
	N_{ST}	0.654	0.701
Gene flow	Nm	0.13	0.11

Table 1. AMOVA results and coefficients of population differentiation. AMOVA was applied at two hierarchical levels (within and among groups). Three group configurations were assessed. Population divergence indexes: Φ_{ST} = analogue of F_{ST} estimated via AMOVA, G_{ST} = differentiation among groups based on haplotype diversity, N_{ST} = differentiation among groups of based on nucleotide diversity. Nm was calculated from N_{ST} . **p<0.001

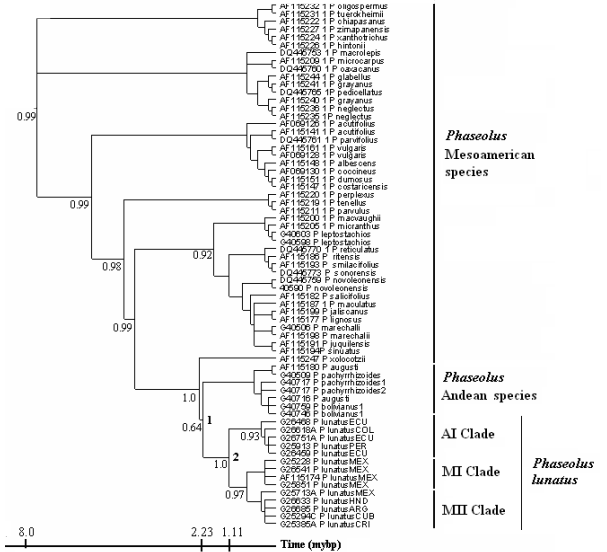


Figure 3. ITS/5.8S major consensus tree obtained from a Bayesian approach for the Millettoid clade. Ages for the MRCA in nodes 1 and 2 were estimated based on a PL "Penalized Likelihood" analysis. mybp (million years before present).

Andean origin of wild Lima bean and recent divergence times

A phylogeny was built for 75 ITS/5.8S sequences retrieved from GenBank which represent several members of the Millettoid crown (5) and 22 accessions of wild Lima beans and allied species sequenced in this study. Divergence times were estimated applying a penalized likelihood (PL) analysis (6,7) and using as age constraints the ages of nodes 58, 67 and 68 reported by (5). Figure 3 shows estimated ages for the most common recent ancestor (MRCA) of *P. lunatus* and Andean sister taxa (node 1), and the MRCA of gene pools within *P. lunatus* (node 2). The present results thus confirm an Andean origin for wild Lima beans as it has been reported previously (2,4). Currently, the Andean allied species are restricted to the central and southern Andes (in Peru, Bolivia and Argentina), while wild Lima beans are more widespread in the Andes and in Mesoamerica. The age estimates place the separation of wild Lima beans from its allied species at or before 2.23 million years. This age correlates with the average age for the eight species clades within *Phaseolus* reported by (8). Since the calculation of divergence times was constrained on the age estimated for fossil records (as reported by (5)), these are minimum divergence ages and not absolute ages. Therefore, the time of origin of wild Lima beans could be older than 2.23 million years and could have occurred in the context of major Andean orogeny, that presumably occurred during the last 10 million years, and the formation of the Panama Isthmus that occurred 7 million years ago. The high genetic differentiation among the three gene pools strongly indicates an early separation of these three gene pools in the Andes. The estimated age for the MRCA of the gene pool AI and gene pools MI+MII is at least 1.1 Ma which may also be in the context of the final uplift of the Andes over the last 10 million years. After the separation of the three gene pools, gene pool AI would have become restricted to the western side of the Andes of Ecuador and northern Peru, and gene pools MI and MII or an early ancestor of them would have migrated outside the ancestral area into Central America and Mesoamerica. The migration would have taken place throughout the Isthmus of Panama by means of the natural dispersal mechanism of the seed by pod dehiscence. Currently, there are some populations of wild Lima bean in the Isthmus of Panama on the Pacific side (9), as witness of this past migration.

References

- MAQUET, A., VEKEMANS, X., BAUDOIN, J.P. (1999) Phylogenetic study on wild allies of Lima bean, *Phaseolus lunatus* (Fabaceae), and implications on its origin. *Plant Systematics and Evolution* 218: 43-54.
- CAICEDO, A.L., GAITAN, E., DUQUE, M.C., TORO, O., DEBOUCK, D.G., TOHME, J. (1999) AFLP fingerprinting of *Phaseolus lunatus* L. and related wild species from South America. *Crop Science* 39: 1497-1507.
- DEBOUCK, D.G. (1996) Colombian Common and Lima beans: views on their origin and evolutionary significance. *Revista Corpoica* 1: 7-15.
- FOFANA, B., BAUDOIN, J.P., VEKEMANS, X., DEBOUCK, D.G., DU JARDIN, P. (1999). Molecular evidence for an Andean origin and a secondary gene pool for the Lima bean (*Phaseolus lunatus* L.) using chloroplast DNA. *Theoretical and Applied Genetics* 98: 202-212.
- LAVIN, M., HERENDEEM, P.S., WOJCIECHOWSKI, M.F. (2005). Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Systematic Biology* 54: 575-594.
- SANDERSON, M.J. (2002). Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* 19: 101-109.
- SANDERSON, M.J. (2004) 8s 1.70 Analysis of rates ("8s") of evolution. Available at <http://ginger.ucdavis.edu/8s/>.
- DELGADO-SALINAS, A., BIBLER, R., LAVIN, M. (2006) Phylogeny of the genus *Phaseolus* (Leguminosae): a recent diversification in an ancient landscape. *Systematic Botany* 31:779-791.
- LACKEY, J.A. and D'ARCY, W.G. (1980). *Phaseolus*. Family 83. Leguminosae. Flora of Panama. *Annals of the Missouri Botanical Garden* 67: 746-751.

Acknowledgments

We are grateful to the Universidad Industrial de Santander, Colombia, and the Banco de la República, Colombia, for providing funding to carry out this research. We are also grateful to the Genetic Resources Unit of CIAT for providing plant material.