

**FILOGEOGRAFÍA DEL FRÍJOL LIMA (*Phaseolus lunatus* L.) SILVESTRE Y  
SU RELACIÓN CON ESPECIES AFINES MESOAMERICANAS Y ANDINAS**

**Martha Liliana Serrano Serrano**

**Universidad Industrial de Santander  
Facultad de Ciencias  
Escuela de Biología  
Bucaramanga  
2008**

**FILOGEOGRAFÍA DEL FRÍJOL LIMA (*Phaseolus lunatus* L.) SILVESTRE Y  
SU RELACIÓN CON ESPECIES AFINES MESOAMERICANAS Y ANDINAS**

**Martha Liliana Serrano Serrano**

**Trabajo de Grado para obtener el título de  
Bióloga**

**Director**

**Dra. Maria Isabel Chacón Sánchez**

**Bióloga PhD.**

**Facultad de Agronomía, Universidad Nacional de Colombia**

**Codirector**

**Dr. Jorge Hernández Torres**

**Biólogo PhD.**

**Laboratorio de Biología Molecular y Biotecnología, Universidad Industrial de  
Santander**

**Universidad Industrial de Santander**

**Facultad de Ciencias**

**Escuela de Biología**

**Bucaramanga**

**2008**



Esta investigación fue financiada por la Dirección de Investigaciones y Extensión de la Facultad de Ciencias, Universidad Industrial de Santander, así como, el Banco de la República de Colombia. La mayor parte del trabajo se efectuó en el Laboratorio de Biología Molecular (CINBIN) de la Universidad Industrial de Santander, Sede Guatiguará, Piedecuesta-Santander. El material vegetal procedió del Centro Internacional de Agricultura Tropical (CIAT).

## **AGRADECIMIENTOS**

A la doctora Maria Isabel Chacón Sánchez por la oportunidad de trabajar a su lado, el apoyo incondicional y enseñanzas en diferentes experiencias durante la realización de esta investigación.

Al biólogo Genis Andrés Castillo Villamizar por el soporte académico y de laboratorio durante este proyecto y los comentarios críticos durante la elaboración del documento.

Al profesor Daniel Rafael Miranda-Esquivel por su orientación durante mi formación académica.

Al profesor Jorge Hernández Torres por su colaboración y apoyo durante el proyecto.

A la Unidad de Recursos Genéticos del CIAT por proveer el material vegetal y asesoría en la selección de entradas y sus datos correspondientes.

A los doctores Daniel Debouck y Matthew Blair, del CIAT, por su apoyo académico y sugerencias para la investigación.

A mi mamá por ser el pilar de mi formación espiritual e intelectual y creer totalmente en mis capacidades.

A mi hermano por el día a día en casa, en la universidad y por darle una mirada crítica pero amorosa a mi vida.

A mi abuela, mi papá y mi tío por su apoyo incondicional en cada camino trazado.

Por último, y con toda el alma, al negro por ser “mi amor, mi cómplice y todo”.

## CONTENIDO

	Pág.
INTRODUCCION	11
1. OBJETIVOS	
1.1    Objetivo general	16
1.2    Objetivos específicos	16
2. MATERIALES Y MÉTODOS	
2.1    Estudio piloto: Marcadores del genoma cloroplástico y del genoma nuclear polimórficos en frijol lima silvestre	17
2.1.1    Material vegetal	17
2.1.2    Amplificaciones PCR en regiones de cpADN y nuclear	18
2.1.3    Purificación de productos PCR y secuenciamiento	19
2.1.4    Selección de regiones de cpADN e ITS/5.8S	19
2.2    Diversidad de haplotipos del cpADN e ITS/5.8S en frijol lima silvestre y su distribución geográfica	20
2.2.1    Haplotipos y distribución geográfica	20
2.2.2    Acervos genéticos	20
2.2.3    Diversidad genética y estructura poblacional	21
2.3    Filogeografía de frijol lima silvestre	22
2.3.1    Historia demográfica	22
2.3.2    Haplotipos y redes haplotípicas	23
2.3.3    Análisis filogenético y tiempos de divergencia	23
3. RESULTADOS	
3.1    Estudio Piloto	26
3.2    Diversidad en frijol lima silvestre	27
3.2.1    Haplotipos y distribución geográfica	27
3.2.2    Acervos genéticos	28
3.2.3    Diversidad Genética y Estructura Geográfica	30
3.3    Filogeografía	31
3.3.1    Historia demográfica	31
3.3.2    Redes haplotípicas y test de contingencia	32
3.3.3    Análisis filogenético y tiempos de divergencia	34
3.3.4    Origen del frijol lima	35
CONCLUSIONES	38
REFERENCIAS	40

**LISTA DE TABLAS**

	Pág.
<b>TABLA 1.</b> Datos del material vegetal	51
<b>TABLA 2.</b> Resultados del estudio piloto de variación genética en cpADN e ITS/5.8S	53
<b>TABLA 3.</b> Secuencias del ITS/5.8S usadas en el análisis de tiempos de divergencia	54
<b>TABLA 4.</b> Distribución geográfica de haplotipos del cpADN	55
<b>TABLA 5.</b> Distribución geográfica de haplotipos del ITS/5.8S	56
<b>TABLA 6.</b> Análisis de diversidad genética de los espaciadores intergénicos de cpADN y la región ITS/5.8S	57
<b>TABLA 7.</b> Diferenciación genética y flujo génico	58
<b>TABLA 8.</b> Análisis de AMOVA	59
<b>TABLA 9.</b> Test de contingencia	60
<b>TABLA 10.</b> Estimación de tiempos de divergencia.	61

**LISTA DE FIGURAS**

	Pág.
<b>FIGURA 1.</b> Ubicación de las regiones de cpADN exploradas	62
<b>FIGURA 2.</b> Topología NJ del cpADN	63
<b>FIGURA 3.</b> Topología NJ del ITS/5.8S	64
<b>FIGURA 4.</b> Distribución geográfica del fríjol lima en América	65
<b>FIGURA 5.</b> Red de haplotipos para cpADN	66
<b>FIGURA 6.</b> Red de haplotipos para ITS/5.8S	67
<b>FIGURA 7.</b> Topología del análisis bayesiano para el clado Milletioide.	68

## RESUMEN

TITULO\* : Filogeografía del frijol lima (*Phaseolus lunatus* L.) silvestre y su relación con especies afines mesoamericanas y andinas

Autor: Martha Liliana Serrano Serrano\*\*

Palabras clave: Filogeografía, *Phaseolus lunatus*, Flora Neotropical, ITS/5.8S, cpADN, Redes haplotípicas.

Esta investigación planteó establecer los patrones filogeográficos en *Phaseolus lunatus* (frijol lima) silvestre, reconociendo los procesos que posiblemente moldearon su distribución geográfica y diversidad genética actual. Para este propósito se identificaron marcadores moleculares del núcleo (ITS/5.8S) y del cloroplasto (espaciadores intergénicos *atpB-rbcL* y *trnL-trnF*), los cuales permitieron la estimación de la diversidad genética de una colección representativa de frijol lima silvestre de varias regiones del continente Americano, desde México hasta Argentina. Las regiones de cpADN e ITS/5.8S identificaron 23 y 28 haplotipos, respectivamente, los cuales presentaron una fuerte asociación geográfica. Con el uso de herramientas de la genética de poblaciones y la filogeografía se reconocieron tres acervos genéticos en frijol lima: el acervo andino AI, distribuido en los Andes de Ecuador-Norte del Perú; el acervo mesoamericano MI distribuido en Mesoamérica en especial en las áreas al occidente y noroccidente del Istmo de Tehuantepec; y el acervo Mesoamericano MII ampliamente distribuido en Mesoamérica, en especial al oriente y suroriente del Istmo de Tehuantepec, en Centroamérica, el Caribe y en Suramérica. Se evidenció una alta subdivisión poblacional ( $N_{ST}>0.5$ ,  $\Phi_{ST}>0.5$ ), así como un flujo génico restringido entre los tres acervos ( $N_m<1.0$ ). Los análisis demográficos sugieren un posible evento de expansión del rango geográfico en el acervo mesoamericano MII. La evidencia filogenética entre el frijol lima silvestre y sus especies aliadas sugiere un origen andino del frijol lima. Así mismo, se estimó que el tiempo de divergencia entre *P. lunatus* y sus especies hermanas andinas es de al menos 2.23 Ma, fecha en la cual el Istmo de Panamá ya se había formado completamente. El análisis filogeográfico apoya un posible evento de fragmentación del rango en los Andes entre el acervo ancestral andino y los acervos mesoamericanos derivados, posiblemente hace 1.11 Ma, correlacionando este resultado con el levantamiento de los Andes durante el Plioceno-Pleistoceno.

---

\* Proyecto de Grado

\*\* Facultad de Ciencias. Escuela de Biología. Director: Maria Isabel Chacón Sánchez PhD. Codirector: Jorge Hernández Torres PhD.

## ABSTRACT

TITLE\* : Phylogeographic analysis in wild lima bean (*Phaseolus lunatus* L.) and its relationship with Mesoamerican and Andean allied species.

Author: Martha Liliana Serrano Serrano\*\*

Key words: Phylogeography, *Phaseolus lunatus*, Neotropical flora, ITS/5.8S, cpDNA, Haplotype Network.

The aim of this research was to study the phylogeographical history of wild *Phaseolus lunatus* (Lima bean) in the Americas, and to assess the processes that could have influenced its current geographical distribution and genetic diversity. For this purpose, nuclear (ITS/5.8S) and chloroplast (intergenic spacers *atpB-rbcL* and *trnL-trnF*) DNA markers were identified and used to estimate the genetic diversity of a representative collection of wild Lima beans from different regions within the Americas. In these DNA regions, 23 cpDNA and 28 ITS/5.8S haplotypes were identified. These haplotypes showed to be geographically structured. On the basis of phylogeographical and population genetic analyses, three gene pools were recognized in wild Lima beans: the Andean pool AI, distributed in the Andes of Ecuador-Northern Peru; the Mesoamerican gene pool MI, distributed in Mesoamerica, specially to the west and north-west of the Isthmus of Tehuantepec; and lastly, the Mesoamerican gene pool MII, which is widespread in Mesoamerica, specially to the east and south-east of the Isthmus of Tehuantepec, in Central America, the Caribbean and in South America. High population differentiation ( $N_{ST}>0.5$ ,  $\Phi_{ST}>0.5$ ) and restricted gene flow ( $Nm<1.0$ ) was found between the three gene pools. The demographic analyses suggest a population expansion in Mesoamerican MII gene pool. The phylogenetic evidence between wild Lima beans and related species from the Andes suggest an Andean origin for wild Lima beans. Age estimates place the divergence of Lima bean from its sister taxa in at least 2.23 Ma, a time when the Isthmus of Panama had already closed. The phylogeographic analysis supports an event of range fragmentation between the Andean gene pool and the Mesoamerican ones. According to age estimates, the fragmentation event would have taken place at or before 1.11 Ma, which might correlate with past events such as the final uplift of the Andes during the Plio-Pleistocene.

---

\* Degree Project

\*\* Science Faculty. Biology Department. Advisors: Maria Isabel Chacón Sánchez PhD, Jorge Hernández Torres PhD.

## INTRODUCCION

Los Neotrópicos poseen gran parte de la biodiversidad del planeta, desde la era Mesozoica, donde su biota adquirió mayor complejidad. Los patrones de distribución de fauna y flora han sido estructurados a lo largo del tiempo, por eventos como la separación geográfica de África y América (90 Ma), el levantamiento de los Andes (15–2 Ma), el reciente cierre del Istmo de Panamá (7-4 Ma) y las fluctuaciones climáticas del Cuaternario (1.6 Ma) (Burnham & Graham 1999, Pennington *et al.* 2004). La formación del Istmo de Panamá permitió la conexión entre Norte y Sur América, y el subsiguiente desplazamiento de plantas y animales a través de los trópicos, en lo que se conoce como el Gran Intercambio Biótico Americano. La estimación del tiempo en el cual se dio este intercambio es actualmente discutida, mediante evidencia basada en geología, registros fósiles y patrones de migración de la biota (Coates 1997).

Las especies mesoamericanas y suramericanas parecen haber estado aisladas al menos hasta el Mioceno, donde los océanos Atlántico y Pacífico constituían una profunda barrera entre las masas terrestres que dieron más tarde origen a Norte América y Suramérica. Con la formación de pequeños archipiélagos, algunas especies capacitadas para “saltar” de isla en isla fueron desplazándose desde Norte América hacia Suramérica, iniciando el intercambio hace unos 8 Ma. De modo que hasta la formación completa del puente terrestre (al final del Plioceno) se inicia el intercambio de un gran número de especies en ambas direcciones (Webb 1997).

Los patrones de migración que ilustran el intercambio a través del Istmo, una vez este estuvo completo, han sido mejor explorados en especies con mecanismos de dispersión a corta distancia (Burnham & Graham 1999). Algunas migraciones desde Suramérica y radiaciones en México han sido documentadas para plantas de la familia *Sapindaceae* (Acevedo-Rodríguez 1993). Sin embargo, pocos estudios han involucrado especies distribuidas a lo largo del Istmo. El análisis filogeográfico de *S.*

*globulifera*, especie de la familia Clusiaceae, cuya distribución es tropical, evidenció el efecto del movimiento tectónico, el surgimiento de sistemas montañosos y la dispersión, como mecanismos formadores de estructura poblacional en los bosques húmedos neotropicales (Dick *et al.* 2003).

Reconstrucciones biogeográficas y filogeográficas en especies neotropicales ampliamente distribuidas, pueden aportar información sobre el efecto de eventos como la formación del Istmo de Panamá (Gentry, 1982), el Intercambio Biótico Americano y el levantamiento de los Andes (Pirie *et al.* 2006) en la distribución geográfica y diversidad de la biota Neotropical. *Phaseolus lunatus* L. representa una de estas especies, dada su distribución en los subcontinentes Mesoamericano y Suramericano.

El género *Phaseolus* posee una taxonomía compleja, Delgado-Salinas *et al.* (1993) presentan el género como un grupo monofilético, con aproximadamente 50 especies de distribución estrictamente neotropical, excepto algunos cultivares presentes en Estados Unidos, Italia y África. *P. lunatus* (frijol lima) es una de las cinco especies domesticadas de su género, perteneciente al clado denominado “Lunatus” que comprende las únicas especies del género que diversificaron en Suramérica. Este clado incluye especies endémicas de los andes (*P. augusti* Harms, *P. bolivianus* Piper y *P. pachyrrhizoides* Harms), islas Galápagos (*P. mollis* Hook) y Bermudas (*P. lignosus* Britton) (Delgado-Salinas 2000; Delgado-Salinas *et al.* 1999).

El frijol lima es una especie anual, auto-compatible con un sistema de apareamiento mixto, predominantemente auto-polinizador (90% autogamia), aunque facultativamente alógamo. Las características morfológicas de las estructuras reproductivas, así como, el tiempo de desarrollo floral facilitan la auto-polinización (Zoro *et al.* 2005). Su mecanismo de fertilización cruzada se encuentra mediado principalmente por la abeja de la miel (*Apis mellifera* L.) (Hardy *et al.* 1997). Un estudio con poblaciones de frijol lima silvestre de Costa Rica reveló que la dispersión

de polen y semillas ocurre en cortas distancias, no excediendo los 6 m (Zoro et al. 2003).

*P. lunatus* ha sido clasificado en dos variedades, var. *silvester* y var. *lunatus*, para designar al material silvestre y domesticado, respectivamente (Baudet 1977). Dentro de cada variedad es posible reconocer dos acervos genéticos, *Mesoamericano* y *Andino* (Gepts 1988). Estos grupos han sido definidos mediante el análisis de caracteres morfológicos, como el tamaño de la semilla (Debouck et al. 1987), características bioquímicas (Gutiérrez et al. 1995; Maquet et al. 1999) y a nivel molecular, por medio de Polimorfismos en Longitud de Fragmentos Amplificados (AFLP) y polimorfismos del genoma cloroplástico (cpADN) (Caicedo et al. 1999; Fofana et al. 1999). El acervo genético denominado *Mesoamericano*, de semilla pequeña, se encuentra distribuido desde Sinaloa, México, hasta Salta, Argentina y algunas islas del Caribe, principalmente presente en alturas por debajo de los 1.600 msnm. En Suramérica se encuentra desde Colombia, distribuido hacia el sur por la pendiente este de los Andes, hasta Argentina. El segundo acervo, denominado *Andino*, con semillas de tamaño mayor, se encuentra en Ecuador y norte de Perú en los valles intra-andinos del lado occidental de los Andes. Este acervo se distribuye cercano a los 1.800 msnm (Baudoin et al. 2004; Debouck 1996).

América Central ha sido considerada como la zona de mayor especiación dentro del género, por ello Fofana et al. (1999) proponen que una forma ancestral del grupo *Lunatus* podría haberse originado en esta región, y posteriormente, haber migrado hacia Suramérica, dejando atrás especies como *P. marechalii*, *P. ritensis*, *P. salicifolius*, *P. polystachyus* y *P. xolocotzii*. En este mismo sentido, esta forma ancestral podría haber derivado en tierras altas andinas a *P. augusti*, *P. pachyrrhizoides*, *P. bolivianus* y *P. lunatus*. Sparvoli et al. (2001), Fofana et al. (1999) y Caicedo et al. (1999) han encontrado mayor cercanía filogenética entre *P. lunatus* y especies aliadas andinas, que con especies mesoamericanas, soportando un origen andino para la especie.

Recientemente, Delgado-Salinas *et al.* (2006) analizaron secuencias nucleares para la región ITS/5.8S y estimaron la edad para el género *Phaseolus* entre 6-8 Ma, correlacionando su diversificación con eventos tectónicos en México, como la formación del Cinturón Trans-Volcánico Mexicano. Así mismo, han indicado edades entre los 1.1 a 1.9 Ma para el grupo *Lunatus*. Estudios recientes describen patrones filogeográficos para el frijol común (*P. vulgaris*), mostrando un posible origen mesoamericano para esta especie, aproximadamente 1.3 Ma atrás, así como dos eventos de migración, uno desde Mesoamérica hacia Suramérica y uno en sentido contrario. Del mismo modo, se infiere una fuerte influencia de eventos geológicos, como el cierre del Istmo de Panamá y el último período del levantamiento de los Andes, en la distribución actual del frijol común (Chacón *et al.* 2007).

*P. lunatus* constituye una oportunidad para analizar patrones de la biota neotropical, como la expansión y diferenciación genética entre linajes. Las especies ampliamente distribuidas pueden haber sido influenciadas por procesos históricos como fragmentación del rango y extinción localizada, para los cuales se espera encontrar una alta divergencia intraespecífica. Por otro lado, una reciente especiación y expansión del rango, podría evidenciar menor estructura genética (Byrne *et al.* 2002; Templeton 1998). Los enfoques genético poblacionales y filogeográficos proveen el marco de evaluación de hipótesis de estos procesos en la evolución de las poblaciones (Avice 1998; Lumaret *et al.* 2002; Schaal *et al.* 2003; Templeton *et al.* 1995; Templeton *et al.* 2000; Templeton 2001; Tomaru *et al.* 1998).

Los análisis filogeográficos se han visto beneficiados por la reciente implementación de herramientas moleculares. Entre éstas se destaca el uso del ADN cloroplástico, cuyas características particulares como son ausencia de recombinación, baja tasa de evolución y herencia maternal en la mayoría de las angiospermas (Birky 1995), así como la disponibilidad de pares de cebadores universales para su amplificación (Taberlet *et al.* 1991, Demesure *et al.* 1995), lo hacen ideal en la determinación de la

variación intraespecífica, la definición de haplotipos y construcción de redes haplotípicas (Chacón *et al.* 2005; Schaal *et al.* 1998; Soltis *et al.* 1989).

Del mismo modo, en el genoma nuclear, la región ITS/5.8S, a pesar de sus limitantes (Álvarez & Wendel, 2003) ha demostrado contener información útil para resolver las relaciones entre poblaciones de especies antiguas y/o ampliamente distribuidas, así como relaciones filogenéticas a nivel intragenérico e historia evolutiva a niveles intraespecíficos (Baldwin *et al.* 1995, Dick *et al.* 2003, Delgado-Salinas *et al.* 2006, Chacón *et al.* 2007).

## 1. OBJETIVOS

Este estudio se planteó investigar los procesos que han dado forma a la distribución actual de fríjol lima silvestre, brindando una visión de los patrones de migración e intercambio de flora en las Américas. Con el uso de herramientas de la filogeografía, genética de poblaciones y filogenética, se propuso:

### 1.1 Objetivo General

Determinar la historia evolutiva y estructura poblacional del fríjol lima silvestre en Mesoamérica y Suramérica y su relación con especies silvestres afines.

### 1.2 Objetivos Específicos

- (i) Evaluar marcadores moleculares útiles para estudios de filogeografía y genética de poblaciones en fríjol lima
- (ii) Determinar los patrones de diversidad genética y distribución de la misma en fríjol lima silvestre
- (iii) Reconocer los procesos históricos que han moldeado la distribución actual de fríjol lima silvestre
- (iv) Establecer los posibles tiempos de divergencia entre *P. lunatus* y sus especies cercanas, y entre los acervos andino y mesoamericano dentro de *P. lunatus*.

## 2. MATERIALES Y METODOS

### 2.1 Estudio Piloto: Marcadores del genoma cloroplástico y del genoma nuclear polimórficos en frijol lima silvestre

Una de las principales limitaciones para realizar estudios filogenéticos y filogeográficos en plantas es la ausencia de marcadores moleculares apropiados, tanto del genoma cloroplástico como del nuclear (Schaal & Olsen, 2000; Hare, 2001). El objetivo de esta fase fue identificar marcadores moleculares que presenten variación intraespecífica, que potencialmente permitan la identificación de haplotipos en diferentes regiones geográficas en frijol lima silvestre, desde una muestra de seis entradas de germoplasma, que representan los dos acervos genéticos. Para ello, se realizó la exploración de 11 regiones no codificantes del genoma cloroplástico (cpADN) y el espaciador interno transcrito del ADN ribosomal (ITS/5.8S).

#### 2.1.1 Material vegetal y extracción de ADN

Dentro del estudio piloto fueron incluidas cuatro especies cercanas a *P. lunatus*, tres de ellas provenientes de Suramérica (*P. augusti*, *P. pachyrrhizoides* y *P. bolivianus*) y una especie de Mesoamérica (*P. novoleonensis*). Todo el material vegetal usado en este estudio proviene del Centro Internacional de Agricultura Tropical-CIAT, Cali, Colombia (Tabla 1) para lo cual se firmó un Acuerdo de Transferencia de Material (ATM). Se germinaron cinco semillas por entrada en cajas de Petri. Las plántulas fueron transferidas a substrato sólido y enriquecido de las cuales se colectó material vegetal almacenado en silicato de sodio (sílica gel) a -20°C. La extracción de ADN se realizó a partir de hojas jóvenes mediante el método CTAB (Doyle & Doyle, 1987).

### 2.1.2 Amplificaciones PCR en regiones de ADN cloroplástico y nuclear

Se exploraron once regiones no-codificantes del genoma cloroplástico, distribuidas a lo largo de la molécula (LSC- Figura 1), entre las cuales están nueve regiones espaciadoras intergénicas y dos intrones. Todas las regiones fueron amplificadas con cebadores universales reportados previamente (Tabla 2). Estas regiones de cpADN fueron escogidas porque han demostrado contener información útil a nivel intraespecífico en diversas especies vegetales, incluyendo frijol (Asmussen & Liston 1998, Bailey & Doyle 1997, Bakker *et al.* 1999, Chacón 2001, Chacón *et al.* 1999, Demesure *et al.* 1995, Dutech *et al.* 2000, Small *et al.* 1998, Taberlet *et al.* 1991, Xu *et al.* 2000, Fofana *et al.* 1999, Fofana *et al.* 2001). El ADN nuclear fue explorado usando el espaciador interno transcrito del ADN ribosomal (ITS/5.8S). Éste se secuenció usando el par de cebadores ITS1 e ITS4 reportados por White *et al.* (1990).

Las reacciones de amplificación por PCR contenían 0.5  $\mu$ M de cada cebador, 1X de Buffer (libre de Mg) para la *Taq* Polimerasa, 2.0 mM  $MgCl_2$ , 0.2 mM dNTPs, 1 unidad de *Taq* DNA polimerasa (Promega, Madison, USA) y 5-10 ng de ADN. Estas condiciones variaron para algunas regiones como se describe a continuación: la concentración de  $MgCl_2$  para las regiones *petA – psbE*, *trnT – trnL* e ITS1 – ITS4 fue 2.5  $\mu$ M, y para la región espaciadora *psbC – tRNASer* y los intrones *trnL* y *rpl16* fue 1.5  $\mu$ M.

El programa de amplificación consistió en: 4 min. a 94°C, seguido por 35 ciclos de 1 min. a 94°C, 1 min. a 50-51°C (dependiendo de cada par de cebadores, ver Tabla 2), y 2 min. a 72°C, por último, una extensión final de 10 min. a 72°C. La región *tRNASer-tRNAfMet* fue tratada mediante un gradiente de  $MgCl_2$  desde 1.0 hasta 3.0  $\mu$ M y un ciclo “step-up” para temperatura de alineamiento desde 45°C hasta 50°C sin obtener amplificaciones. Para la región *rpl16 intron* se trabajó una temperatura de extensión final de 65°C. Los productos PCR fueron visualizados por tinción con bromuro de etidio en geles de agarosa al 1% en 0.5X Buffer TBE.

### 2.1.3 Purificación de productos PCR y secuenciamiento

Los productos PCR fueron purificados con el kit wizard SV Gel and PCR Clean-up System de Promega Corporation (Madison, WI, USA). Se enviaron muestras de 15  $\mu$ l, a una concentración de 35-50 ng/ $\mu$ l de producto PCR purificado, para secuenciamiento automatizado en un ABI 3700 Macrogen Inc. (Seoul, Korea). Se analizaron muestras por duplicado como control de calidad de secuencia.

Todas las regiones fueron secuenciadas en sentido 5'-3', excepto las regiones *petA* – *psbE*, *psbC* – *tRNA<sup>Ser</sup>* y *trnT* – *trnL* para el estudio piloto y *atpB* – *rbcL* para el total de las muestras, las cuales fueron secuenciadas en sentido 5'-3' y 3'-5', de modo que, dos secuencias complementarias fueron obtenidas para cada producto PCR. Las secuencias obtenidas fueron editadas en Bioedit (Versión 5.0.6 – Hall 2001), donde se construyeron las matrices para análisis posteriores. Se realizó una prueba para evaluar la posible existencia de polimorfismos intra-entrada, para lo cual se amplificó y secuenció el ITS/5.8S de 20 individuos para cada una de dos entradas, una perteneciente al acervo genético mesoamericano y una al andino.

### 2.1.4 Selección de regiones de cpADN e ITS/5.8S

Durante la selección de regiones informativas a nivel filogeográfico en el cpADN se estimaron parámetros como el número de sitios polimórficos, inter e intra específicos, sitios informativos a nivel de parsimonia (PI= se refiere a aquellos caracteres que son compartidos por al menos dos terminales) y la diversidad nucleotídica  $P_i$  ( $\pi$ ) (Nei & Miller, 1990) para cada región, usando el software DNAsp (Rozas et al. 2003). Las regiones que mostraron un mayor nivel de variación intraespecífica y estructura geográfica (es decir, que permiten la detección de diferentes haplotipos en diversas regiones geográficas) fueron preferidas para los fines de este trabajo. Estos indicadores, así como la amplificación exitosa en el total de las muestras de la fase

piloto permitieron la determinación de las regiones de cpADN útiles para la realización del estudio.

## **2.2 Diversidad de haplotipos del cpADN e ITS/5.8S en *Phaseolus lunatus* silvestre y su distribución geográfica.**

### *2.2.1 Haplotipos y distribución geográfica*

Se realizó secuenciamiento de los espaciadores intergénicos *atpB-rbcL* y *trnL-trnF* del cpADN seleccionados del estudio piloto y la región ITS/5.8S en una muestra de 59 entradas de frijol lima silvestre (Tabla 1). Estas muestras representan: (i) el rango de distribución geográfico del frijol lima silvestre y (ii) los dos acervos genéticos dentro de la especie. En adición, se seleccionaron 11 entradas de las 4 especies aliadas mencionadas previamente, y dos especies mesoamericanas adicionales (*P. leptostachyus* y *P. marechalli*) que pertenecen al acervo genético secundario de *P. lunatus*.

Los haplotipos para cada región de ADN se definieron colapsando las secuencias según el número de pasos mutacionales y el nivel de conectividad en el software TCS (Clement *et al.* 2000). La codificación de los gaps (presencia/ausencia) en las secuencias se realizó según el método propuesto por Simmons & Ochotorena (2000).

### *2.2.2 Acervos genéticos en *P. lunatus silvestre**

Las muestras de frijol lima usadas en este estudio representan entradas de germoplasma presentes a lo largo del rango de distribución de la especie y no poblaciones naturales, por ello, para este estudio se definieron poblaciones discretas teniendo en cuenta los acervos genéticos descritos para la especie (Caicedo *et al.* 1999; Fofana *et al.* 1999). Para corroborar la subdivisión dentro del frijol lima en dos

acervos genéticos, se obtuvo una matriz de distancias genéticas a partir de las secuencias del cpADN e ITS/5.8S y bajo el modelo de sustitución nucleotídica de Jukes-Cantor (frecuencias nucleotídicas iguales y un solo tipo de sustitución). A partir de esta matriz se construyó un dendrograma por medio del algoritmo de Neighbour-Joining (NJ) en el software PAUP\* (Swofford, 1998). El soporte de los grupos en el dendrograma se validó mediante análisis de bootstrap con 1000 réplicas (Giribert 2003). Diferentes aproximaciones, tanto en método de reconstrucción, como modelo de evolución de secuencias, fueron realizadas para el conjunto de datos, obteniendo el mismo tipo de relaciones entre las OTUs (Unidad Taxonómica Operacional) analizadas.

### 2.2.3 *Diversidad genética y Estructura Poblacional*

La diversidad genética dentro de poblaciones se estimó con base en varios índices usando el software DNAsp (Rozas *et al* 2003). Estos índices incluyeron: el número de sitios segregantes o polimórficos (S), la diversidad nucleotídica  $P_i$  ( $\pi$ ) (que corresponde al número promedio de diferencias nucleotídicas por sitio entre pares de secuencias) (Nei 1987, Nei & Miller 1990) y la diversidad haplotípica ( $H_d$ ) (Nei 1987). La subdivisión poblacional a nivel nucleotídico se evaluó mediante el estadístico  $N_{ST}$  (Lynch & Crease, 1990). Como un método indirecto de estimación del flujo génico entre las poblaciones se calculó el número de migrantes por generación ( $N_m$ ) a partir del  $N_{ST}$ .

Una segunda exploración del nivel de estructura poblacional fue realizada mediante el Análisis de Varianza Molecular (AMOVA – Excoffier *et al.* 1992) usando el software Arlequín 2.00 (Schneider *et al.* 2000). La partición de la variación se realizó dentro y entre las agrupaciones definidas por el análisis de distancia. Se realizaron tres tipos de comparación, la primera de estas teniendo en cuenta los tres grupos obtenidos (AI, MI y MII). La segunda comparando solo los grupos mesoamericanos (MI y MII) y la

última de estas, comparando el acervo andino (AI) con el previamente planteado acervo mesoamericano (MI+MII). El nivel de significancia del Índice de Fijación  $\Phi_{ST}$  se estimó mediante 10000 permutaciones.

## **2.3 Filogeografía de *Phaseolus lunatus* silvestre**

### *2.3.1 Historia demográfica*

La historia demográfica del frijol lima silvestre se investigó mediante dos tests estadísticos, el test estadístico de neutralidad D (Tajima 1989) y el test  $R_2$  (Ramos-Onsins & Rozas, 2002), con el fin de evaluar la hipótesis nula de tamaños poblacionales constantes versus incremento reciente del tamaño poblacional.

El test D de Tajima evalúa las frecuencias de las variantes polimórficas en una secuencia en particular. Bajo neutralidad y tamaños poblacionales constantes se espera que el valor D sea cercano a cero. Valores negativos para D indican que hay un exceso de sitios polimórficos en muy bajas frecuencias, patrón que puede ser causado por reciente selección o incremento reciente del tamaño poblacional (expansión). Para determinar el nivel de significancia de D, se realizaron simulaciones por coalescencia con 5000 réplicas.

El test  $R_2$  está basado en la diferencia entre el número de mutaciones únicas y el número promedio de diferencias nucleotídicas. Se espera que el índice  $R_2$  tenga valores significativamente más bajos en un escenario de reciente crecimiento poblacional. Este índice es sugerido para tamaños de muestra pequeños y/o cuando los eventos recombinación son considerados. La significancia de  $R_2$  se midió mediante 5000 simulaciones por coalescencia.

### 2.3.2 *Redes haplotípicas y Test de contingencia*

La historia filogeográfica del fríjol lima silvestre se investigó mediante la construcción de redes de haplotipos y el test de contingencia para evaluar la asociación entre los haplotipos y su distribución geográfica.

Las redes haplotípicas se construyeron según el criterio de Parsimonia propuesto por Templeton (1987), implementado en el software TCS (Clement *et al.* 2000), con un nivel de conectividad del 95%. Los haplotipos en la red fueron agrupados en clados anidados de acuerdo al algoritmo de Templeton *et al.* (1992). En este algoritmo, los haplotipos separados por un paso mutacional son agrupados en clados de primer orden, luego los clados de primer orden son agrupados en clados de segundo orden, y así sucesivamente hasta que todos los clados se agrupan en un cladograma total.

Una vez realizado el anidamiento, se llevaron a cabo las pruebas de asociación entre los haplotipos y sus distancias geográficas mediante tests de contingencia implementados en el software GeoDis (Posada *et al.* 2000). Para este fin se calcularon dos distancias,  $D_c$  y  $D_n$ .  $D_c$  indica la dispersión geográfica de los individuos anidados dentro de un mismo clado, mientras  $D_n$  indica la dispersión geográfica de los individuos dentro de un mismo clado con respecto al centro geográfico del clado de orden mayor donde éste está anidado. En todos los casos la hipótesis nula es de no asociación geográfica de los haplotipos dentro de los clados. En los casos en que la probabilidad  $\chi^2$  es significativa ( $p < 0.05$ ) se concluye que hay evidencia de estructura geográfica de los haplotipos. La significancia del test de contingencia se evaluó mediante permutaciones ( $n=1000$ ).

### 2.3.3 *Análisis filogenético y tiempos de divergencia*

El análisis filogenético se realizó utilizando 97 secuencias del ITS/5.8S: 23 entradas secuenciadas entre *P. lunatus* y especies afines, y 54 secuencias obtenidas del

GenBank (Tabla 3), las cuales son representativas del clado Milletioide (Figura 3 en Lavin *et al.* 2005). El alineamiento múltiple fue realizado en Muscle (versión 3.6 Edgar, 2005) el cual desarrolla un algoritmo rápido y eficiente, con respecto a otros programas. Se realizó una exploración de costos para los gaps, utilizando valores desde -20 hasta 0 (siendo 1 el valor por defecto) para generar los alineamientos. De cada alineamiento producido se realizaron búsquedas heurísticas rápidas (TBR) usando el criterio de parsimonia, en Winclada (Versión 1.00.08, Nixon, 2002). Los grupos recuperados en cada topología fueron comparados entre sí.

El modelo de evolución molecular fue seleccionado usando el programa MrModeltest (Nylander, 2004). Una primera evaluación se realizó para el total del alineamiento, es decir la región ITS/5.8S, que comprende los espaciadores intergénicos (ITS1 e ITS2) y el gen ribosomal 5.8S. Dado que regiones codificantes y no codificantes pueden evolucionar de forma independiente (Page & Holmes, 1998), una segunda exploración se efectuó separando dichas regiones y realizando la selección del modelo de manera independiente. Búsquedas por medio de análisis bayesiano se realizaron usando Mr.Bayes (Huelsenbeck & Ronquist, 2001) bajo el modelo GTR + G + I, usando dos cadenas y 400.000 generaciones, mediante un muestreo cada 100 generaciones. Los árboles correspondientes al primer 25% del análisis fueron descartados (100.000 generaciones, 1000 árboles). Se realizó un consenso de la mayoría sobre los árboles restantes.

Para estimar el tiempo de divergencia de la especie, la primera hipótesis planteada constituye el comportamiento de reloj molecular en las secuencias. Esta hipótesis se evaluó mediante el Likelihood Ratio Test (LRT, Muse & Weir 1992) que compara los valores de likelihood de la topología estimados con (L1) y sin (L0) reloj molecular. El test se realizó comparando dos veces la diferencia en el logaritmo de los likelihood

con una distribución de chi-cuadrado con  $n-2$  grados de libertad, siendo  $n$  el número de terminales.

En la estimación de tiempos de divergencia de la especie y los acervos genéticos, se aplicó un análisis Penalized Likelihood (PL – Sanderson, 2002) usando el software R8S (Sanderson, 2004), el cual permite la variación de la tasa de substitución entre ramas, penalizando aquellas tasas que cambian muy rápido. El valor de “smoothing”<sup>2</sup> se determinó mediante el proceso de validación cruzada (cross-validation). De un muestreo de 100 topologías obtenidas por MrBayes con longitudes de rama estimadas en PAUP\* se hizo constricción de la edad de tres ancestros comunes recientes, *Xeroderris stuhlmannii-Phaseolus* (40.2 – 48.4 Ma), *Wajira albescens-Phaseolus* (8.7 – 13.2 Ma) y *Vigna unguiculata-Phaseolus* (6.4 – 10.4 Ma), usando las edades mínimas y máximas calculadas (nodos 58, 67 y 68 respectivamente) en Lavin et al. (2005). Estas edades provienen de la constricción en 55 Ma como edad mínima del registro fósil de *Barnebyanthus buchananensis*, representando el clado Papilionoideae.

---

<sup>2</sup> Smoothing ( $\lambda$ ): parámetro que controla el ajuste de los datos a un modelo,  $\lambda \approx 0$  permite gran variación entre las tasas de substitución,  $\lambda \approx \infty$  tasas de substitución iguales entre las ramas

### 3. RESULTADOS Y DISCUSION

#### 3.1 Estudio Piloto

De las 11 regiones del cloroplasto exploradas, solo se realizó el secuenciamiento de 9 de ellas, dado que, las regiones *tRNASer-tRNAfMet* y *rpl16 Intron* no fue posible amplificarlas y/o la calidad de su secuencia no fue óptima. El secuenciamiento de 9 regiones del cpADN en la muestras del estudio piloto de frijol lima silvestre produjo un total de 5193 nucleótidos. La amplificación interespecífica de estas regiones mostró poca variación en tamaño ( $\pm 10$  bp), debido principalmente a la presencia de pequeños indels<sup>3</sup>. La diversidad nucleotídica  $\pi$  de estas nueve regiones se muestra en la Tabla 2. En total fueron encontrados el doble de polimorfismos interespecíficos (101) que intraespecíficos (50). Basados en los criterios propuestos previamente, se identificaron tres regiones de cpADN que mostraron ser útiles para estudios filogenéticos en este género. Estas regiones fueron *petA-psbE*, *psbC-tRNAser* y *trnS-trnG*, las cuales contienen 1.8%, 1.3% y 1% caracteres PI, respectivamente.

Las regiones que resultan mas útiles para estudios genético-poblaciones y filogeográficos son aquellas que son de fácil amplificación y que contienen altos niveles de diversidad genética estructurada geográficamente. Dos regiones del cloroplasto mostraron estas propiedades, los espaciadores intergénicos *atpB-rcbL* y *trnL-trnF*, que aunque contienen niveles moderados de diversidad intraespecífica ( $\pi_{intra}$ ) comparados con otras regiones (ver tabla 2), sus polimorfismos son parsimoniosamente más informativos a nivel intraespecífico (definen linajes intraespecíficos) y además son de fácil y exitosa amplificación en frijol Lima. Estas dos regiones del cloroplasto fueron seleccionadas para el análisis posterior en el total de las muestras de frijol lima.

---

<sup>3</sup> Indel: Acrónimo de los dos tipos de mutación genética, que involucran la inserción o delección de uno o varios pares de bases, entre una secuencia y otra

El ITS/5.8S presentó poca variación interespecífica en tamaño. Se detectaron un total de 84 polimorfismos, 25 de los cuales fueron intraespecíficos y 59 fueron interespecíficos. Esta región mostró una diversidad nucleotídica intraespecífica e interespecífica de 0.00817 y 0.02846, respectivamente (Tabla 2). En resumen, el estudio piloto de polimorfismos mostró que algunas regiones cpADN y el ITS/5.8S son útiles para detectar variación inter e intraespecífica. Esta variación es apropiada para estudios filogenéticos y filogeográficos en esta especie. Sin embargo, análisis más detallados en otras regiones de cpADN serían útiles para reconocer su utilidad en estudios futuros.

### **3.2 Diversidad de haplotipos del cpADN e ITS/5.8S en *Phaseolus lunatus* silvestre y su distribución geográfica.**

#### *3.2.1 Haplotipos y distribución geográfica*

Se realizó secuenciamiento de los espaciadores intergénicos *atpB-rbcL* y *trnL-trnF* del cpADN en una muestra de 59 entradas de frijol lima silvestre. En total se secuenciaron 1324 pb por cada entrada y se observaron un total de 29 sitios polimórficos, 16 de los cuales fueron informativos bajo el criterio de parsimonia. Los sitios polimórficos permitieron la identificación de 23 haplotipos (denominados con las letras A hasta W), 21 de los cuales se encuentran restringidos a una sola región geográfica y solo dos de ellos (haplotipos G y T) se encuentran distribuidos entre Mesoamérica y Suramérica (Tabla 4). Estos resultados sugieren que la diversidad del cpADN dentro del frijol lima silvestre se encuentra estructurada geográficamente, lo cual permite el uso de estos marcadores en estudios filogeográficos dentro de la especie. Para la región ITS/5.8S, se obtuvo secuencia para 673 pb en cada una de las 59 entradas analizadas y se observó un total de 16 sitios polimórficos, 10 de éstos fueron informativos bajo el criterio de parsimonia. Estos polimorfismos permitieron

la diferenciación de 28 haplotipos, 26 de los cuales están geográficamente localizados y solo dos (U y Z) poseen una amplia distribución (Tabla 5).

### 3.2.2 *Acervos genéticos en P. lunatus silvestre*

Los dendrogramas derivados de cpADN e ITS/5.8S se muestran en las figuras 2 y 3, respectivamente. Valores de bootstrap para los nodos menores al 50% no se muestran. La topología de cpADN sugiere la diferenciación de tres grupos principales dentro de *P. lunatus* silvestre (grupo AI, MI y MII) aunque los valores bootstrap del nodo que separa el grupo AI de los grupos MI + MII, y del nodo que separa los grupos MI y MII son bajos (51% y <50%, respectivamente). El grupo AI podría corresponder al acervo genético andino y los otros dos al acervo genético mesoamericano. El acervo genético andino (grupo AI) se encuentra restringido a Ecuador y norte de Perú (Figura 4A), distribuyéndose principalmente en zonas con altitudes entre los 800 a 1800 msnm, excepto en las entradas G26721, G26468 y G26606 que corresponden al muestreo más occidental del acervo andino, presentes entre los 400 a 800 msnm.

Dentro del acervo genético mesoamericano se reconocen dos grupos, el primero de estos, denominado grupo MI, comprende casi el total de las entradas de *P. lunatus* presentes en México (al occidente y sur occidente del Istmo de Tehuantepec) y un par de entradas en Guatemala, Honduras y Colombia. Por otro lado, el grupo MII, de este mismo acervo, posee una distribución mayor, que comprende las islas del Caribe, la península de Yucatán, Mesoamérica (al oriente y suroriente del Istmo de Tehuantepec) y Suramérica, incluyendo Colombia, sur de Perú y norte de Argentina.

Para la región ITS/5.8S (Figura 3) se encontró el mismo patrón observado en cpADN, sugiriendo la existencia de tres acervos principales, uno andino (AI) y dos mesoamericanos (MI-MII) aunque con valores de bootstrap bajos para los nodos (<50%) al igual que se observó con el cpADN. En el caso de la región ITS/5.8S, el

grupo MI, de distribución menor dentro del acervo mesoamericano, aparece aun más restringido, exclusivamente en la zona de la Sierra Madre del Sur y el Cinturón Volcánico Trans-Mexicano (Figura 4B).

En Colombia y el norte de Ecuador se encontraron algunas ambigüedades en cuanto a que algunas entradas aparecen genéticamente más cercanas al acervo mesoamericano según la evidencia de ITS/5.8S y más cercanas al acervo andino según su genoma cloroplástico (ej. G26606). Ambigüedades también se presentan entre los dos grupos mesoamericanos, donde entradas silvestres como G26704 y G25819 son ubicadas por cpADN en el grupo MI y el ITS/5.8S las sitúa en el grupo MII. Las ambigüedades encontradas en las entradas de Colombia y norte de Ecuador sugieren, al igual que Fofana *et al.* (2001), la existencia de una zona de encuentro entre los acervos andino y mesoamericano en esta región.

En resumen, el análisis de distancia en *P. lunatus* silvestre sugiere la separación de tres grupos, aunque las relaciones entre estos no son del todo claras debido a los bajos valores de bootstrap en los nodos que separan los grupos ( $\leq 50\%$ ), soportando débilmente la subdivisión en los acervos mesoamericano y andino, como ha sido sugerido anteriormente (Gutiérrez Salgado *et al.* 1995; Maquet *et al.* 1999, Caicedo *et al.* 1999; Fofana *et al.* 1999).

Adicionalmente se observó que el frijol lima está más cercanamente relacionado con las especies afines andinas (AA), es decir, con *P. pachyrrhizoides*, *P. augusti* y *P. bolivianus*, que con aquellas de distribución mesoamericana (AM) (Figuras 2 y 3). Este resultado nos permite sugerir un origen andino para el frijol lima silvestre, específicamente en el área de Ecuador-Norte del Perú, y se plantearía una posterior migración hacia diferentes regiones en Suramérica y Mesoamérica. Los resultados de este análisis de distancia muestran las mismas agrupaciones propuestas por Delgado-Salinas *et al.* (1999, 2006), donde el grupo Lunatus comprende a *P. lunatus* y

especies de distribución en Suramérica, mientras que especies como *P. marechalli*, *P. leptostachios* y *P. novoleonensis* pertenecen a otros grupos.

### 3.2.3 Diversidad Genética y estructura poblacional

Los índices de diversidad estimados reflejan para ambos marcadores moleculares (cpADN e ITS/5.8S), una mayor diversidad en el acervo andino (grupo AI) con respecto al mesoamericano (0.00263 vs. 0.00236 para cpADN y 0.00686 vs. 0.00353 para ITS/5.8S, respectivamente) (Tabla 6). Así mismo, dentro del acervo mesoamericano, se encuentra que el grupo MI (geográficamente más restringido) es más variable que su complemento mesoamericano ampliamente distribuido (MII), especialmente en secuencias ITS/5.8S ( $P_i = 0.00145$  vs. 0.00132 en cpADN, 0.00341 vs. 0.00198 en ITS/5.8S, para los grupos MI y MII, respectivamente).

Los valores de  $N_{ST}$  para el cpADN y el ITS/5.8S indican altos niveles de subdivisión poblacional ( $N_{ST} > 0.5$  en todos los casos, ver tabla 7). Para cpADN se encuentra alta divergencia entre el acervo andino y mesoamericano ( $N_{ST} = 0.64345$ ), así como entre los grupos mesoamericanos ( $N_{ST} = 0.65393$ ), siendo mayor el nivel de subdivisión entre estos últimos. El ITS/5.8S muestra una estructura similar para los grupos evaluados, indicando una alta diferenciación entre los grupos del acervo mesoamericano ( $N_{ST} = 0.70125$ ). La estimación del número de migrantes entre cada par de poblaciones ( $N_m$ ) revela un escenario de flujo génico restringido ( $N_m < 1.0$ , Tabla 7), corroborando el grado de subdivisión poblacional encontrado.

El nivel de estructura genética encontrada entre los acervos de frijol lima silvestre, Andino versus Mesoamericano, así como, entre los grupos mesoamericanos, evidencia un flujo génico restringido y alta divergencia entre poblaciones. Este patrón de estructura y dinámica poblacional se produce como un efecto sinérgico entre la biología reproductiva de la especie, autogamia y dispersión corta, y las barreras geográficas que separan dichos acervos.

El nivel de divergencia entre los grupos del acervo genético mesoamericano ( $N_{ST}=0.654$  y  $0.701$  para cpADN e ITS/5.8S respectivamente), es aproximadamente de igual o mayor magnitud que entre el acervo andino y mesoamericano del frijol lima. Este patrón podría evidenciar una separación muy temprana de éstos, durante la historia de la especie, sugiriendo así, el tratamiento de los grupos MI y MII como acervos genéticos independientes.

Del mismo modo, los resultados del análisis de AMOVA en cpADN presentan estructura poblacional en los tres enfoques realizados (Tabla 8), recuperando el mayor porcentaje de variación entre poblaciones, que dentro de ellas. Se encuentra un índice de fijación mayor cuando se hace la comparación entre los tres grupos definidos por el análisis de distancia, que cuando se evalúa el acervo mesoamericano como un todo. Se encontró alta subdivisión poblacional para cpADN e ITS/5.8S entre los dos grupos mesoamericanos (MI y MII), con valores de  $\Phi_{ST}$  de  $0.596$  y  $0.763$  respectivamente.

Es así como los índices de fijación y el análisis de varianza dan una fuerte evidencia a la estructuración del frijol lima en tres grupos, uno de distribución andina, uno ampliamente distribuido desde México hasta Argentina y uno distribuido principalmente en México al norte del Istmo de Tehuantepec.

### **3.3 Filogeografía de *P. lunatus silvestre***

#### *3.3.1 Historia Demográfica*

El test de Tajima en todos los casos mostró valores negativos, pero solo fue consistentemente significativo ( $p<0.05$ ) para el acervo mesoamericano MII tanto para el cpADN como el ITS/5.8S. Este valor podría indicar un proceso de expansión demográfica reciente en el área de Mesoamérica, al oriente y suroriente del Istmo de

Tehuantepec, y en Suramérica donde este acervo está principalmente distribuido. Adicionalmente,  $R_2$  se encontró significativamente bajo en el análisis total del acervo mesoamericano ( $R_2= 0.0562$ ,  $p=0.03400$  en cpADN y  $R_2= 0.0521$ ,  $p=0.00800$  en ITS/5.8S), siendo el grupo MII, el que nuevamente contiene valores mas bajos (Tabla 6).

El acervo genético andino de frijol lima silvestre, parece encontrarse en equilibrio demográfico, no encontrando huellas de expansión. Su distribución restringida permite el aislamiento geográfico e impide el flujo génico con su contraparte mesoamericana, contribuyendo a su diferenciación. Por otro lado, el acervo mesoamericano, subdividido en dos grupos, el cual parece haber pasado por un proceso de expansión del rango (valores del test de Tajima y  $R_2$  significativos, especialmente el grupo MII), actualmente se encuentra ampliamente distribuido.

### 3.3.2 *Redes haplotípicas y tests de contingencia*

La red de cpADN para frijol lima silvestre (Figura 5) contiene tres niveles de anidamiento (clados de grado 1, 2 y 3). El cladograma total se divide en tres grupos, el clado 3-1 contiene todos los haplotipos pertenecientes al acervo andino, el clado 3-2 y 3-3 poseen los haplotipos correspondientes al acervo genético mesoamericano. El clado 3-2 corresponde al grupo MII, ampliamente distribuido, y el clado 3-3 corresponde al grupo mesoamericano (MI) que comprende en su mayoría entradas de México (Figura 2 y Figura 4A). La red haplotípica para el ITS/5.8S presenta mayor grado de anidamiento (Figura 6), con clados hasta de cuarto orden. El cladograma total contiene dos grupos, el 4-1 que corresponde al acervo genético andino, separado por al menos 4 pasos mutacionales del clado 4-2 que corresponde al acervo mesoamericano. Dentro del clado 4-2 se distinguen tres clados de tercer orden, el clado 3-3 corresponde al acervo MI y los clados 3-4 y 3-5 corresponden al acervo MII (Figura 3 y Figura 4B). En la Tabla 9 se muestran los clados en los cuales se encontró asociación significativa ( $p<0.05$ ).

El test de contingencia para el cladograma total de cpADN muestra que tanto el clado 3-1 (acervo andino y presumiblemente ancestral) como el clado 3-3 (acervo mesoamericano MI y presumiblemente derivado) tienen distancias  $D_c$  significativamente pequeñas. Bajo el modelo de fragmentación del rango se espera que las distancias  $D_c$  sean significativamente pequeñas tanto para clados ancestrales como para clados derivados. Por lo tanto, los patrones de distancias geográficas encontradas para el cpADN apoyan la hipótesis de un evento histórico de fragmentación del rango en la especie, posiblemente en los Andes, dando origen a los acervos andino y mesoamericano.

El clado 3-2 del cpADN (acervo mesoamericano) contiene a su vez el clado 2-6 (ampliamente distribuido en Centroamérica y el norte de Suramérica), el cual es presumiblemente ancestral al clado 2-5 (restringido a Perú y Argentina y presumiblemente derivado). Las distancias  $D_c$  y  $D_n$  del clado ancestral son significativamente más pequeñas de lo esperado, lo cual contrasta con la distancia  $D_n$  del clado derivado la cual es significativamente más grande de lo esperado. Este patrón sugiere un evento de expansión del rango para el acervo mesoamericano en donde los clados ancestrales se encuentran más cerca de su área de origen, mientras que los clados derivados se encuentran ampliamente distribuidos y alejados de su origen.

Las redes de haplotipos permiten inferir patrones de migración y divergencia en el frijol lima silvestre. A partir de un acervo ancestral en el área de Ecuador-Norte del Perú (acervo AI, Figuras 5 y 6), el acervo mesoamericano (MI+MII) habría divergido posiblemente por un evento de fragmentación y habría migrado hacia el norte y sur de Suramérica, hacia Centroamérica y posteriormente hacia Mesoamérica. A partir del linaje que entró a Mesoamérica, un segundo evento de divergencia habría ocurrido resultando en la separación de los acervos MI y MII, los cuales actualmente se encuentran casi totalmente aislados geográficamente por el Istmo de Tehuantepec.

### 3.3.3 Análisis filogenético y tiempos de divergencia

En la exploración de costos para los gaps se encontró que valores muy bajos ( $< -5$ ) generaban alineamientos de mayor longitud, abundantes en gaps, mientras que valores como 0, 1 y -2, producen resultados estables. Las topologías resultantes (figuras no mostradas) mantienen el número y disposición de los grupos obtenidos. El alineamiento usado en los análisis posteriores fue producido con los costos por defecto para el programa (1 apertura, 1 extensión). La selección del modelo de evolución molecular para el total de la región ITS/5.8S mediante el criterio de Akaike (AIC) y el Test Jerárquico (hLRT), presentó a GTR + G + I como modelo mas probable, para ambos casos. Al realizar la búsqueda por regiones se obtiene que para los dos espaciadores se selecciona GTR + G + I y SYM+G+I para el gen 5.8S, por lo que el análisis bayesiano fue realizado con GTR+G+I.

El test jerárquico rechaza la hipótesis nula de evolución bajo un modelo de reloj molecular para los 100 árboles evaluados ( $\chi^2= 450-600$ , g.l.=95  $P<0.001$ ). La figura 7 representa el consenso de la búsqueda por Análisis Bayesiano involucrando el clado Millettoid (Lavin et al. 2005). El análisis filogenético realizado recupera en su mayoría los grupos propuestos por Delgado-Salinas et al. (1999), encontrando a *P. lunatus* mas cercano filogenéticamente a especies de distribución estrictamente andina. De este modo se reconoce el denominado clado *Lunatus* que contiene a *P. augusti*, *P. bolivianus*, *P. pachyrrhizoides* y *P. lunatus* (ver figura 7) propuesto como la única diversificación del genero en Suramérica (Delgado-Salinas et al. 2006). Este agrupamiento da soporte al origen andino para la especie.

Las edades estimadas para el ancestro común mas reciente de *P. lunatus* y sus especies hermanas (nodo 1, figura 7), y entre *P. lunatus* mesoamericano y *P. lunatus* andino (nodo 2, figura 7) se muestran en la tabla 10. Los análisis realizados indican una tiempo de divergencia de al menos 2.23 Ma entre *P. lunatus* y las especies de

distribución andina y una posterior divergencia entre los acervos andino y mesoamericano, hace al menos 1.11 Ma.

La estimación de la divergencia de *P. lunatus* indicando una edad de al menos 2.23 Ma, es mayor que la propuesta por Delgado-Salinas et al. (2006) para el grupo Lunatus ( $1.1 \pm 0.2$  para *ITS/5.8S* y  $1.9 \pm 0.5$  para *trnK*), que contiene las especies de *Phaseolus* de distribución en Suramérica. De este modo, *P. lunatus* posiblemente se habría originado en Suramérica una vez formado el Istmo de Panamá (7-4 Ma), de modo que durante la evolución del género se evidenciaría, como ha propuesto Fofana et al. (1999), una migración desde mesoamérica de una forma ancestral de la cual habrían derivado las especies del grupo Lunatus.

La distribución geográfica y la diversidad genética en la vegetación suramericana se encuentran influenciados por la historia geológica de la región (Gentry 1982) y los patrones encontrados corresponden a un efecto combinado, difícil de elucidar, entre la orogenia andina (Mioceno al presente), el cierre del Istmo de Panamá (Plioceno) y los cambios climáticos del Pleistoceno (Burnham & Graham 1999). En la historia de la especie se puede observar una fuerte influencia de la orogenia Andina en este patrón, dado que los acervos genéticos, andino y mesoamericano, se encuentran separados según la estimación desde hace al menos 1.11 Ma. Esta fragmentación podría resultar como consecuencia del último levantamiento de la zona occidental del norte de los andes (Ecuador y norte de Perú) hacia el final del Plioceno (Simpson, 1975).

#### 3.3.4 Origen del frijol lima

Se podrían considerar dos hipótesis alternas en cuanto a la región de origen de la especie. La primera de ellas indica que el frijol lima silvestre es de origen andino (Sparvoli et al., 2001; Fofana et al. 1999; Caicedo et al. 1999) y la segunda indica un origen mesoamericano, como se ha encontrado para algunas especies del mismo género (Chacón et al. 2007).

Los resultados de esta investigación también sugieren un origen andino para el frijol lima silvestre. La evidencia proviene, principalmente del análisis filogenético del frijol lima silvestre y especies aliadas andinas y mesoamericanas. En este estudio encontramos que los acervos genéticos de *P. lunatus* aparecen relacionados con especies como *P. bolivianus*, *P. pachyrrhizoides* y *P. augusti*, las cuales se encuentran estrictamente distribuidas en la zona andina de Ecuador, Perú y Bolivia. Dos estudios filogenéticos anteriores dentro del género apoyan también esta asociación (Delgado-Salinas et al. 1999; Delgado-Salinas et al. 2006), en donde *P. lunatus* se agrupa con especies aliadas andinas en un clado con un soporte bootstrap mayor al 98%. Adicionalmente en apoyo a la hipótesis de un origen andino del frijol lima silvestre, en especial en la región del norte de los Andes, se observó en este análisis filogeográfico que el haplotipo C del cpADN del acervo andino es compartido con algunas entradas de *P. augusti* (datos no mostrados), una especie aliada de distribución andina. De acuerdo al modelo de flujo génico restringido, se espera que los haplotipos ancestrales presenten distribuciones geográficas mayores que los haplotipos derivados. En este caso, el conjunto de haplotipos ancestrales del norte de los Andes presentan una distribución geográfica mucho más restringida que los haplotipos mesoamericanos, presumiblemente derivados. El escenario más probable que podría explicar esta distribución corresponde a una fragmentación del rango seguido de aislamiento geográfico y expansión del rango en el acervo mesoamericano. El evento de fragmentación puede verse influenciado por el levantamiento de los Andes en la zona occidental de Ecuador y norte de Perú, hacia el final del Plioceno e inicio del Pleistoceno (3.5–1.8 Ma) (Simpson, 1975).

La hipótesis de un origen mesoamericano podría verse apoyada por los patrones de diversidad reportados previamente para los acervos genéticos andino y mesoamericano, donde Gutiérrez et al. (1995) observó un mayor número de variantes electroforéticas en proteínas de semilla para el acervo mesoamericano (nueve en comparación contra cuatro en el acervo andino) y Fofana et al. (1999, 2001) encontró mayor diversidad genética dentro del grupo mesoamericano estimada a partir de

Polimorfismos en Longitud de Fragmentos de Restricción (RFLPs) del cpADN. En el análisis filogeográfico, la única evidencia que podría apoyar la hipótesis de un origen mesoamericano para el frijol lima es la amplia distribución de los haplotipos mesoamericanos T (cpADN) y Z (ITS/5.8S). Sin embargo, este patrón de distribución podría considerarse como producto de la reciente expansión del rango del grupo mesoamericano y no a persistencia de polimorfismos ancestrales en una amplia región geográfica.

## CONCLUSIONES

La exploración de múltiples marcadores de ADN, tanto organelares como nucleares, permiten la identificación de regiones informativas, según la pregunta de investigación. En este caso, las regiones espaciadoras del genoma cloroplástico *atpB-rbcL* y *trnL-trnF*, así como el ITS/5.8S, representando el genoma nuclear, demostraron ser fuente de información apropiada para estudios filogeográficos y de domesticación en *P. lunatus*. Dichos marcadores moleculares poseen alto número de polimorfismos intra e interespecíficos, que en compañía de una alta reproducibilidad permiten su utilización. Los haplotipos identificados en estas regiones de ADN, presentan gran estructura geográfica, permitiendo el estudio filogeográfico dentro de la especie.

El análisis filogenético de *P. lunatus* silvestre con sus especies cercanas, indica un origen andino para la especie hace al menos 2.23 Ma. Los haplotipos presentes en *P. lunatus* muestran una fuerte estructura geográfica, permitiendo la identificación de tres acervos genéticos dentro de la especie. El grupo andino, posee una distribución restringida en la región occidental del norte de los andes y muestra una mayor diversidad genética que sus contrapartes mesoamericanas. Este acervo parece haberse diferenciado de los mesoamericanos desde hace al menos 1.11 Ma, esta edad indica una posible influencia del levantamiento andino en la zona de Ecuador y norte de Perú, en el aislamiento e interrupción del flujo génico para este acervo.

Se reconocen dos grupos mesoamericanos los cuales se encuentran fuertemente estructurados, sugiriendo su tratamiento como acervos independientes. El primero de ellos, distribuido desde el suroriente del Istmo de Tehuantepec, el caribe y en Suramérica hasta Argentina (pendiente oriental de los Andes), parece haber pasado por un evento de expansión del rango reciente, dada su amplia distribución y relativamente bajos niveles de diversidad genética. El segundo acervo mesoamericano propuesto se encuentra distribuido al nor-occidente del Istmo de

Tehuantepec, principalmente en la zona de la Sierra Madre del Sur y el Cinturón Trans-Volcánico Mexicano, y posee un nivel de diversidad mayor.

El alto nivel de diferenciación entre los acervos AI, MI y MII sugiere que éstos se formaron relativamente temprano durante la historia evolutiva de la especie. En principio, a partir de una población ancestral se habría diferenciado el acervo AI y el ancestro de los acervos MI y MII en la zona entre Ecuador y norte del Perú. Posteriormente, a partir de este ancestro se habría dado una diferenciación muy temprana de los acervos MI y MII los cuales habrían alcanzado su distribución actual por diferentes procesos (por ejemplo expansión del rango en el acervo MII) y por posiblemente por patrones de migración para los cuales no se posee evidencia. La alta diferenciación genética entre los acervos sugiere que la historia evolutiva de éstos ha transcurrido en relativo aislamiento el uno del otro. Un análisis más detallado con respecto a estos dos acervos, así como la inclusión de material de la zona del Caribe, podría brindar mayor información sobre la historia reciente de la especie.

Los tiempos de divergencia entre el fríjol lima y sus especies hermanas, así como entre las entradas mesoamericanas y andinas, sugieren que eventos biogeográficos como la formación del Istmo de Panamá y el levantamiento de los Andes parecen haber influenciado la distribución y estructura de la especie. Posteriores investigaciones en especies de distribución neotropical que involucren las barreras mencionadas contribuirían al entendimiento de la configuración actual y los procesos que dieron forma a la flora neotropical.

## REFERENCIAS

- Acevedo-Rodríguez P (1993) Systematics of *Serjania* (Sapindaceae) Part I: A revision of *Serjania* sect. *Platycoccus*. *Memoirs of The New York Botanical Garden* 67, 1-94.
- Alvarez A, Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*. 29: 417-434.
- Asmussen CB, Liston A (1998) Chloroplast DNA characters, phylogeny and classification of *Lathyrus* (Fabaceae). *American Journal of Botany*. 85: 387-401
- Avice JC (1998) The history and purview of phylogeography: a personal reflection. *Molecular Ecology*. 7, 371-379.
- Bailey CD, Doyle JJ (1997) A comparison of the potential genetic utility of an intron of the MADS box gene *pistillata* to the nrDNA ITS and the chloroplast *trnL* intron in *Sphaerocardamum Schauer* (Brassicaceae). *American Journal of Botany*. 84: 175
- Bakker FT, Culham A, Daugherty LC, Gibby M (1999) A *trnL-F* based phylogeny for species of *Pelargonium* (Geraniaceae) with small chromosomes. *Plant Systematics and Evolution*. 216:309-324
- Baldwin, BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ (1995). The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden* 82: 247-277.
- Baudet JC (1977) Origine et classification des espèces cultivées du genre *Phaseolus*. *Bull. Soc. Roy. Bot. Belgique* 110: 65-76.

- Baudoin J, Rocha J, Degreef A, 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 12. International Plant Genetic Resources Institute, Rome, Italy.
- Beebe S, Skroch PW, Tohme J, Duque MC, Pedraza F, Nienhuis J (2000) Structure of genetic diversity among common bean landraces of Middle American origin based on correspondence analysis of RAPD. *Crop Science*. 40: 264-273.
- Birky W Jr. (1995) Uniparental inheritance of mitochondrial and chloroplast genes: Mechanisms and evolution. *Proc. Natl. Acadm. Sci. USA*. Vol. 92, 11331-11338.
- Burnham R, Graham A (1999) The History of Neotropical Vegetation: New Developments and Status. *Annals of the Missouri Botanical Garden*. Vol. 86, No. 2. pp. 546-589.
- Byrne M, Macdonald B, Coates D (2002) Phylogeographical patterns in chloroplast DNA variation within the *Acacia acuminata* (Leguminosae: Mimosoideae) complex in Western Australia. *Journal Evolution Biology* 15, 576-587.
- Caicedo AL, Gaitan E, Duque MC, Toro O, Debouck DG, Tohme J (1999) AFLP fingerprinting of *Phaseolus lunatus* L. and related wild species from South America. *Crop. Science*. 39: 1497-1507.
- Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9(10): 1657-1660
- Chacón M (2001) Chloroplast DNA polymorphisms and the evolution and domestication of the common bean (*Phaseolus vulgaris* L.). PhD Thesis-The University of Reading, Reading, England.

- Chacón MI, Pickersgill B, Debouck DG (1999) Intraspecific chloroplast DNA diversity in common bean (*Phaseolus vulgaris*) for domestication studies. Annual Report of the Bean Improvement Cooperative. 42: 79-80.
- Chacón MI, Pickersgill B, Debouck DG (2005) Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. Theoretical and Applied Genetics. 110: 432-444.
- Chacón MI, Pickersgill D, Debouck DG, Arias S (2007) Phylogeographic analysis of the chloroplast DNA variation in wild common bean (*Phaseolus vulgaris* L.) in the Americas. Plant Systematics and Evolution.
- Coates A (1997) Central America. A Natural and Cultural History United States of America. Yale University Press. USA.
- Debouck DG, Liñan J, Campanas S, De La Cruz A, Rojas JH (1987) Observations on the domestication of *Phaseolus lunatus* L. P1. Genet. Resources Newslett. 70:26-32.
- Debouck DG (1996) Colombian Common and Lima beans: views on their origin and evolutionary significance. Revista Corpoica 1: 7-15
- Delgado-Salinas A (2000) New Species of Mexican *Phaseolus* (Fabaceae). Systematic Botany. 25,414-436.
- Delgado-Salinas A, Bruneau A, Doyle J (1993) Chloroplast DNA phylogenetic studies in New World *Phaseolinae* (Leguminosae: Papilionoideae: Phaseoleae). Systematic Botany, 18, 6-17.

- Delgado-Salinas A, Turley T, Richman A, Lavin M (1999) Phylogenetic analysis of the cultivated and wild species of *Phaseolus* (Fabaceae). *Systematic Botany*. 24 (3):438-460.
- Delgado-Salinas A, Bibler R, Lavin M (2006) Phylogeny of the genus *Phaseolus* (Leguminosae): a recent diversification in an ancient landscape. *Systematic Botany* 31(4):779-791.
- Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology*. 4: 129-131
- Dick C, Abdul-Salim K, Bermingham E (2003) Molecular Systematic Analysis Reveals Cryptic Tertiary Diversification of a Widespread Tropical Rain Forest Tree. *The American Naturalist*, 162: 691-703.
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem.Bull.* 19, 11-15.
- Dutech C, Maggia L, Joly L (2000) Chloroplast diversity in *Vouacapoua americana* (Caesalpiniaceae), a neotropical forest tree. *Molecular Ecology*. 9: 1427-1432
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, Vol. 32, No. 5 1792-1797.
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131, 479-491.
- Felsenstein J (1989) PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* 5:164-166.

- Fofana B, du Jardin P, Baudoin JP (2001) Genetic diversity in the Lima bean (*Phaseolus lunatus* L.) as revealed by chloroplast DNA (cpDNA) variations. *Genetic Resources and Crop Evolution*. 48(5): 437-445.
- Fofana B, Baudoin JP, Vekemans X, Debouck DG, 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.
- Gentry AH (1982) Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Annals of the Missouri Botanic Gardens*. 69: 557-593
- Gepts P (1988) Genetic Resources of *Phaseolus* beans. *Current Plant Science and Biotechnology in Agriculture*. Kluwer Academic Publishers, Dordrecht. pp. 613.
- Giribert G (2003) Stability in phylogenetic formulations and its relationship to nodal support. *Syst. Biol.* 52 (4): 554-564.
- Gutiérrez A, Gepts P, Debouck DG (1995) Evidence for two gene pools of the Lima bean, *Phaseolus lunatus* L., in the Americas. *Genetic Resources and Crop Evolution*. 42: 15-28.
- Hall T (2001) Bioedit Version 5.0.6. Biological sequence alignment editor. Copyright 1997-2001.
- Hare MP (2001) Prospects for nuclear gene phylogeography. *Trends in ecology and evolution*. 16(12): 700-706.
- Hardy OS, Dubois I, Zoro B, 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. Newslett.* 109:1-6.

- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17:754-755. Disponible en <http://mrbayes.csit.fsu.edu/download.php>
- Khairallah MM, Sears BB, Adams MW (1992) Mitochondrial restriction fragment length polymorphisms in wild *Phaseolus vulgaris* L.: insights on the domestication of the common bean. *Theoretical and Applied Genetics*. 84: 915-922.
- Ladizinsky G (1998) Plant evolution under domestication. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 254.
- Lavin M, Herendeen P, Wojciechowski MF (2005) Evolutionary rates analysis of *Leguminosae* implicates a rapid diversification of lineages during the Tertiary. *Systematic Biology*. 54 (4): 530-549. Data sets: rbcL (TreeBase study accession numbers S578)
- Lynch M, Crease T (1990) The analysis of population survey data on DNA sequence variation. *Mol. Biol. Evol.* 7 (4): 377-394
- Lumaret R, Mir C, Michaud H, Raynal V (2002) Phylogeographical variation of chloroplast DNA in holm oak (*Quercus ilex* L.). *Molecular Ecology*, 11, 2327-2336
- Maquet A, Vekemans X, Baudoin JP (1999) Phylogenetic study on wild allies of Lima bean, *Phaseolus lunatus* L. (Fabaceae), and implications on its origin. *Plant Syst. Evol.* 218:43-54
- Muse SV, Weir BS (1992) Testing for Equality of Evolutionary Rates. *Genetics* 132: 269-276.

- Nei M, Miller JC (1990) A simple method for estimating average number of nucleotide substitutions within and between populations from restriction data. *Genetics* 125: 873–879.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Nylander JA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University
- Nixon KC (2002) WinClada ver. 1.00.08 Published by the author, Ithaca, NY.
- Page RD, Holmes EC (1998) *Molecular Evolution: A Phylogenetic Approach*. Blackwell Science: Oxford, pp 346. ISBN 0-86542-889-1.
- Pennington T, Dick C (2004) The role of immigrants in the assembly of the South American rainforest tree flora. *Philosophical Transactions The Royal Society B*, 359, 1611-1622.
- Pirie MD, Chatrou LW, Mols JB, Erkens RHJ, Oosterhof J (2006) Andean-centred genera in the short-branch clade of Annonaceae: testing biogeographical hypothesis using phylogeny reconstruction and molecular dating. *J. Biogeogr.* 33: 31-46.
- Posada D, Crandall KA, Templeton AR (2000) GeoDis: A program for the Cladistic Nested Analysis of the Geographical Distribution of Genetic Haplotypes. *Molecular Ecology* 9(4):487-488.
- Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* 19 (12): 2092-2100.

- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496-2497.
- Sanderson MJ (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol.Biol.Evol.* 19: 101-109.
- Sanderson MJ (2004) r8s ver.1.70. Analysis of rates of evolution. Disponible en: <http://loco.biosci.arizona.edu/r8s/>.
- Schaal B, Hayworth D, Olsen K, Rausher J, Smith W (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology*, 7, 465-474.
- Schaal BA, Olsen KM (2000) Gene genealogies and population variation in plants. *Proceedings of the National Academy of Sciences of the United States of America.* 97(13): 7024-7029.
- Schaal B, Gaskin J, Caicedo A (2003) Phylogeography, Haplotype Trees, and Invasive Plant Species. *Journal of Heredity*, 94, 197-204.
- Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Tamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugital M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M. (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO Journal*, 5 (9): 2043-2049.
- Singh SP, Gepts P, Debouck DG (1991) Races of common bean (*Phaseolus vulgaris*, Fabaceae). *Economic Botany*. 45: 379-396

- Simmons M, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analysis. *Syst. Biol.* 49(2): 369-381.
- Simpson BB (1975) Pleistocene changes in the flora of the high tropical Andes. *Paleobiology* 1:273-294.
- Small RL, Ryburn JA, Cronn RC, Seelanan T, Wendel J (1998) The tortoise and the hare: choosing between noncoding plastome and nuclear Adh sequences for phylogeny reconstruction in a recently diverged plant group. *American Journal of Botany.* 85: 1301-1315
- Schneider, S., Roessli, D., and Excoffier, L. (2000) Arlequin: A software for population genetics data analysis. Ver 2.000. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.
- Sparvoli F, Lanave C, Santucci A, Bollini R, Lioi L (2001) Lectin and lectin-related proteins in Lima bean (*Phaseolus lunatus* L.) seeds: biochemical and evolutionary studies. *Plant Mol. Biol.* 45:587-597.
- Soltis D, Soltis P, Ranker T, Ness B (1989) Chloroplast DNA variation in a wild plant, *Tolmie menziessi*. *Genetics*, 121, 819-826.
- Swofford DL (1998) *Phylogenetic Analysis Using Parsimony*, version 4.0. Sinauer Associates, Sunderland (USA).
- Tajima F (1989) Statistical Method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 123: 585-595.
- Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology.* 17: 1105-1109.

- Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132: 619-633.
- Templeton A, Routman C, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander. *Ambystoma tigrinum*. *Genetics* 140, 767-782
- Templeton A (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. . *Molecular Ecology*, 7, 381-397.
- Templeton A, Maskas S, Cruzan M (2000) Gene trees: a powerful tool for exploring the evolutionary biology of species and speciation. *Plant Species Biology*, 15, 211-222.
- Templeton A (2001) Using phylogeographic analyses of gene trees to test species status and processes. *Molecular Ecology*, 10, 779-791.
- Templeton A (1987) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics*. 134: 659-669.
- Tomaru N, Takahashi M, Tsumura Y, Takahashi M, Ohba K (1998) Intraspecific variation and phylogeographic patterns of *Fagus crenata* (Fagaceae) mitochondrial DNA. *American Journal of Botany*, 85, 629-636.
- Webb SD (1997) The great American faunal interchange. In: A. G. Coates (ed.) *Central America: a natural and cultural history*. Yale University Press, Michigan. pp. 97-122.
- White T, Bruns J, Lee T, Taylor W (1990) Amplification and direct sequencing of fungal ribosomal genes for phylogenetics. Pages 315-322 in: *PCR Protocols: A Guide to*

Methods and Applications. M. A. Innis, D. H. Gelfand, J. J. Shinsky, and T. J. White, eds. Academic Press Inc., London.

Xu DH, Abe J, Sakai M, Kahazawa A, Shimamoto Y (2000) Sequence variation of non-coding regions of chloroplast DNA of soybean and related wild species and its implications for the evolution of different chloroplast haplotypes. *Theoretical and Applied Genetics*. 101: 724-732

Zoro B, Maquet A, Baudoin JP (2005) Mating system of wild *Phaseolus lunatus* L. and its relationship to population size. *Heredity* 94: 153-158.

Zoro B, Maquet A, Baudoin JP (2003) Population genetic structure of wild *Phaseolus lunatus* (Fabaceae), with special reference to population size. *American Journal of Botany* 90 (6): 897-904.

Tabla 1. Datos del material vegetal usado en el estudio, cpADN-Hap: Haplotipo definido según cpADN, ITS-Hap: haplotipo definido según región ITS/5.8S. ARG: Argentina, BLZ: Belice, COL: Colombia, CRI: Costa Rica, CUB: Cuba, ECU: Ecuador, GTM: Guatemala, HND: Honduras, MEX: México, PER: Perú, SLV: El Salvador. \* = Entradas usadas en el estudio piloto.

Número G-CIAT	Especie	País	Departamento	Altitud (msnm)	cpADN Hap	ITS Hap.	Estado Biológico
G40716*	<i>P. augusti</i>	ARG	TUCUMAN	1520	-	-	Silvestre
G40758	<i>P. augusti</i>	ECU	AZUAY	1530	-	-	Silvestre
G40774	<i>P. augusti</i>	ECU	AZUAY	1660	-	-	Silvestre
G40746*	<i>P. bolivianus</i>	BOL	COCHABAMBA	2830	-	-	Silvestre
G40759	<i>P. bolivianus</i>	BOL	COCHABAMBA	2650	-	-	Silvestre
G40590*	<i>P. novoleonensis</i>	MEX	NUEVO LEON	2830	-	-	Silvestre
G40603	<i>P. leptostachyus</i>	GTM	JALAPA	1040	-	-	Silvestre
G40598	<i>P. leptostachyus</i>	CRI	SAN JOSE	1210	-	-	Silvestre
G40506	<i>P. marechalli</i>	MEX	MORELOS	-	-	-	Silvestre
G40509	<i>P. pachyrrhizoides</i>	PER	CUZCO	2200	-	-	Silvestre
G40717*	<i>P. pachyrrhizoides</i>	PER	JUNIN	2390	-	-	Silvestre
G26685	<i>P. lunatus</i>	ARG	SALTA	650	M	Z	Silvestre
G26404*	<i>P. lunatus</i>	ARG	CHACO	-	K	Z	Silvestre
G26749	<i>P. lunatus</i>	ARG	JUJUY	-	K	Z	Silvestre
G25233	<i>P. lunatus</i>	BLZ	CAYO DISTRICT	40	T	S	Silvestre
G25819	<i>P. lunatus</i>	COL	MAGDALENA	580	J	U	Silvestre
G26309	<i>P. lunatus</i>	COL	MAGDALENA	50	O	U	Silvestre
G26515	<i>P. lunatus</i>	COL	ATLANTICO	45	T	Z	Silvestre
G26527	<i>P. lunatus</i>	COL	CORDOBA	70	S	Z	Silvestre
G26530*	<i>P. lunatus</i>	COL	BOYACA	1600	T	U	Silvestre
G26618A	<i>P. lunatus</i>	COL	BOYACA	1440	T	A	Silvestre
G26687	<i>P. lunatus</i>	COL	CUNDINAMARCA	1680	T	Z	Silvestre
G26704	<i>P. lunatus</i>	COL	CALDAS	1485	G	U	Silvestre
G25227	<i>P. lunatus</i>	CRI	HEREDIA	1100	T	Z	Silvestre
G25385A	<i>P. lunatus</i>	CRI	SAN JOSE	1171	U	X	Silvestre
G25584A	<i>P. lunatus</i>	CRI	GUANACASTE	100	T	Z	Silvestre
G25963	<i>P. lunatus</i>	CRI	SAN JOSE	1450	T	T	Silvestre
G25966	<i>P. lunatus</i>	CRI	CARTAGO	1320	V	Z	Silvestre
G26726	<i>P. lunatus</i>	CRI	ALAJUELA	1280	T	T	Silvestre
G26727	<i>P. lunatus</i>	CRI	PUNTA ARENAS	1100	T	Z	Silvestre
G25294C	<i>P. lunatus</i>	CUB	MATANZAS	3	N	U	Silvestre
G26459	<i>P. lunatus</i>	ECU	LOJA	1580	C	C	Silvestre
G26468	<i>P. lunatus</i>	ECU	BOLIVAR	870	B	F	Silvestre
G26469	<i>P. lunatus</i>	ECU	IMBADURA	1390	C	G	Silvestre
G26606*	<i>P. lunatus</i>	ECU	CHIMBORAZO	1550	C	U	Silvestre
G26608	<i>P. lunatus</i>	ECU	EL ORO	800	C	D	Silvestre
G26609	<i>P. lunatus</i>	ECU	AZUAY	1570	C	C	Silvestre
G26721	<i>P. lunatus</i>	ECU	AZUAY	470	A	E	Silvestre
G26751A	<i>P. lunatus</i>	ECU	PICHINCHA	1740	C	C	Silvestre

Número G-CIAT	Especie	País	Departamento	Altitud (msnm)	cpADN Hap	ITS Hap.	Estado Biológico
G25273A	<i>P. lunatus</i>	SLV	AHUACHAPAN	754	T	Q	Silvestre
G25222	<i>P. lunatus</i>	GTM	ZACAPA	50	T	V	Silvestre
G25272	<i>P. lunatus</i>	GTM	SANTA ROSA	893	T	R	Silvestre
G25911	<i>P. lunatus</i>	GTM	RETALHULEU	240	T	Z	Silvestre
G25912*	<i>P. lunatus</i>	GTM	SACATEPEQUEZ	1400	T	Z	Silvestre
G25977	<i>P. lunatus</i>	GTM	ESCUINTLA	490	I	Z	Silvestre
G26508	<i>P. lunatus</i>	GTM	PETEN	150	P	W	Silvestre
G26653	<i>P. lunatus</i>	GTM	HUEHUETENANGO	590	T	Z	Silvestre
G26683	<i>P. lunatus</i>	GTM	QUEZALTENANGO	1750	G	A1	Silvestre
G26628	<i>P. lunatus</i>	HND	EL PARAISO	840	R	Y	Silvestre
G26629	<i>P. lunatus</i>	HND	OLANCHO	550	W	Y	Silvestre
G26631	<i>P. lunatus</i>	HND	YORO	210	G	Y	Silvestre
G26632	<i>P. lunatus</i>	HND	SANTA BARBARA	260	T	B1	Silvestre
G26633	<i>P. lunatus</i>	HND	COPAN	860	Q	W	Silvestre
G25228	<i>P. lunatus</i>	MEX	NAYARIT	600	E	H	Silvestre
G25230	<i>P. lunatus</i>	MEX	COLIMA	3	H	J	Silvestre
G25704*	<i>P. lunatus</i>	MEX	JALISCO	1390	G	M	Silvestre
G25713A	<i>P. lunatus</i>	MEX	CAMPECHE	100	G	N	Silvestre
G25785	<i>P. lunatus</i>	MEX	CAMPECHE	140	G	Y	Silvestre
G25816	<i>P. lunatus</i>	MEX	YUCATAN	30	T	Y	Silvestre
G25843	<i>P. lunatus</i>	MEX	MORELOS	900	G	O	Silvestre
G25851	<i>P. lunatus</i>	MEX	MORELOS	1250	G	K	Silvestre
G26360	<i>P. lunatus</i>	MEX	PUEBLA	1600	G	L	Silvestre
G26516	<i>P. lunatus</i>	MEX	CHIAPAS	40	D	P	Silvestre
G26518	<i>P. lunatus</i>	MEX	GUERRERO	240	F	I	Silvestre
G26535	<i>P. lunatus</i>	MEX	CHIAPAS	1160	G	Z	Silvestre
G26541	<i>P. lunatus</i>	MEX	PUEBLA	820	I	L	Silvestre
G26741	<i>P. lunatus</i>	MEX	OAXACA	20	G	L	Silvestre
G25913*	<i>P. lunatus</i>	PER	CAJAMARCA	3	C	C	Silvestre
G26348	<i>P. lunatus</i>	PER	CAJAMARCA	1100	C	B	Silvestre
G26547	<i>P. lunatus</i>	PER	JUNIN	980	L	Z	Silvestre

Tabla 2. Resultados del estudio piloto de variación genética en cpADN e ITS. Diversidad intraespecífica se refiere a la diversidad dentro del frijol lima silvestre. PI: caracteres parsimonia informativos (ver texto), PU: caracteres parsimonia no-informativos, Intra: Intraespecífico, Inter: Interespecífico. Tm, Temperatura de alineamiento.

Región	Tm (°C)	No. de bases secuenciadas	No. de sitios polimórficos				Diversidad Nucleotídica (Pi)		Fuente
			PI	PU	Intra	Inter	Intra	Inter	
<b>cpADN</b>									
<i>rps14-psaB</i>	50	529	0	0	0	2	0.00000	0.00151	2
<i>trnL</i> intron	51	478	2	2	1	4	0.00139	0.00377	3
<i>accD-psaI</i>	51	632	1	14	2	14	0.00213	0.00832	4
<i>trnS-trnG</i>	51	605	6	17	7	18	0.00223	0.01043	5
<i>atpB-rbcL</i>	50	894	6	9	6	12	0.00294	0.00503	1
<i>trnL-trnF</i>	50	433	2	9	6	9	0.00431	0.01215	3
<i>psbC-tRNA<sup>Ser</sup></i>	50.5	633	8	5	9	9	0.00586	0.00681	2
<i>trnT-trnL</i>	50	657	5	18	11	20	0.00916	0.01329	3
<i>petA-psbE</i>	50	332	6	11	8	13	0.01121	0.01436	2
<b>Total</b>		<b>5193</b>	<b>36</b>	<b>85</b>	<b>50</b>	<b>101</b>	<b>0.00436</b>	<b>0.00841</b>	
ITS / 5.8S	50	695	15	46	25	59	0.00817	0.02846	6

Fuente de secuencias de cebadores: 1. Chiang *et al.* (1998), 2. Fofana *et al.* (1999), 3. Taberlet *et al.* (1991), 4. Small *et al.* (1998), 5. Hamilton (1999), 6. White (1990).

Tabla 3. Secuencias del ITS/5.8S usadas en el análisis de tiempos de divergencia, tomadas del GenBank.

Número GenBank	Especie	Número GenBank	Especie
AF467015	<i>Abrus precatorius</i>	AF115205	<i>Phaseolus micranthus</i>
AF417018	<i>Amphicarpaea bracteata</i>	AF115209	<i>Phaseolus microcarpus</i>
AF467019	<i>Apios americana</i>	AF115236	<i>Phaseolus neglectus</i>
AF467020	<i>Austrosteenisia blackii</i>	AF115235	<i>Phaseolus neglectus</i>
AF467021	<i>Brachypterum robustum</i>	DQ445759	<i>Phaseolus novoleonensis</i>
AF467043	<i>Dalbergiella nysaeae</i>	DQ445760	<i>Phaseolus oaxacanus</i>
AF467046	<i>Derris laxiflora</i>	AF115232	<i>Phaseolus oligospermus</i>
AY583515	<i>Dipogon lignosus</i>	DQ445761	<i>Phaseolus parvifolius</i>
AF467048	<i>Fordia splendidissima</i>	AF115211	<i>Phaseolus parvulus</i>
AF467049	<i>Galactia striata</i>	DQ445765	<i>Phaseolus pedicellatus</i>
U60551	<i>Glycine max</i>	AF115220	<i>Phaseolus perplexus</i>
AY583516	<i>Lablab purpureus</i>	DQ445770	<i>Phaseolus reticulatus</i>
AF467057	<i>Lonchocarpus lanceolatus</i>	AF115186	<i>Phaseolus ritensis</i>
AF467474	<i>Millettia grandis</i>	AF115182	<i>Phaseolus salicifolius</i>
AF467481	<i>Millettia thonningii</i>	AF115194	<i>Phaseolus sinuatus</i>
AF467482	<i>Mundulea sericea</i>	AF115193	<i>Phaseolus smilacifolius</i>
AF467483	<i>Neodunnia richardiana</i>	DQ445773	<i>Phaseolus sonorensis</i>
AF467486	<i>Paraderris elliptica</i>	AF115219	<i>Phaseolus tenellus</i>
AF115141	<i>Phaseolus acutifolius</i>	AF115231	<i>Phaseolus tuerckheimii</i>
AF069126	<i>Phaseolus acutifolius</i>	AF115161	<i>Phaseolus vulgaris</i>
AF115148	<i>Phaseolus albescens</i>	AF115161	<i>Phaseolus vulgaris</i>
AF115180	<i>Phaseolus augusti</i>	AF069128	<i>Phaseolus vulgaris</i>
AF115222	<i>Phaseolus chiapasanus</i>	AF115224	<i>Phaseolus xanthotrichus</i>
AF069130	<i>Phaseolus coccineus</i>	AF115247	<i>Phaseolus xolocotzii</i>
AF115147	<i>Phaseolus costaricensis</i>	AF115227	<i>Phaseolus zimapanensis</i>
AF115151	<i>Phaseolus dumosus</i>	AF467487	<i>Philenoptera eriocalyx</i>
AF115244	<i>Phaseolus glabellus</i>	AF467488	<i>Philenoptera laxiflora</i>
AF115240	<i>Phaseolus grayanus</i>	AF467490	<i>Piscidia piscipula</i>
AF115241	<i>Phaseolus grayanus</i>	AF467491	<i>Platycyamus regnellii</i>
AF115226	<i>Phaseolus hintonii</i>	AF467494	<i>Pongamiopsis amygdalina</i>
AF115199	<i>Phaseolus jaliscanus</i>	AF338217	<i>Pueraria montana</i>
AF115191	<i>Phaseolus juquilensis</i>	AY583502	<i>Sphenostylis angustifolia</i>
AF115177	<i>Phaseolus lignosus</i>	AY508800	<i>Strophostyles umbellata</i>
AF115174	<i>Phaseolus lunatus</i>	AF467497	<i>Tephrosia heckmanniana</i>
DQ445753	<i>Phaseolus macrolepis</i>	DQ445736	<i>Vigna unguiculata</i>
AF115187	<i>Phaseolus maculatus</i>	AY583508	<i>Wajira albescens</i>
AF115200	<i>Phaseolus macvaughii</i>	AF467485	<i>Xeroderris stuhlmannii</i>
AF115198	<i>Phaseolus marechalii</i>		

Tabla 4. Distribución geográfica de haplotipos del cpADN de frijol lima silvestre. Los números corresponden al número de entradas con cada haplotipo en cada región, los grupos (AI, MI, MII) están definidos según la topología NJ (Figura 2).

Región		Grupo																			Total							
		AI			MI						MII																	
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S		T	U	V	W			
Mesoamérica	México				1	1	1	8	1	1											1				14			
	Guatemala								1		1						1							5	8			
	Belice																							1	1			
	Honduras							1										1	1					1	5			
	El Salvador																							1	1			
	Costa Rica																							5	1	1	7	
Caribe	Cuba													1											1			
Suramérica	Colombia							1		1					1						1	4			8			
	Argentina												2		1										3			
	Ecuador	1	1	6																					8			
	Perú			2									1												3			
<b>Total</b>		1	1	8	1	1	1	11	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1	18	1	1	1	59

Tabla 5. Distribución geográfica de haplotipos del ITS/5.8S de frijol lima silvestre. Los números corresponden al número de entradas con cada haplotipo en cada región, los grupos (AI, MI, MII) están definidos según la topología NJ (Figura 3).

Región	Grupo																				Total								
	AI							MI					MII																
	Haplotipo																												
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A1	B1	
México								1	1	1	1	3	1	1	1	1								2	1				14
Guatemala																		1				1	1		4	1			8
Belice																		1											1
Mesoamérica																							1	3			1		5
El Salvador																1													1
Costa Rica																				2			1	4					7
Caribe																						1							1
Suramérica																							4		3				8
Colombia	1																							3					3
Argentina																									3				3
Ecuador																						1							8
Perú																									1				3
<b>Total</b>																													
	1	1	4	1	1	1	1	1	1	1	1	3	1	1	1	1	1	1	1	2	6	1	2	1	5	16	1	1	59

Tabla 6. Análisis de diversidad de los espaciadores intergénicos del cpADN (*atpB-rbcL* y *trnL-trnF*) y la región ITS/5.8S en fríjol Lima silvestre. Los grupos están definidos según la topología de NJ (Figuras 2 y 3). S: Sitios polimórficos, PI: Sitios informativos bajo el criterio de parsimonia, H: número de haplotipos, Hd: Diversidad Haplótipica,  $\pi$ : Diversidad Nucleotídica, SD: desviación estándar, D-Tajima= test neutralidad, Prob.  $|DT|>0$ , si este número es  $<0,05$  entonces  $DT$  es significativamente diferente de cero,  $R_2$ = índice de Ramos-Onsins & Rozas (2002) y su significancia= Prob.  $(|R_2|>0)$ . Valores significativos en *Itálica*. OTU, No. De Unidades Taxonómicas Operacionales.

Grupo	OTU	No. de sitios	S	PI	H	Hd $\pm$ SD	$\pi$ $\pm$ SD	D-Tajima	Prob. $( DT >0)$	$R_2$	Prob. $( R_2 >0)$
<b>cpADN</b>											
Acervo Andino (AI)	10	1319	11	5	6	0.867 $\pm$ 0.085	0.00263 $\pm$ 0.00076	-0.6361	0.28080	0.1273	0.13620
Acervo Mesoamericano (MI+MII)	49	1319	21	11	18	0.804 $\pm$ 0.046	0.00236 $\pm$ 0.00022	-1.4565	0.05160	0.0562	<i>0.03400</i>
Grupo MI	18	1319	9	4	9	0.758 $\pm$ 0.106	0.00145 $\pm$ 0.00033	-1.2453	0.10300	0.0845	<i>0.04560</i>
Grupo MII	31	1314	13	7	12	0.630 $\pm$ 0.102	0.00132 $\pm$ 0.00029	-1.8583	<i>0.01380</i>	0.0523	<i>0.00180</i>
<b>Total</b>	59	1324	29	16	21	0.849 $\pm$ 0.033	0.00323 $\pm$ 0.00029	-	-	-	-
<b>ITS</b>											
Acervo Andino (AI)	10	670	13	7	7	0.867 $\pm$ 0.107	0.00686 $\pm$ 0.00125	-0.0348	0.52140	0.1443	0.32400
Acervo Mesoamericano (MI+MII)	49	673	14	6	12	0.720 $\pm$ 0.064	0.00353 $\pm$ 0.00057	-1.6278	<i>0.02760</i>	0.0521	<i>0.00800</i>
Grupo MI	8	673	7	1	5	0.786 $\pm$ 0.151	0.00341 $\pm$ 0.00100	-1.1414	0.10500	0.1167	<i>0.02580</i>
Grupo MII	41	671	8	3	8	0.611 $\pm$ 0.082	0.00198 $\pm$ 0.00037	-1.6941	<i>0.01460</i>	0.0563	<i>0.01580</i>
<b>Total</b>	59	673	16	10	17	0.789 $\pm$ 0.051	0.00498 $\pm$ 0.00066	-	-	-	-

Tabla 7. Diferenciación genética y flujo génico en *P. lunatus* silvestre. Poblaciones definidas según las figuras 2 y 3.  $\pi$  D, valor promedio de sustituciones nucleotídicas por sitio.  $N_{ST}$ , Distancia genética promedio entre poblaciones, valores entre 0 y 1, indicando nula y completa subdivisión poblacional, respectivamente. Nm, número de migrantes estimado desde  $N_{ST}$ , flujo génico restringido cuando ( $Nm < 1$ ).

<b>Población 1</b>	<b>Población 2</b>	<b><math>\pi</math> D (= <math>K_{xy}</math>)</b>	<b><math>N_{ST}</math></b>	<b><math>Nm</math></b>
<b>cpADN</b>				
Andina (AI)	Mesoamérica (MI + MII)	6.91020	0.643*	0.14
Mesoamérica (MI)	Mesoamérica (MII)	4.08065	0.655*	0.13
<b>ITS</b>				
Andina (AI)	Mesoamérica (MI + MII)	5.27551	0.549*	0.21
Mesoamérica (MI)	Mesoamérica (MII)	6.90000	0.701*	0.11

Significancia  $F_{ST}$  \* $p < 0.001$

Tabla 8. Resultados del análisis AMOVA dentro y entre los grupos obtenidos por el análisis de distancia.

Acervos	% de variación		Índices de Fijación ( $\Phi_{ST}$ )
	Entre poblaciones	Dentro de las poblaciones	
<b>cpADN</b>			
AI – MI – MII	64.63	35.37	0.646**
MI – MII	59.61	40.39	0.596**
AI – (MI+MII)	56.78	43.22	0.568**
<b>ITS/5.8S</b>			
AI – MI – MII	70.30	29.70	0.730**
MI – MII	76.28	23.72	0.763**
AI – (MI+MII)	51.57	48.43	0.516**

Significancia  $\Phi_{ST}$  \*\* $p < 0.001$

Tabla 9. Test de contingencia en *P. lunatus* silvestre. Los clados de menor grado son excluidos. Dc, dispersión del clado. Dn, distancia del clado desde el centro del clado anidado (desplazamiento). Significancia  $p < 0.05$ . Distancias significativas se indican en negrilla y el superíndice S, cuando las distancias son significativamente menores a lo esperado ( $P < 0.05$ ); y el superíndice L, cuando las distancias son significativamente mayores a lo esperado ( $P < 0.05$ ). Los clados ancestrales están resaltados con sombra gris.

<b>cpADN</b>				
Clado	$X^2 (P)$	Clados Anidados	Dc	Dn
3-1	10 (0.000)	2-1	89.39	100.66
		2-2	221.74	222.26
		I-T	132.35	121.59
3-2	30 (0.000)	2-5	808.01	<b>3597.34<sup>L</sup></b>
		2-6	<b>713.36<sup>S</sup></b>	<b>813.15<sup>S</sup></b>
		I-T	-94.64	<b>-2784.19<sup>S</sup></b>
3-3	18 (0.000)	2-3	705.68	711.04
		2-4	1050.06	1044.42
		I-T	344.38	333.38
Total	116 (0.000)	3-1	<b>197.94<sup>S</sup></b>	1575.86
		3-2	1091.57	<b>1069.97<sup>S</sup></b>
		3-3	<b>785.12<sup>S</sup></b>	1468.39
<b>ITS/5.8S</b>				
Clado	$X^2 (P)$	Clados Anidados	Dc	Dn
4-1	10 (0.000)	3-1	380.76	385.14
		3-2	144.95	178.16
		I-T	-235.80	-206.98
4-2	96 (0.000)	3-3	<b>384.27<sup>S</sup></b>	1630.04
		3-4	420.85	904.09
		3-5	1064.34	1036.97
Total	58 (0.000)	I-T	-513.30	-240.71
		4-1	<b>323.04<sup>S</sup></b>	1568.09
		4-2	<b>1124.75<sup>S</sup></b>	1220.99

Tabla 10. Estimación de tiempos de divergencia basada en una muestra de 100 árboles por inferencia bayesiana de ITS/5.8S. Ma, Millones de años. SD, desviación estándar.

<b>Nodo</b>	<b>MCRA</b>	<b>Edad media (Ma)</b>	<b>SD (Ma)</b>	<b>Mínimo (Ma)</b>	<b>Máximo (Ma)</b>
A	<i>Xeroderris – Phaseolus</i>	45.2	1.7	40.2	48.4
B	<i>Wajira – Phaseolus</i>	10.7	0.9	8.7	13.2
C	<i>Vigna - Phaseolus</i>	8.0	0.8	6.4	10.4
1	<i>P. lunatus</i> – especies aliadas	2.23	0.32	1.74	3.01
2	<i>P. lunatus</i> acervo andino – <i>P. lunatus</i> acervo mesoamericano	1.11	0.16	0.85	1.51

Figura 1. Ubicación de las regiones de cpADN exploradas, gráfica basada en el mapa físico del genoma cloroplástico de *Nicotiana tabacum* (Shinozaki et al. 1986).

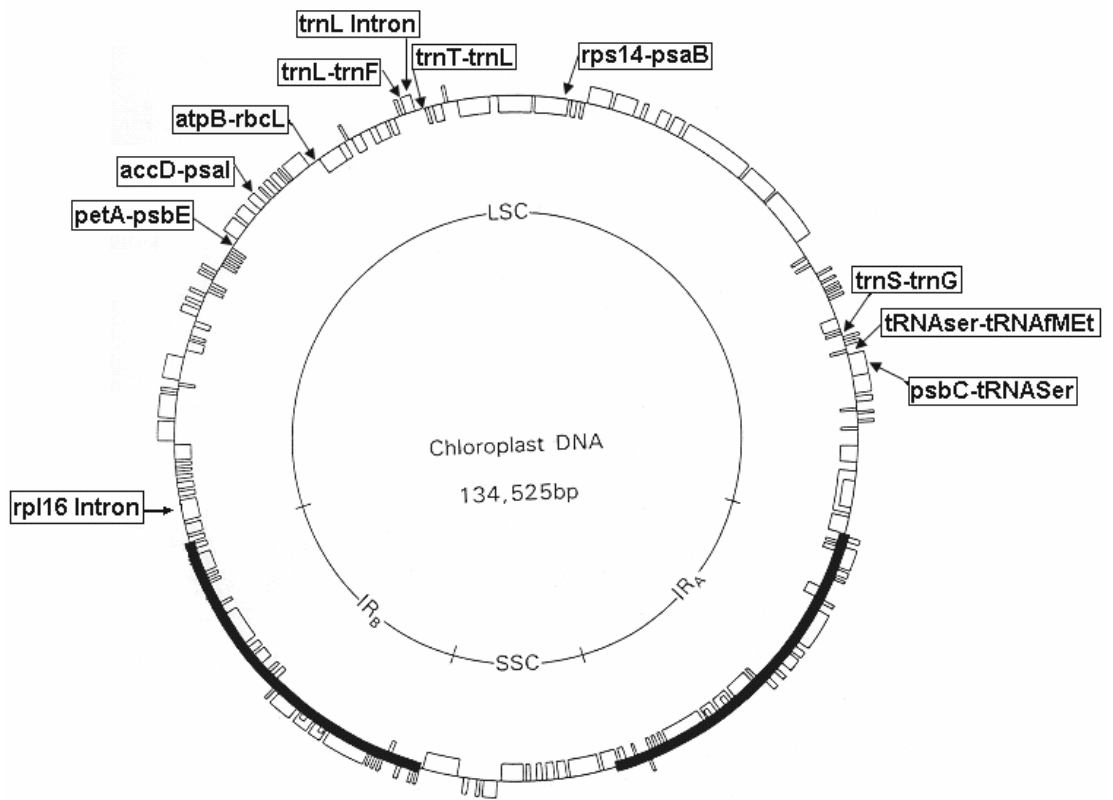


Figura 2. Topología NJ del cpADN para *P. lunatus* silvestre (las entradas están indicadas con los códigos 25- y 26-, seguido de la inicial del país de origen). MEX=México, GTM=Guatemala, BLZ= Belice, HND= Honduras, SLV= El Salvador, CRI=Costa Rica, COL= Colombia, CUB= Cuba, ECU= Ecuador, PER= Perú, BOL= Bolivia, BRA= Brasil, ARG= Argentina. Grupos definidos dentro de fríjol lima silvestre AI, MI, MII. AA: Especies Aliadas Andinas, AM: Especies Aliadas Mesoamericanas. Los números sobre las ramas corresponden a los valores de bootstrap (%).

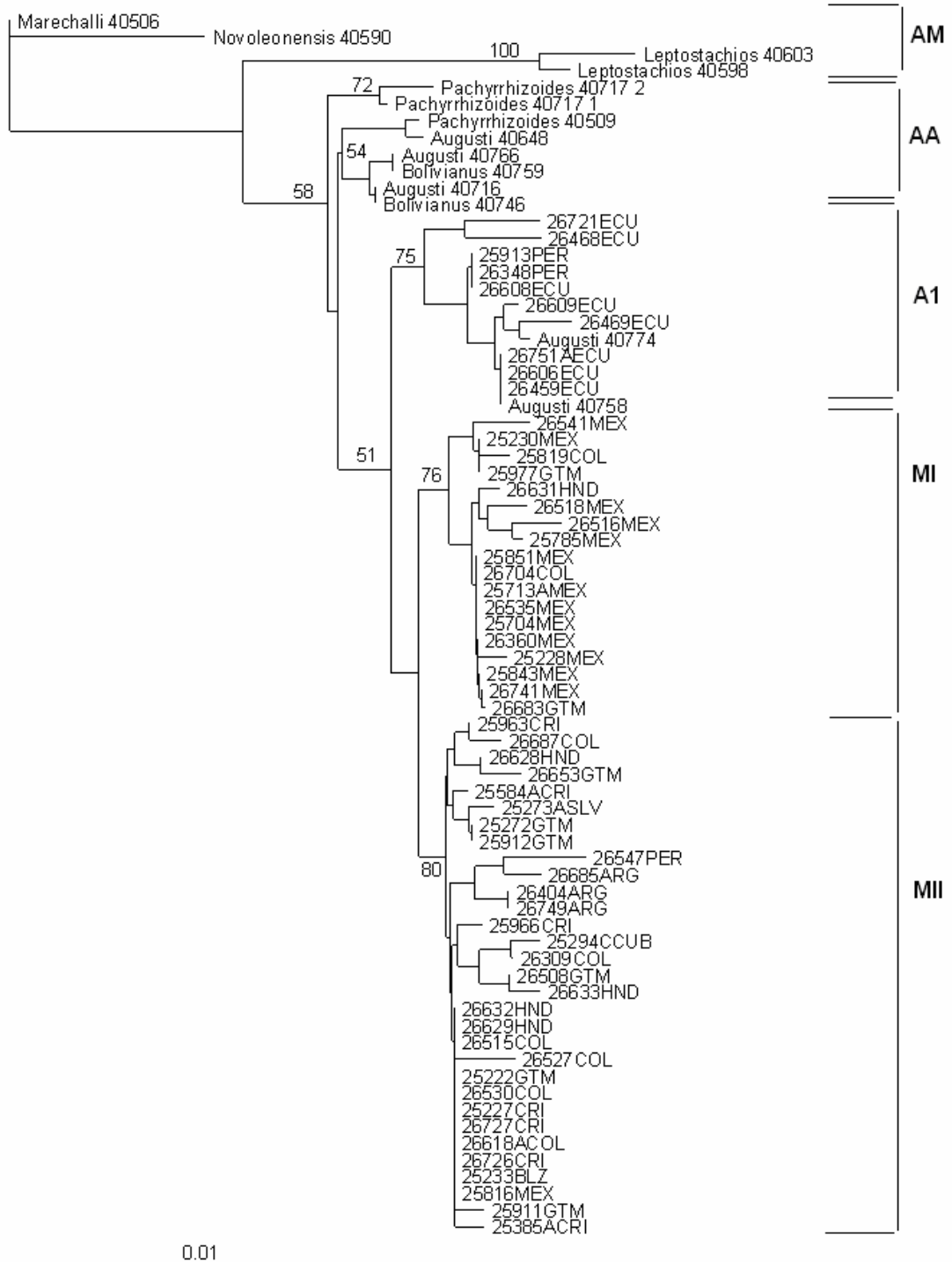


Figura 3. Topología NJ del ITS para *P. lunatus* silvestre (las entradas están indicadas con los códigos 25- y 26-, seguido de la inicial del país de origen). MEX=México, GTM=Guatemala, BLZ= Belice, HND= Honduras, SLV= El Salvador, CRI=Costa Rica, COL= Colombia, CUB= Cuba, ECU= Ecuador, PER= Perú, BOL= Bolivia, BRA= Brasil, ARG= Argentina. Grupos definidos dentro de fríjol lima silvestre AI, MI, MII. AA: Especies Aliadas Andinas, AM: Especies Aliadas Mesoamericanas. Los números sobre las ramas corresponden a los valores de bootstrap (%).

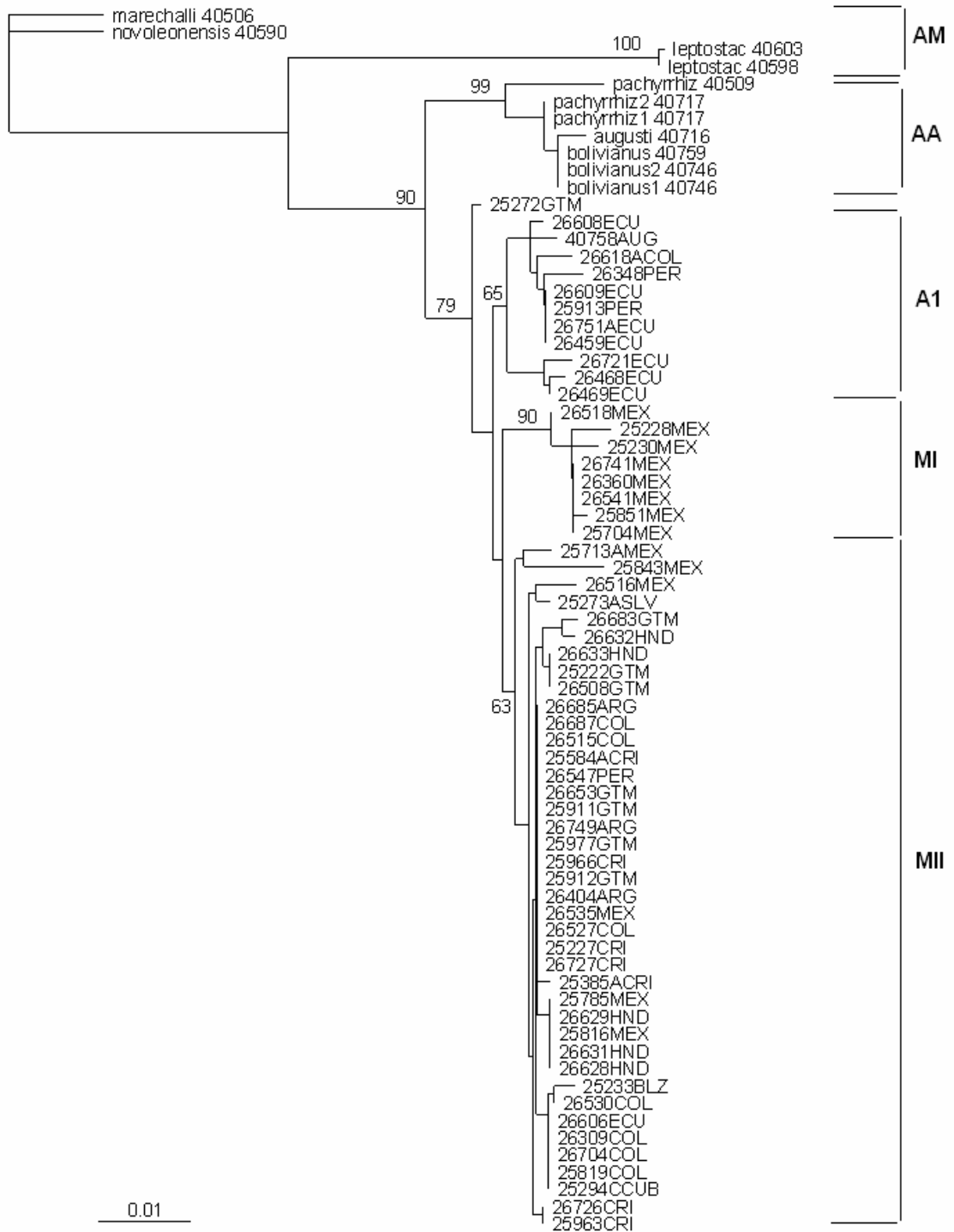


Figura 4. Distribución geográfica de *P. lunatus* silvestre en América. A, grupos definidos por NJ mediante marcadores cpDNA (ver figura 2). B, grupos definidos por NJ mediante marcadores ITS (ver figura 3).

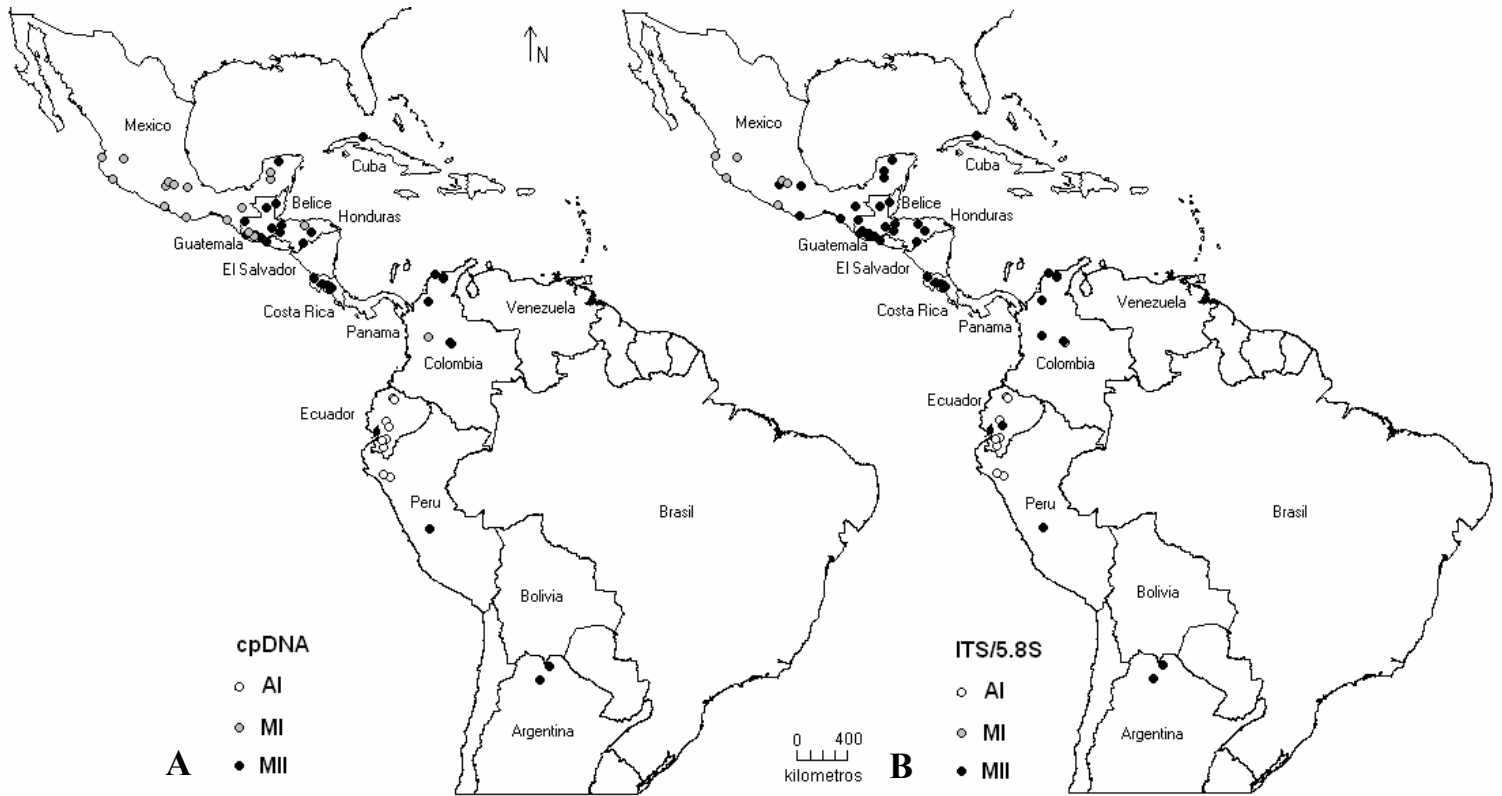


Figura 5. Red de haplotipos para cpADN de frijol lima silvestre. Haplotipos indicados en letras (A-W) encerrados en círculos o cuadrados cuyo tamaño es directamente proporcional a la frecuencia del haplotipo. Haplotipos intermedios no observados se muestran como puntos negros. Tres categorías de anidamiento presentes, clados de primer grado (1-1 al 1-15), segundo grado (2-1 al 2-6), tercer grado (3-1 al 3-3). AI, MI y MII según figura 2.

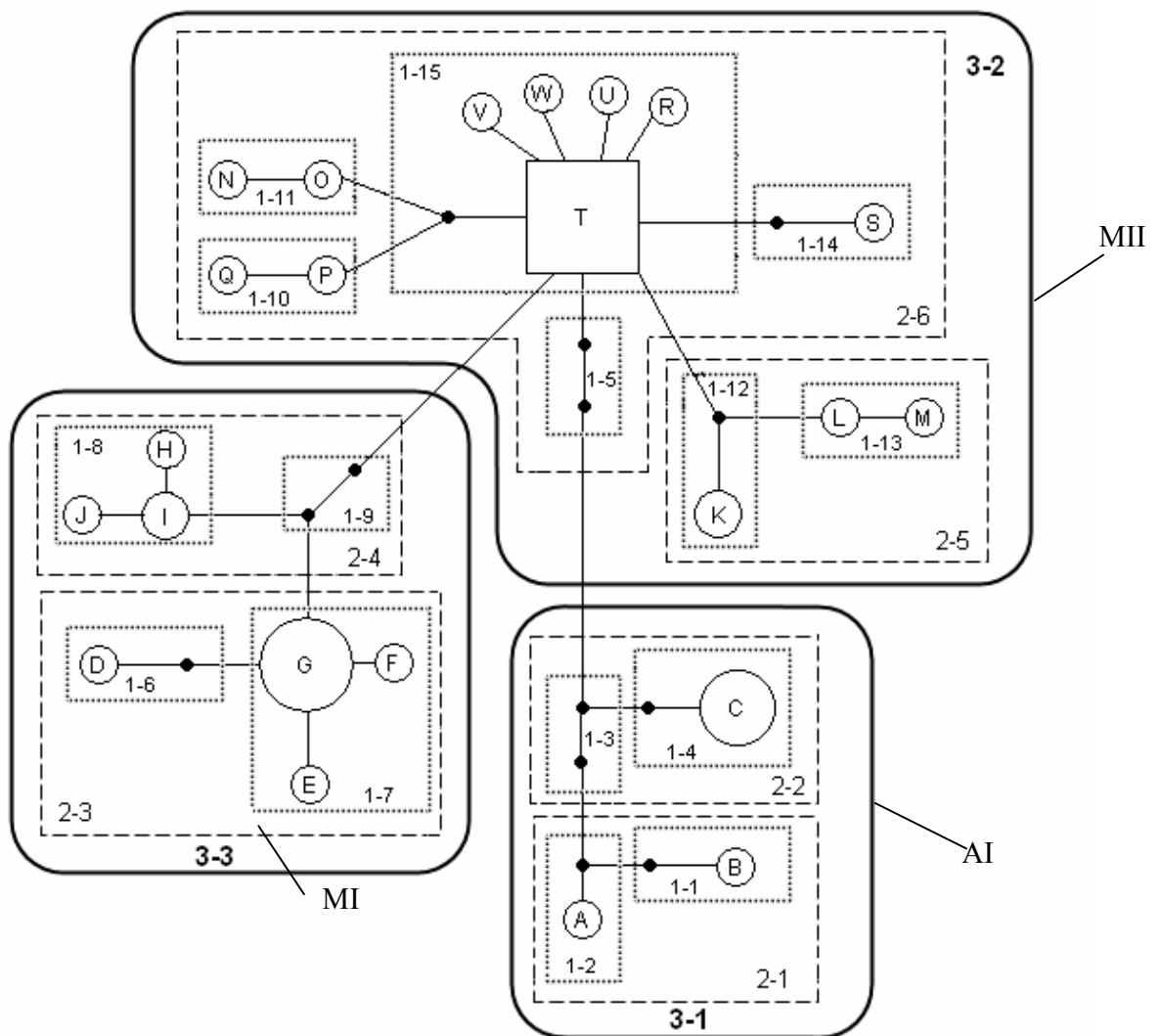


Figura 6. Red de haplotipos para ITS de fríjol lima silvestre. Haplotipos indicados en letras (A-B1) encerrados en círculos o cuadrados cuyo tamaño es directamente proporcional a la frecuencia del haplotipo. Haplotipos intermedios (no observados) se muestran como puntos negros. Cuatro categorías de anidamiento presentes, clados de primer grado (1-1 al 1-28), segundo grado (2-1 al 2-13), tercer grado (3-1 al 3-5), cuarto grado (4-1, 4-2). AI, MI y MII según figura 3.

