



UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales

Departamento de Ecología, Genética y Evolución

**ESTUDIOS DE RELACIONES FILOGENÉTICAS Y PATRONES
DE DIVERSIFICACIÓN EN EL GÉNERO *PROSOPIS* EN BASE A
MARCADORES MOLECULARES**

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires en el área: Ciencias Biológicas

SANTIAGO ANDRÉS CATALANO

Director de tesis: Dr. Juan César Vilardi
Dra. Beatriz Ofelia Saidman
Consejera de Estudios: Dra. Beatriz Ofelia Saidman

Lugar de trabajo: Departamento de Ecología Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Buenos Aires

Buenos Aires, 2009

Quisiera dedicar esta tesis a dos personas que hoy no están entre nosotros:

A Carlos Naranjo

A mi viejo

RESUMEN

El género *Prosopis* (Leguminosae) comprende 44 especies de árboles y arbustos que son importantes integrantes de ambientes áridos y semiáridos alrededor del mundo. Con el objeto de estudiar la diversificación y evolución de este género se realizó un estudio que combinó análisis de filogenia molecular, análisis de datación molecular y optimización de caracteres ambientales. Las relaciones filogenéticas se infirieron a partir de las secuencias de ADN de 5 marcadores moleculares (*matK-trnK*, *trnL-trnF*, *trnS-psbC*, *G3pdh*, *NIA*). El muestreo involucró 30 especies *Prosopis* que representaron todas las secciones y series del género además de estar también representada toda la distribución geográfica del género. Los resultados sugirieron que *Prosopis* no sería monofilético. Esto está apoyado por (a) la posición filogenética de *Xerocladia viridiramis* que apareció junto con las especies americanas de *Prosopis* (b) la posición filogenética de las especies de *Prosopis* del viejo mundo que en ningún análisis formaron un grupo monofilético con el resto de las especies de *Prosopis*. La sección Strombocarpa fue monofilética, no así la sección Algarobia ya que *P. argentina*, perteneciente a la sección Monilicarpa, apareció entre especies de la sección Algarobia. Las especies de las series Chilenses, Pallidae y Ruscifoliae (clado de los mesquites o de los algarrobos) formaron un grupo altamente apoyado, apareciendo como grupo hermano de éste un clado formado por el resto de las especies de la sección Algarobia más *P. argentina*. El análisis de datación molecular realizado mediante el método de Penalized Likelihood indicó que la divergencia entre las secciones Strombocarpa y Algarobia habría ocurrido en el Oligoceno, contrastando con una divergencia mucho más reciente (Mioceno Tardío) para los linajes dentro de cada uno de estos grupos. La diversificación del grupo de los algarrobos habría comenzado en el Plioceno, mostrando una alta tasa de diversificación. El momento de diversificación de los principales linajes de las especies americanas de *Prosopis* coincidiría con el desarrollo de ambientes áridos en América, sugiriendo un control ambiental de la diversificación del grupo. La optimización de parámetros ambientales sobre la filogenia obtenida de *Prosopis* sugirió una ocupación antigua de ambientes áridos por parte de los ancestros de las especies actuales. En esta tesis se han incluidos dos estudios complementarios a los previamente mencionados. Uno de ellos estuvo relacionado al análisis de la evolución de una inversión encontrada durante el proceso de secuenciación de la región del genoma de cloroplasto entre los genes *trnS* y *psbC*. En un segundo análisis se realizó una

evaluación crítica de los métodos utilizados para optimizar variables ambientales en un contexto filogenético. Dicho desarrollo teórico fue tomado en cuenta a la hora de realizar la optimización de caracteres ambientales en *Prosopis*

Palabras Claves: *Prosopis*, filogenia, diversificación, aridez, optimización

ABSTRACT

The genus *Prosopis* (Fabaceae) comprises 44 species of trees and shrubs that are important members of arid and semiarid environments around the world. To study *Prosopis* diversification and evolution, a combined approach including molecular phylogeny, molecular dating, and character optimization analysis was applied. Phylogenetic relationships were inferred from five different molecular markers (*matK*-*trnK*, *trnL-trnF*, *trnS-psbC*, G3pdh, NIA). Taxon sampling involved a total of 30 *Prosopis* species that represented all sections and series and the complete geographical range of the genus. The results suggest that *Prosopis* is not a natural group. This result is supported by (a) the phylogenetic position of *Xerocladia viridiramis*, together with American species of *Prosopis*, (b) the phylogenetic position of the Old world species of *Prosopis* that never formed a monophyletic groups with the rest of the species of *Prosopis*. Section *Stombocarpa* appeared as monophyletic, but not Section *Algarobia* since *P. argentina* of Section *Monilicarpa* appeared intermingled with species of the former Section. Species of series *Chilenses*, *Pallidae*, and *Ruscifoliae* (mesquite or algarrobos clade) formed a strongly supported group with a sister clade formed by the rest of the species of *Algarobia* plus *P. argentina*. Molecular dating analysis indicated that the divergence between section *Strombocarpa* and section *Algarobia* plus section *Monilicarpa* occurred in the Oligocene, contrasting with a much recent diversification (Late Miocene) within each of these groups. The diversification of the mesquite group was inferred to have started in the Pliocene, showing a high diversification rate. The moment of diversification within the major lineages of American species of *Prosopis* is coincident with the spreading of arid areas in the Americas, suggesting a climatic control for diversification of the group. Optimization of habitat parameters on the *Prosopis* phylogeny suggests an ancient occupation of arid environments by *Prosopis* species. Two complementary analyses were conducted in this thesis. One of them related to the study of the evolution of an inversion found during the sequencing process of the *trnS-psbC* region of chloroplast genome. The second analysis was associated with a critical evaluation of the methods used to optimize environmental variables in a phylogenetic context. This theoretical development was taken into account in the optimization of environmental characters in *Prosopis*.

Keywords: *Prosopis*, phylogeny, diversification, aridity, optimization

AGRADECIMIENTOS

A mis Directores Juan y Beatriz por abrirme las puertas de su laboratorio, dándome la oportunidad para realizar esta tesis y además por haberme acompañado en cada uno de mis caprichosos proyectos.

A mis amadas Ana y Juli por hacerme feliz día a día

A mi familia: mis hermanos Tin, Nics y Leo por haberme ayudado durante este tiempo; mi vieja por haberme dado, junto con mi viejo, la posibilidad de emprender una carrera universitaria, y por su constante interés sobre mis emprendimientos; al resto de la familia en Trelew y Mar del Plata que también ha dado su incondicional apoyo.

A la Dra. Viviana Confalonieri por su apoyo durante todos estos años.

A Dr. Pablo Goloboff por su constante enseñanza y asesoramiento en temas filogenéticos y de programación.

A Fer Biganzoli por haberme abierto las puertas de su casa una y mil veces haciéndome sentir siempre muy cómodo. Sin su ayuda habría sido muchísimo más difícil terminar con esta tesis.

A la Dra. Melissa Luckow por haberme facilitado secuencias no publicadas de *Prosopis* y *Xerocladia* además de haberme enviado ADN de *Xerocladia* para su amplificación y secuenciación.

A mis compañeras y amigas de Sistemática Teórica y de los cursos, Shirley y Marce por su compañía y por los buenos momentos que me hicieron pasar.

A las “chicas” del labo, que me bancaron en mi interminable trabajo de mesada: Noe, Nati, Caro.

A Pep y Javi por su amistad durante estos años en la facu

A mis compañeras del 58 Andre, Pauli, Ceci y Lau.

A aquellas que compartieron un rato con nosotros pero que después tomaron otros rumbos: Sara, Amy, Vero, Carly.

A Alexandra Gottlieb por haberme ayudado en el labo y en diferentes etapas de esta tesis.

A los chicos cladistas, Marcos Mirande, Marquiños Llanho, Boris, Pancho y Amelia, Nacho Escapa, Ceci Kopuchian y toda la comunidad cladistica Argentina.

A los “tucus” por hacerme sentir como en casa cada vez que fui “pal norte”

Al Dr. Iusem y la Dra. Laura Maskin por haber colaborado en el clonado del gen NIA

A Dani Tosto por ayudarme con los clonados

A los Doctores Alejandra Vilela, Oscar Ruiz, Pedro Steibel, Héctor Troiani, por la colaboración prestada en para la obtención de material biológico

A los curadores de los herbario CTE, SI, LIL y LP por permitirme analizar el material de herbario para el estudio de las distribuciones de las especies estudiadas.

A Emilio Ulibarri por su colaboración en el Instituto de Botánica Darwinion

Al Dr. Luis Scolaro por prestar amablemente la Savanth SpeedVac

Esta tesis fue realizada con fondos aportados por subsidios otorgados por la Agencia Nacional de Ciencia y Tecnología y la Universidad de Buenos Aires a la Dra Beatriz Saidman y por subsidios de CONICET otorgados al Dr. Juan Vilardi.

Durante el desarrollo de esta tesis el doctorando fue becario de la Universidad Nacional de La Plata y del CONICET.

INDICE

1.- INTRODUCCIÓN.....	1
1.1. Generalidades.....	2
1.2. Distribución geográfica.....	2
1.3. Historia taxonómica del género.....	5
1.4. Diversidad morfológica en <i>Prosopis</i>	7
1.5. Importancia económica y social del género.....	11
1.6. Estudios previos sobre evolución del género.....	14
1.7. Análisis filogenéticos previos.....	16
1.8. Ambientes ocupados por especies de <i>Prosopis</i>	18
1.9. Antecedentes de estudios de diversificación en zonas áridas y semiáridas.....	20
1.10. Optimización de caracteres ambientales.....	21
1.11. Objetivos.....	23
1.12. Bibliografía.....	24
2.- ANALISIS FILOGENÉTICO.....	30
2.1. Introducción.....	31
2.2. Materiales y Métodos.....	33
2.2.1. Muestreo de genes y especies	33
2.2.2. Características de los genes analizados.....	38
2.2.3. Técnicas moleculares.....	40
2.2.3.1. <i>Extracción de ADN</i>	40
2.2.3.2. <i>Amplificación y purificación</i>	41
2.2.3.3. <i>Clonado de fragmentos de ADN</i>	44
2.2.3.4. <i>Secuenciación</i>	46
2.2.4. Material biológico.....	46
2.2.3. Metodología del Análisis filogenético.....	47
2.2.3.1 <i>Homologías estáticas vs homologías dinámicas</i>	47
2.2.3.2. <i>Criterios de optimalidad en análisis filogenéticos</i>	48
2.2.3.3. <i>Búsquedas heurísticas en Parsimonía</i>	51
2.2.3.4. <i>Evaluación de apoyo de grupos</i>	53
2.2.3.5. <i>Análisis bayesianos</i>	54

2.3. Resultados.....	54
2.3.1. Características de las regiones estudiadas.....	54
2.3.2. Resultado de las búsquedas.....	57
2.3.2.1. <i>Análisis de dos marcadores</i>	57
2.3.2.2. <i>Análisis de tres marcadores</i>	60
2.3.2.3. <i>Análisis de cinco marcadores</i>	63
2.4. Bibliografía.....	66
3- ANÁLISIS DE INVERSIONES.....	70
3.1. Introducción.....	71
3.2. Materiales y Métodos.....	74
3.2.1. Optimización de inversiones.....	74
3.2.2. Comparación filogenética.....	78
3.2.3. Optimización de energía libre y largo de tallo.....	80
3.2.4. Conjunto de datos.....	83
3.3. Resultados.....	83
3.3.1. La inversión en <i>Prosopis</i>	83
3.3.2. La inversión en Angiospermas.....	84
3.3.2.1. <i>Caracterización de la estructura de tallo y lazo</i>	84
3.3.2.2. <i>Optimización de inversiones, energía libre y largo de tallo en Angiospermas</i>	85
3.3.2.3. <i>Relación entre la ocurrencia de inversiones y los valores de energía libre ó largo de RI</i>	88
3.3.2.3.1. <u>Energía libre</u>	88
3.3.2.3.2. <u>Largo del tallo</u>	90
3.4. Discusión.....	92
3.4.1. Estabilidad de la estructura de tallo-lazo y los eventos de inversión.....	93
3.4.2. Asignaciones ancestrales y homología.....	96
3.5. Bibliografía.....	98
4- TIEMPOS DE DIVERGENCIA Y DIVERSIFICACIÓN.....	102
4.1. Introducción.....	103
4.1.1 Tiempos de divergencia.....	103
4.1.2. Diversificación en zonas áridas.....	105
4.2. Materiales y Métodos.....	107

4.2.1. Cálculos de tiempos de divergencia en <i>Prosopis</i>	107
4.2.2. Análisis de diversificación.....	108
4.3. Resultados.....	110
4.3.1. Datación Molecular.....	110
4.3.2. Análisis de diversificación.....	112
4.4. Bibliografía.....	115
5-. OPTIMIZACIÓN DE CARACTERES AMBIENTALES.....	119
5.1. Optimización de condiciones ambientales: análisis crítico de la teoría subyacente a dichos estudios.....	120
5.1.1 Introducción	120
5.1.2. Caracteres ecológicos, nicho y herencia	121
5.1.3. Teoría de nicho y herencia	123
5.1.4. Cuando el Nicho Realizado es diferente al Nicho Potencial	127
5.1.5. Variables: más problemas en la estimación del nicho ancestral....	129
5.1.6. Problemas en los Test propuestos para evaluar la heredabilidad de nicho.....	131
5.1.7. Posibles correcciones o mejoras a la metodología de estimación de condiciones ambientales ancestrales.....	132
5.1.8. Consideraciones finales.....	133
5.2. Análisis de condiciones ambientales ancestrales en <i>Prosopis</i>.....	134
5.2.1. Introducción.....	134
5.2.2. Resultados.....	138
5.2.3. Evaluación de los resultados en función de las críticas hechas en el Inciso 5.1.....	141
5.2.3.1. <i>Variable analizada e interacción con otras variables</i>	141
5.2.3.2. <i>Condiciones ambientales ancestrales</i>	142
5.2.3.3. <i>Nicho Realizado vs. Nicho Potencial</i>	143
5.3. Bibliografía.....	145
6-. DISCUSIÓN Y CONCLUSIONES.....	148
6.1. Relaciones filogenéticas	149
6.2. Genes utilizados, ventajas y desventajas.....	158
6.3. Diversificación, tiempos de divergencia y restos fósiles.....	161
6.4. Diversificación del grupo de los algarrobos.....	164
6.5. Hibridación y Evolución de mecanismos	

de aislamiento en Algarobia.....	166
6.5.1. Barreras de Aislamientos Precigóticos.....	168
6.5.2. Barreras de Aislamientos Postcigóticos.....	169
6.6. Un escenario posible para la evolución de <i>Prosopis</i> en América.....	170
6.6.1. Evidencias obtenidas en esta tesis.....	170
6.6.2. Diversificación de taxones en zonas áridas: evidencias para un patrón común de diversificación.....	173
6.7. Conclusiones.....	180
6.8. Bibliografía.....	181
ANEXOS.....	189
Anexo I. Scripts usados en el análisis de inversiones	190
Anexo II. Lista de especies incluidas en el análisis de inversiones.....	211
Anexo III. Scripts usados en el análisis de caracteres ambientales.....	214

CAPÍTULO 1- INTRODUCCIÓN

1.1. Generalidades

El género *Prosopis* L. (Fabaceae-Mimosoideae) comprende alrededor de 45 especies que son elementos importantes en ecosistemas áridos y semiáridos alrededor del mundo (Figura 1). La mayoría de sus miembros son especies arbóreas o arbustivas, existiendo además algunas formas rastreras. Muchas de estas especies presentan una gran importancia económica, principalmente como fuente de madera para muebles y leña aunque también con un sinnúmero de utilidades secundarias. La importancia económica de estas especies ha atraído el interés de científicos desde hace mucho tiempo. Tal interés se ve claramente representado en la publicación de libros completos referidos a especies de este género (Simpson, 1977; Habit, 1985; Pasiecznik et al., 2001).

Las cualidades productivas que presentan muchas de las especies de *Prosopis* junto con la capacidad para sobrevivir y prosperar en suelos pobres y bajo condiciones de sequía hacen de ellas elementos ideales para su inclusión en programas de forestación y en sistemas agrosilvopastoriles en zonas áridas y semiáridas (Blasco et al., 2005). La explotación irracional que ha sufrido este recurso ha llevado a la desaparición de parte de los bosques naturales de estas especies (Hunziker et al., 1986). La situación actual de este recurso hace que sea necesario realizar estudios tendientes a aumentar el conocimiento de sus características biológicas para poder tomar decisiones acertadas para la conservación y explotación sustentable del mismo.

1.2. Distribución geográfica

La distribución de *Prosopis* está correlacionada generalmente con ambientes áridos. Algunas especies de este género son características de las regiones más áridas del planeta. Por ejemplo, *P. tamarugo* es una de las pocas especies arbóreas capaces de sobrevivir en el extremadamente árido desierto de Atacama, en el Norte de Chile. Otras especies del género habitan los grandes desiertos del Norte de México y Sur de EEUU (Sonora, Tamaulipas, Chihuahua) mientras que en Sudamérica son importantes miembros de diferentes

comunidades de regiones áridas y semiáridas (Chaco, Monte, Patagonia y Puna). Sin embargo algunos de sus representantes también están presentes en regiones más húmedas. Tal es el caso de *P. africana* que habita ambientes subtropicales ubicados al sur del desierto de Sahara. Las especies del género *Prosopis* se hallan representadas en diversas zonas de Argentina, desde Santa Cruz al Sur hasta Jujuy al Norte y desde Buenos Aires al Este hasta los pies de la cordillera de los Andes al Oeste. Sus representantes se encuentran en las provincias biogeográficas de: Chaco, Monte, Espinal, Patagonia y Puna siendo muy características de las primeras dos provincias. Para una descripción de las provincias biogeográficas puede consultarse a Cabrera & Willink (1973).

La distribución geográfica original del género ha sido fuertemente modificada en los últimos siglos debido a la actividad humana. Algunas de sus especies disminuyeron su distribución de la mano de la destrucción de bosques nativos y la sobreexplotación de este recurso forestal. En Argentina sólo hasta el año 1975 se había perdido la mitad de la superficie original ocupada por *Prosopis* (D'Antoni & Solbrig, 1977). Un segundo punto que ha conducido a cambios en la distribución de las especies ha sido la generación de ambientes degradados, que han resultado ser propicios para expansión de algunas especies de *Prosopis*. Tal es el caso de *P. ruscifolia* en el Chaco argentino (Astrada & Adámoli, 2005)

La utilización de este género con fines productivos llevó a su introducción en regiones donde no había especies nativas del mismo, tales como el sur de África y Australia. En estos lugares es frecuente que las especies de *Prosopis* presenten una característica fuertemente invasiva llegando en algunos casos a ser considerada una verdadera plaga (Shiferaw et al., 2004; Tiwari, 1999).

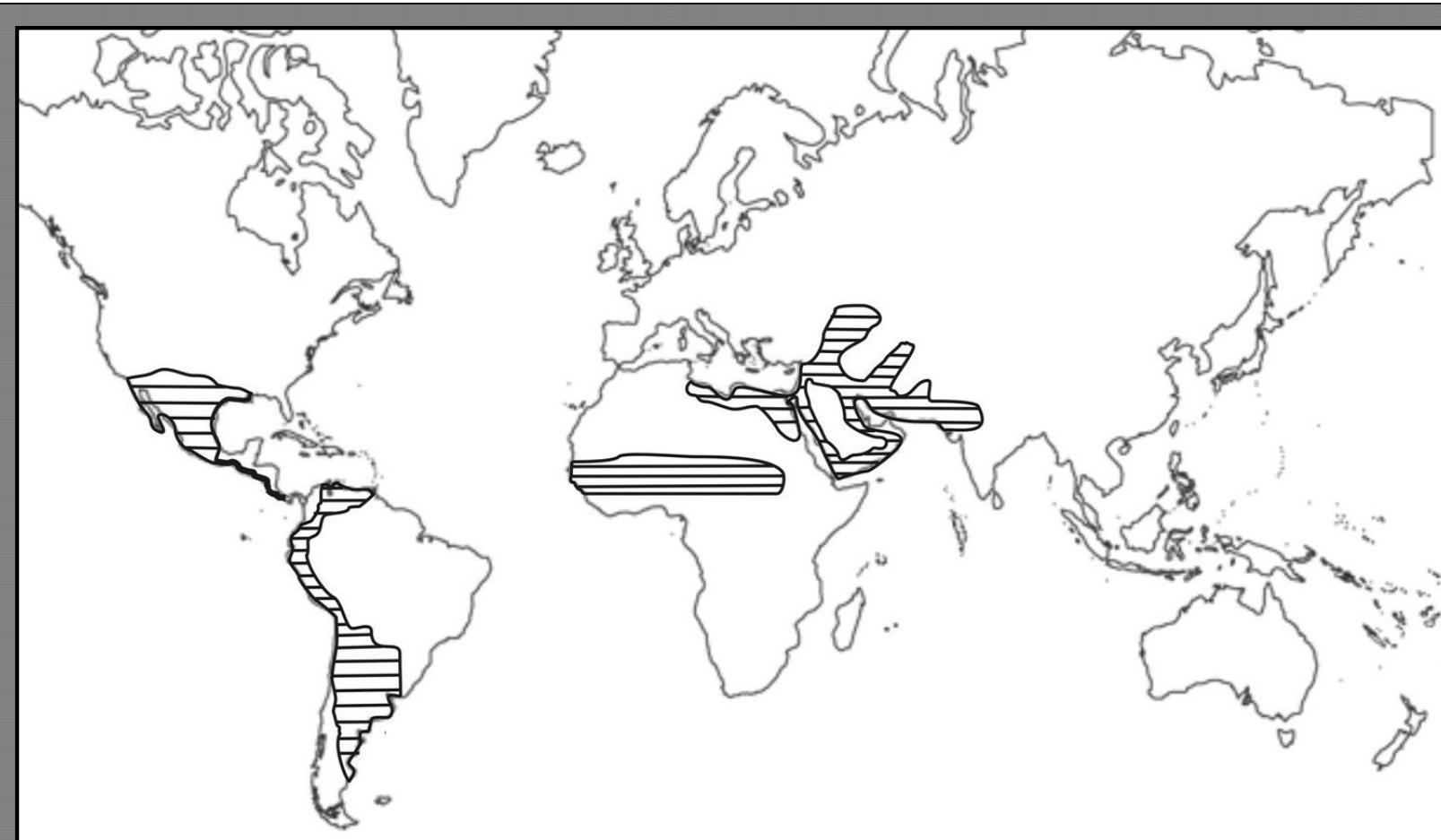


Figura 1. Distribución natural de *Prosopis* según Pasiecznik et al. (2001)

1.3. Historia taxonómica del género

Este apartado está basado principalmente en los desarrollos de Burkart (1976) y Pasiecznik et al. (2001). La historia taxonómica del género *Prosopis* es compleja y extensa. El nombre *Prosopis* fue originalmente utilizado por Linneo (1767) para describir una especie asiática (*P. spicigera*) la cual es actualmente sinónima de *P. cineraria*, la especie tipo del género (Burkart, 1976). La segunda especie descripta dentro del género fue *Ceratonia chilensis*, que corresponde a la especie actualmente válida *P. chilensis*, mientras que la tercera especie descripta fue *Mimosa juliflora* Swartz, recién incorporada a *Prosopis* por De Candolle en 1825. La primera especie de la sección *Strombocarpa* (*sensu* Burkart, 1976) fue descripta por Lamarck (1789) como *Mimosa strumbulifera*.

Los límites y las subdivisiones actuales dentro del género fueron fundados por De Candolle y Bentham en el siglo XIX. De Candolle (1825) incorporó la sección *Algarobia* con 13 especies pero no reconoció la sección *Strombocarpa*, que fue definida posteriormente por Bentham (1842). Al incorporar las especies americanas, las especies asiáticas del género fueron asignadas a una nueva sección *Adenopsis* que más tarde fue llamada sección *Prosopis* por Burkart (1976).

Posteriormente se fundaron nuevos géneros con especies que actualmente son asignadas a *Prosopis*. Rafinesque Schmaltz (1838) creó el género *Neltuma* con *P. juliflora* como tipo. Bentham (1839) consideró que *Prosopis* sección *Algarobia* DC merecía un status genérico. Según Burkart (1976), Bentham habría basado tal conclusión en la descripción parcialmente incorrecta que De Candolle (1825) hizo del grupo. Más tarde Bentham (1842) aceptó el nombre *Prosopis* con límites bastante cercanos a los que luego consideraría Burkart (1976) en su monografía del género. Posteriormente Bentham (1875) consideró que todas las especies actualmente asignadas a la sección *Algarobia* serían en realidad una sola especie politípica (*P. juliflora*).

En tiempos más recientes Britton & Rose (1928) separaron a *Prosopis* en 3 géneros: *Soprosopis*, *Strombocarpa* y *Neltuma* (Algarobia). Según Burkart (1976) la sección *Soprosopis* (que incluye únicamente a *P. palmeri*) sería sólo una “forma extraordinaria de la sección *Strombocarpa*”. Burkart (1976) ataca tal separación intragenérica planteando que las diferencias entre las secciones *Prosopis*, *Algarobia* y *Strombocarpa* son sólo de caracteres vegetativos, mereciendo entonces sólo el rango subgenérico. Burkart (1976) además consideró que la sección *Lomentaria* definida por Britton & Rose (1928) debía ser suprimida ya las especies que la integran no corresponderían al género. De hecho éstas se consideran actualmente pertenecientes al género *Prosopidastrum* Burkart.

Finalmente Burkart (1976) separó al género en 5 secciones basándose principalmente en el tipo de espinas y la estructura del tallo. La sección **Algarobia** comprende la mayor parte de las especies del género y es estrictamente americana. Sus representantes se hallan tanto en Sudamérica como en Norteamérica y se caracterizan por tener espinas axilares, y ramas diferenciadas en braquiblastos y macroblastos aunque algunas de sus especies carecen de espinas. La sección **Strombocarpa**, también presenta especies distribuidas por Norteamérica y Sudamérica y se caracteriza por presentar espinas foliares. La sección **Anonychium** presenta una sola especie, *P. africana* que carece de espinas y habita la zona Sudano-Guineana y áreas vecinas de África, desde Senegal en el Oeste hasta Sudan y Kenya en el este (Pasiecznik et al., 2001). La sección **Prosopis** comprende 3 especies que son nativas del Norte de África y Asia, llegando al Este hasta la India, al Norte hasta Georgia y Turkmenistan y al Oeste hasta Argelia. Sus especies se caracterizan por tener espinas caulinares. La sección **Monilicarpa** presenta sólo una especie, *P. argentina*, que carece de braquiblastos, tiene un fruto moniliforme y presenta espinas axilares y apicales, uninodales y solitarias. Las secciones *Algarobia* y *Strombocarpa* fueron a su vez divididas en series. La primera presenta seis series: *Ruscifoliae*, *Sericanthaes*, *Denudantes*, *Humiles*, *Pallidae*, *Chilenses*. Por su parte la sección *Strombulifera* presenta las series *Cavenicarpeae* y *Strombocarpeae* (Tabla 1).

Durante esta tesis seguiremos la nomenclatura derivada de la monografía de Burkart (1976) puesto que los cambios surgidos con posterioridad o bien son sólo menores o no cuentan con consenso entre los especialistas. Uno los cambios propuestos ha sido separar a *Prosopis* en tres subgéneros basados principalmente las características polínicas (Guinet & Bessedik, 1984). El Subgénero *Neoprosopis* correspondería a todas las especies americanas, incluyendo a las secciones *Strombocarpa*, *Algarobia* y *Monilicarpa* sensu Burkart (1976). El segundo subgénero correspondería a las especies incluidas por Burkart (1976) en la sección *Prosopis* mientras que el subgénero restante sólo incluiría a *P. africana* (sección *Anonychium* sensu Burkart 1976).

La labor taxonómica dentro de la sección *Algarobia* ha sido dificultosa desde su propia definición. Los principales conflictos se relacionan con el grupo de los mesquites en Norteamérica y los algarrobos en Sudamérica. La presencia de morfologías intermedias sería el principal obstáculo para la delimitación específica. Con posterioridad a la monografía de Burkart (1976) fueron descriptas nuevas especies (ej. Palacios, 2006; Schinini, 1981). Los estudios de marcadores genéticos han aportado amplia información al respecto, y han indicado que la mayoría de las especies reconocidas morfológicamente también se diferencian genéticamente aún en casos en las que las poblaciones de diferentes especies sean simpátricas (Bessegå et al., 2000a).

Tabla 1. Clasificación infragenérica de *Prosopis* según Burkart (1976)

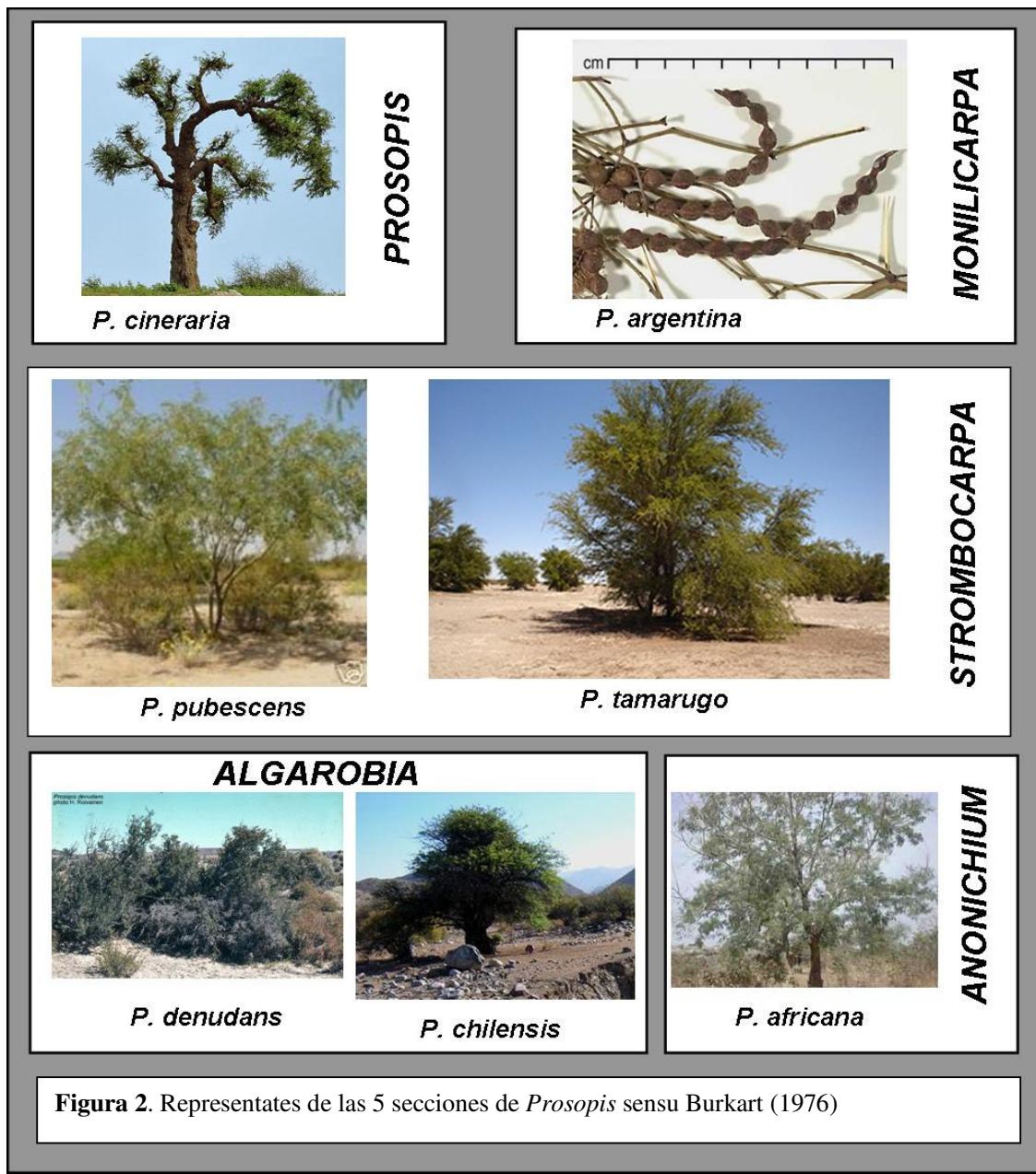
sección	serie	Especie	sección	serie	Especie
Algarobia	Chilenses	<i>P. chilensis</i>		Ruscifoliae	<i>P. ruscifolia</i>
		<i>P. juliflora</i>			<i>P. fiebrigii</i>
		<i>P. nigra</i>			<i>P. vinalillo</i>
		<i>P. caldenia</i>			<i>P. hassleri</i>
		<i>P. laevigata</i>		Sericantheae	<i>P. sericantha</i>
		<i>P. flexuosa</i>			<i>P. kuntzei</i>
		<i>P. glandulosa</i>		Humiles	<i>P. humilis</i>
		<i>P. alpataco</i>			<i>P. rojasiana</i>
		<i>P. alba</i>	Strombocarpa	Strombocarpae	<i>P. palmeri</i>
		<i>P. velutina</i>			<i>P. strombocarpa</i>
		<i>P. pugionata</i>			<i>P. reptans</i>
Pallidae		<i>P. rubriflora</i>			<i>P. abbreviata</i>
		<i>P. campestris</i>			<i>P. torquata</i>
		<i>P. pallida</i>			<i>P. pubescens</i>
		<i>P. affinis</i>			<i>P. burkartii</i>
		<i>P. articulata</i>		Cavenicarpae	<i>P. ferox</i>
		<i>P. elata</i>			<i>P. tamarugo</i>
		<i>P. tamaulipana</i>	Monilicarpa		<i>P. argentina</i>
Denudantes	<i>P. denudans</i>		Prosopis		<i>P. cineraria</i>
		<i>P. ruizleali</i>			<i>P. farcta</i>
		<i>P. castellanosii</i>			<i>P. koelziana</i>
		<i>P. calingastana</i>	Anonichium		<i>P. africana</i>

1.4. Diversidad morfológica en *Prosopis*

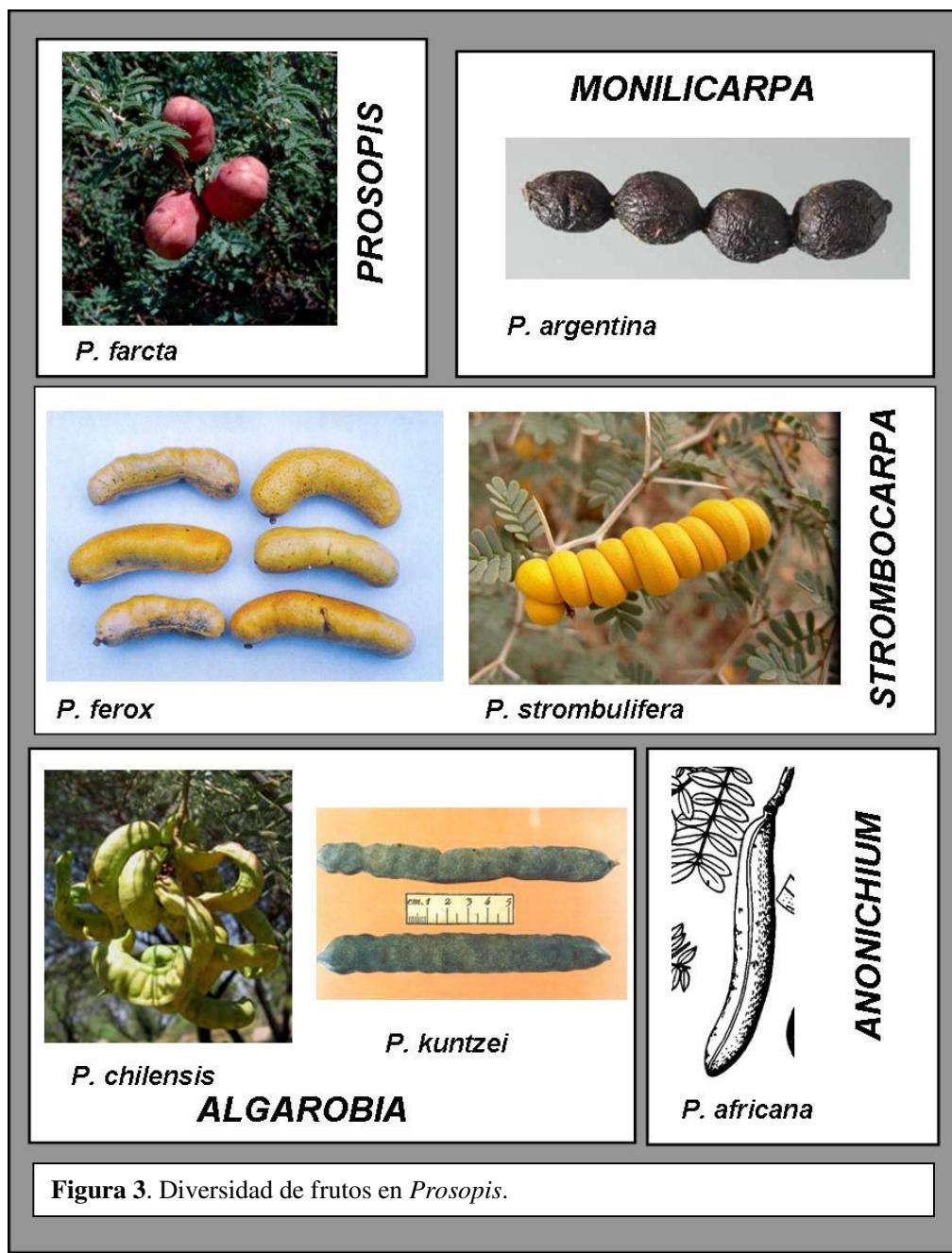
Las especies de *Prosopis* presentan una amplia diversidad de morfologías y hábitos. A pesar que las más conocidas del género son árboles (Figura 2), existen también formas decumbentes y rastreras. La mayoría de las especies arbóreas tienen portes que van de los 8 a 12 metros de altura, alcanzando en algunos casos alrededor de 20 metros (*P. africana* y *P. tamarugo*). Entre las de menor porte existe un conjunto de especies arbustivas tales como *P. elata*, *P. alpataco*, *P. strombulifera* y *P. argentina*. Otras especies son rastreras y apenas se levantan del suelo: *P. campestris*, *P. sericantha*.

Una de las características distintivas del género, es la presencia de un tipo de fruto particular (Figura 3): el lomento drupácea (Burkart, 1976). Éste es una legumbre indehiscente modificada y que se segmenta en fragmentos con una sola semilla cada uno. En general el mesocarpo es carnoso y el endocarpo es fibroso. El exocarpo varía tanto en color como en

grosor. Los frutos varían ampliamente tanto en tamaño como en forma. Algunas especies presentan frutos espiralados, condición que alcanza su mayor expresión en la serie Strombocarpae. *P. denudans* también presenta frutos espiralados aunque el espiral es mucho menos cerrado. Los frutos de mayor tamaño se encuentran en la sección Algarobia, en la que algunas especies presentan frutos aproximadamente lineales de más de 40 cm. En algunas especies los frutos maduros se separan en artejos individuales. Tal es el caso de *P. calingastana*.



Las hojas de *Prosopis* son compuestas, difiriendo en el grado de pinnado de mono a tetrapinnnada. Algunas especies, (ej. *P. denudans*) presentan unas hojas muy reducidas tanto en número como en tamaño de foliolulos. En otros casos los folioluos son extremadamente grandes llegando a tener hasta 2 cm. de ancho y 5 cm. de largo. (*P. ruscifolia*). Existen especies que pierden parcial o totalmente sus hojas, presentando en estos casos tallos fotosintéticos (*P. kuntzei*, *P. sericantha*, *P. castellanosii*, *P. humilis*).



Las flores son muy similares en todo el grupo, en general comparten las siguientes características: pentámeras, actinomorfas, decandras, hermafroditas, con anteras con glándulas. Las mismas se agrupan formando inflorescencias que en general presentan una forma racimos alargado.

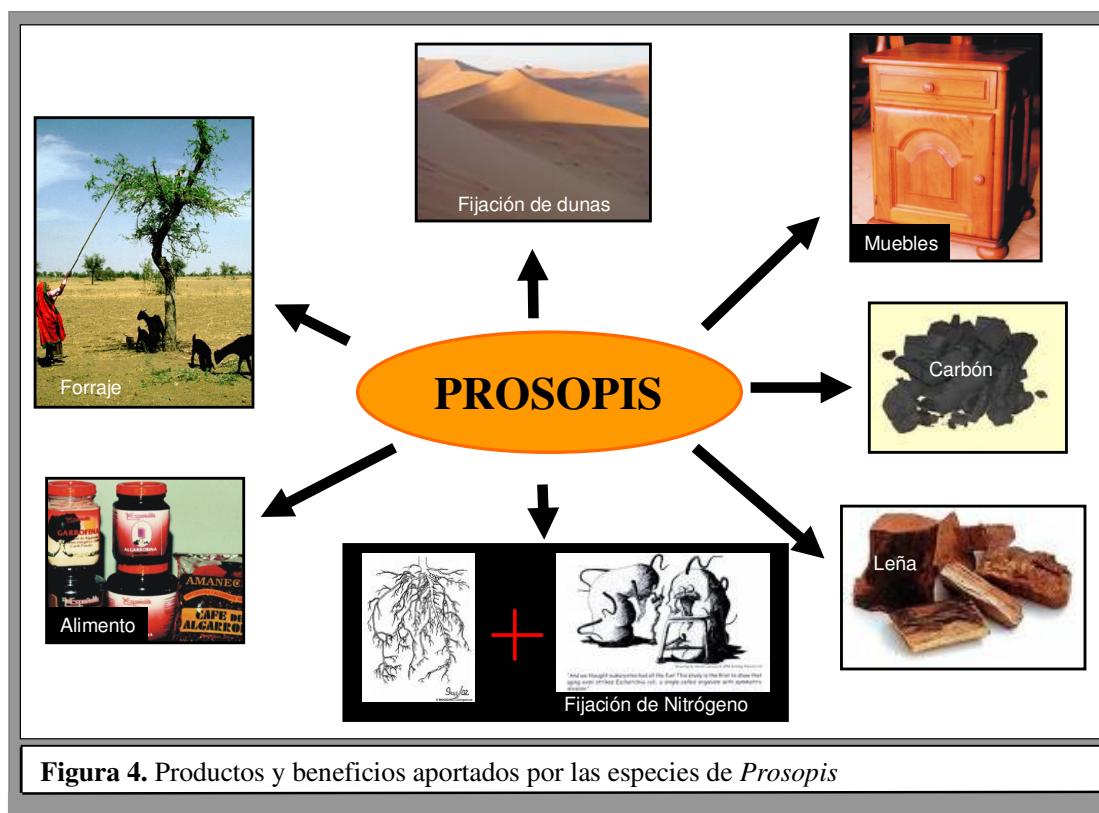
1.5. Importancia socio-económica del género

Muchas de las especies del género *Prosopis*, especialmente las pertenecientes a la sección Algarobia, presentan una gran importancia económica debido principalmente a la posibilidad de extraer múltiples beneficios de las mismas (Figura 4). La utilización de productos derivados de especies de este género por parte del hombre tiene una extensa historia, existiendo sitios arqueológicos de más de 6000 años de antigüedad que atestiguan tal hecho (Newton, 2005; Rodríguez, 2005). Los pueblos originarios americanos han continuado haciendo uso de estas especies hasta el presente. La importancia de las mismas en su vida cotidiana se ve demostrado en el modo los quichuas se refieren al mismo: lo llaman *tacu*, *taco* ó *tacco*, que traducido sería “El árbol”. La importancia de los mismos se mantiene hasta el presente en áreas rurales (ej. Ladio & Lozada, en prensa)

Ya desde un punto de vista productivo las principales características que presentan las especies de este género son: el alto valor maderero; la producción de frutos (vainas o chauchas) de utilidad como alimento humano y forraje; y la producción de diversos productos secundarios tales como goma arábiga, taninos, tintes y fibras (Pasiecznik et al., 2001). La madera es el principal producto comercial que puede obtenerse de las especies de este género ya sea para combustible como para muebles, vigas y postes. Una de las propiedades más características de la madera de *Prosopis* es su duramen pesado, durable y fuerte (Pometti et al., en prensa). Tanto la madera como el carbón producido a partir de ella presentan altos niveles calóricos y no producen ni fuertes olores ni chispas (Pasiecznik et al., 2001). Una ventaja adicional de esta madera para su uso como combustible es que la misma se puede

utilizar tanto seca como verde. Los troncos de estas especies son muy útiles para la producción de postes, pilares y marcos de puertas. Las tablas son utilizadas principalmente para la para la confección de muebles.

Las vainas producidas por especies de *Prosopis* tienen un alto valor alimenticio debido a los importantes niveles de azúcares, carbohidratos y proteínas que presentan (Felker et al., 2003). A diferencia de otras especies de árboles, el volumen de producción de vainas en estas especies es particularmente alto. En el caso de *P. nigra* el rinde promedio por árbol es cercano a los 40 Kg./año (Ing. Mauricio Ewen comunicación personal). Otros productos aprovechables son ceras, gomas y taninos. Las gomas son exudadas al producirse una herida en el tronco. La calidad de las gomas producidas por especies de *Prosopis* es muy alta, similar a la producida por *Acacia senegal*, una especie que es tomada como parámetro de alta calidad de gomas (Pasiecznik et al., 2001). Los taninos son extraídos de las raíces, madera y corteza. Los principales usos son como agente de tintura y para curtiembre.



Sumados a los diversos productos que pueden aprovecharse de estos árboles, algunas de las especies presentan ventajas adicionales como lo son la capacidad de fijar dunas y de mejorar la fertilidad del mismo colaborando con la fijación de nitrógeno atmosférico mediante sus simbiosis con bacterias del género *Rhizobium* (Felker & Clark, 1980; Torres, 1985). Otra utilidad recientemente señalada es que algunas de sus especies serían útiles para la extracción de metales pesados de suelos contaminados (Senthilkumar et al., 2005).

Otras especies del género, por contrapartida, son consideradas plaga debido a su enorme capacidad de invasión y rápido crecimiento (Tiwari, 1999). En la mayoría de los casos el comportamiento invasivo se ve evidenciado en regiones geográficas donde las especies no son nativas: Australia (Robinson et al., 2008); Sudáfrica (Zimmermann, 1991); Etiopía (Berhanu & Tesfaye, 2006); Brasil (Pegado et al., 2006) Estados Árabes Unidos (El-Keblawy & Al-Rawai, 2005) e India (Tiwari, 1999) entre otros. Shiferaw et al. (2004) sugirieron algunas de las características biológicas que producirían el comportamiento invasivo de *Prosopis*. Entre ellas se destacan: 1-. producción de alto número de semillas pequeñas que sobreviven en el banco de semillas en condiciones muy extremas y además sobreviven al pasaje por el aparato digestivo de los animales 2-. Un fruto atractivo para los dispersores animales 3-. Presencia de importante material de reserva en las semillas 4-. Gran capacidad de regeneración a partir de árboles dañados o podados.

La especie invasiva por excelencia en Argentina es *P. ruscifolia* (Astrada & Adámoli, 2005) que ha tenido un excepcional aumento en su distribución en los últimos siglos asociado a los cambios ambientales provocados por el hombre. Dichos cambios llevaron al colmataje de los cauces, la salinización de los suelos y el aumento de la frecuencia de incendios, todas modificaciones que llevaron a aumentar de manera apreciable la distribución de sitios propicios para su establecimiento (Roig, 1993). Una de las caracterizaciones más radicales de esta especie fue dada por (Bernardi, 1984). Según este autor *P. ruscifolia* es “...un monstruo de ferocidad vegetal, con espinas “mata-venado” de hasta 30 cm. de largo...”. Además el

autor no dudó en calificarla como “... *este arbusto samurai, pinchador de ruedas como pocos...* “.

1.6. Estudios previos sobre evolución del género

El género *Prosopis* ha sido objeto de un gran numero de estudios citogenéticos, genético-poblacionales y morfológicos que han colaborado a aumentar la comprensión de los procesos y patrones evolutivos dentro del grupo. Burkart (1976) considera que *Prosopis* es un género antiguo dentro de las Mimosoideae y basa su interpretación en caracteres morfológicos que serían primitivos para la subfamilia: granos simples, pétalos principalmente libres y la condición decandra. Además propone que el género se habría originado en África, ya que es ahí donde habita la especie que, según su criterio, sería la menos especializada (*P. africana*). Burkart (1976), basándose en la teoría de Wegener (1912) de deriva continental, considera que el género *Prosopis* se habría dispersado hacia el Este y el Oeste. La dispersión hacia lo que actualmente es América habría sido posible debido a que en la época en que Burkart (1976) sitúa la divergencia del género (Mesozoico-Terciario temprano), todavía existían conexiones entre África y América. La dispersión geográfica habría estado acompañada de una adaptación creciente a zonas áridas. Ramírez et al. (1999) aceptan la idea de Burkart (1976), y sitúan al origen del género hace 130 millones de años. Con respecto a la diversificación de *Prosopis* en América Burkart (1976) considera que una posibilidad es la existencia de una antigua flora desértica común a Sudamérica y Norteamérica y que recién a posteriori tal distribución se habría dividido en los dos centros de diversificación: Mexicano-Texano y Argentino-Paraguayo-Chileno. Roig (1993) considera que el grupo de especies Argentinas se habría originado en la región del Chaco y habría ocupado posteriormente zonas mas xéricas hasta alcanzar regiones como las que hoy son Atacama y Patagonia.

Los primeros estudios citogenéticos mostraron que las especies de *Prosopis* tendrían un numero cromosómico somático $2n= 28$. Aunque hay varios registros de poliploidía

(Cherubini, 1954; Hunziker et al., 1975; Burkart, 1976; Bukhari, 1997) recientemente se confirmó que *P. juliflora* sería la única especie descripta como tetraploide ($2N= 56$) hasta el momento (Trenchard et al., 2008).

La mayor parte de los estudios evolutivos en especies de *Prosopis* han estado centrados en estudios genético-poblacionales. Para tales estudios se han utilizado un sinnúmero de técnicas que arrojaron en general resultados concordantes. Entre ellas se incluyen datos de cromoatografía (Naranjo et al., 1984; Palacios & Bravo, 1984; Zallochi et al., 1990), proteínas seminales (Burghardt, 1995; Burghardt & Palacios, 1997), isoenzimas (Solbrig & Bawa, 1975; Saidman, 1985; Verga, 1995; Saidman & Vilardi, 1987; Bessega et al., 2000a; Julio, 2000), estudios polimorfismos de restricción de sitios de ADN (RFLP) (Saidman et al., 1995) y polimorfismos de loci amplificados al azar (RADP) (Bessega et al., 2000b; Juárez-Muñoz et al., 2002; Ferreyra et al., 2004; Vega & Hernández, 2005). La mayor parte de estos estudios estuvieron centrados en especies de la sección Algarobia, casi exclusivamente en las series Chilenses, Pallidae y Ruscifoliae. La principal conclusión a la cual se ha llegado es que las especies de estas series presentan una composición genética muy similar. Por otra parte se ha observado que en las mismas la mayor parte de la variabilidad genética se encuentra dentro de las poblaciones y la que la divergencia entre poblaciones de la misma especie es baja. A pesar de (a) la alta similitud interespecífica (b) que las especies se hallan en simpatría (c) de la existencia de híbridos interespecíficos (Palacios & Bravo, 1981; Naranjo et al., 1984; Hunziker et al., 1986), las poblaciones de una misma especie se agrupan generalmente de acuerdo a su estatus específico. Parte de los estudios genético-poblacionales han estado centrados en evaluar la diversidad genética a lo largo de una gran parte de la distribución geográfica de las especies. Tal es el caso del estudio realizado por Verga (1995) sobre *P. chilensis* y *P. flexuosa* como también el de Julio (2000) realizado sobre *P. chilensis*.

La alta variabilidad genética no estaría solamente asociada a marcadores neutros o casi neutros como los utilizados en los estudios previamente indicados. En un análisis de progenie-

origen, Cony (1996) encontró una alta variabilidad genética para caracteres morfológicos tales como altura y nivel de espinescencia.

Los estudios genético-poblacionales en especies de las demás secciones del género son escasos, destacándose los trabajo realizado por (Saidman et al., 1996; Burghardt et al., 2000). En el primero de estos trabajos se estudió la divergencia genética y la variabilidad en 5 especies de la sección *Strombocarpa*. Los resultados fueron muy relevantes ya que mostraron un patrón contrastante con respecto a las especies de la sección *Algarobia* tanto en la diversidad genética, que fue baja en las especies de la sección *Strombocarpa*, como en el nivel de divergencia entre las especies que fue mucho mayor en las especies de la sección *Strombocarpa*.

1.7. Análisis filogenéticos previos

Son pocos los estudios que han intentado analizar las relaciones filogenéticas entre especies de este género (Figura 5). El primer análisis fue el realizado por (Ramírez et al., 1999). Dicho trabajo incluyó un total de 15 especies y se basó en marcadores moleculares tipo RAPD (Random Amplified Polymorphic Loci). Independientemente de la importancia del mismo como pionero en el tema, sus resultados deben ser evaluados teniendo en cuenta que los marcadores RAPD, de clara utilidad en estudios a nivel intraspecífico o de especies muy cercanas, son fuertemente desaconsejados para estudios a nivel de especies más lejanamente emparentadas. Un análisis más reciente evaluó las relaciones entre especies del género *Prosopis* en base a frecuencias isoenzimáticas y utilizando métodos fenéticos (Bessegå et al., 2005). Dicho estudio estuvo circunscrito a 7 especies americanas incluyendo una especie de la sección *Strombocarpa*, una especie de la sección *Monilicarpa* y 5 especies de la sección *Algarobia*. Las principales conclusiones derivadas por los autores de este estudio fueron que las series de *Algarobia* no serían monofiléticas y que *P. argentina* debería ser incluida en una sección independiente ya que la diferenciación de la misma con especies de la sección

Algarobia y con *P. reptans* de la sección Strombocarpa era similar. En el año 2006 se publicó el primer análisis cladístico en el género (Bessegá et al., 2006). Uno de los principales datos que arrojó este estudio fue que ni la sección Algarobia ni la sección Strombocarpa serían monofiléticas.



Figura 5. Estudios fenéticos y cladísticos en *Prosopis*. A- Burghardt & Espert (2007), consenso estricto de los árboles más parsimoniosos obtenidos en uno de los análisis realizados a partir de una matriz combinada de caracteres morfológicos y bioquímicos. B-. Ramírez et al (1999), árbol de Neighbor joining obtenido a partir de una matriz de distancias derivadas de bandas RAPD. C- . Bessegá et al. (2005), agrupamientos obtenidos a partir de datos de frecuencias alozímicas y analizado con UPGMA. D-. Bessegá et al. (2006), árbol más parsimonioso obtenido a partir de secuencias de ITS.

Posterior a este análisis se ha publicado un estudio filogenético en base a caracteres morfológicos y bioquímicos (Burghardt & Espert, 2007) que abarcó especies de todas las series y secciones americanas del género. A pesar de que los resultados difirieron en función de los parámetros analíticos, dicho estudio mostró una correspondencia mucho mayor con los grupos definidos tradicionalmente dentro de *Prosopis*.

1.8. Ambientes ocupados por especies de *Prosopis*

Uno de los principales controles ambientales de la distribución de especies vegetales a escala regional es la disponibilidad de agua en el suelo. A pesar de que la mayoría de las especies de *Prosopis* habitan en ambientes áridos (hiperáridas a semiáridos), algunas de ellas también se encuentran en zonas en las cuales existe una mayor disponibilidad de agua. Tal es el caso por ejemplo de *P. africana* en África subsahariana y de *P. affinis* en Argentina. Otras especies tales como *P. nigra* y *P. alba* aunque toleran un mayor nivel de sequía, viven en zonas con un balance de agua positivo (Vilela & Palacios, 1997). El nivel de precipitaciones en las áreas habitadas por el género es muy variable, yendo desde los 50 mm/año a más de 1300 mm/ año. Entre las especies de *Prosopis* existen además diferencias en la amplitud de los valores de aridez/humedad que pueden habitar. Mientras que especies tales como *P. castellanosii* y *P. argentina* sólo habitan en zonas áridas y *P. africana* sólo en regiones húmedas, otras especies tienen distribuciones que incluyen regiones tanto semihúmedas como áridas. Tal es el caso de *P. flexuosa* y *P. chilensis*. Estas especies presentan lo que se denomina distribución azonal y zonal (Roig, 1993). La vegetación *zonal* se desarrolla en sitios en donde la influencia del clima sobre la vegetación se ve poco alterada por las condiciones del relieve y el suelo. La vegetación *azonal* está fuertemente influenciada en su presencia y distribución por las características del suelo y/o del sustrato geológico. En el caso de *Prosopis* la distribución *azonal* esta relacionado con la existencia de agua extra de la obtenida a partir de las lluvias. Especies como *P. flexuosa* o *P. chilensis* pueden sobrevivir en regiones donde

el régimen pluviométrico no permitiría su existencia pero los aportes extras de agua hacen posible su existencia (agua subterránea). Este comportamiento freatófito es compartido por muchas especies de *Prosopis* (Mooney et al., 1977). Para ello cuentan con dos sistemas radiculares, uno de ellos superficial para absorber el agua proveniente de las precipitaciones y un segundo sistema más profundo que alcanza la mayor profundidad y permite a la planta aprovechar el agua retenida en el suelo.

Dos especies características de regiones áridas son *P. ferox* y *P. tamarugo* de la serie *Cavencarpae*. Mientras que la distribución de *P. ferox*, asociada a cursos de agua temporarios del Norte de Argentina, es una clara evidencia de su carácter freatófito, la forma mediante la cual *P. tamarugo* logra sobrevivir en el extremadamente árido desierto de Atacama ha sido y continúa siendo causa de controversia. Algunos autores han indicado que dicha especie sería capaz de absorber agua del aire a partir de las hojas (Sudzuki, 1985). Tal proceso se vería ayudado mediante la existencia de ceras epicuticulares que permitirían atrapar y retener la humedad del ambiente (Hull & Bleckmann, 1977). Sin embargo la ocurrencia de este proceso ha sido fuertemente cuestionada principalmente porque existir una explicación alternativa a este proceso para comprender la supervivencia de esta especie en tal ambiente: la misma lograría llegar con sus raíces al agua de la napa freática que se haya presente en el desierto de Atacama (Pasiczniak et al., 2001). Independientemente de que tal mecanismo estuviera actuando o no en *P. tamarugo*, el resto de las especies de *Prosopis* que habitan en regiones áridas presentan una serie de caracteres que permitirían la existencia de las mismas en dichas regiones. Además del comportamiento freatófito previamente indicado también existen características anatómicas y funcionales. Muchas especies presentan caracteres foliares que le permitirían vivir en tales ambientes (Vilela & Palacios, 1997): estomas hundidos, principalmente en la cara abaxial, cutículas gruesas y ceras, clorénquima dorsiventral en empalizada, y presencia de mucílagos en las células que permitirían retener la humedad. También se han citados cambios metabólicos que permitirían a las especies de *Prosopis*

soportar períodos de sequía (Pasiczniak et al., 2001). Por ejemplo Alarcón & Díaz (1993) encontraron que la concentración de solutos en las células de las hojas aumentaría la concentración de todos los nutrientes produciéndose un ajuste osmótico. Otra característica de varias especies de *Prosopis* es la existencia de tallos fotosintéticos. Esta característica es común a especies áfilas o subáfilas de la sección Algarobia (*P. sericantha*, *P. kuntzei*, *P. castellanosii*, *P. humilis*, etc.) y también de *P. argentina* de la sección Monilicarpa. Otras especies (ej. *P. juliflora*) tiene tallos fotosintéticos de manera temporaria, lo que permite que las mismas pierdan las hojas produciéndose una disminución concomitante en la evapotranspiración (El Fadl, 1997).

1.9. Antecedentes de estudios de diversificación en zonas áridas y semiáridas

La utilización de datación molecular ha permitido definir un marco temporal de la diversificación de grupos con escaso o nulo registro fósil. Uno de los principales avances ha sido el desarrollo de métodos que permiten dataciones moleculares con tasas de sustituciones variables (Sanderson, 2002; Sanderson, 2003), es decir cuando no se cumple el supuesto de reloj molecular. Otro avance metodológico importante ha sido el diseño de métodos para evaluar cambios en la tasa de diversificación en función del tiempo. Gracias a esto es posible determinar si han ocurrido cambios de tasas de diversificación en paralelo para diversos grupos presentes en una región y su posible correlación con cambios ambientales. Estos avances metodológicos han generado avances importantes en el terreno empírico. Esto se hace evidente estudios referidos a la diversificación en zonas áridas y semiáridas, donde recientemente se han publicado una cantidad importante de trabajos (ej. Klak et al., 2003; Verboom et al., 2003; Good-Avila et al., 2006). A pesar de que los resultados no son concluyentes, es posible observar que en varios de estos grupos existiría una correlación bastante pronunciada entre el desarrollo histórico de ambientes áridos y semiáridos y la diversificación de los grupos estudiados. La asociación de *Prosopis* con ambientes áridos

junto con el extenso conocimiento sobre las mismas acumuladas a lo largo de muchos años brindan una oportunidad única para comprender los procesos y controles asociados a la diversificación de grupos vegetales en zonas áridas.

1.10. Optimización de caracteres ambientales

Los recientes cambios climáticos producidos por la actividad humana han afectado de manera notable la biodiversidad. La influencia de tales cambios puede verse, entre otras manifestaciones, por aumentos en la tasa de extinción de las especies así como también en cambios en su distribución (Hughes et al., 2000; Parmesan & Yohe, 2003). El impacto de los cambios climáticos en la distribución de las especies han sido estudiado mediante el método conocido como “bioclimatic envelope modelling” ó “modelado del sobre bioclimático”. En este análisis se define la relación entre las distribuciones actuales de las especies y las variables climáticas para posteriormente predecir las distribuciones de las mismas en nuevos escenarios climáticos (Araújo & New, 2006). Una extensión de estos estudios ha sido su combinación con estudios filogenéticos en una aproximación que ha sido denominada “Modelaje Filoclimático” (Yesson & Culham, 2006). Estos estudios parten de definir un “modelo climático” para especies actuales en las cuales se definen los rangos de diferentes variables climáticas. Posteriormente los valores de dichas variables son optimizados sobre la filogenia del grupo y se establecen las condiciones en las cuales habrían vivido los ancestros de estas especies. Estos estudios han despertado un fuerte interés ya que se consideran que serían muy útiles a la hora de evaluar los posibles efectos de cambios climáticos en la distribución de las especies así como también de su potencial extinción. Sin embargo no ha habido un análisis a conciencia de los métodos empleados, por lo cual no son claras las respuestas ante preguntas tales como ¿Es posible obtener información de las condiciones ambientales en que vivían los ancestros del grupo de estudio a partir de la optimización de los caracteres ambientales? Una respuesta negativa a esta cuestión pondría en duda la utilidad de

los métodos anteriormente citados para abordar soluciones a los problemas ambientales originados por el cambio climático mundial.

1.11. Objetivos

Objetivo general

- Analizar el proceso de diversificación en género *Prosopis*, estudiando las relaciones filogenéticas entre sus especies, el marco temporal y los posibles factores extrínsecos que afectaron la misma.

Objetivos particulares

- Establecer las relaciones filogenéticas dentro del género *Prosopis*, a partir de la secuenciación de ADN, evaluando tanto la monofilia del género como la de los grupos infragenéricos tradicionalmente reconocidos en la taxonomía.
- Establecer un marco temporal para la diversificación de *Prosopis*. Realizar una datación molecular y estimar la tasa de diversificación a lo largo del tiempo
- Analizar el marco histórico en el cual ocurrió dicha diversificación, tanto desde un punto de vista geológico como también paleo climático. Comparar la diversificación de este género con otros de zonas áridas
- Inferir las condiciones ambientales en las cuales habrían diversificado estas especies a partir de la utilización de técnicas de optimización de caracteres ambientales sobre la filogenia del género, haciéndose además una evaluación crítica sobre la utilización de dichas metodologías para cumplir este propósito.

1.12. Bibliografía

- Alarcón, C. & M. Díaz. 1993. Hydric and nutritional relations of *Prosopis juliflora* DC (Fabaceae, Mimosoideae) in semiarid parts of Falcon state, Venezuela. Revista de Biología Tropical 41: 433-441.
- Araújo, M. B. & M. New. 2006. Ensemble forecasting of species distributions. Trends in Ecology and Evolution 22: 42-47.
- Astrada, E. & J. Adámoli. 2005. Ecología y manejo de vinalares: Perspectiva regional y aplicaciones en el centro de Formosa. En M. Arturi, J. Frangi y J. Goya (eds), Ecología y manejo de bosques de Argentina.. Editorial de la Universidad Nacional de La Plata, La Plata.
- Bentham, G. 1839. Plantae Hartwegianae.
- Bentham, G. 1842. Hooker's Journal of Botany 4: 346-352.
- Bentham, G. 1875. Revision of the suborder Mimosae. Transactions of the Linnean Society 30: 335-664.
- Berhanu, A. & G. Tesfaye. 2006. The *Prosopis* dilemma, impacts on dryland biodiversity and some controlling method. Journal of the Drylands 1: 158-164.
- Bernardi, L. 1984. Contribucion a la dendrologia paraguaya I. Boissiera 35: 1-341.
- Bessega, C., L. I. Ferreyra, J. C. Vilardi & B. O. Saidman. 2000a. Unexpected low genetic differentiation among allopatric species of Section Algarobia of *Prosopis* (Leguminosae). Genetica 109: 255-266.
- Bessega, C., B.O. Saidman & J.C. Vilardi. 2000b. Isozyme and RAPD studies in *Prosopis glandulosa* and *P. velutina* (Leguminosae, Mimosoideae). Genetics and Molecular Biology 23: 1-5.
- Bessega, C., B. O. Saidman & J. C. Vilardi. 2005. Genetic relationships among American species of *Prosopis* (Leguminosae) based on enzyme markers. Genetics and Molecular Biology 28: 277-286.
- Blasco, C., S. Carenzo & E. Astrada. 2005. Evaluación de un sistema silvopastoril sobre vinalares en Formosa, Argentina. Revista Iberoamericana de Economía Ecológica 2: 57-68.
- Britton, N. & J. Rose. 1928. Mimosaceae. North American Flora. 23: 1 -194.
- Bukhari, Y. 1997. Cytoevolution of Taxa in *Acacia* and *Prosopis* (Mimosaceae). Australian Journal of Botany 45: 879-891.

- Burghardt, A. D. 1995. La identidad de cuatro especies de *Prosopis* L. expresada a través de sus patrones electroforeticos. *Mendeliana* 12: 38-50.
- Burghardt, A.D. 2000. Estudio electroforético de proteínas de semilla en *Prosopis* (Leguminosae) I: sección *Strombocarpa*. *Boletín de la Sociedad Argentina de Botánica* 35, 149–156.
- Burghardt, A. & R. Palacios. 1997. Electrophoretic characterization of the American sections of *Prosopis* L. (Leguminosae: Mimosoidae). *Bulletin of the International Group for the study of Mimososidae* 20: 71-83.
- Burghardt, A. & S. Espert. 2007. Phylogeny of *Prosopis* (Leguminosae) as shown by morphological and biochemical evidence. *Australian Journal of Botany* 20: 332–339.
- Burkart, A. 1976. A monograph of the genus *Prosopis* (Leguminosae subFam Mimosoideae). *Journal of the Arnold Arboretum* 57: 219-249.
- Cabrera, A. & A. Willink. 1973. Biogeografía de América Latina. OEA, Washington D.C.
- Candolle, A. P. d. 1825. *Prodromus systematis naturalis regni vegetabilis*. Treuttel et Würtz. Paris 2: 1-644.
- Cherubini, C. 1954. Numeros de cromosomas de algunas especies del genero *Prosopis* (Leguminosae-Mimosoidae). *Darwiniana* 10: 637–643.
- Cony, M. 1996. Genetic variability in *Prosopis flexousa* D.C. a native tree of Monte phytogeography province, Argentina. *Forest Ecology and Management* 87: 41-49.
- D'antoni, H. & O. Solbrig. 1977. Algarrobos in South American cultures: past and present. En B. B. Simpson (Ed) *Mesquite, Its Biology in Two Desert Ecosystems*. Dowden, Hutchinson and Ross Inc., Stroudsburg, Pennsylvania, USA, pp. 189-200.
- El Fadl, M. A. 1997. Management of *Prosopis juliflora* for use in agroforestry systems in the Sudan, *Tropical Forestry Reports* 6. University of Helsinki, Helsinki, Finland.
- El-Keblawy, A. & A. Al-Rawai. 2005. Effects of salinity, temperature and light on germination of invasive *Prosopis juliflora* (Sw.) D.C. *Journal of Arid Environments* 61: 555–565.
- Felker, P. & P. Clark. 1980. Nitrogen fixation (acetylene reduction) and cross inoculation in 12 *Prosopis* (mesquites) species. *Plant and Soil* 57: 177-186.
- Felker, P., N. Gradosc, G. Cruzc & D. Prokopiukd. 2003. Economic assessment of production of flour from *Prosopis alba* and *P. pallida* pods for human food applications. *Journal of Arid Environments*. 53: 517-528.

- Ferreyra, I., C. Bessega, J. C. Vilardi & B. O. Saidman. 2004. First report on RAPDs patterns able to differentiate some argentinean species of section algarobia (*Prosopis*, Leguminosae). *Genetica* 121: 33-42.
- Good-Avila, S. V., V. Souza, B. S. Gaut & L. E. Eguiarte. 2006. Timing and rate of speciation in *Agave* (Agavaceae). *Proceedings of the National Academy of Sciences* 103: 9124–9129.
- Guinet, P. H. & M. Bessedik. 1984. Presence de Genre *Prosopis* (Leguminosae-Mimosoideae) a l'Aquitainien Basal dans l'aude (Languedoc-france). *Pollen et spores* 26: 101-108.
- Habit, M. 1985. The current state of knowledge on *Prosopis tamarugo*. Food and Agriculture Organization, Santiago, Chile.
- Hughes, J. B., G. C. Daily & P. R. Ehrlich. 2000. Conservation of insect diversity: a habitat approach. *Conservation Biology* 23: 1788–1797.
- Hull, H. M. & C. A. Bleckmann. 1977. An unusual epicuticular wax ultrastructure on leaves of *Prosopis tamarugo* (Leguminosae). *American Journal of Botany* 64: 1083-1091.
- Hunziker, J., L. Poggio, C. Naranjo, R. Palacios & A. Andrade. 1975. Cytogenetics of some species and natural hybrids in *Prosopis* (Leguminosae). *Canadian Journal of Cytology* 17: 253-262.
- Hunziker, J. H., C. A. Naranjo, R. A. Palacios, L. Poggio & B. O. Saidman. 1986. Hybridization and Genetic variation of Argentine Species of *Prosopis*. *Forest Ecology and Management* 16: 301-315.
- Juárez-Muñoz, J., G. Carrillo-Castañeda, R. Arreguin & A. Rubluo. 2002. Inter- and intra-genetic variation of four wild populations of *Prosopis* using rapd-pcr fingerprints. *Biodiversity and Conservation* 11: 921–930.
- Julio, N. 2000. Estudios alozímicos sobre variabilidad, estructura y diferenciación genética en *Prosopis chilensis* (Leguminosae, Mimosoideae) y especies relacionadas. Tesis Doctoral. Universidad Nacional de Córdoba.
- Klak, C., G. Reeves & T. Hedderson. 2003. Unmatched tempo of evolution in Southern African semi-desert ice plants. *Nature* 427: 63-65.
- Ladio, A. & M. Lozada. En prensa. Human ecology, ethnobotany and traditional practices in rural populations inhabiting the Monte region: Resilience and ecological knowledge. *Journal of Arid Environments*.
- Lamarck, J. 1789. *Encyclopédie méthodique: Botanique*.

- Linnaeus, C. 1767. *Mantissa plantarum altera Generum editionis VI and specierum editions II.*
- Mooney, H. A., B. B. Simpson & O. T. Solbrig. 1977. Phenology, morphology, physiology. In B. B. Simpson (ed) *Mesquite, Its Biology in Two Desert Ecosystems*. Dowden, Hutchinson and Ross Inc., Stroudsburg, Pennsylvania, USA, pp. 26-43.
- Naranjo, C. A., L. Poggio & S. Enus Zeiger. 1984. Phenol chromatography, morphology and cytogenetics in three species and natural hybrids of *Prosopis* (Leguminosae-Mimosoideae). *Plant Systematic and Evolution* 144: 257-276.
- Newton, C. 2005. Upper Egypt: vegetation at the beginning of the third millennium BC inferred from charcoal analysis at Adaïma and Elkab. *Journal of Archaeological Science* 32: 355–367.
- Palacios, R. 2006. Los Mezquites Mexicanos: Biodiversidad y Distribución Geográfica. *Boletín de la Sociedad Argentina de Botánica* 41: 99-121.
- Palacios, R. A. & L. D. Bravo. 1981. Hibridación natural en *Prosopis* (Leguminosae) en la región chaqueña argentina. Evidencias morfológicas y cromatográficas. *Darwiniana* 23: 3-35.
- Palacios, R. A. & L. D. Bravo. 1984. On some Argentine species of *Prosopis*. *Bull. IGSM* 12: 103-112.
- Parmesan, C. & G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.
- Pasiecznik, N. M., P. Felker, P. J. C. Harris, L. N. Harsh, G. Cruz, J. C. Tewari, K. Cadoret & L. J. Maldonado. 2001. The *Prosopis juliflora*-*Prosopis pallida* Complex: A Monograph. HDRA, Coventry UK.
- Pegado, C., L. Andrade, L. Félix & I. Pereira. 2006. Efeitos da invasão biológica de algaroba - *Prosopis juliflora* (Sw.) DC. sobre a composição e a estrutura do estrato arbustivo-árboreo da caatinga no Município de Monteiro, PB, Brasil. *Acta Botanical Brasilica*. 20: 887-898.
- Pometti, C., B. Pizzo, M. Brunetti, N. Macchioni, M. Ewens & B. Saidman. En prensa. Argentinean native wood species: Physical and mechanical characterization of some *Prosopis* species and *Acacia aroma* (Leguminosae; Mimosoideae). *Bioresource Technology*
- Rafinesque Schmaltz, C. 1838. *Sylva Telluriana*.

- Ramírez, L., A. De La Vega, N. Razkin, V. Luna & P. J. C. Harris. 1999. Analysis of the relationships between species of the genus *Prosopis* revealed by the use of molecular markers. *Agronomie* 19: 31-43.
- Robinson, T., R. van Klinkenb & G. Metternicht. 2008. Spatial and temporal rates and patterns of mesquite (*Prosopis* species) invasion in Western Australia. *Journal of Arid Environments* 72: 175–188.
- Rodríguez, M. 2005. Human evidence from the mid-Holocene in the salty Argentine Puna: analysis of the archaeobotanical record. 132: 15–22.
- Roig, F. A. 1993. Informe nacional para la selección de germoplasma en especies de *Prosopis* en la República Argentina. IADIZA-CRICYT.
- Saidman, B.O. & J.C. Vilardi. 1987. Analysis of the genetic similarities among seven species of *Prosopis* (Leguminosae: Mimosoideae). *Theoretical and Applied Genetics* 75: 109-116.
- Saidman, B.O., J.C. Vilardi, S. Montoya, M. Dieguez & H. Hopp. 1995. Molecular Markers: A tool for the understanding of the relationships among species of *Prosopis* (Leguminosae, Mimosoideae). In S. Puri (ed) *Tree Improvement : Applied Research and Technology Transfer*. Oxford & IBH Co.
- Saidman, B. O. 1985. Estudio de la variación alozímica en el género *Prosopis*, Estudio de la variación alozímica en el género *Prosopis*. Tesis Doctoral. Universidad de Buenos Aires, Buenos Aires, Argentina.
- Saidman, B. O., J. C. Vilardi, M. I. Pocovi & N. Acreche. 1996. Genetic divergence among species of section *Strombocarpa*, genus *Prosopis* (Leguminosae). *Journal of Genetics* 75: 139-149.
- Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* 19: 101-109.
- Sanderson, M. J. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19: 301-302.
- Schinini, A. 1981. Contribución a la flora de Paraguay. *Bonplandia* 5: 101-108.
- Senthilkumar, P., W. S. P. M. Prince, S. Sivakumar & C. V. Subbhuraam. 2005. *Prosopis juliflora* : A green solution to decontaminate heavy metal (Cu and Cd) contaminated soils. *Chemosphere* 60: 1493-1496.
- Shiferaw, H., D. Teketay, S. Nemomissa & F. Assefa. 2004. Some biological characteristics that foster the invasion of *Prosopis juliflora* (Sw.) DC. at Middle Awash Rift Valley Area, north-eastern Ethiopia. *Journal of Arid Environments* 58: 135–154.

- Simpson, B. 1977. Mesquite, Its Biology in Two Desert Ecosystems. Dowden, Hutchinson and Ross Inc., Stroudsburg, Pennsylvania, USA.
- Solbrig, O. & K. Bawa. 1975. Isozyme variation in species of *Prosopis* (Leguminosae). *Journal of the Arnold Arboretum, Harvard University* 56: 398-412.
- Suzuki, F. 1985. Environmental moisture utilization by *Prosopis tamarugo* Phil. In M. A. Habit (ed) The Current State of Knowledge on *Prosopis tamarugo* Phil. FAO, Rome, Italy., pp. 33-48.
- Tiwari, J. 1999. Exotic weed *Prosopis juliflora* in Gujarat and Rajasthan, India—boon or bane. *Tigerpaper* 26: 21–25.
- Torres, M. 1985. Biological Nitronge Fixation in *Prosopis tamarugo* and *Prosopis alba*. In M. E. Habit (Ed) The Current State of Knowledge on *Prosopis tamarugo* Phil. Food and Agriculture Organization, Santiago, Chile, pp. 251-256.
- Trenchard, L., P. Harris, S. Smith & N. Pasiecznik. 2008. A review of ploidy in the genus *Prosopis* (Leguminosae). *Botanical Journal of the Linnean Society* 156: 425-438.
- Vega, M. & P. Hernández. 2005. Molecular evidence for natural interspecific hybridization in *Prosopis*. *Agroforestry Systems* 64: 197–202.
- Verboom, G. A., H. P. Linder & W. D. Stock. 2003. Phylogenetics of the grass genus *Ehrharta*: Evidence for radiation in the summer-arid zone of the South African cape. *Evolution* 57: 1008–1021.
- Verga, A. R. 1995. Genetische untersuchungen an *Prosopis chilensis* und *Prosopis flexuosa* (Mimosaceae) im trockenen Chaco Argentiniens. Tesis Doctoral. Universität Göttingen.
- Vilela, A. E. & R. A. Palacios. 1997. Adaptive features in leaves of South American species of the genus *Prosopis* (Leguminosae: Mimosoideae). *Bulletin of the International Group for the Study of Mimosoideae* 20: 62-70.
- Wegener, A. 1912. Die Entstehung der Kontinente. Dr. A. Petermanns Mitteilungen aus Justus Perthes' Geographischer Anstalt 63: 185–195.
- Yesson, C. & A. Culham. 2006. Phyloclimatic Modeling: Combining Phylogenetics and Bioclimatic Modeling. *Systematic Biology* 55: 785–802.
- Zallochi, E., R. Palacios & M. Brizuela. 1990. Interpopulation. variation in *Prosopis flexuosa* DC. from northern Chile. *Bulletin I.G.S.M.* 18: 135-149.
- Zimmermann, H. 1991. Biological control of mesquite, *Prosopis* spp. (Fabaceae), in South Africa. *Agriculture, Ecosystems and Environment* 37: 175-186.

CAPITULO 2 – ANÁLISIS FILOGENÉTICO

2.1. Introducción

Uno de los puntos indicados en el capítulo introductorio fue el marcado incremento en los últimos años del conocimiento de las relaciones filogenéticas en *Prosopis* (Ramírez et al., 1999; Bessega et al., 2006; Burghardt & Espert 2007). Sin embargo estos estudios presentan ciertas limitaciones que en mayor o menor medida representan un impedimento para una total comprensión de la filogenia de este género. Entre las mismas se puede destacar que en ninguno de los casos se incluyó más de un taxón como grupo externo y en algunos dicho taxón estaba lejanamente emparentado con *Prosopis*, imposibilitando la evaluación de la monofilia del género. Con respecto al muestreo, en la mayoría de los casos estaban muy bien representadas las series Chilenses, Ruscifoliae y Pallidae de la sección Algarobia, pero la representación del resto de los grupos era incompleta. Por otra parte, el único análisis de secuencias de ADN se baso en un sólo marcador (Bessega et al., 2006). Teniendo esto en cuenta estas limitaciones, en la presente tesis se buscó incrementar el conocimiento de las relaciones filogenéticas del género a partir de: (A) Aumentar el número de genes analizados, pasando de 1 a 5 genes. (B) Incrementar la cantidad de especies (de 20 a 31) secuenciadas. (C) Aumentar la representación taxonómica en nuestro conjunto de datos incluyendo especies de cada una de las series (pasando de 5 a 8) y secciones del género (de 4 a 5). (D) Incluir especies de géneros afines con el objeto de poner a prueba la monofilia de *Prosopis*. En uno de los análisis realizados en esta tesis se incluyeron 3 especies de géneros afines: *Acacia*, *Prosopidastrum* y *Xerocladia* además de *Lotus* como grupo externo, mientras que en un segundo análisis se incluyeron 77 especies de Leguminosas.

Como fue previamente indicado, el único análisis filogenético publicado sobre *Prosopis* realizado en base a secuencias de ADN estuvo restringido a un solo marcador (Bessega et al., 2006): el espaciador interno transcripto de los genes del ARN ribosómico

(ITS). Aun cuando la utilidad de este marcador para resolver relaciones a nivel infragenérico en plantas ha sido demostrada extensamente en la bibliografía (Baldwin, 1992; Yuan et al., 1996; Pansarin et al., 2008), su uso presenta varias desventajas. Entre éstas se encuentra la existencia de copias parálogas y la falta de una completa evolución concertada entre las diferentes copias de esta región del genoma (Alvarez & Wendel, 2003). Marcadores alternativos a los ITS para la evaluación de las relaciones entre especies estrechamente emparentadas generalmente se restringen a ADN de cloroplasto. El principal problema de los mismos es que suelen presentar muy poca divergencia. Es así que es difícil encontrar regiones genómicas que permitan llevar a cabo estudios filogenéticos a niveles taxonómicos bajos. Una posible alternativa es la utilización de genes nucleares de bajo número de copias. Estos genes han comenzado a tener gran importancia en los estudios de sistemática molecular en plantas (Sang & Zhang, 1999; Evans et al., 2000; Cronn et al., 2002; Malcomber, 2002; Roalson & Friar, 2004), principalmente con el desarrollo de cebadores universales que facilitaron de manera sustancial la amplificación de estos genes en grupos donde se conoce poco o nada de su genoma. En la presente tesis se decidió incluir 2 genes nucleares de baja copia (NIA y g3pdh) y 3 genes de cloroplasto (*trnS-psbC*, *trnK-matK*, *trnL-trnF*).

El diseño original del estudio filogenético involucraba la secuenciación de tres genes (nitrato reductasa, *NIA*; gliceraldehído-3-fostato-deshidrogenasa, *g3pdh*; y región intergénica entre los genes *trnS-psbC* de cloroplasto) para las 30 especies de *Prosopis* y cuatro grupos externos (*Acacia caven*, *Prosopidastrum angusticarpum*, *Xerocladia viridiramis* y *Lotus japonicus*). Este estudio será denominado durante esta tesis “análisis de tres marcadores”. Debido a que los resultados filogenéticos obtenidos a partir de este conjunto de datos pusieron en duda la monofilia del género, se decidió hacer un estudio con un muestro mucho más exhaustivo de taxones relacionados con el género, incluyendo 77 especies de mimosoideas. La secuenciación de estas 77 especies para los genes incluidos en el estudio de tres genes

habría tomado un tiempo no compatible con los tiempos de becas y además habría requerido de recursos económicos que excedían el presupuesto con el que se contaba a la hora de encarar este trabajo. Es así que la se decidió realizar un análisis con una aproximación denominada “supermatriz”, es decir la combinación de conjuntos de datos con diferentes muestras de especies para diferentes genes. La utilidad de dicha aproximación ha sido claramente demostrada (ej. Matthee et al., 2004; McMahon & Sanderson, 2006). Para generar esta supermatriz se combinó la información de la matriz de los tres genes (*NIA*, *g3pdh* y *trnS-psbC*) con una matriz que incluyó secuencias de 5 especies de *Prosopis* y 77 especies de leguminosas para los genes *trnL-trnF* y *trnK-matK*. Las 5 especies de *Prosopis* fueron secuenciadas en nuestro laboratorio y representaron a cada una de las secciones del género (Tabla 2). El resto de las secuencias fueron obtenidas de GenBank. El análisis de la supermatriz conteniendo los cinco marcadores será referido en esta tesis como el “análisis de 5 marcadores” mientras que el estudio que solo incluyó a *matK-trnK* y *trnL-trnF* se denominará “análisis de dos marcadores”

2.2. Materiales y Métodos

2.2.1. Muestreo de genes y especies

Las especies incluidas en el estudio de tres marcadores se indican en la Tabla 2. En la Tabla 3 se indican las especies cuyas secuencias fueron obtenidas de GenBank e incluidas en los estudios de 2 y 5 marcadores.

2 – ANÁLISIS FILOGENETICO

Especie	Sección-serie	Nº de Herbario/ Accesión banco de germoplasma	Dist. Natural	Número de Accesion en GenBank				
				G3pdh	trnS-psbC	NIA	trnL-trnF	trnK/matK
<i>P. cineraria</i> (L.) Druce	Prosopis	DANIDA Nº 01089/82	AF	EF165221	EF165309	EF165269	EF165292	EF165287/ EF165248
<i>P. africana</i> (Guill., Perr., & Rich.) Taubert	Anonychium	HDRA Nº63	AF	EF165209	EF165297	EF165253	EF165291	EF165286/ EF165251
<i>P. strombulifera</i> (Lam.) Bentham	Strombocarpa-Strombocarpaceae	Hunziker & Gamerro 11359; SI	SA	EF165241	EF165327	EF165281	-	-
<i>P. reptans</i> Bentham	Strombocarpa-Strombocarpaceae	Saidman 36; BAFC	SA, NA	EF165237	EF165333	-	-	-
<i>P. torquata</i> (Cavanilles ex Lagasca)	Strombocarpa-Strombocarpaceae	Hunziker 9733 (SI)	SA	EF165244	EF165330	-	-	-
<i>P. pubescens</i> Bentham	Strombocarpa-Strombocarpaceae	Evans Nº 15 (USDA-USA)	NA	EF165236	EF165323	-	-	-
<i>P. palmeri</i> S. Watson	Strombocarpa-Strombocarpaceae	-	NA	EF165234	EF165322	EF165278	-	-
<i>P. ferox</i> Grisebach	Strombocarpa-Cavencarpaceae	Hunziker et al. 10451; SI	SA	EF165223	EF165311	-	-	-
<i>P. tamarugo</i> F. Philippi	Strombocarpa-Cavencarpaceae	DANIDA Nº 01215/83	SA	EF165242	EF165328	-	EF165293	EF165289/ EF165250
<i>P. argentina</i> Burkart	Monilicarpa	B.N.G. Prosopis Nº 15-88.	SA	EF165212	EF165300	EF165261	EF165290	EF165288/ EF165249
<i>P. sericantha</i> Gilles ex Hooker & Arnott	Algarobia - Sericanthae	Saidman & Vilardi 961; BAFC	SA	EF165240	EF165326	EF165280	-	-
<i>P. kuntzei</i> Harms	Algarobia - Sericanthae	Saidman & Vilardi 516,521; BAFC	SA	EF165230 EF165231	EF165318 EF165319	EF165276 EF165277	-	-
<i>P. ruscifolia</i> Grisebach	Algarobia-Ruscifoliae	Saidman & Vilardi 420; BAFC	SA	EF165239	EF165325	EF165257	-	-
<i>P. vinalillo</i> Stuckert	Algarobia-Ruscifoliae	Saidman & Vilardi 505, 511; BAFC	SA	EF165245 EF165246	EF165331 EF165332	EF165283 EF165284	-	-
<i>P. denudans</i> Bentham	Algarobia-Denudantes	Catalano 33; BAFC	SA	EF165222	EF165310	EF165270	-	-
<i>P. ruizleali</i> Burkart	Algarobia-Denudantes	Catalano 13; BAFC	SA	EF165238	EF165324	EF165279	-	-
<i>P. castellanosii</i> Burkart	Algarobia-Denudantes	Catalano 8; BAFC	SA	EF165218	EF165306	EF165267	-	-
<i>P. humilis</i> Gilles ex Hooker & Arnott	Algarobia-Humiles	Catalano 1; BAFC	SA	EF165227	EF165315	EF165275	-	-
<i>P. campestris</i> Grisebach	Algarobia -Pallidae	Catalano 18,20; BAFC	SA	EF165216 EF165217	EF165304 EF165305	EF165265 EF165266	-	-
<i>P. affinis</i> Sprengel	Algarobia – Pallidae	Saidman et al. 526; BAFC	SA	EF165208	EF165296	EF165256	-	-

Especie	Sección-serie	Nº de Herbario/ Accesión banco de germoplasma	Dist. Natural	Número de Accesión en GenBank				
				G3pdh	trnS-psbC	NIA	trnL- trnF	trnK/matK
<i>P. articulata</i> S. Watson	Algarobia – Pallidae	-	NA	EF165213	EF165301	EF165262	-	-
<i>P. tamaulipana</i> Burkart	Algarobia – Pallidae	-	NA	EF165243	EF165329	EF165282	-	-
<i>P. chilensis</i> (Molina) Struntz emend. Burkart	Algarobia- Chilenses	Saidman & Vilardi 911, 922; BAFC	SA	EF165219 EF165220	EF165307 EF165308	EF165268 EF165285	-	-
<i>P. juliflora</i> (Swartz) DC	Algarobia- Chilenses	Hunziker 10039, 10040; SI	NA, AC, SA	EF165228 EF165229	EF165316 EF165317	EF165258 EF165274	-	-
<i>P. nigra</i> (Grisebach) Hieronymus	Algarobia- Chilenses	Saidman & Vilardi 315, 435; BAFC	SA	EF165232 EF165233	EF165320 EF165321	EF165259 EF165260	-	-
<i>P. caldenia</i> Burkart	Algarobia- Chilenses	Saidman & Vilardi 350,138; BAFC	SA	EF165214 EF165215	EF165302 EF165303	EF165263 EF165264	-	-
<i>P. flexuosa</i> DC	Algarobia- Chilenses	Saidman & Vilardi 302; BAFC, Catalano 10; BAFC	SA	EF165224 EF165225	EF165312 EF165313	EF165271 EF165272	-	-
<i>P. glandulosa</i> Torrey	Algarobia- Chilenses	Evans Nº 5 (USDA- USA)	NA	EF165226	EF165314	EF165273	-	-
<i>P. alpataco</i> R.A. Philippi	Algarobia- Chilenses	Catalano 34; BAFC	SA	EF165211	EF165299	EF165255	-	-
<i>P. alba</i> Grisebach	Algarobia- Chilenses	Saidman & Vilardi 413; BAFC	SA	EF165210	EF165298	EF165254	-	-
<i>Prosopidastrum angusticarpum</i> <i>Palacios & Hoc</i>	NA	Peter 244; BAFC	SA	EF165235	EF165295	-	-	-/EF165252
<i>Xerocladia viridiramis</i> Taub.	NA		AF	EF585492	EF585493		Tabla 3	Tabla 3
<i>Acacia caven</i> (Mol.) Molina	NA	Saidman et al. 789; BAFC	SA	EF165207	EF165294	EF165206	Tabla 3	Tabla 3
<i>Lotus japonicus</i> L.	NA	-	SA	EF165247	Tabla 3	-	Tabla 3	Tabla 3

Tabla 2. Secuencias obtenidas durante el desarrollo de la tesis.

Tabla 3. Secuencias obtenidas de la base de datos GenBank

Species	Gen <i>trnL</i> , intergen <i>trnL-trnF</i> y gen <i>trnF</i>	Region <i>matK</i>	Región <i>trnK</i>
<i>Acacia adoxa</i>	AF195703.	AF523076.1	AF523076.1
<i>Acacia ampliceps</i>	AF522983.1	AF523074.1	AF523074.1
<i>Acacia berlandieri</i>	AF522978.1	AF274145.1	AF274145.1
<i>Acacia boliviiana</i>	AF522981.1	AF274144.1	AF274144.1
<i>Acacia caven</i>	AF522967.1	AF274131.1	AF274131.1
<i>Acacia chiapensis</i>	AY574114.1	AY574097.1	AY574097.1
<i>Acacia cochlicantha</i>	AF522968.1	AF274133.1	AF274133.1
<i>Acacia colei</i>	AF522987.1	AF274215.1	AF274215.1
<i>Acacia collinsii</i>	AY574107.1	AY574092.1	AY574092.1
<i>Acacia constricta</i>	AF522969.1	AY574104.1	AY574104.1
<i>Acacia cornigera</i>	AY574120.1	AY574105.1	AY574105.1
<i>Acacia coulteri</i>	AF522990.1	AF523105.1	AF523105.1
<i>Acacia dolichostachya</i>	AF522989.1	AF523190.1	AF523190.1
<i>Acacia drummondii</i>	AF195704.1	AF523106.1	AF523106.1
<i>Acacia elata</i>	AF195702.1	AF523084.1	AF523084.1
<i>Acacia farnesiana</i>	AY574119.1	AF523115.1	AF523115.1
<i>Acacia glomerosa</i>	AF522980.1	AF274147.1	AF274147.1
<i>Acacia hindsii</i>	AY574116.1	AY574099.1	AY574099.1
<i>Acacia karroo</i>	AF522972.1	AF274137.1	AF274137.1
<i>Acacia longifolia</i>	AF195698.1/ AF195679.1	AF523086.1	AF523086.1
<i>Acacia lycopodiifolia</i>	AF195705.1	AF523077.1	AF523077.1
<i>Acacia macracantha</i>	AY574117.1	AY574100.1	AY574100.1
<i>Acacia mearnsii</i>	AF195694.1/ AF195675.1	AF523110.1	AF523110.1
<i>Acacia melanoxylon</i>	AF195699.1/AF195680.1	AF274166.1	AF274166.1
<i>Acacia modesta</i>	AF522975.1	AF274222.1	AF274222.1
<i>Acacia neovernicosa</i>	AF522970.1	AF523113.1	AF523113.1
<i>Acacia nilotica</i>	AF522973.1	AF274139.1	AF274139.1
<i>Acacia pennatula</i>	AY574111.1	AF274134.1	AF274134.1
<i>Acacia platycarpa</i>	AF522985.1	AF274223.1	AF274223.1
<i>Acacia pulchella</i>	AF195692.1/AF195673.1	AF523100.1	AF523100.1
<i>Acacia roemeriana</i>	AF522977.1	AF523099.1	AF523099.1
<i>Acacia schottii</i>	AF522971.1	AF274136.1	AF274136.1
<i>Acacia schweinfurthii</i>	AF522979.1	AF523101.1	AF523101.1
<i>Acacia senegal</i>	AF522976.1	AF274143.1	AF274143.1
<i>Acacia spinescens</i>	AF195706.1	AF523082.1	AF523082.1
<i>Acacia tortilis</i>	AF522974.1	AF274140.1	AF274140.1
<i>Acacia translucens</i>	AF522984.1	AF523087.1	AF523087.1
<i>Acacia tumida</i>	AF522986.1	AF523111.1	AF523111.1
<i>Acacia visco</i>	AF522982.1	AF523116.1	AF523116.1
<i>Adenanthera pavonina</i>	AF278486.1	AF521808.1	-
<i>Alantsilodendron alluaudianum</i>	AF278523.1	AF521809.1	-
<i>Alantsilodendron humbertii</i>	AF278522.1	AF521810.1	-
<i>Alantsilodendron pilosum</i>	AY125844.1	AF521811.1	-
<i>Albizia adinocephala</i>	AF522995.1	-	AF523092.1
<i>Albizia harveyi</i>	AF522991.1	AF523075.1	AF523075.1
<i>Albizia kalkora</i>	AF522945.1	AF523083.1	AF523083.1
<i>Albizia plurijuga</i>	AF522993.1	AF523080.1	AF523080.1
<i>Albizia sinaloensis</i>	AF522946.1	AF274121.1	AF274121.1
<i>Albizia tomentosa</i>	AF522994.1	AF523093.1	AF523093.1
<i>Amblygonocarpus andongensis</i>	AF278487.1	AF521812.1	-
<i>Anadenanthera colubrina</i>	AF278481.1	AF521813.1	AF521813.1
<i>Anadenanthera peregrina</i>	AF278480.1	AF521814.2	-
<i>Calliandra carbonaria</i>	AF278516.1	AF521815.1	-
<i>Calliandropsis nervosus</i>	AF278520.1	AF521816.1	-
<i>Calpocalyx dinklagei</i>	AF278483.1	AY944551	AY944551
<i>Calpocalyx heitzii</i>	AF278482.1	AF521817.1	-

<i>Cathormion umbellatum</i>	AF522949.1	AF274122.1	-
<i>Cedrelina cateniformis</i>	AF278511.1	AF521818.1	-
<i>Cercidium andicola</i>	AY944536.1	AY944552	AY944552
<i>Chloroleucon mangens</i>	AF522950.1	AF523072.1	AF523072.1
<i>Cyclicodiscus gabunensis</i>	AY125845.1	AF521819.1	-
<i>Desmanthus acuminatus</i>	AF278490.1	AF521820.1	-
<i>Desmanthus balsensis</i>	AF278531.1	AF521821.1	-
<i>Desmanthus bicornutus</i>	AF522939.1	AF523108.1	AF523108.1
<i>Dichrostachys paucifoliolata</i>	AF278526.1	AF521822.1	-
<i>Dichrostachys richardiana</i>	AF278519.1	AF521823.1	-
<i>Dichrostachys spicata</i>	AF278524.1	AF521824.1	-
<i>Dichrostachys unijuga</i>	AF278525.1	AF521825.1	-
<i>Dichrostachys venosa</i>	AF278521.1	AF521826.1	-
<i>Dinizia excelsa</i>	AF278479.1	AF521827.1	-
<i>Ebenopsis ebano</i>	AF522951.1	AF274123.1	AF274123.1
<i>Elephantorrhiza elephantina</i>	AF278484.1	AF521828.1	-
<i>Entada abyssinica</i>	AY125846.1	AF521829.1	-
<i>Enterolobium contortisiliquum</i>	AF522952.1	AF274124.1	AF274124.1
<i>Enterolobium cyclocarpum</i>	AF522953.1	AF523096.1	AF523096.1
<i>Faidherbia albida</i>	AF522954.1	AF523081.1	AF523081.1
<i>Fillaeopsis discophora</i>	AY125847.1	AF521832.1	-
<i>Gagnebina bakoliae</i>	AF278527.1	AF521834.1	-
<i>Gagnebina bernieriana</i>	AY125848.1	AF521835.1	-
<i>Gagnebina commersoniana</i>	AF278529.1	AF521836.1	-
<i>Gagnebina pervilleana</i>	AF278528.1	AF521837.1	-
<i>Gagnebina pterocarpa</i>	AF278530.1	AF521838.1	-
<i>Havardia albicans</i>	AF522956.1	AF523085.1	AF523085.1
<i>Havardia pallens</i>	AF522955.1	AF274125.1	AF274125.1
<i>Inga edulis</i>	AF522957.1	AF523078.1	AF523078.1
<i>Kanaloa kahoolawensis</i>	AF278489.1	AF521839.1	-
<i>Leucaena greggii</i>	AF278492.1	AF521840.1	-
<i>Leucaena leucocephala</i>	AF522942.1	AF523094.1	AF523094.1
<i>Lotus japonicus</i>	NC002694	NC002694	NC002694
<i>Lysiloma acapulcense</i>	AF522958.1	AF274126.1	AF274126.1
<i>Lysiloma divaricatum</i>	AF522940.1	AF523088.1	AF523088.1
<i>Lysiloma tergeminum</i>	AF522959.1	AF523089.1	AF523089.1
<i>Microlobius foetidus</i>	AF522960.1	AF523095.1	AF523095.1
<i>Mimosa aculeaticarpa</i>	AF278513.1	AF521843.1	-
<i>Mimosa quitensis</i>	AF278514.1	AF521844.1	-
<i>Mimosa tenuiflora</i>	AF522943.1	AF274120.1	AF274120.1
<i>Mimozyganthus carinatus</i>	AY944539.1	AY944556.1	AY944556.1
<i>Neptunia gracilis</i>	AF278494.1	AF521845.1	-
<i>Neptunia monosperma</i>	AF522944.1	AF523090.1	AF523090.1
<i>Newtonia buchananii</i>	AF278501.1	AF521847.1	-
<i>Newtonia hildebrandtii</i>	AF278502.1	AF521848.1	-
<i>Parapiptadenia rigida</i>	AF278505.1	AF521849.1	-
<i>Pararchidendron pruinosum</i>	AF522961.1	AF274127.1	AF274127.1
<i>Paraserianthes lophantha</i>	AF522962.1	AF274128.1	AF274128.1
<i>Parkia biglandulosa</i>	AF278498.1	AF521850.1	-
<i>Parkia speciosa</i>	AF278499.1	AF521851.1	-
<i>Parkia timoriana</i>	AF195701.1/ AF195682.1	AF523091.1	AF523091.1
<i>Pentaclethra eetveldtiana</i>	AY125850.1	AF521852.1	-
<i>Pentaclethra macrophylla</i>	AF278485.1	AF521853.1	-
<i>Piptadeniastrum africanum</i>	AF278488.1	AF521857.1	-
<i>Piptadenia flava</i>	AY574110.1	AY574095.1	AY574095.1
<i>Piptadenia moniliformis</i>	AF278496.1	AF521854.1	-
<i>Piptadenia obliqua</i>	AY125855.1	AF521855.1	-
<i>Piptadenia viridiflora</i>	AF522963.1	AF521856.1	-
<i>Piptadeniopsis lomentifera</i>	AY944541.1	AY944559.1	AY944559.1
<i>Plathymenia reticulata</i>	AF278509.1	AF521858.1	-
<i>Prosopidastrum mexicanum</i>	-	AY386919.1	-

<i>Prosopis juliflora</i>	AY574115.1	AY574098.1	AY574098.1
<i>Prosopis nigra</i>	AY944545	AY944562	AY944562
<i>Prosopis glandulosa</i>	-	AY386851.1	-
<i>Pseudoprosopis gilletii</i>	AY125851.1	AF521861.1	-
<i>Pseudosamanea guachapele</i>	AF522964.1	AF523079.1	AF523079.1
<i>Samanea saman</i>	AF522964.1	AF523073.1	AF523073.1
<i>Schleinitzia insularum</i>	AF278491.1	AF521862.1	-
<i>Tetrapleura tetrapteria</i>	AY125852.1	AF521864.1	-
<i>Xerocladia viridiramis</i>	EU004653	EU000438	EU000438
<i>Xylia africana</i>	AY125849.1	AF521866.1	-
<i>Zapoteca tetragona</i>	AF522966.1	AF523097.1	AF523097.1

2.2.2. Características de los genes analizados

Gen *g3pdH*: Nombre completo: Gliceraldehído-3-fostato-deshidrogenasa. Esta región nuclear utilizada previamente para estudios filogeográficos (Olsen & Schaal, 1999) no había sido empleada aun para estudios filogenéticos en plantas. Las secuencias de los cebadores fueron obtenidas de la bibliografía (Strand et al., 1997). Debido a la dificultad que se encontró al intentar amplificar algunas especies, se decidió generar cebadores específicos para el grupo teniendo en cuenta las secuencias de aquellas especies que lograron ser amplificadas y secuenciadas. La confección de estos cebadores, al igual que los generados para amplificar otras regiones, se realizó con la ayuda del programa *Primer* (Rozen & Skaletsky, 2000). Este programa busca el par (o los pares) de cebadores más convenientes para la amplificación de una secuencia dada a partir de considerar una serie de parámetros tales como temperatura de desnaturalización, posibilidad o no de generar estructuras secundarias, largo de los cebadores, largo de la secuencia, etc.

Gen *NIA*: Nombre completo: Nitrato Reductasa. La región secuenciada involucró el intrón 3 de dicho gen. Esta región del genoma nuclear fue elegida por el alto nivel de variación encontrada en otros grupos y además por existir cebadores universales para su amplificación (Howarth & Baum, 2002). A diferencia de los cebadores utilizados para amplificar las otras regiones de ADN, los cebadores universales eran degenerados, es decir, eran una mezcla de oligonucleótidos similares en su secuencia pero que difieren entre ellos en pocas bases y que tienen la ventaja de permitir amplificar un fragmento de DNA sin conocer la secuencia exacta

del sitio de hibridación de los cebadores. Debido a la imposibilidad de utilizar los cebadores degenerados para la secuenciación de este fragmento fue necesario clonar el fragmento amplificado y generar cebadores no degenerados específicos (ver Inciso 2.2.3.).

Intergen *psbC-trnS*: Región del genoma de cloroplasto que se encuentra entre los genes que codifican para el ARN de transferencia S (codon UAG) y región intergénica del gen que codifica para la apoproteína *C* del fotosistema II. Esta región había sido analizada previamente mediante la técnica de PCR-RFLP (ej. Parani et al., 2000) pero nunca había sido secuenciada para ser incluida en un análisis filogenético. Un análisis preliminar realizado sobre 5 especies del género indicó que la mayor parte de la variación estaba ubicada en los primeros 500 pb a partir de la región 3' del gen *trnS*. En consecuencia, sólo esta región fue subsecuentemente secuenciada para el resto de las especies.

Región *trnL-trnF*: Comprende el intergen entre los genes *trnL* y *trnF* y el intrón del gen *trnF* del genoma de cloroplasto. La región fue elegida por haber sido secuenciada en un gran número de leguminosas (Wojciechowski, 2003). Esta región fue amplificada en dos reacciones. En la primera se amplificó el intergen *trnL-trnF* con el par de cebadores *e – f* (Taberlet et al., 1991) mientras que en una segunda reacción se amplificó el intrón del gen *trnL* con el par de cebadores *c – d* (Taberlet et al., 1991).

Región *trnK-matK*: Esta región del genoma de cloroplasto comprende el intrón del gen *trnK* en el cual esta inmerso el gen *matK*, existiendo una amplia gama de cebadores para amplificar diferentes zonas de esta región. Para esta tesis se amplificaron dos fragmentos de esta región que se que en otras leguminosas presentaron los mayores niveles de variación nucleotídica y que además habían sido secuenciada un número importante de leguminosas (Miller & Bayer, 2001; Wojciechowski et al., 2004).

2.2.3. Técnicas moleculares

2.2.3.1. *Extracción de ADN.* La extracción de ADN se realizó a partir de diferentes tejidos dependiendo de la disponibilidad de los mismos en cada especie. De ser posible el ADN se extrajo de tejido vegetativo con el objeto de que el ADN extraído coincidiera con el material referencia. En la mayoría de los casos el material provino de hojas, excepto en caso de las especies áfilas donde el ADN fue extraído de los tejidos fotosintéticos de tallos jóvenes. Sólo se utilizaron semillas cuando era el único material disponible. Para la extracción de ADN se utilizó el kit de extracción DNAeasy Plant mini Kit (Qiagen) siguiendo el protocolo indicado por el fabricante (Cuadro 1).

Cuadro 1. Protocolo de Extracción de ADN utilizando el kit de Extracción DNAeasy

Plant Minikit (Quiagen)

- 1-. Moler el tejido de la planta o el hongo bajo nitrógeno líquido usando un mortero. Transferir el tejido en polvo y el nitrógeno líquido a un tubo de tamaño apropiado y dejar que el aire líquido se evapore. No permitir que la muestra se descongele. Continuar inmediatamente con el paso 2.
- 2-. Agregar 400 µl del Buffer AP1 y 4 µl de la solución de stock de la RNasa A (100 mg/ml) a un máximo de 100 mg de del tejido molido (peso fresco) o 20 mg (seco) y “vortexearlo” de manera vigorosa.
- 3-. Incubar la mezcla por 10 min a 65°C. Mezclar 2–3 veces durante el período de incubación invirtiendo el tubo.
- 4-. Agregar 130 µl de buffer AP2 para lisar, mezclar e incubar por 5 min en hielo.
- 5-. Volcar el lisado en la columna para centrifugación “QIAshredder” introducida en un tubo de 2 µl y centrifugar a máxima velocidad por 2 min.
- 6-. Transferir el líquido recolectado en el tubo de 2 ml a un nuevo tuvo sin mover el pellet.
- 7-. Agregar 1.5 volúmenes de Buffer AP3/E al lisado y mezclar pipeteando
- 8-. Agregar 650 µl de la mezcla del paso 7, incluyendo cualquier precipitado la columna de centrifugación DNeasy mini colocada en un tubo de 2 ml. Centrifugar por 1 min a 6000 x g y descartar el líquido recolectado en el tubo.
- 9-. Repetir el paso 8 con la muestra remanente y descartar el líquido recolectado en el tubo.
- 10-. Colocar la columna DNeasy en un tubo Nuevo de 2 ml, agregarle 500 µl de Buffer AW a la columna y centrifugar por 1 min a 6000x g. Descartar el líquido recolectado

11-. Agregar 500 μ l de Buffer AW a la columna y centrifugar por 2 min a maxima velocidad para secar la membrana.

12-. Transferir la columna a un tubo de 1.5 ml y pipetear 1000 μ l de Buffer AE precalentado (65°C) directamente en la membrana. Incubar por 5 min. a temperatura

La cuantificación de ADN genómico se realizó mediante la comparación en una corrida electroforética de la intensidad de bandas de las muestras y de un marcador *Lambda Eco RI-Hind III*. Las muestras fueron corridas durante 2 horas en un gel de agarosa al 0,8 %, sembrándose 5 μ l de ADN genómico eluído y de 2 μ l del marcador. Todos los geles de agarosa utilizados en la tesis tuvieron la misma composición (Cuadro 2) variando solamente el porcentaje de agarosa agregados a cada uno de ellos. Para disolver la agarosa se utilizó un horno microondas. El buffer de la cuba electroforética fue TAE 1X (Cuadro 2).

Cuadro 2. Preparación de gel de agarosa al 0,8% (100 mL)

2 mL TAE 50X (Tris-Acetate-EDTA)*
0.8 grs. de agarosa
98 mL de Agua destilada

*Composición del buffer TAE 50X:
242 g Tris Base
57.1 mL Ácido Acético Glacial
100 mL 500 mM EDTA, pH 8.0
600 mL ddH₂O

Para la visualización de las bandas en la preparación de los geles se incluyeron 0.5 μ g/ml de Bromuro de Etidio. Una vez corrido el gel, las bandas de ADN se observaron en un transiluminador UV. Los geles se fotografiaron con una cámara digital Olympus.

2.2.3.2. Amplificación y purificación.: Las condiciones de amplificación de los genes utilizados en el análisis filogenético se indican en la Tabla 4. Los cebadores utilizados para las reacciones de amplificación se indican en la Tabla 5. Todas las reacciones de PCR se realizaron en un volumen final de 50 μ l, e incluyendo 80 ng del ADN genómico, 1X del Buffer Taq, 2.5 mM de MgCl₂, 0.25 U de Taq Polimerasa, 0.2 mM de cada DNTP y 0.04 μ M

de cada cebador. Se buscó optimizar las condiciones de amplificación modificando los parámetros de amplificación (temperatura de apareamiento, tiempo de extensión y número de ciclos de amplificación). En el caso de los cebadores degenerados (como era el caso de los cebadores originales de NIA) se utilizó una estrategia de “touchdown” en el cual la temperatura de apareamiento es muy alta al principio de la amplificación y en sucesivos ciclos la temperatura disminuye hasta llegar a condiciones de amplificación más relajadas.

Tabla 4. Programas de amplificación utilizadas en cada uno de los marcadores. A: paso de desnaturalización previo al ciclado. B paso del ciclado propiamente dicho. C: paso final de extensión.

Fragmento	Paso	Numero de ciclos	Tiempo	Temperatura
G3pdh	A	35	2 min	94°C
	B1		1 min	94°C
	B2		1 min	52°C
	B3		2 min	72°C
	C		7 min	72°C
psbC-trnS	A	35	6 min	94°C
	B1		1 min	94°C
	B2		1.5 min	56°C
	B3		2 min	72°C
	C		7 min	72°C
NIA	A	10	5 min	94°C
	B1		1 min	94°C
	B2		1 min	52°C -1 p/ciclo
	B3		2 min	72°C
	B1	30	1 min	94°C
	B2		1 min	45°C
	B3		2 min	72°C
	C	1	7 min	72°C
trnK-matK	A	35	5 min	94°C
	B1		1 min	94°C
	B2		1.3 min	56°C
	B3		2 min	72°C
	C	1	7 min	72°C
trnL-trnF	A	35	5 min	94°C
	B1		1 min	94°C
	B2		1.5 min	56°C
	B3		2 min	72°C
	C	1	7 min	72°C

Los termocicladores utilizados fueron: Eppendorf Mastercycle (gentilmente prestado por la Dra. Lidia Poggio) y Techne Progene thermal cycler. En algunos genes, principalmente en los nucleares, el rendimiento de la amplificación fue bajo. Los primeros ensayos indicaron que el rendimiento luego de la amplificación era menor a 25 ng por tubo de 50 µl de reacción. Debido a esto, el número de reacciones necesarias llegar a la cantidad de producto requerida en el caso de *NIA* y *g3pdH* fue de alrededor de 10 tubos por individuo. Las 10 muestras fueron posteriormente colectadas en un solo tubo y llevado a su liofilización en una máquina Savanth SpeedVac, amablemente prestada por el Dr. Luis Scolaro. Posteriormente en producto se resuspendió en 50 µl y se purificó como el resto de las muestras.

Tabla 5. Secuencias de los cebadores utilizados para la amplificación de las regiones analizadas.

Gen	Cebador	Cita	Secuencia (sentido 5'-3')
G3pdh	G3pdhintF	Original	GAATGGAGGGGTGGAAGAG
	GPDX7F	Strand et al. ,1997	GATAGATTGGAATTGTTGAGG
	GPDXR9	Strand et al. ,1997	AAGCAATTCCAGCCTTGG
<i>psbC-trnS</i>	<i>psbCint</i>	Original	CGTTCTGCCAAGGCTGTAT
	<i>trnS</i>	Demesure et al., 1995	GGTCGAATCCCTCTCTCTC
	<i>psbC</i>	Demesure et al., 1995	GGTCGTGACCAAGAAACAC
NIA			
	<i>NIA_i3_F</i>	Howarth & Baum, 2002	AARTAYTGGTGYTGGTGYTTYTGGTC
	<i>NIA_i3_R</i>	Howarth & Baum, 2002	GAACCACARTTGTTCATCATDCC
	<i>NIAPRoR</i>	Original	GAACCAGCAGTTGTTCATCAT
matK-trnK	<i>NIAPRoF</i>	Original	ACTGGTGCTGGTGTGTTTGG
	<i>Ac283R</i>	Miller & Bayer, 2001	CACTGACGGCAAGCCCCCTCTG
	<i>trnK 3914F</i>	Miller & Bayer, 2001	GGGGTTGCTAACTAGG
	685F	Lavin et al., 2000	
<i>trnF-trnL</i>	1265R	Lavin, et al., 2000	
	<i>c</i>	Taberlet et al., 1991	CGAAATCGGTAGACGCTACG
	<i>d</i>	Taberlet et al., 1991	GGGGATAGAGGGACTTGAAC
	<i>e</i>	Taberlet et al., 1991	GGTTCAAGTCCCTCTATCCC
	<i>f</i>	Taberlet et al., 1991	ATITGAACGGTGACACGAG

Para purificar el producto de amplificación, las muestras se corrieron en geles de agarosa al 1.5% durante 2.5-3 horas. Posteriormente las bandas fueron cortadas con un cubre

objeto y colocadas en tubos de 1.5 ml para su posterior purificación. Si las muestras no eran purificadas en el momento, los tubos se guardaban en un freezer a -20 °C. Para la purificación de las bandas se utilizó el kit QIAquick Gel extraction. (Cuadro 3)

Cuadro 3. Protocolo de purificación de bandas cortadas a partir de un gel de agarosa utilizando el kit QIAquick Gel extraction

- 1-. Cortar el fragmento de ADN del gel de agarosa con un escalpelo limpio y filoso.
- 2-. Pesar el trozo de gel en un tubo y agregar 3 volúmenes de QG a un volumen de gel (100 mg ~ 100 µl).
- 3-. Incubar a 50°C por 10 min (o hasta que el trozo se haya disuelto totalmente). Para colaborar en su disolución, mezclar por vortex el tubo cada 2-3 minutos de incubación.
- 4-. Luego que el fragmento del gel se haya disuelto completamente, chequear si la mezcla del color es amarilla
- 5-. Agregar 1 volumen igual al del fragmento del gel de isopropanol y mezclar.
- 6-. Ubicar la columna QIAquick en un tubo de 2 ml
- 7-. Para que se pegue el ADN, aplicar la muestra a la columna QIAquick y centrifugar por 1 min.
- 8-. Descartar el líquido recuperado en el tubo y ubicar nuevamente la columna en el tubo de 2 ml.
- 9-. Agregar 0.5 ml de Buffer QG a la columna QIAquick column y centrifugar por 1 min.
- 10-. Para limpiar, agregar 0.75 ml de buffer PE a la columna QIAquick y centrifugar por 1 min.
- 11-. Descartar líquido recuperado en el tubo y centrifugar por un min a 13,000 rpm
- 12-. Colocar la columna en un tubo de 1.5 ml
- 13-. Para eluir el ADN, agregar 50 µl de EB Buffer (10 mM Tris·Cl, pH 8.5) al centro de la membrana de la columna y centrifugar por 1 min.

2.2.3.3. Clonado de fragmentos de ADN

Durante el proceso de amplificación y secuenciación fue necesario en algunos casos recurrir al clonado del fragmento amplificado por dos motivos. Uno de ellos fue que, debido a la baja eficiencia de algunas amplificaciones, la cantidad de producto no alcanzaba la

mínima cantidad requerida por la empresa de secuenciación para realizar esta tarea. Clonar los productos de reacciones con poco rendimiento fue una solución muy eficiente a este problema, ya que una vez inserto en el plásmido y transformadas las bacterias, el fragmento se podía amplificar de una manera muy sencilla utilizando cebadores con sitio de reconocimiento en el plásmido. La otra causa por la cual se clonó fue la intención de evaluar posibles copias parálogas del gen que se pueden reconocer al obtener más de 2 secuencias diferentes del gen analizado por individuo.

El primer paso para el clonado de los fragmentos fue someterlo a una reacción de “*Fill in*” en la cual la enzima *Klenow* extrae los residuos de Timina que deja la enzima *Taq polimerasa* utilizada para la amplificación. Las condiciones de reacción se detallan en la Tabla 6. Este paso es necesario para la combinación entre el fragmento a clonar y el plásmido. Posteriormente se realizó una reacción de ligación para combinar el fragmento a clonar dentro del plásmido *pbluescriptSK* (Startek). Las condiciones de reacción de detallan en la Tabla 7. Posteriormente estos plásmidos con la inserción fueron utilizados para transformar bacterias competentes. Las bacterias transformadas fueron sembradas en cajas de Petri con medio de cultivo LB Agar con ampicilina IPTG y X-Gal (Tabla 8). Las bacterias transformadas con el plásmido incluyendo la inserción fueron seleccionadas mediante el método de IPTG y Xgal. De esta manera las bacterias que no son transformadas no pueden crecer (no son resistentes a la ampicilina), las bacterias transformadas pero sin la inserción son de color azul y aquellas colonias que tienen la inserción presentan un color blanco. Se tomaron varias colonias blancas y posteriormente se generaron varios cultivos en medio líquido de 5 cm³. Luego de 16 horas de crecimiento se realizó una “miniprep” para obtener el plásmido contenido la inserción. Dicha purificación se realizó utilizando el kit *Genieprep* de la empresa Ambion. Para asegurarse de que el plásmido purificado presentara la inserción, se realizó una restricción con enzimas que permitieran liberar la inserción, corriéndose posteriormente la restricción en un gel de agarosa. Las plásmidos que contenían la inserción,

presentaban 2 bandas (la inserción y el plásmido linealizado). A partir de las secuencias obtenidas se generó un par de cebadores específicos con el fin de utilizarlos para la secuenciación directa de esta región intrónica

Tabla 6. Condiciones de la reacción de “fill-in”

Agua	4 μ l
DNTP’s (10mM)	1 μ l
Buffer Klenow	2 μ l
Klenow (5U/ μ l)	1 μ l
DNA (30ng/ μ l)	12 μ l
Volumen total	20μl

Tabla 7. Condiciones de la reacción de ligación

Vector	1,5 μ l (25ng)
Inserción	2 μ l
Enzima Eco RV	0,5 μ l
Ligasa	1,5 μ l
Buffer ligasa	2 μ l
Agua	2,5 μ l
Total	10 μl

Tabla 8. Condiciones del Medio de cultivo

Para 1 litro de medio de cultivo	
Bactrotriptona	10 g
Extracto levadura	5 g
NaCl	10 g
Agar	1,5 g
NaOH	0,2 ml

2.2.3.4. *Secuenciación:* Los productos purificados fueron enviados para su secuenciación a la empresa Macrogen Inc. (Corea).

2.2.4. Material Biológico

Se realizó un viaje de colección en el mes de marzo de 2004 que abarcó las provincias de La Pampa, Mendoza, San Luis y Neuquén. Durante el mismo se obtuvo material de especies de la sección Algarobia que no habían sido colectadas previamente en nuestro laboratorio, incluyendo: *P. campestris*, *P. castellanosii*, *P. ruizleali* y *P. humilis*.

2.3.3. Metodología del Análisis Filogenético

2.3.3.1 *Homologías estáticas vs homologías dinámicas*

Debido a la existencia de delecciones e inserciones en el genoma, el establecimiento de la correspondencia entre los sitios nucleotídicos de diferentes secuencias puede resultar una labor no trivial, más aun cuando también han ocurrido sustituciones entre las secuencias. La manera tradicional de trabajar en filogenias moleculares ha sido realizar alineamientos “a priori” en los que se establecen las líneas de correspondencia (qué sitio de una secuencia se corresponde con qué sitio de otra secuencia) y luego se analizan las secuencias alineadas de la misma manera que se trabajaba con caracteres “convencionales”. A mediados de la década pasada Wheeler (1996) propuso que el paso del alineamiento previo al análisis filogenético no era necesario y que en realidad la manera más correcta de analizar las secuencias sería evaluando la existencia de las inserciones y delecciones durante las búsquedas filogenéticas. Así, cada fragmento de ADN analizado sería un “carácter” cuyos cambios entre estados estarían dados por sustituciones, inserciones y delecciones. Una consecuencia de este modo de analizar el problema es que se puede prescindir de alineamientos múltiples. En conjunto estas aproximaciones se las denomina “Optimización Directa” ú “Optimization alignment” (Wheeler, 1996). Desde un punto de vista teórico considero este modo es el más adecuado para analizar secuencias de ADN que presentan inserciones y delecciones. Sin embargo la implementación de estas ideas choca contra un problema práctico: la complejidad computacional. En un comienzo el programa diseñado para realizar estas búsquedas (POY) (Wheeler et al., 2002) podía analizar sólo un número muy limitado de secuencias. Eso fue mejorando con el tiempo. Sin embargo cuando realicé los análisis de esta tesis los tiempos de búsqueda eran tales que negaban la posibilidad de encontrar soluciones que se acerquen lo suficiente a soluciones óptimas. Es así que para los análisis realizados en esta tesis se decidió utilizar una aproximación de homologías estáticas alineando las secuencias utilizando el programa Clustal X (Thompson et al., 1997), procediendo luego a pequeños arreglos

manuales. Los “gaps” o delecciones fueron tratados como quinto estado o como estado ausente.

2.2.3.2. Criterios de optimalidad en análisis filogenéticos

En la actualidad existe una variada gama de métodos de reconstrucción filogenética. Desde los años 80 los métodos basados en parsimonia jugaron un rol central entre los métodos elegidos por los investigadores a la hora de realizar sus análisis filogenéticos. La lógica de la utilización de parsimonia para reconstrucciones filogenéticas proviene de elegir aquella hipótesis (=filogenia) que permita explicar la mayor cantidad de observaciones (similitudes) ¿Cómo permite una filogenia explicar observaciones? Lo hace siempre que una similitud pueda ser explicada por ancestralidad común (Farris, 1984; Farris, 2008). Toda similitud no explicada es denominada homoplasia y la hipótesis filogenética más parsimoniosa es aquella en la cual la homoplasia es mínima. La idea de buscar una hipótesis que permita explicar la mayor cantidad de observaciones (pesadas) es una cuestión no solamente asociada a los análisis filogenéticos sino a cualquier rama de la ciencia: si existen dos hipótesis A y B, y la primera permite explicar una mayor cantidad de observaciones entonces A es preferida sobre B. Esto debe ser así ya que si se prescinde de esto la elección de una hipótesis no podría ser cuestionada en función de las observaciones. Así cualquier configuración posible del mundo se vuelve intrascendente para la elección de nuestra hipótesis ya que sin importar la cantidad de evidencia en contra, la misma se podría sostener.

A finales de la década del 90 comenzaron a tomar relevancia los análisis filogenéticos basados en el criterio de Máxima Verosimilitud (ML). La descripción original de este método ya había sido publicada mucho antes (Felsenstein, 1981) pero fue recién en 1998 con la implementación de este método en el programa PAUP vers 4.0 (Swofford, 1998) que el mismo logró una gran difusión. Una de las principales diferencias teóricas con los métodos basados en parsimonia es que dentro del marco de ML se considera que es posible modelar

estadísticamente y de una manera más o menos detallada los procesos evolutivos que ocurrieron a lo largo de la historia del grupo. Una de las ventajas que tendría este método es que el mismo presentaría lo que se denomina “consistencia estadística”, es decir que el resultado (árbol) sería el correcto para una cantidad de infinita de datos (caracteres) (ver abajo). Una desventaja práctica de los métodos de ML es el tiempo requerido para completar búsquedas aun para conjuntos de datos de tamaños apenas modestos (50 taxa). Esto aunque todavía cierto, ha comenzado a cambiar con el desarrollo de programas mucho más eficientes (ej. RaxML; Stamatakis, 2006). Como una solución a este problema y además siguiendo la corriente que había comenzado a surcar otras áreas de la biología, en el año 1996 (Li 1996; Mau, 1996; Rannala & Yang, 1996) tres grupos de trabajo desarrollaron independientemente aproximaciones para reconstrucciones filogenéticas basadas en estadística Bayesiana. Fue recién a partir de la publicación del programa MrBayes (Huelsenbeck & Ronquist, 2001) que este método de reconstrucción filogenética comenzó a ser utilizado y su uso se incrementó exponencialmente a lo largo de los años. Desde el punto de vista conceptual la principal diferencia con ML es que los métodos bayesianos permitirían introducir de una manera directa el grado de certeza que el investigador considera que tienen sobre determinado detalles de la evolución, definiendo probabilidades a priori (“priors”) de los diferentes parámetros. Sin embargo la principal (sino la única) causa de la gran aceptación de este método entre la comunidad científica es que el tiempo de cómputo es mucho menor que para el criterio de ML.

Ha habido muchísima discusión sobre si los métodos basados explícitamente en modelos evolutivos estadísticos son mejores o no que los basados en parsimonia. Uno de los trabajos más importantes, a mi entender, que ha posibilitado comparar los métodos de ML y parsimonia ha sido el de Farris (2000). En este trabajo, Farris plantea que el enfrentamiento entre parsimonia y métodos de ML surge sólo de que los métodos actuales de ML utilizan modelos no-realistas y que los árboles de máxima parsimonia (más corroborados por la

evidencia) serían también los de máxima verosimilitud bajo modelos más realistas de evolución. Este punto sumado a la fuerte justificación epistemológica que presenta parsimonia hizo que consideremos utilizar este método para los análisis filogenéticos de esta tesis.

Ante la existencia de los diferentes métodos filogenéticos previamente indicados algunos autores consideran que una buena práctica científica es analizar cada conjunto de datos con diferentes métodos, definiendo como “robustos” a aquellos resultados que no difieran entre los diferentes métodos (ver Giribet 2003). Esta idea es bastante pobre por diferentes razones. La primera de ellas es que existe la posibilidad que todos los métodos desarrollados hasta ahora tengan el mismo sesgo con lo cual podría haber una convergencia a una hipótesis filogenética falsa. El segundo punto es que si hay métodos que no están teóricamente bien fundados no tiene sentido que sean tenidos en cuenta.

A pesar de esta crítica al “pluralismo” considero que sí hay una posibilidad de una aproximación similar sustentada en una cuestión totalmente diferente que surge de la necesidad de comunicación entre la sociedad científica. Dos científicos ubicados en diferentes escuelas de pensamiento filogenético verán con cierta desconfianza los resultados obtenidos por el otro ya que considerarán que es posible que tales resultados no sean totalmente válidos. Aquí entonces se genera una necesidad de que los resultados puedan ser compartidos y aceptados por la comunidad científica. Si yo considero que parsimonia es el mejor método debo usar ese método. Ahora también podría usar por ejemplo métodos bayesianos ya no con el objeto de decir “mis resultados son más robustos ya que ambos métodos dan el mismo resultado” sino con la idea de decir “mis resultados son también válidos para aquellos que consideran que parsimonia no es un método válido”. Si los dos métodos dan resultados diferentes entonces la decisión es clara, solo derivaré conclusiones a partir de los resultados obtenidos con el método que considero como más correcto. Es esta visión restrictiva de “pluralidad” la que consideraremos en esta tesis. Así además de hacer un estudio basándome en un criterio de parsimonia, también realicé un análisis bayesiano.

2.2.3.3. Búsquedas heurísticas en Parsimonia

Idealmente sería deseable que durante la elección del árbol más parsimonioso se evaluaran cada una de las posibles topologías (i.e. búsqueda exacta). Dadas las actuales limitaciones de software y hardware esto es sólo posible para casos con un número de taxones muy limitado (ca. 20). Debido a ello para cualquier conjunto real de datos es necesario evaluar sólo una fracción de todas las topologías posibles. Las búsquedas heurísticas no aseguran llegar a la/s topología/s óptima/s pero permiten llegar a la/s misma/s en la mayoría de los casos. Existen dos principales búsquedas heurísticas para encontrar árboles más parsimoniosos: árboles de Wagner y técnicas de intercambio de ramas o “branch swapping”. En el primer caso se construye un árbol con un número crecientes de taxones y en cada paso se va agregando un taxón en la posición del árbol parcial que hace mínima la cantidad de pasos. El orden de incorporación de los taxones al árbol es al azar y por eso esta técnica también es conocida como secuencias de adición de al azar o RAS (Random Addition Sequence). Esta técnica permite construir árboles de manera muy rápida pero que generalmente están lejos de los óptimos. Una manera de mejorar estos árboles es mediante las técnicas de “branch swapping”. En las mismas se van cortando secuencialmente diferentes ramas del árbol que son reconectadas en otras ramas. Si el árbol con la nueva posición de la rama es más largo que el original se prueba conectar la rama en otro nodo hasta haber probado todas las posibilidades. Luego se pasa a cortar otra rama. Si el árbol con la nueva posición de la rama es más corto que el original, se lo guarda y se comienza de nuevo el proceso. Existen dos técnicas principales: SPR (Subtree Prunning and Regrafting) y TBR (Tree Bisection and Reconnection), la diferencia es que la segunda permite que los subárboles cortados puedan conectarse con el resto del árbol con enraizamientos diferentes de los que tenía, mientras que el primero solo se consideran re-arreglos con enraizamientos iguales al los que originalmente tenía en el árbol. Estas técnicas suelen permitir llegar a los árboles más parsimoniosos,

aunque en algunos casos eso no es posible. Esto se debe a que para llegar a un óptimo global desde una determinada topología se requiere en algunos casos “pasar” por árboles subóptimos, cosa que no es posible utilizando los algoritmos de TBR y SPR. Para resolver este problema existen dos posibilidades: aumentar el número de orígenes independientes de la búsqueda (secuencias de adición al azar) o aceptar durante la búsqueda árboles subóptimos. Esto último es lo que hacen métodos como Parsimony Ratchet (Nixon, 1999) y Tree Drifting (Goloboff, 1999). Cuando las matrices son más grandes (mayor a 150 taxones) la estrategia de aceptar subóptimos puede no ser suficiente. Esto se debe a que en árboles tan grandes es posible que existan problemas para encontrar resoluciones óptimas ya que se requiere que todos los sectores del árbol estén en una configuración óptima. Para resolver esto, en el método conocido como Sectorial Searches (Goloboff, 1999) se eligen sectores del árbol y se crea una matriz reducida con los taxones incluidos en ese sector. Luego se hace una búsqueda y si se encuentra un árbol más corto para ese subset que el original se reemplaza en el árbol completo. Esto se repite varias veces. Su eficiencia suele ser mayor cuando se está lejos del óptimo.

Todos los análisis filogenéticos fueron realizadas con TNT (Goloboff et al., 2008). En el análisis de tres genes la estrategia de búsqueda estuvo compuesta por 100 secuencias de adición al azar (RAS) guardando hasta 6 árboles por réplica, seguidos por 10 ciclos de “Tree Drifting”. Para el caso de los análisis de 2 marcadores y 5 marcadores, la estrategia de búsqueda involucró 200 RAS seguidas por Tree Drifting y Random Sectorial Searches. Las perturbaciones en Tree Drifting se realizaron cada 10 ciclos, aceptando re-arreglos cuando la diferencia en el ajuste o “*fit*” absoluto era menor que 1, o 20 en valores relativos, con un índice de rechazo de 30. La etapa perturbación finalizó cuando el número de reemplazos era de 10 ó se había finalizado el 99% de los re-arreglos posibles. En Sectorial Searches la estrategia utilizada para la elección de los sectores fue al azar (Random Sectorial Searches), con límites del tamaño de tales sectores entre 30 y 70 taxones. Si el sector elegido al azar era

mayor que 50 nodos, la búsqueda en dicho sector involucraba 5 ciclos de Tree Drifting. Una vez finalizada cada búsqueda. Los árboles resultantes se sometieron a una ronda de TBR guardando hasta 10.000 árboles en memoria. Para evaluar el largo de árboles que presentaban los grupos de interés que no aparecieron en el consenso estricto (por ejemplo todas las especies de un continente o un grupo taxonómico) se realizaron búsquedas bajo constreñimientos (constraints). Esto implica una búsqueda en la cual se descartan todos árboles que no presentan el grupo de interés. La estrategia de dichas búsquedas fueron las mismas que en las búsquedas originales.

2.2.3.4. Evaluación de apoyo de grupos

Un paso importante en estudios filogenéticos es la evaluación de la evidencia que apoya a cada uno de los grupos obtenidos. Si los caracteres pudieran clasificarse entre aquéllos que apoyan un cierto agrupamiento y aquéllos que los contradicen, el apoyo de la evidencia a la hipótesis de agrupamiento podría calcularse de un modo directo. Sin embargo esta clasificación no es posible ya que en algunos casos tal apoyo surge de la interacción de caracteres y no de la evaluación independiente de cada uno de ellos (Goloboff et al., 2003). Es así que se requieren de métodos indirectos de evaluación de apoyo. Uno de los métodos más utilizados es el índice de Bremer que mide la diferencia de optimalidad entre el árbol más corto con el grupo y el árbol más corto sin el grupo (Bremer, 1994). Cuanto mayor es este valor mayor será la evidencia que apoya tal agrupamiento. Otra serie de medidas de apoyo son las técnicas de remuestreo. Originalmente consideradas como una manera de evaluar estadísticamente la verosimilitud o estabilidad de los resultados (Efron, 1979; Felsenstein, 1985) más recientemente se han considerado también como una manera de evaluar la evidencia a favor o en contra de los agrupamientos (Farris et al., 1996). Existen diferentes técnicas de remuestreo que difieren en que porcentaje de los caracteres son elegidos para aumentarles o disminuirles su peso en cada réplica. El método de Jackniffing es una de estas

técnicas que asocia a cada carácter una probabilidad igual de eliminación. Así la matriz obtenida en cada caso presenta un número menor de caracteres que los que originalmente tenía la matriz.

En el presente análisis se utilizaron dos medidas de apoyo, el índice de Bremer y la técnica de Jackniffing. Para calcular el índice de Bremer, se hicieron búsquedas guardando árboles subóptimos en forma sucesiva de 1 a 12 pasos más largo que el óptimo, guardando en cada caso 10.000 árboles. En el análisis Jackniffing se consideró una probabilidad de remoción para cada carácter de 0.36 y se calcularon las frecuencias absolutas de cada grupo. En el análisis de tres marcadores, la estrategia de búsqueda en cada réplica de Jackniffing fue la misma que en la búsqueda original, mientras que para las matrices más grandes el número de réplicas fue menor (50).

2.2.3.5. Análisis bayesianos

Los análisis filogenéticos Bayesianos fueron realizados con el algoritmo de Metropolis-coupled Markov chain Monte Carlo implementado en el programa MrBayes versión 3.1.2 (Ronquist & Huelsenbeck, 2003). El modelo óptimo fue seleccionado con el Modeltest (Posada & Crandall, 1998) y los valores para cada parámetro fueron determinados durante la búsqueda. El análisis involucró dos corridas independientes de 2×10^6 generaciones muestreadas cada 100 generaciones luego de un período de *burnin* de 0.50×10^6 generaciones. Cada corrida involucró tres “cadenas calientes”. La convergencia se evaluó mediante la graficación de los valores de verosimilitud y los Desvíos Estándar de las frecuencias de las particiones.

2.3. Resultados

2.3.1. Características de las regiones genómicas estudiadas

El proceso de secuenciación de las muestras derivó en el depósito de 130 accesiones en la base de datos GenBank. En el caso de la región *NIA* se obtuvieron un total de 31

secuencias pertenecientes a 25 especies de *Prosopis* además de una secuencia para *Acacia caven*. La amplificación de esta región la amplificación de las especies de la sección Strombocarpa fue muy dificultosa aunque fue posible incluir secuencias de dos de sus especies: *P. palmeri* y *P. strombulifera*. La región amplificada varió según la especie entre 508 y 700 pb. de largo (Figura 6, izquierda). Al correr las muestras amplificadas en geles de agarosa en todas las especies se observó una segunda banda de 380 pb. Un análisis mediante Blast en la base de datos GenBank indicó que esta secuencia también correspondería una copia del gen Nitrato Reductasa. Para el análisis filogenético solamente utilizamos el mayor de los fragmentos.

El largo de la región amplificada del gen *g3pdh* con los cebadores G3pdhintF-G3pdhR varió entre 580 y 610 pb (Figura 6, Derecha). Todas las especies de *Prosopis* incluidas en el estudio y todos los grupos externos fueron amplificados exitosamente.

El fragmento amplificado correspondiente al intergen entre los genes *trnS-psbC* presentó ~1650 bp. Al comparar las secuencias obtenidas para esta región se pudo observar que el nivel de divergencia era muy bajo existiendo pocas sustituciones entre las especies analizadas. Sin embargo en algunas comparaciones entre especies aparecía una región de diez pares de bases donde se observaba lo que parecían ser múltiples sustituciones. Sin embargo tales secuencias eran el complemento reverso de las demás, sugiriendo que la diferencia entre las secuencias se debería a una inversión de este pequeño fragmento. Al analizar con mayor detenimiento fue posible determinar que dicha región se hallaba flanqueada con secuencias repetidas invertidas. Un análisis bibliográfico indicó que todas las pequeñas inversiones en cloroplasto están flanqueadas por estas repeticiones invertidas, apoyando la posibilidad que en nuestro caso las secuencias difieran por una inversión. Esta región fue excluida del análisis filogenético pero fue objeto de un análisis detallado, que se presenta en el Capítulo 3. En el mismo estudió la ocurrencia de la inversión en *Prosopis* y en Angiospermas, buscando además determinar los factores que afectarían su ocurrencia.

La región *trnL-trnF* fue amplificada en la mayoría de los casos en dos reacciones: una de ellas con el par de cebadores *e – f* mientras que para la segunda reacción se utilizó el par *c – d*. La región amplificada con el par *e – f* tuvo aproximadamente 350 pb. mientras que la región amplificada con el par *c – d* tuvo 500 pb. La amplificación no presentó dificultades en ninguna muestra. La secuencia de *P. cineraria* resultó ser mucho más corta que las demás especies secuenciadas. Una comparación con otras secuencias de Mimosoideas permitió reconocer que el menor tamaño de la región amplificada en esta especie se debía a una delección de 300 pb en la región intergénica. La presencia de tal delección en diversos grupos de Leguminosas indica claramente que sería un evento recurrente dentro de la familia y que se vería favorecido por la existencia de secuencias repetidas en los extremos de tal región.

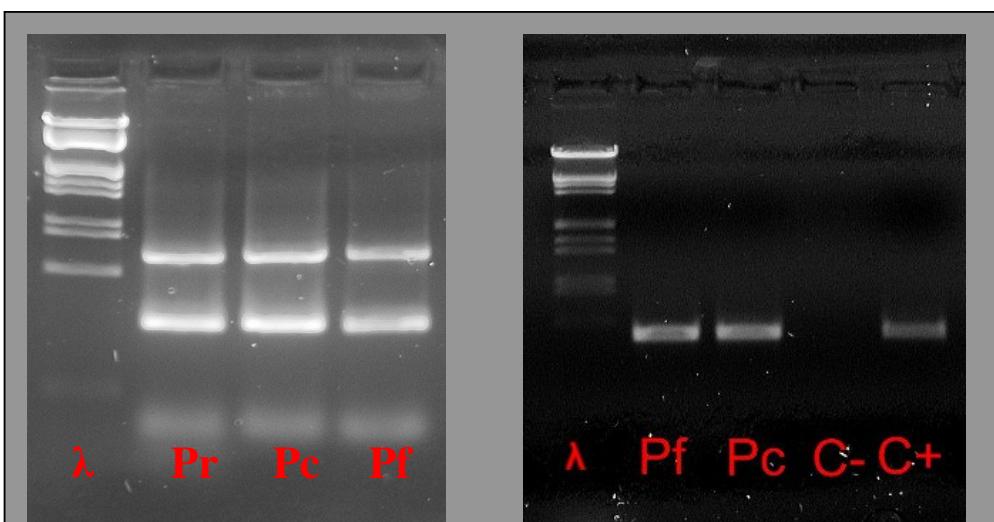


Figura 6. Amplificaciones de las regiones *NIA* (izquierda), *g3pdh* (derecha. Pr: *P. ruizleali*, Pc: *P. caldenia*, Pf: *P. flexuosa*, C-: control negativo. C+: Control positivo. λ : Marcador de peso molecular.

La región *trnK-matK* también fue amplificada en dos reacciones pero en este caso las regiones involucradas no fueron contiguas. Uno de los pares de cebadores elegidos amplificó la región codificante del gen *matK* (~500 pb) mientras que el otro par fue utilizado para amplificar la región intrónica del gen *trnK* (~410 pb) (Ver Tabla 5).

2.3.2. Resultado de las búsquedas filogenéticas

Los resultados obtenidos en los diferentes análisis filogenéticos fueron altamente congruentes entre sí y mostraron además una congruencia en términos generales con la separación en secciones del Género. Sin embargo los tres conjuntos de datos analizados coinciden en indicar la no monofilia de *Prosopis* debido a: (i) la posición de *X. viridinamis* junto con las especies americanas de *Prosopis*, (ii) *P. africana* no aparece junto con el resto de las especies del género en ninguno de los análisis, (iii) *P. cineraria* no aparece junto con las demás especies de *Prosopis*.

2.3.2.1. Análisis de dos marcadores. El consenso estricto de los árboles más parsimoniosos (AMP) indicó que el género *Prosopis* no sería monofilético (Figura 7, Tabla 9). El árbol más parsimonioso obtenido al forzar la monofilia de *Prosopis* tuvo 9 pasos más que el óptimo indicando que existiría fuerte evidencia en contra de esta hipótesis. Las especies Americanas de *Prosopis* (EAP) y *Xerocladia viridiramis* formaron un clado monofilético altamente apoyado (Apoyo de Bremer, $B = 4$; Jackniffing, $J = 95$; grupo llamado subsecuentemente EAP + *Xerocladia*). El grupo de las EAP no fue monofilético ya que *X. viridiramis* apareció como grupo hermano de *P. tamarugo*. Sin embargo este grupo estuvo pobemente apoyado ($B = 1$; $J = 66$). A pesar de que el grupo EAP no fue monofilético, el árbol más parsimonioso que contenía tal grupo (es decir cuando las búsquedas se realizaron forzando la monofilia del mismo) tuvo tan solo un paso más (Tabla 10). *P. cineraria*, la única especie asiática incluida en el estudio apareció como grupo hermano del clado EAP + *Xerocladia*, aunque con bajo apoyo ($B = 1$; $J < 50$). *P. nigra* de la sección Algarobia formó un grupo altamente apoyado ($B = 5$; $J = 98$) con *P. argentina* de la sección Monilicarpa. *P. africana* se ubicó en el árbol lejos del grupo formado por el resto de las especies de *Prosopis*, en una posición más basal. El árbol más parsimonioso en el cual se encuentra *P. africana* con el resto de las especies de *Prosopis* + *X. viridinamis* es 4 pasos más largo que el AMP indicando que existe una relativa

evidencia en contra de este agrupamiento. Es así que según este análisis *Prosopis* no sería un grupo natural, siendo de hecho parafilético en términos de *X. viridinamis* y polifilético en términos de *P. africana*.

El análisis Bayesiano dio resultados idénticos con respecto a la no monofilia de *Prosopis* y la relación entre las EAP. Sin embargo, a diferencia del análisis de parsimonia *P. cineraria* no formó un clado monofilético con el grupo EAP + *Xeroacladia*. Las probabilidades a posteriori se indican en la Figura 7.

Tabla 9. Estadísticos de los diferentes análisis filogenéticos. AMP, árboles mas parsimoniosos

	Análisis de 2 marcadores	Análisis de 3 marcadores	Análisis de 5 marcadores
Número de AMP	224	9	4680
Largo	5079	1552	6710
Nº de caracteres	4031(1444-2587) ¹	1764 (648-701-415) ²	5795
Nº caracteres inf.	1357 (609-748) ¹	391 (193-183-15) ²	1458
CI / RI	0.496/ 0.655	0.695 / 0.846	0.535 /0.700

CI, índice de consistencia (excluyendo caracteres no informativos); RI, índice de retención

¹ Estadísticos de los caracteres de la región *trnL-trnF* y *trnK-matK* respectivamente.

² Estadísticos de los caracteres de la región G3pdh, NIA y *trnS-psbC* respectivamente

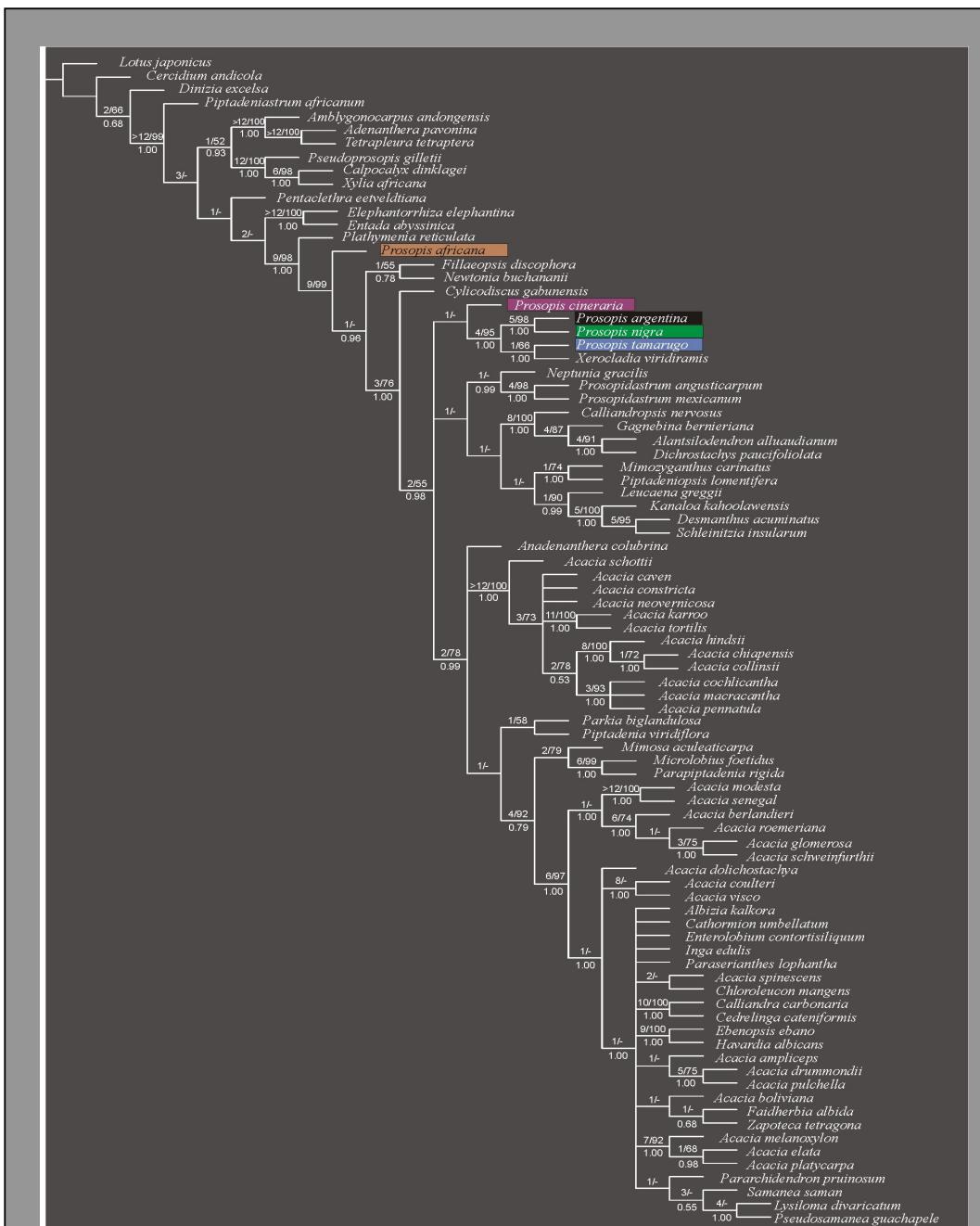


Figura 7. Consenso estricto de 224 árboles óptimos en el análisis de dos marcadores (*trnL-trnF* y *matK-trnK*) considerando a parsimonia como criterio de optimalidad. Se muestran arriba de las ramas el apoyo de Bremer absoluto (izquierda) y los valores Jackknifing por arriba del 50% (derecha) y debajo de las ramas las probabilidades a posteriori estimadas en el análisis bayesiano Azul: sec. Strombocarpa; Verde sec. Algarobia; Negro: sec. Monilicarpa; Naranja: sec. Anonichium; Bordo sec. *Prosopis*

Tabla 10. Pasos extras obtenidos cuando las búsquedas se realizaron forzando la monofilia de algunos grupos que no aparecieron en los árboles más parsimoniosos. Entre paréntesis se da el largo óptimo sin restricciones

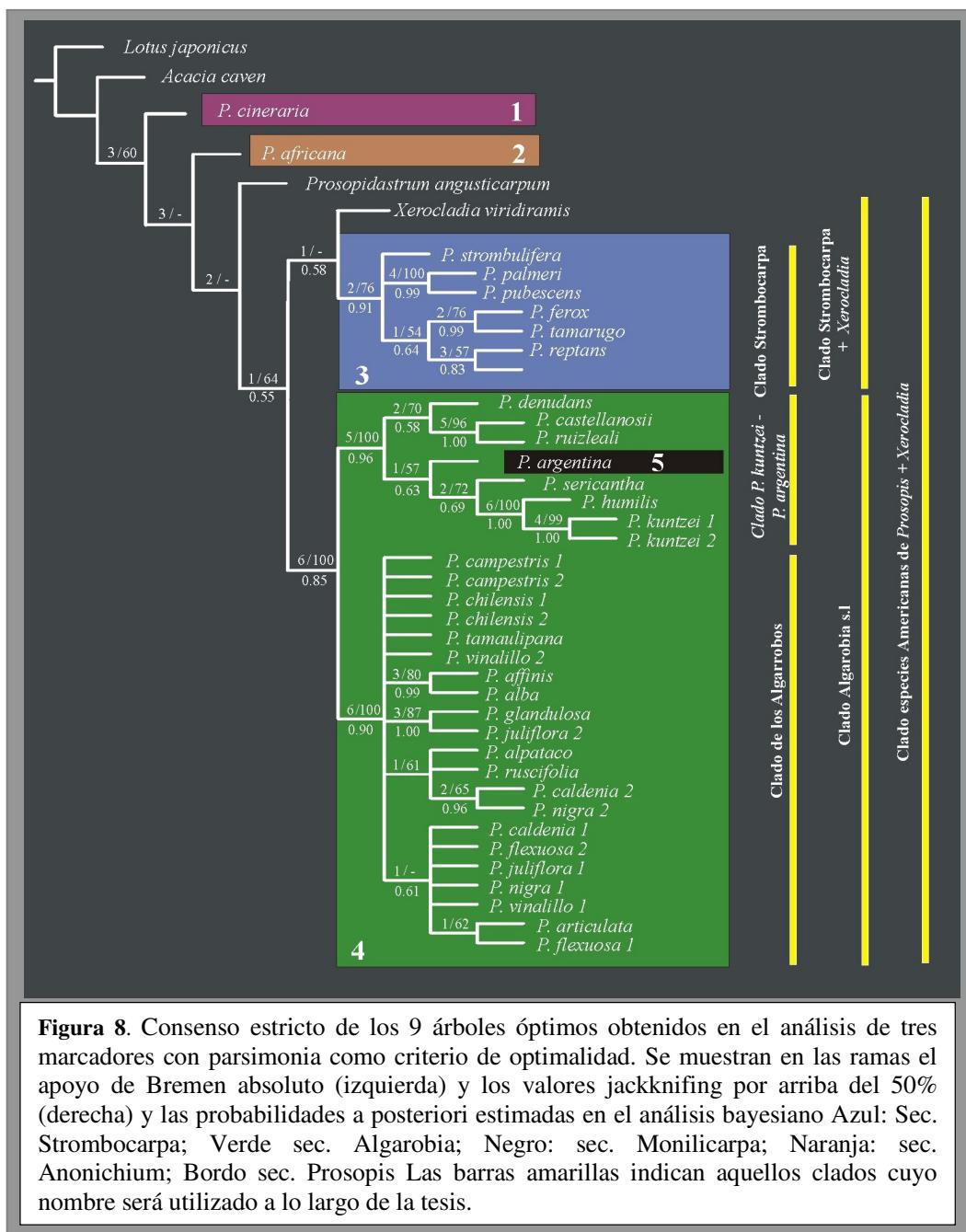
GRUPO	Análisis de 2 marcadores (5079)	Análisis de 3 marcadores (1552)	Análisis de 5 marcadores (6710)
sección <i>Strombocarpa</i> + <i>P. cineraria</i>	6	63	68
sección <i>Algarobia</i>	NA	10	10
<i>Sp. VM de Prosopis</i>	4	5	11
<i>Sp. Am. de Prosopis</i>	1	1	3
<i>Prosopis</i>	9	6	11
<i>Prosopis</i> + <i>Xerocladia</i> + <i>Prosopidastrum</i>	5	0	3
<i>Prosopis</i> + <i>Xerocladia</i>	4	5	5
<i>Prosopis</i> + <i>Prosopidastrum</i>	14	3	14
<i>Prosopidastrum</i> + <i>Xerocladia</i>	6	1	9
<i>Sp. Am. de Prosopis</i> + <i>Xerocladia</i> + <i>Prosopidastrum</i>	1	1	3
<i>Sp. VM de Prosopis</i> + <i>Xerocladia</i>	8	6	11
<i>Sp. Am. de Prosopis</i> + <i>Prosopidastrum</i>	10	1	9

Sp. VM = Especies del Viejo Mundo. **Sp. Am.** = Especies Americanas. NA = no aplicable.

2.3.2.2. Análisis de tres marcadores

El consenso estricto obtenido en este análisis estuvo de acuerdo con los resultados obtenidos en el análisis de 2 marcadores al contradecir la monofilia de *Prosopis* (Figura 8). Al forzar la monofilia del género se obtuvo un árbol seis pasos más largo que el árbol óptimo (Tabla 10). En este análisis al igual que el análisis de 2 marcadores *Prosopis* no fue monofilético debido a la posición de *X. viridinamis* entre especies de *Prosopis*. A diferencia del estudio de 2 marcadores, *P. cineraria* no formó un grupo monofilético con las EAP + *Xerocladia* sino que en este caso el grupo hermano fue *Prosopidastrum angusticarpum*. Otra diferencia entre el estudio de dos marcadores y el estudio de tres marcadores es la posición relativa entre *P. cineraria*, *P. africana* y el grupo formado por las EAP + *Xerocladia*. Mientras en el primero las relaciones relativas entre ellos fueron (*P. africana* (*P. cineraria* “EAP + *Xerocladia*”)) en el caso de tres marcadores las relaciones implicadas por

el consenso estricto fue (*P. cineraria* (*P. africana* “EAP + *Xerocladia*”)). Sin embargo el apoyo de estos agrupamientos fue muy bajo tanto al ser evaluado mediante el índice de Bremer como mediante Jackniffing. El apoyo del grupo formado por EAP + *Xerocladia* fue bajo ($B= 1$; $J = 64$) y la monofilia del grupo formado por las EAP se logró en un árbol un paso más largo que el óptimo. Dentro de las EAP + *Xerocladia* se formaron dos clados. Uno de ellos conformado por las especies de la sección *Strombocarpa* más *X. viridinamis*, mientras que el segundo clado estuvo formado por las especies de la sección *Algarobia* más *P. argentina* de la sección *Monilicarpa* (subsecuentemente llamado *Algarobia* sensu lato: *Algarobia* s.l.). Este grupo presentó un fuerte apoyo ($B= 6$; $J = 100$). La sección *Algarobia* fue parafilética ya que *P. argentina* formó un clado fuertemente apoyado ($B= 5$; $J = 100$) con especies de las secciones *Humiles*, *Sericantha* y *Denudantes* de la sección *Algarobia* (subsecuentemente llamado clado *P. argentina-P. kuntzei*). Los árboles más parsimoniosos que presentaron a la sección *Algarobia* como monofilética tuvieron diez pasos más que los árboles óptimos. La serie *Denudantes* (representada aquí por *P. denudans*, *P. ruizleali* y *P. castellanosii*) fue monofilética, no así la serie *Sericantha* (*P. kuntzei* y *P. sericantha*) ya que *P. humilis* de la serie *Humiles* apareció como grupo hermano de los dos terminales de *P. kuntzei*. El grupo hermano del clado *P. kuntzei* - *P. argentina* estuvo formado por especies de las series *Chilenses*, *Pallidae* y *Ruscifoliae*, denominado de aquí en más clado de los *Algarrobos*. Este clado estuvo fuertemente apoyado ($B= 6$, $J = 100$), sin embargo las relaciones entre especies dentro de este grupo no pudieron resolverse debido a que el consenso estricto estuvo altamente irresuelto y que incluso los diferentes individuos de la misma especie no aparecieron formando un clado monofilético.



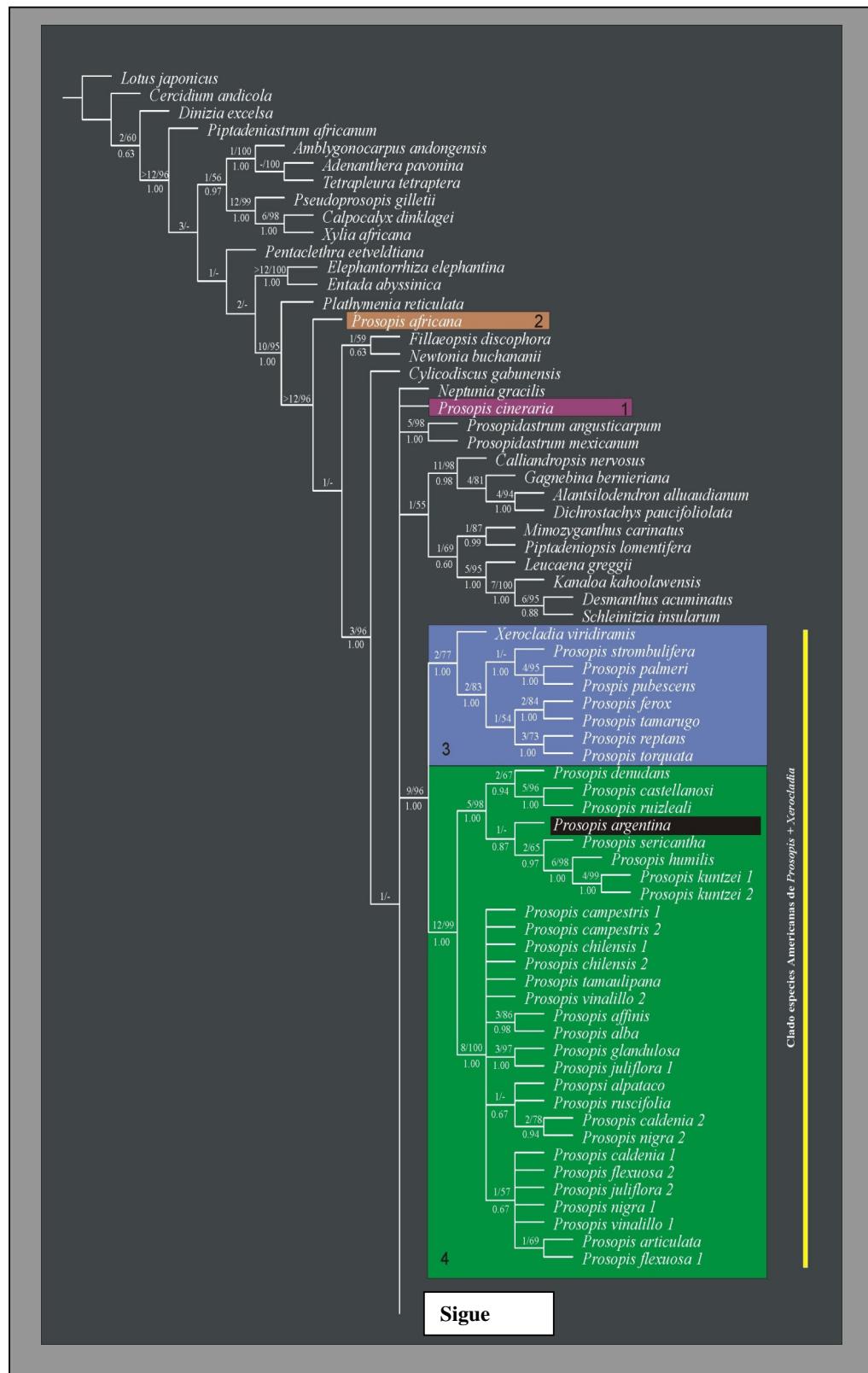
La sección Strombocarpa fue monofilética con *X. viridiramis* como grupo hermano.

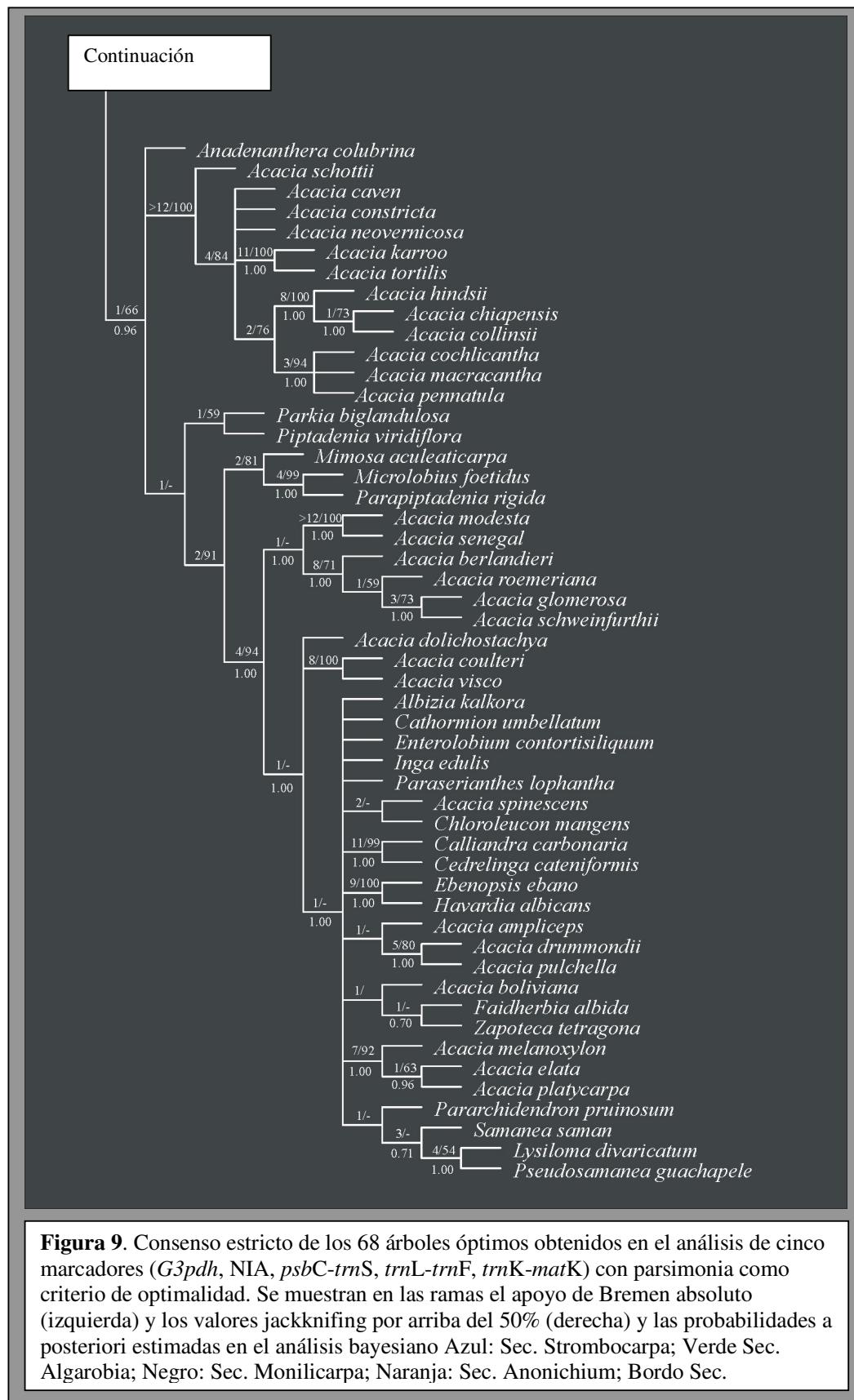
Dentro del grupo formado por especies de esta sección, aparecieron dos grupos fuertemente apoyados. Uno de ellos correspondía a las especies de la serie Cavenicarpeae (*P. ferox* y *P. tamarugo*) mientras que el otro grupo estuvo formado por dos especies americanas de la serie Strombocarpae: *P. pubescens* y *P. palmeri*.

El análisis Bayesiano mostró resultados muy similares al estudio de parsimonia. La principal diferencia estuvo relacionada con el grupo hermano de EAP + *X. viridiramis* ya que en el análisis Bayesiano estuvo formado por *Acacia caven* y *P. angusticarpum*.

2.3.2.3. Análisis de cinco marcadores

El resultado de este análisis fue totalmente concordante con el de los otros dos análisis (Figura 9). El principal punto a destacar aquí es que el análisis de los 5 marcadores de forma combinada incrementó el apoyo de varios grupos relevantes: el clado EAP + *Xerocladia* (B= 9; J = 96); *Strombocarpa* (B= 2; J = 83); clado *Strombocarpa* + *X. viridiramis* (B= 2; J = 77). Sin embargo, a diferencia del estudio de dos marcadores, *P. cineraria* no apareció como grupo hermano del clado EAP + *Xerocladia* aunque las posiciones relativas entre *P. africana* – *P. cineraria* y EAP + *Xerocladia* fueron las mismas. El consenso estricto indicó relaciones idénticas dentro del clado EAP + *Xerocladia* que en el estudio de tres marcadores aunque con menor resolución para los grupos dentro del clado de los Algarrobos.





2.4. Bibliografía

- Alvarez, I. & J. F. Wendel. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29: 417-434.
- Baldwin, B. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the Compositae. *Molecular Phylogenetic and Evolution* 1: 3–16.
- Bessegá, C., J.C. Vilardi & B.O. Saidman. 2006. Genetic relationships among American species of the genus *Prosopis* (Mimosoideae, Leguminosae) inferred from ITS sequences: evidence for long-distance dispersal. *Journal of Biogeography* 33: 1905-1915.
- Bremer, K. 1994. Branch support and tree stability. *Cladistics* 10: 295-304.
- Burghardt, A. & S. Espert. 2007. Phylogeny of *Prosopis* (Leguminosae) as shown by morphological and biochemical evidence. *Australian Journal of Botany* 20: 332–339.
- Cronn, R., R. Small, T. Haselkorn & J. Wendel. 2002. Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. *American Journal of Botany* 89: 707-725.
- Demesure, B., N. Sodzi & R. J. Petit. 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.
- Efron, B. 1979. Bootstrap methods: another look at the jackknife. *Annals of Statistics* 7: 1-26.
- Evans, R., L. Alice, C. Campbell, E. Kellogg & T. Dickinson. 2000. The Granule-Bound Starch Synthase (GBSSI) Gene in the Rosaceae: Multiple Loci and Phylogenetic Utility. *Molecular Phylogenetics and Evolution* 17: 388-400.
- Farris, J. 1984. The logical basis of phylogenetic analysis. *Conceptual Issues in Evolutionary Biology*. MIT Press, Cambridge, MA: 675–702.
- Farris, J. 2000. Corroboration versus “Strongest Evidence”. *Cladistics* 16: 385–393.
- Farris, J. 2008. Parsimony and explanatory power. *Cladistics* 24: 825-847.
- Farris, J., V. Albert, M. Källersjö, D. Lipscomb & A. Kluge. 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* 12: 99-124.
- Felsenstein, J. 1981. Evolutionary trees from DNA sequences: A maximum-likelihood approach. *Journal of Molecular Evolution* 17: 368-76.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783-791.

- Giribet, G. (2003). Stability in Phylogenetic Formulations and Its Relationship to Nodal Support. *Systematic Biology* 52:554-564
- Goloboff, P.A, J. Farris, M. Källersjo, B. Oxelman, M. Ramírez & C. Szumik. 2003. Improvements to resampling measures of group support. *Cladistics* 19: 324–332.
- Goloboff, P. A. 1999. Analyzing large data sets in reasonable times: solutions for composite optima. *Cladistics* 15: 415-428.
- Goloboff, P. A., J. S. Farris & K. Nixon. 2008. TNT: Tree Analysis Using New Technology. Program and documentation available at <http://www.zmuc.dk/public/phylogeny>.
- Howarth, D. G. & D. A. Baum. 2002. Phylogenetic utility of a nuclear intron from nitrate reductase for the study of closely related plant species. *Molecular Phylogenetics and Evolution* 23: 525-528.
- Huelsenbeck, J. & F. Ronquist. 2001. MrBayes: Bayesian inference of phylogeny. *Bioinformatics* 17: 754–755.
- Lavin, M., M. Thulin, J. Labat & R. T. Pennington. 2000. Africa, the odd man out: molecular biogeography of Dalbergioid legumes (Fabaceae) suggests otherwise. *Systematic Botany* 25: 449–467.
- Li , S. 1996. Phylogenetic tree construction using Markov chain Monte Carlo, Phylogenetic tree construction using Markov chain Monte Carlo. University of Columbus, Ohio State, Ph.D. Dissertation.
- Malcomber, S. 2002. Phylogeny of *Gaertnera* Lam. (Rubiaceae) based on multiple DNA markers: evidence of a rapid radiation in a widespread, morphologically diverse genus. *Evolution* 56: 42-57.
- Matthee, C. A., B. Jansen Van Vuuren, D. Bell & T. J. Robinson. 2004. A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene. *Systematic Biology* 53:433–447.
- Mau, B. 1996. Bayesian phylogenetic inference via Markov chain Monte Carlo methods., Bayesian phylogenetic inference via Markov chain Monte Carlo methods. Ph.D. Dissertation, Univ. Wisconsin, Madison., .
- McMahon, M. & M. Sanderson. 2006. Phylogenetic Supermatrix Analysis of GenBank Sequences from 2228 Papilionoid Legumes. *Systematic Biology* 55.
- Miller, J. T. & R. J. Bayer. 2001. Molecular phylogenetics of *Acacia* (Fabaceae: Mimosoideae) based on the chloroplast matK coding sequence and flanking trnK intron spacer regions. *American Journal of Botany* 88: 697-705.

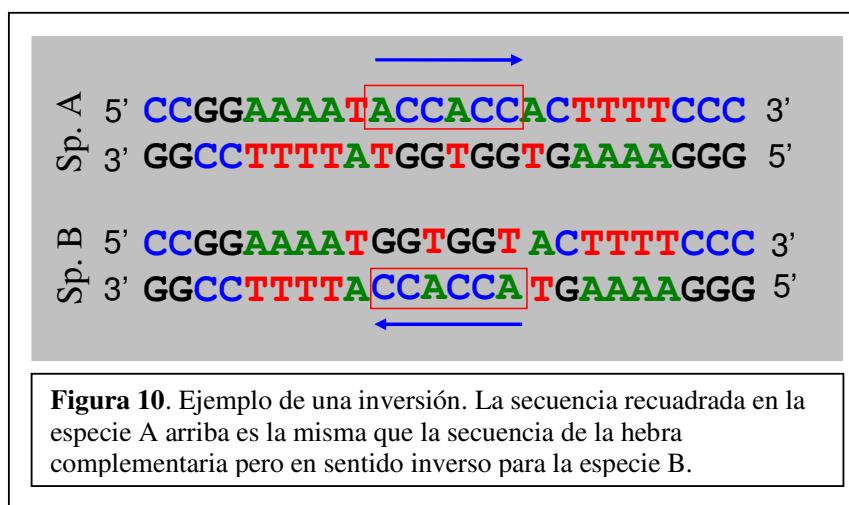
- Nixon, K. 1999. The Parsimony Ratchet, a New Method for Rapid Parsimony Analysis. *Cladistics* 15: 407–414.
- Olsen, K. & B. Schaal. 1999. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences* 96: 5586-5591.
- Parani, M., M. Lakshmi, B. Ziegenhagen, M. Fladung, P. Senthilkumar & A. Parida. 2000. Molecular phylogeny of mangroves VII. PCR-RFLP of *trnS-psbC* and *rbcL* gene regions in 24 mangrove and mangrove-associate species. *Theoretical and Applied Genetics* 100: 454-460.
- Pansarin, E., A. Salatino & M. Salatino. 2008. Phylogeny of South American Pogonieae (Orchidaceae, Vanilloideae) based on sequences of nuclear ribosomal (ITS) and chloroplast (psaB, *rbcL*, *rps16*, and *trnL-F*) DNA, with emphasis on *Cleistes* and discussion of biogeographic implication. *Organisms Diversity & Evolution* 8: 171-181.
- Posada, D. & K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14: 817-818.
- Ramírez, L., A. De La Vega, N. Razkin, V. Luna & P. J. C. Harris. 1999. Analysis of the relationships between species of the genus *Prosopis* revealed by the use of molecular markers. *Agronomie* 19: 31-43.
- Rannala, B. & Z. Yang. 1996. Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. *Journal of Molecular Evolution* 43: 304-311.
- Roalson, E. & E. Friar. 2004. Phylogenetic analysis of the nuclear alcohol dehydrogenase (Adh) gene family in *Carex* section *Acrocystis* (Cyperaceae) and combined analyses of Adh and nuclear ribosomal ITS and ETS sequences for inferring species relationships. *Molecular Phylogenetics and Evolution* 33: 671-686.
- Ronquist, F. & J. Huelsenbeck. 2003. Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 19: 1572-1574.
- Rozen, S. & H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. *Primer3 on the WWW for general users and for biologist programmers*. Humana Press, pp. 365-386.
- Sang, T. & D. Zhang. 1999. Reconstructing hybrid speciation using sequences of low copy nuclear genes: hybrid origins of five *Paeonia* species based on Adh gene phylogenies. *Systematic Botany* 24: 148-163.
- Stamatakis, A. 2006. RAxML-VI-HPC: Maximum Likelihood-based Phylogenetic Analyses with Thousands of Taxa and Mixed Models. *Bioinformatics* 22: 2688–2690.

- Steel, M., L. Szekely & M. Hendy. 1994. Reconstructing trees when sequence sites evolve at variable rates. *Journal of Computation Biology* 1: 153-163.
- Strand, A. E., J. Leebens-Mack & B. G. Milligan. 1997. Nuclear DNA-based markers for plant evolutionary biology. *Molecular Ecology* 6: 113-118.
- Swofford, D. 1998. PAUP: Phylogenetic Analysis Using Parsimony (*and Other Methods). Ver. Version 4.0. Sinauer Associates, Inc. Publishers.
- Taberlet, P., L. Gielly, G. Pautou & J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105-1109.
- Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin & D. G. Higgins. 1997. The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acid Research* 25.
- Wheeler, W. 1996. Optimization alignment: The end of multiple sequence alignment in phylogenetics? *Cladistics* 12: 1-9.
- Wheeler, W. C., D. S. Gladstein & J. De Laet. 2002. POY. Version 3.0 Source code and binaries available at <ftp://ftp.amnh.org/pub/molecular/poy>. Documentation by D. Janies and W. Wheeler.
- Wojciechowski, M. 2003. Reconstructing the phylogeny of legumes (Leguminosae): an early 21st century perspective. *En* B. Klitgaard and A. Bruneau (eds), *Advances in legume systematics*, Royal Botanic Gardens, Kew, pp. 5-35.
- Wojciechowski, M., M. Lavin & M. Sanderson. 2004. A phylogeny of legumes (Leguminosae) based on analysis of the plastid matK gene resolves many well-supported subclades within the family. *American Journal of Botany* 91: 1846-1862.
- Yuan, Y., P. Kupfer & J. Doyle. 1996. Infrageneric phylogeny of the genus *Gentiana* (Gentianaceae) inferred from nucleotide sequences of the internal transcribed spacers (ITS) of nuclear ribosomal DNA. *American Journal of Botany* 83: 641-652

CAPITULO 3 –ANÁLSIS DE INVERSIONES

3.1. Introducción

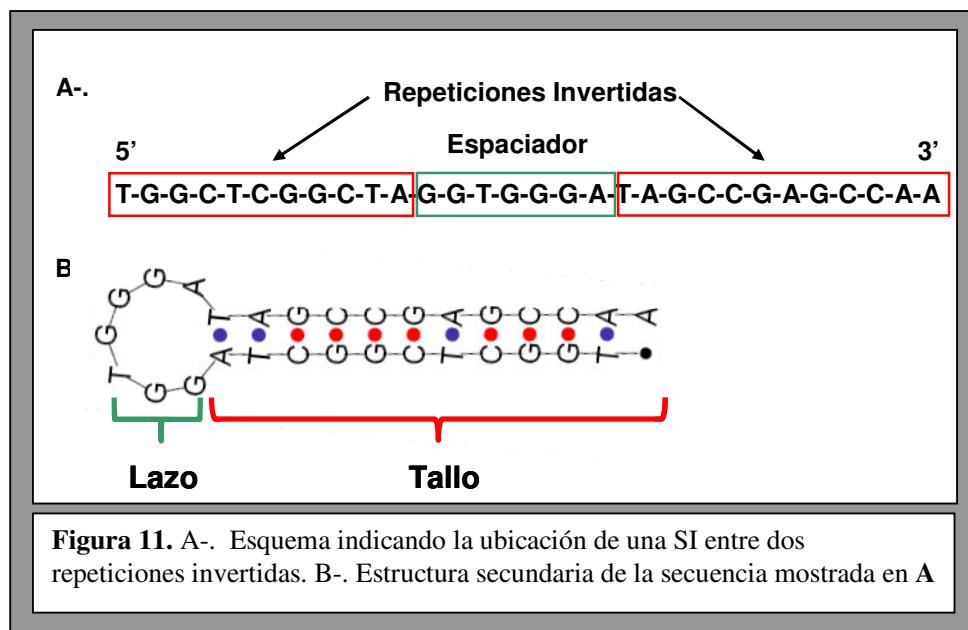
En este capítulo se presenta un estudio que en un principio no formaba parte del plan original de Tesis pero tuvo su origen en un descubrimiento realizado durante la realización de la misma. Como fue indicado en el capítulo II, durante el proceso de secuenciación de la región intergénica *psbC-trnS* se descubrió una pequeña inversión de 10 pares de bases. La misma fue reconocida debido a que el intergen *psbC-trnS* presentaba muy bajos niveles de divergencia excepto en un segmento de 10 pares de bases en los que se observaban muchas diferencias nucleotídicas entre algunas de las especies. Al realizar un análisis más detallado fue posible observar que si se tomaba el fragmento de una de las especies y se calculaba su complemento reverso (es decir se tomaba la secuencia de la hebra complementaria pero en sentido inverso) la secuencia resultante era idéntica a las demás secuencias. Esto es lo que se espera al ocurrir una inversión en el ADN ya que debido a direccionalidad (5'- 3') de esta molécula las inversiones en la misma sólo pueden ocurrir de manera tal que la secuencia que sufrió la inversión es el complemento reverso de la secuencia original, siendo imposibles inversiones directas del tipo ACTG a GTCA (Figura 10).



El estudio que aquí se presenta estuvo dirigido a investigar cuestiones metodológicas y empíricas sobre la evolución de esta inversión. En los últimos años se ha publicado un número considerable de trabajos referidos a estas pequeñas inversiones en el genoma de cloroplasto para diversos linajes de Angiospermas (Kelchner & Wendel, 1996; Graham & Olmstead, 2000; Mes et al., 2000; Kim & Lee, 2004; Bain & Jansen, 2006; Swangpol et al., 2007). Dichas inversiones están ubicadas en diferentes regiones de este genoma, apareciendo siempre flanqueadas por repeticiones invertidas (RI), llamadas también palíndromes, que generalmente tienen entre 8 a 50 pb. (Figura 11A). Cuando esta región está en cadena simple, se forma una estructura en la cual las repeticiones invertidas se aparean formando un “tallo” y el segmento presente entre ellas (el espaciador) forma un “lazo” (Figura 11B). Las estructuras de tallo-lazo son muy comunes en el extremo 3' de los genes de tARN y en regiones entre genes con una configuración de 3'-3'. Se piensa que estas estructuras estarían relacionadas con la estabilización de las moléculas de ARN transcriptas a partir de esta región (Stern & Gruissem, 1987; Stern et al., 1989). Si bien la mayoría de las inversiones asociadas con palíndromes se descubrieron en el genoma de cloroplasto, no son exclusivas del mismo. También se han hallado pequeñas inversiones en genomas mitocondriales de plantas (Dumolin-Lapègue et al., 1998) y en mitocondrias humanas, asociadas en este último caso con la ocurrencia de enfermedades (Musumeci et al., 2000; Blakeley et al., 2006).

Al estudiar la distribución de la inversión sobre la filogenia de *Prosopis* surgió el interrogante si la misma también ocurría en otros grupos de plantas. Para responder a esta pregunta se obtuvieron de GenBank alrededor de 150 secuencias para la región *psbC-trnS* del plastoma. A pesar de que una comparación de secuencias entre algunos grupos de especies emparentadas permitió reconocer la existencia de inversiones, en otros casos la ocurrencia de *indels* (inserciones-delecciones) y de sustituciones en la

región analizada dificultaban el establecimiento de un posible evento de inversión. Es así que para evaluar la ocurrencia de inversiones a lo largo de la filogenia de Angiospermas se requería de una aproximación algorítmica. Los métodos propuestos para analizar inversiones a nivel genómico sólo permiten reconocer inversiones cuando el dato de partida es la ubicación relativa de genes en un genoma y su sentido (e.g. Blanchette et al., 1997; Moret et al., 2002). Sin embargo tales aproximaciones no son útiles para analizar inversiones a nivel nucleotídico. Debido a ello desarrollamos un método para poder evaluar la ocurrencia de inversiones a lo largo de una filogenia.



Otra cuestión que surgió del análisis de esta inversión fue que la misma no presentaba una distribución uniforme a lo largo de la filogenia de Angiospermas sino que su ocurrencia se hallaba agrupada en ciertos linajes. Kelchner & Wendel (1996) sugirieron que la ocurrencia de pequeñas inversiones (SI) estaría afectada por la estabilidad termodinámica de la estructura de tallo-lazo (medida a partir de su correspondiente valor de energía libre). Estructuras con mayores valores de estabilidad tendrían mayores oportunidades para la formación de la estructura de tallo-lazo y,

concomitantemente, mayor frecuencia de inversiones. Además, Kim & Lee (2005) propusieron que las SI estarían correlacionadas con el largo de los palíndromes. Es así que la ocurrencia de inversiones podría estar afectada por la distribución de estas variables a lo largo de la filogenia. Su evaluación en un contexto filogenético requiere del uso de un método comparativo así como de un método para optimizar las inversiones sobre el árbol.

Como resumen de los objetivos de este capítulo se puede indicar que intentamos evaluar diferentes cuestiones metodológicas y empíricas sobre la evolución de las SI, tomando como caso de estudio la inversión previamente indicada. En particular nosotros analizamos:

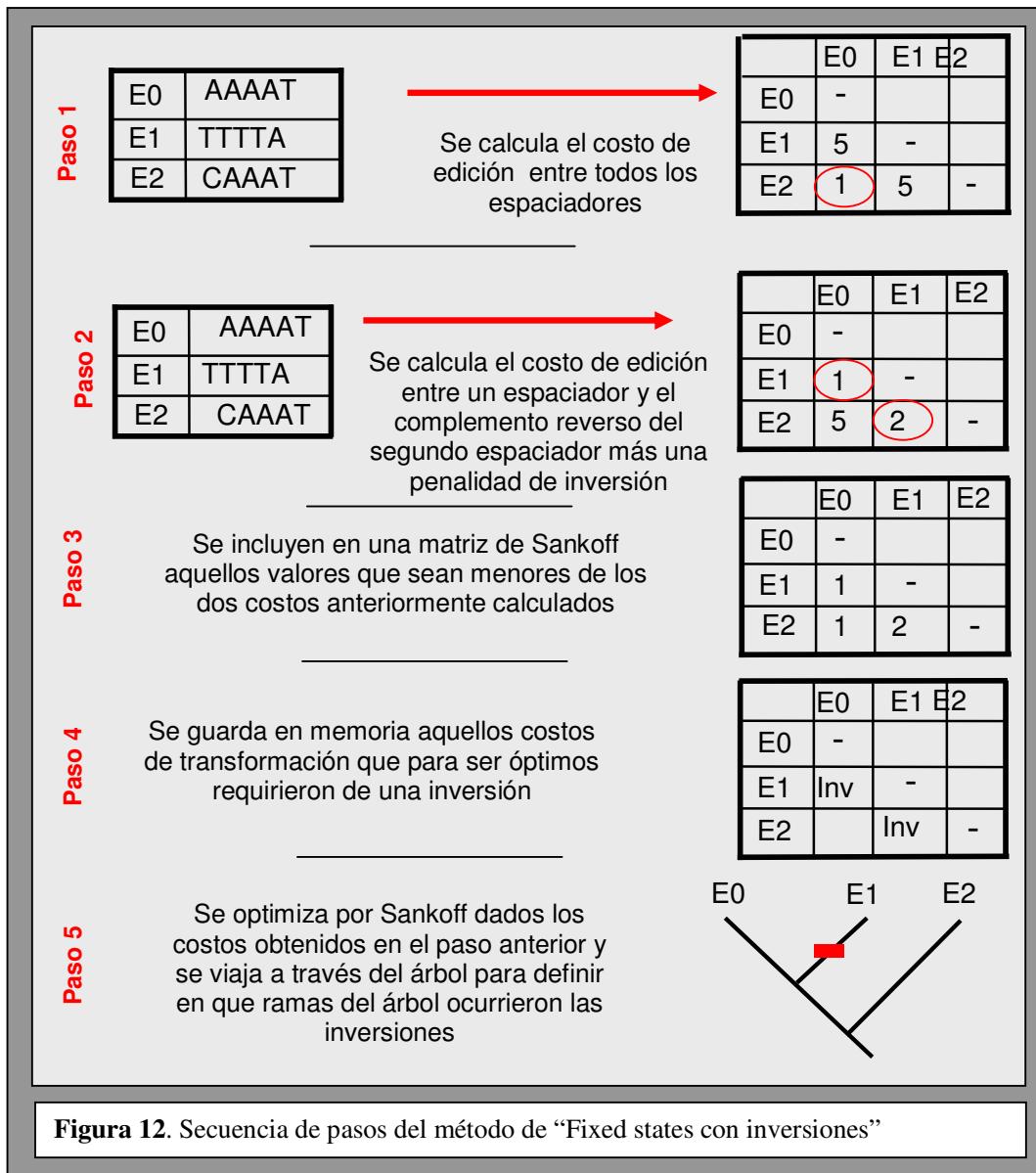
- 1-. Cómo pueden optimizarse las inversiones en casos en las cuales el segmento invertido también sufre de inserciones, delecciones y sustituciones.
- 2-. La ocurrencia de una inversión en el integen de *psbC-trnS* en *Prosopis* en particular y en Angiospermas en general.
- 3-. La posible asociación de los eventos de inversión con una mayor estabilidad de la estructura de tallo-lazo y / o repeticiones invertidas más largas.
- 4-. El efecto de diferentes tipos de codificación y delimitación de los caracteres en la evaluación de dicha asociación.

3.2. Materiales y Métodos

3.2.1. Optimización de inversiones

El análisis de las SI dentro de un contexto filogenético ha estado restringido generalmente a su codificación como un carácter binario (llamado aquí método de codificación binaria), en el cual ambos sentidos de la secuencia invertida representaban a cada uno de los estados (Graham & Olmstead, 2000). En esta tesis presentamos una metodología para estudiar inversiones que también permite tener en cuenta sustituciones

e *indels* en la región analizada. El enfoque aplicado se basa en la optimización conocida como “Fixed States” (Wheeler, 1999). En la versión original de este método (Figura 12, Paso 1), cada segmento de las secuencias analizadas es considerado un carácter, calculándose el costo de edición entre todos los posibles pares de “estados” observados. Este nuevo carácter es subsecuentemente analizado por optimización de “Sankoff” (Sankoff & Rousseau, 1975). El método propuesto (llamado aquí optimización de *Fixed States con inversiones FSI*, Figura 12), considera a las inversiones de la misma manera que la optimización de Fixed States trata a las *indels* y a las sustituciones. El primer paso de optimización de FSI involucra el cálculo del costo de edición entre cada par de secuencias de los “espaciadores” usando el paso de “wavefront” del algoritmo de Needleman & Wunsch (1970). Este es un algoritmo que calcula el costo de edición óptimo *via* programación dinámica, es decir tomando subproblemas, solucionando de manera óptima estos subproblemas de manera recursiva y utilizando estas soluciones óptimas para construir una solución óptima para el problema original. Posteriormente se calcula del mismo modo un segundo costo pero comparando en este caso una de las secuencias observadas de los espaciadores con el complemento reverso de un segundo espaciador; el costo final de la inversión es calculada como el costo de edición más una penalidad por la inversión. De estos dos costos, es decir el calculado con y sin la inversión del segmento, se escoge el que sea menor para construir una matriz de costos. Así el carácter compuesto es optimizado sobre el árbol considerando los costos previamente calculados. Finalmente, se visita cada nodo para evaluar cuales transformaciones implicaron inversiones.



Existen límites lógicos a la asignación de costos de transformación en la optimización de FSI. Los costos deben ser dados de tal manera que si una secuencia de largo L es el complemento reverso exacto de una segunda secuencia, entonces una inversión debe ser siempre aceptada como óptima en vez de L eventos de substitución. Una posible manera de definir tal límite para los costos es considerar penalidades de inversión que sean menores que el menor costo calculado entre cada secuencia y su

complemento reverso. Los conjuntos de costos utilizados en el presente trabajo cumplen con tal condición.

Durante la construcción de la matriz de costos de transformación entre fragmentos, es posible que ocurran empates cuando el costo de edición entre dos fragmentos es igual al costo de edición entre una secuencia y el complemento reverso de la segunda secuencia más un costo de penalidad de inversión. Esta ambigüedad fue tomada en cuenta al visitar los nodos del árbol para definir en qué ramas habría ocurrido una inversión. En el análisis se consideraron dos conjuntos de costos diferentes:

1-. Penalidad de indels (g) = 10, penalidad de substitución (s) = 1 y penalidad de inversión (i)= 1.

2- Penalidad de indels (g) = 10, penalidad de substitución (s) = 1 y penalidad de inversión (i)= 3.

La optimización por FSI requiere que los límites del espaciador sean los mismos para todas las secuencias analizadas. De hecho, en la inversión analizada la mayoría de las secuencias tenían el mismo límite. Las secuencias que presentaron otros límites (aproximadamente el 10% de los casos) fueron excluidas del análisis.

La optimización de FSI fue implementada en una secuencia de instrucciones ó “script” (Anexo I) de TNT. Un problema con la implementación en el lenguaje de macros de TNT es que número máximo de estados permitidos es 31. Así sólo se pudieron analizar simultáneamente 31 motivos diferentes de espaciadores. Como el conjunto de datos completo presentaba 35 motivos diferentes, decidimos excluir al azar cuatro de ellos. Se evaluó si la exclusión de diferentes motivos afectaba los resultados finales. Las conclusiones finales no cambiaron.

3.2.2. Comparación filogenética

La evaluación de la hipótesis sobre la correlación entre la ocurrencia de inversiones y estructuras de tallo-lazo con tallos más largos o con valores de estabilidad más altos (evaluado mediante sus valores de energía libre) están relacionados a aquellos análisis que intentan establecer si los cambios en un carácter (variable dependiente) están concentrados en ramas con determinado estado en una segunda variable (variable independiente). Con el objeto de evaluar estas hipótesis se realizó un análisis basando en el método de “Concentrated Changes Test” (CCT; Maddison, 1990) pero con algunas modificaciones. Este método fue originalmente diseñado para analizar dos caracteres binarios, y no se han diseñado tests específicos para tratar con una variable predictora continua y una variable respuesta discreta. Dado esto se discretizaron las variables ΔG y el largo de las RI, asignando los valores de las mismas a dos posibles estados: largo y corto para largo de RI y bajo y alto para la energía libre.

En el análisis de CCT, los cambios en la variable respuesta se diferencian en ganancias y pérdidas, es decir las transiciones de 0 a 1 son diferentes que las transiciones de 1 a 0. En el caso de las inversiones, todos los cambios son equivalentes. Así la hipótesis nula en nuestro caso debería ser replanteada como (cf. Maddison, 1990): C cambios en el primer carácter están distribuidas al azar sobre el árbol independientemente del valor del segundo carácter. La probabilidad de una distribución particular sobre el árbol es igual a B / A donde A es el número de posibles maneras de tener el número observado de inversiones en el árbol y B es el número de posibles distribuciones que presentan el mismo número de inversiones sobre el árbol en ramas con baja y alta energía (o IRS cortas/largas). En el CCT donde las ganancias y las pérdidas se diferencian, esta probabilidad no puede ser calculada simplemente contando las ramas, siendo necesario tener en cuenta la topología durante los cálculos (Harvey &

Pagel, 1991; Maddison, 1990; Pagel & Lutzoni, 2001). Esto ocurre porque si la variable respuesta se pierde en una rama no puede volver a perderse en la rama siguiente. Este no es el caso al tratar con las inversiones: una secuencia puede invertirse en una rama y luego volver a perderse en la rama siguiente. Consecuentemente, una vez que se cuenta el número de ramas que presenta cada valor de ΔG y se verifican las ramas donde ocurrieron las inversiones, la topología puede ser dejada de lado y la probabilidad puede calcularse utilizando una distribución hipergeométrica. Esta distribución es la misma que usa la prueba exacta de Fisher (1934). Así cuando ganancias y pérdidas no se diferencian, el CCT es equivalente a realizar un análisis con una tabla de contingencia de 2x2 con las ramas clasificadas en 1-. Inversión / no inversión, (2) alto ΔG / bajo ΔG .

Con el objeto de evaluar la hipótesis nula (las inversiones ocurren al azar sin importar el valor de ΔG o el largo de las IRs) es también necesario computar la probabilidad de cualquier patrón más extremo que el observado. Así se debió utilizar una probabilidad hipergeométrica acumulada.

Cuando las probabilidades se calculan de la forma previamente indicada, la distribución nula incluye todas aquellas reconstrucciones que presentan el mismo número de inversiones independientemente si estas reconstrucciones son reconstrucciones de máxima parsimonia (MPR) (Nota: en la descripción original del CCT se incluyen ambas). Wenzel & Carpenter (1994) indicaron que sólo las MPRs deberían ser incluidas en la distribución nula. Esta probabilidad no puede calcularse mediante la distribución hipergeométrica. Así con el objeto de comparar ambas probabilidades, hicimos un análisis de simulación en el cual la hipótesis nula incluyó solamente MRP con el mismo número de cambios que el patrón observado. Esta simulación se escribió en un script de TNT. En el método de CCT y el análisis comparativo hecho en este trabajo, aquellas ramas donde la variable predictora cambia

no son informativas. Así estas ramas fueron excluidas del los cálculos. Para evaluar el efecto de diferentes MPRs en los resultados, se consideraron hasta 50 diferentes reconstrucciones para cada carácter (fácilmente implementado con la opción *iterrecons* de TNT). Además para tener en cuenta la ambigüedad en la topología considerada, nosotros evaluamos hasta 50 diferentes combinaciones de resoluciones de las politomías.

3.2.3. Optimización de energía libre y largo de tallo

Con el objeto de evaluar si las inversiones ocurrían en ramas que presentaban estructuras de tallo-lazo con alta estabilidad termodinámica, se optimizaron los valores de ΔG en tres modos diferentes. Primero, los valores de ΔG fueron clasificados en dos clases: alta y baja con un límite entre las clases de $-10.0 \text{ kcal mol}^{-1}$ considerando el histograma de valores de ΔG que presentaron las especies incluidas en el análisis (Figura 13Izq). Luego este carácter se optimizó como si fuera un carácter binario. En la segunda estrategia, ΔG fue directamente optimizada como un carácter continuo, usando la optimización de Farris (1969) implementado en TNT (Goloboff et al., 2003). Una estrategia similar fue seguida para el largo de las RIs. Aquí se evaluaron dos límites entre las clases corto y largo: >10 y $>11 \text{ bp}$ (Figura 13Der).

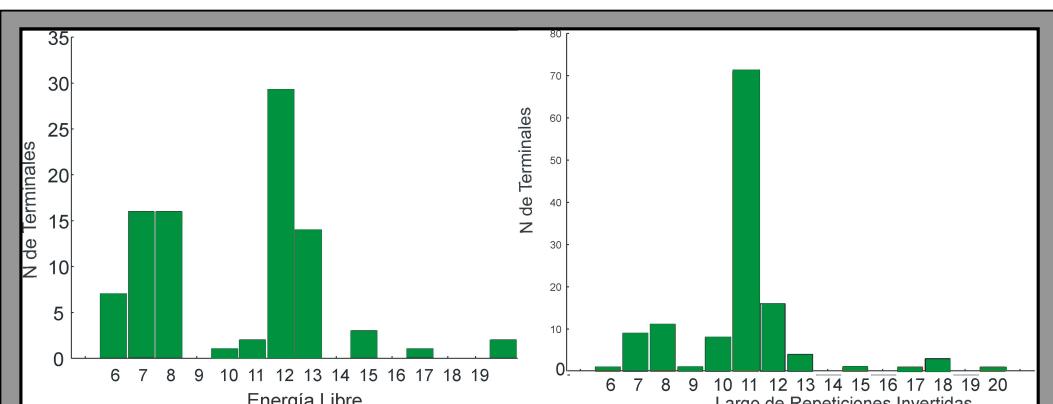
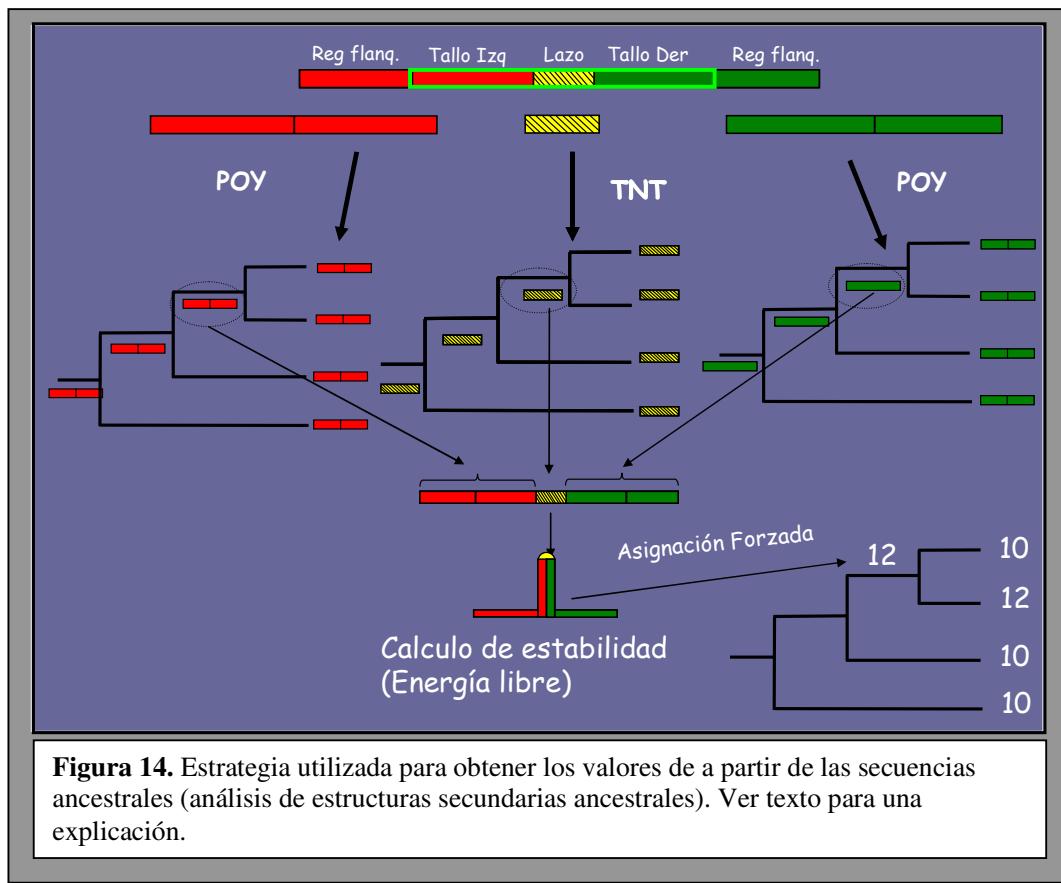


Figura 13. Histograma los valores de Energía libre (Izquierda) y largo de las repeticiones invertidas (Derecha) presentes en las especies analizadas

Una tercera estrategia para la asignación de ΔG a los nodos internos involucró la evaluación de las estructuras secundarias ancestrales (llamado a partir de aquí Análisis de estructura secundaria ancestral, Figura 14). En este caso, en vez de optimizar los valores de ΔG , primero se infirieron las secuencias ancestrales y luego se calcularon los valores de ΔG para cada secuencia ancestral del mismo modo que cuando se calcularon los valores de ΔG para los terminales.



La estrategia para establecer las secuencias ancestrales involucró la separación de las mismas en 3 segmentos: (a) el espaciador (b) 25 pb río abajo del espaciador (c) 25 pb río arriba del espaciador. Las regiones flanqueantes fueron optimizadas con POY IV Beta (Varón et al., 2007) utilizando el algoritmo denominado “Direct Optimization” (DO; Wheeler, 1996) y una combinación de DO e Iterative Pass Optimization (IP; Wheeler, 2003). Las secuencias ancestrales para cada nodo fueron obtenidas mediante

el comando *report (diagnosis)*. Posteriormente, y en cada nodo, las regiones flanqueantes fueron combinadas con la asignación óptima del motivo del espaciador obtenido en la optimización FSI hecha en TNT. Subsecuentemente, la estructura secundaria y el valor de ΔG asociado fueron calculados para cada secuencia ancestral. Estos valores entonces fueron forzados en cada nodo para obtener una reconstrucción de los valores de ΔG para el árbol completo (cómo estas asignaciones no fueron obtenidas mediante la optimización de los valores de ΔG a partir de los terminales, estas reconstrucciones no tienen que coincidir con las reconstrucciones más parsimoniosas para el valor de ΔG). La misma estrategia se siguió para estimar los valores ancestrales para el largo de las RI.

Dado las heurísticas utilizadas en POY IV para la reconstrucción de secuencias ancestrales, la optimización de las secuencias originales y la optimización de sus complementos reversos (y su posterior restitución a la secuencia original) puede dar resultados diferentes, un fenómeno que fue corroborado en pequeños grupos de datos generados para evaluar este comportamiento. Así en el estudio de la inversión encontrada en la región *psbC-trnS* evaluamos tanto la optimización del complemento reverso como de las secuencias ancestrales. Durante las optimizaciones se consideraron diferentes criterios de asignación de costos: (a) Penalidad de *indel* (g) = 12, Penalidad de sustitución (s) = 1 y (b) g = 2, s = 1. Como algunas secuencias presentaban menos de 25 pb. río abajo y río arriba del espaciador, realizamos un segundo análisis con sólo 17 pares de bases, que era el mayor largo de las secuencias flanqueantes que no presentaba datos ausentes. La estructura secundaria y los valores de ΔG asociados fueron calculados con *Quickfold* (Markham & Zuker, 2005) usando las opciones “default”.

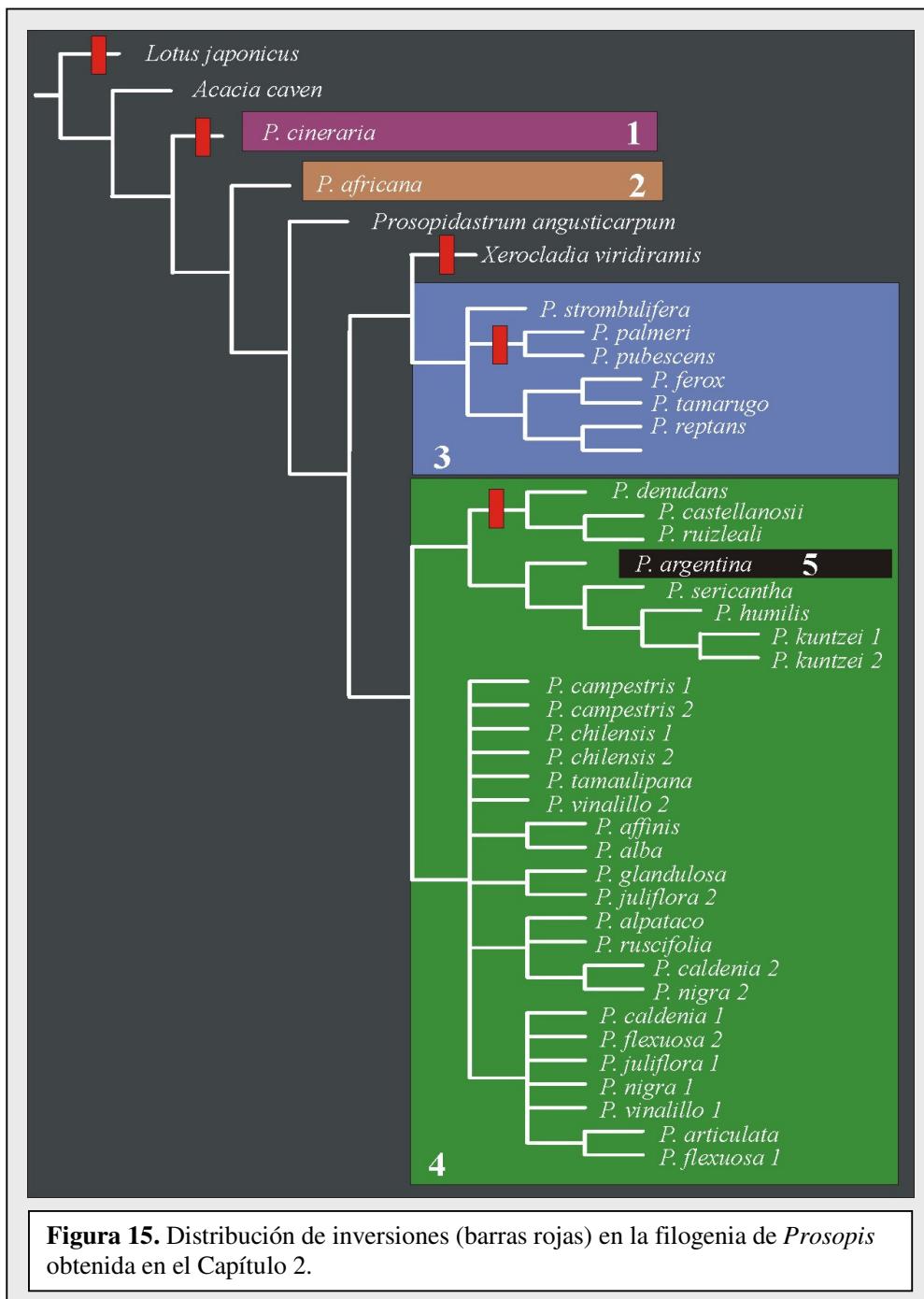
3.2.4. Conjunto de datos

En total se obtuvieron de GenBank datos de secuencias de 150 especies pertenecientes a 104 géneros de angiospermas. (Anexo II). Las secuencias fueron editadas con el programa Bioedit (Hall 1999). En los casos en que la misma especie presentaban más de una secuencia, se eligió al azar sólo una de ellas. Como algunos géneros estaban representados por más de una especie, existía la posibilidad que los resultados se vieran influenciado de manera muy fuerte por aquellos géneros con mayor número de especies. Es así que además de realizar un análisis con todas las especies de cada género, se realizó un segundo análisis en el cual se incluyeron como máximo dos especies por género. El esqueleto del árbol filogenético utilizado para optimizar las inversiones fue establecido a partir de las relaciones definidas en APG II (2003). Otros estudios (Olmstead et al., 1999; Barker et al., 2001; Bremer et al., 2002; Lavin et al., 2005) se utilizaron para definir relaciones dentro de grupos particulares. Además se utilizó una topología alternativa donde las monocotiledóneas (Liliopsida) era el grupo hermano de las eudicotiledóneas.

3.3. Resultados

3.3.1. La inversión en *Prosopis*

La optimización de la inversión sobre la filogenia obtenida en el análisis de tres marcadores se muestra en la Figura 15. En la misma se observa que las inversiones habrían ocurrido al menos dos veces en el grupo de las especies americanas de *Prosopis*. Además habrían ocurrido otros dos eventos de inversión, uno en el linaje que conduce a *P. cineraria* y otro a *X. viridinamis*. La estructura de la región invertida fue idéntica en todas las especies de *Prosopis* analizada, presentando lazos y tallos de 10 pb.



3.3.2. La inversión en Angiospermas

3.3.2.1. Caracterización de la estructura de tallo-lazo

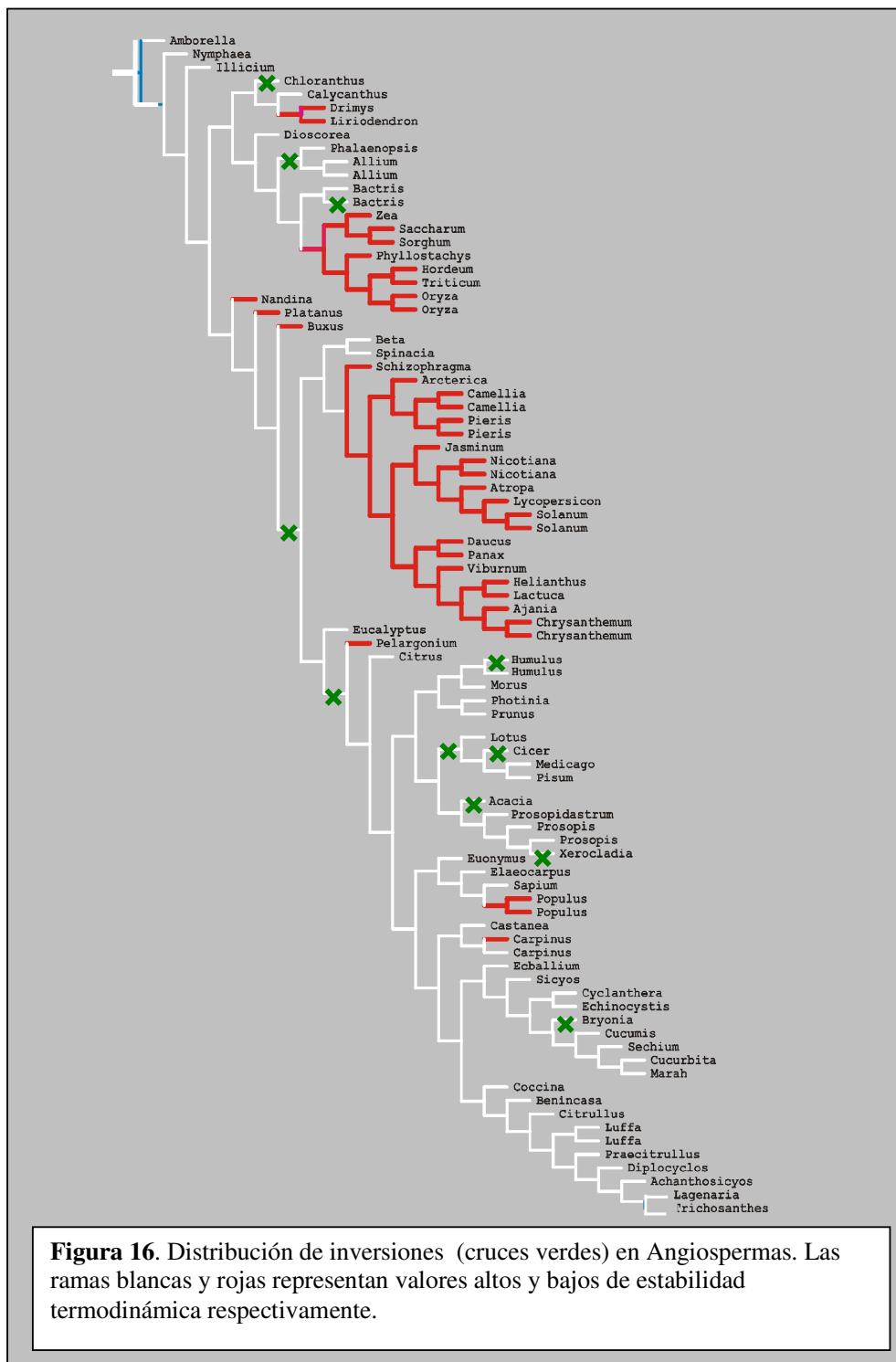
La estructura de tallo-lazo presente en el intergen *psbC-trnS* estuvo conservada en la mayoría de las especies analizadas. Sin embargo en algunas especies la estructura

o bien estaba ausente o estaba fuertemente modificada por numerosas sustituciones y re-arreglos (Apéndice 1). Dentro de las secuencias consideradas en el análisis, el largo del espaciador varió de 8 a 11 pb. mientras que el largo de las RI estuvo entre 9 y 17 pares de bases. Sin embargo el 97% de las especies presentaron RI entre 10 y 12 pares de bases y el 63% presento un tallo de 11 pb. La mayoría de las especies presentaron valores de energía libre (ΔG) entre -8.8 y -13.3 kcal mol⁻¹.

3.3.2.2. Optimización de inversiones, energía libre y largo de tallo en Angiospermas

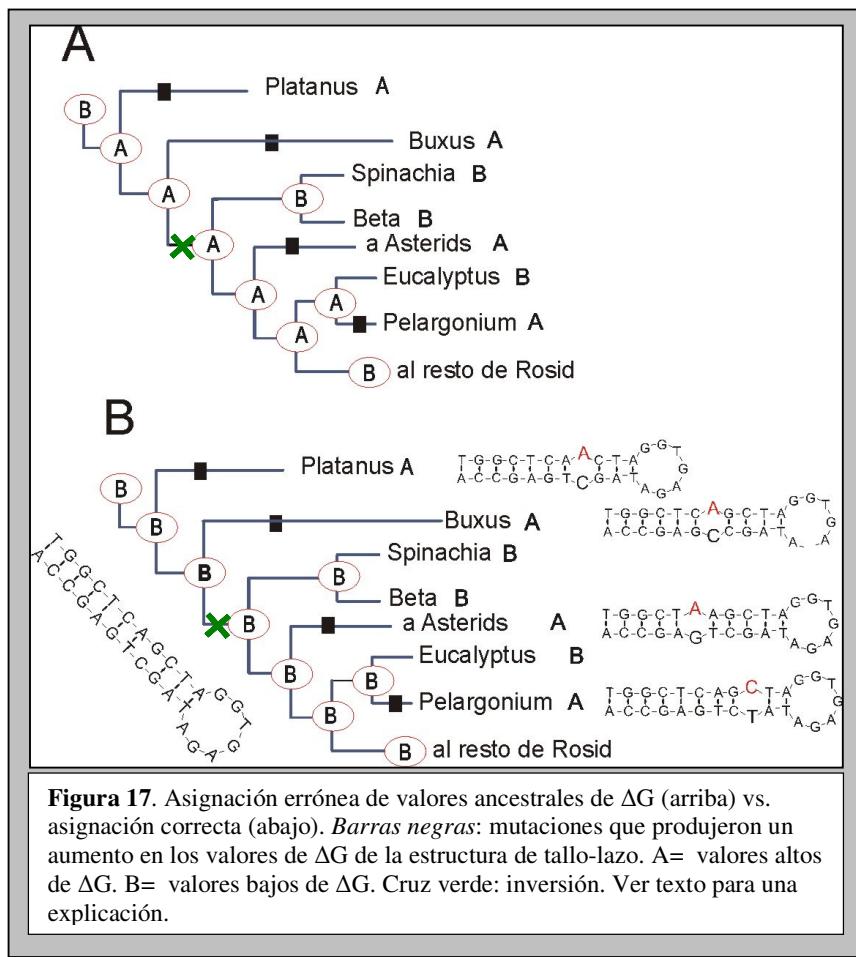
Los diferentes análisis de optimización realizados indicaron que la inversión en el intergen *psbC-trnS* habría ocurrido al menos 10 veces en la historia de las angiospermas (Figura 16). Los resultados de la optimización de la inversión cuando la misma fue codificada como un carácter binario y cuando la misma fue analizada mediante la optimización de FSI fueron los mismos en nodos superficiales (cercaos a los terminales) apareciendo diferencias en nodos basales.

Las reconstrucciones ancestrales de los valores de ΔG ya sea como carácter binario o como carácter continuo dieron valores muy similares. Sin embargo los resultados obtenidos en el análisis de las estructuras secundarias ancestrales difirieron substancialmente de los resultados obtenidos mediante la codificación binaria y continua. La causa de esta discrepancia se muestra en la Figura 17.



En la optimización binaria y continua algunos de los nodos más profundos presentaron una asignación ambigua de valores de ΔG (Figura 17A), y uno de las reconstrucciones más parsimoniosas (MPR) implicaba altos valores de ΔG para todos

estos nodos. Sin embargo cuando las secuencias mismas fueron optimizadas y se calcularon los valores de ΔG para las secuencias ancestrales obtenidas, la reconstrucción indicó que cada clado había derivado independientemente altos valores de ΔG por mutaciones en diferentes partes del tallo (Figura 17B). Consecuentemente, la asignación correcta para estos nodos ancestrales es una secuencia que implicaba bajos valores de ΔG (i.e. alta estabilidad termodinámica). Una situación similar se observa en el ACMR de *Drymis* y *Liriodendron* en el cual los altos valores de ΔG fueron adquiridos independientemente por ambos taxa.



Como fue previamente indicado la estimación del los valores del largo de las repeticiones invertidas se hizo de diferentes modos: optimizando los valores como caracteres discretos, como caracteres continuos o estimando las secuencias ancestrales y

evaluando el largo de las repeticiones invertidas. Los análisis realizados indicaron que los largos inferidos por los diferentes métodos de optimización fueron muy similares.

Las secuencias ancestrales obtenidas mediante la optimización de las secuencias difirieron al considerar diferentes costos de transformación, diferentes algoritmos de optimización (DO sólo ó DO + IP) o diferentes largos de los segmentos analizados (Ver Inciso 3.2.2. Sin embargo estos cambios nunca ocurrieron en la región de las inversiones repetidas. Así, las estructuras ancestrales obtenidas no variaron entre los diferentes parámetros analíticos. Las estructuras ancestrales fueron también las mismas al optimizar el complemento reverso de las secuencias analizadas.

3.3.2.3. Relación entre la ocurrencia de inversiones y los valores de energía libre ó largo de RI

3.3.2.3.1. Energía libre. La significación de la relación entre la estabilidad de la estructura de tallo-lazo y la ocurrencia de inversiones fue afectada fuertemente por las ambigüedades en la optimización de los valores de ΔG en ambas aproximaciones: binaria y continua. En algunos casos las probabilidades llegaron a diferir en 55 veces entre reconstrucciones (Tabla 11). La principal causa de estas diferencias fue que la estimación incorrecta de los valores ancestrales de ΔG previamente indicada (Figura 17) produjo valores no significativos. Cuando los valores ancestrales de ΔG se obtuvieron al evaluar la estructura secundaria ancestral, la comparación filogenética indicó que la ocurrencia de inversiones estaba fuertemente correlacionada con la estabilidad de la estructura de tallo-lazo (Tabla 11), con todas las inversiones situadas en ramas con alta estabilidad de la estructura. Los resultados fueron muy similares en las diferentes topologías utilizadas en el análisis.

Tabla 11. Relación entre estabilidad de la estructura de tallo-lazo (ΔG) y la ocurrencia de inversiones. El rango de valores en cada celda representa el valor mínimo y máximo obtenido para diferentes reconstrucciones y topologías. Hiper= probabilidades calculadas usando una distribución hipergeométrica acumulada para el muestreo que involucró 2 especies por género.

Optimización		Nº Ramas		Nº Inversiones		Probabilidad	
ΔG	Inversión	Alta ΔG	Baja ΔG	En ramas con alta ΔG	Total	Hiper	Simulación
Binaria	Binaria	102-112	58-68	0-2	11-12	0.005-0.091	0.002 -0.066
	FSI	102-112	58-68	0-3	10-13	0.003-0.169	0.001-0.125
Continua	Binaria	102-112	58-68	0-2	11-12	0.005-0.091	0.002 - 0.130
	FSI	102-112	58-68	0-3	11-13	0.003-0.169	0.001-0.127
Sec.*	FSI	113	56	0	11	0.004	0.004

Los diferentes conjuntos de costos y métodos para optimizar las inversiones no afectaron a los resultados del análisis comparativo. Las probabilidades calculadas mediante la simulación difieren parcialmente de aquellas calculadas mediante la distribución hipergeométrica sin embargo ambos presentaron la misma tendencia (Tabla 11). Los resultados fueron similares también al considerar todas las especies de cada género (Tabla 12) o sólo dos especies por género (Tabla 11).

Tabla 12. Relación entre estabilidad de la estructura de tallo-lazo (ΔG) y la ocurrencia de inversiones. El rango de valores en cada celda representa el valor mínimo y máximo obtenido para diferentes reconstrucciones y topologías. Hiper= probabilidades calculadas usando una distribución hipergeométrica acumulada para el muestreo que involucró todas las especies muestreadas de cada género

Optimización		Nº Ramas		Nº Inversions		Probabilidad	
ΔG	Inversión	Alta ΔG	Baja ΔG	En ramas con alta ΔG	Total	Hiper	Simulación
Binaria	Binaria	160-170	68-78	0-2	13-15	0.005-0.012	0.002-0.091
	FSI	160-170	68-78	0-3	12-16	0.003-0.22	0.0014-0.192
Continua	Binaria	160-170	68-78	0-2	13-15	0.048-0.121	0.002-0.096
	FSI	160-170	68-78	0-3	13-16	0.003-0.22	0.018-0.191
Sec.*	FSI	166	77	0	15	0.007*	0.0027*

El patrón de correlación observado con el análisis solo incluyó las especies con RI de 11 pb. (Tabla 13), indicando que existiría una correlación entre los valores de energía libre y la ocurrencia de inversiones que es independiente del largo de las RI.

Tabla 13. Relación entre estabilidad de la estructura de tallo-lazo (ΔG) y la ocurrencia de inversiones. El rango de valores en cada celda representa el valor mínimo y máximo obtenido para diferentes reconstrucciones y topologías. Hiper= probabilidades calculadas usando una distribución hipergeométrica acumulada. El muestreo incluye todas las especies muestreadas para cada género que presentan inversiones repetidas de 11 pb *Análisis de estructura secundaria ancestral.

Optimización		Nº Ramas		Nº Inversions		Probabilidad	
ΔG	Inversión	Alta ΔG	Baja ΔG	En ramas con alta ΔG	Total	Hiper	Simulación
Binaria	Binaria	58-64	80-86	0-1	7-8	0.014-0.165	0.007-0.152
	FSI	58-64	80-86	0-1	7-8	0.014-0.102	0.008-0.086
Continua	Binaria	58-64	80-86	0-1	7-8	0.014-0.102	0.008-0.086
	FSI	58-64	80-86	0-1	7-8	0.014-0.102	0.009-0.088
Sec.*	FSI	54	85	0	7	0.029	0.022

Los resultados no se vieron afectados al considerar un conjunto de costos de transformación diferentes al optimizar la inversión (Tabla 14)

Tabla 14. Relación entre estabilidad la estructura de tallo-lazo (ΔG) y la ocurrencia de inversiones. El rango de valores en cada celda representa el valor mínimo y máximo obtenido para diferentes reconstrucciones y topologías. Costos de transformación alternativos para la optimización de la inversión (inversión= 3; sustitución= 1; gap= 10) Hiper= probabilidades calculadas usando una distribución hipergeométrica acumulada. El muestreo incluye dos especies por género. *Análisis de estructura secundaria ancestral.

Optimización de valores de energía	Nº Ramas		Nº Inversiones		Probabilidad	
	Alta ΔG	Baja ΔG	En ramas con alta ΔG	Total	Hiper	Simulación
Continua	58-68	102-112	0-3	11-13	0.003-0.229	0.001 -0.130
Sec*	56	113	0	11	0.004	0.004

3.3.2.3.2. Largo del tallo. Ninguno de los análisis arrojó resultados significativos para la correlación entre la ocurrencia de inversiones y el largo del tallo (Tabla 15, 16). Cuando el límite entre RI cortas y largas se definió en 11 bp y el muestreo incluyó todas las

especies de cada género (Tabla 15), se evidenció una tendencia hacia mayores frecuencias de inversiones en las ramas con RI largas. Sin embargo este patrón fue no significativo. La optimización del largo de las RI a partir de los terminales (ya sea codificación binaria o continua) y el análisis que derivó los largos de las RI a partir de las estructuras ancestrales dieron resultados muy similares. Las probabilidades calculadas mediante la distribución hipergeométrica fueron similares a las calculadas mediante la simulación. Los diferentes regímenes de costos y métodos para optimizar la inversión no afectaron los resultados.

Tabla 15. Relación entre el largo de las IRs y la ocurrencia de inversiones. El rango de valores en cada celda representa el valor mínimo y máximo obtenido para diferentes reconstrucciones y topologías. *Limite* indica el valor tomado para separar entre las clases “Largo” y “Corto”. Hiper= probabilidades calculadas usando una distribución hipergeométrica acumulada. Probabilidades calculadas para el muestreo que involucró 2 especies por género

Optimización			Nº Ramas		Nº Inversiones		Probabilidad	
Largo	Limite	Inversión	RI cortas	RI largas	En RI cortas	Total	Hiper	Simulación
Binaria	10-11	Binaria	64-68	174-17	4-5	13-15	0.648-0.868	0.663-0.849
		FSI	64-68	173-177	4-5	13-16	0.577-0.857	0.582-0.893
	11-12	Binaria	229-232	10-13	11-14	12-15	0.094-0.574	0.088 - 0.523
		FSI	231-234	8-11	12-15	13-16	0.071-0.536	0.064-0.590
Continua	10-11	Binaria	64-68	174-178	4-5	13-15	0.648-0.868	0.662-0.858
		FSI	64-68	174-178	4-5	13-16	0.582-0.868	0.585-0.889
	11-12	Binaria	229-232	10-13	11-14	12-15	0.094-0.574	0.094-0.623
		FSI	229-232	10-13	12-15	13-16	0.107-0.599	0.107-0.649

Tabla 16. Relación entre el largo de las IRs y la ocurrencia de inversiones. El rango de valores en cada celda representa el valor mínimo y máximo obtenido para diferentes reconstrucciones y topologías. *Limite* indica el valor tomado para separar entre las clases “Largo” y “Corto”. Hiper= probabilidades calculadas usando una distribución hipergeométrica acumulada. Probabilidades calculadas para el muestreo que involucró todas las especies de cada género

Optimización			Nº Ramas		Nº Inversiones		Probabilidad	
Largo	Limite	Inversión	RI cortas	RI largas	En RI cortas	Total	Hiper	Simulación
Binaria	>10	Binaria	11-12	162-166	2	8-12	0.96-0.99	0.974 - 0.999
		FSI	11-13	162-166	2	8-12	0.951-0.991	0.968-0.998
	>11	Binaria	158-164	9-11	9-11	10-12	0.123-0.698	0.123-0.752
		FSI	158-164	10-12	9-11	11-13	0.144-0.728	0.150-0.791
Continua	>10	Binaria	11-12	162-166	1-2	8-12	0.95-0.99	0.978-0.999
		FSI	11-13	162-166	1-2	8-12	0.96-0.99	0.972-0.998
	>11	Binaria	158-164	9-11	9-11	10-12	0.120-0.69	0.123-0.756
		FSI	158-164	10-12	9-11	11-13	0.144-0.69	0.147-0.759

3.4. Discusión

Una característica distintiva de la inversión analizada en el presente estudio fue su múltiple ocurrencia en diversos linajes de Angiospermas (monocotiledóneas, Magnoliids, rosids). La mayoría de las secuencias invertidas (SI) previamente documentadas presentaron menores niveles de recurrencia y en general involucraban taxones cercanamente emparentados (Kelchner & Wendel, 1996; Kelchner & Clark, 1997; Mes et al., 2000; Kim & Lee, 2005; Swangpol et al., 2007). La única SI con una distribución filogenética y con un nivel de recurrencia similar fue documentada por Bain & Jansen (2006) en la región intergénica entre los genes *psbA* y *trnH*. Sin embargo estos autores analizaron la inversión comparando pares de especies y no mediante optimización sobre un árbol.

La SI analizada en esta tesis pudo ser reconocida a lo largo de otros linajes de angiospermas debido al alto nivel de conservación que presentó la estructura de tallo-lazo, un hecho inesperado dada la condición de ADN no codificante que presenta esta región. Graham & Olmstead (2000) observaron una situación similar en una SI

encontrada en linajes basales de angiospermas y sugirieron que tal nivel de conservación estaría asociada con el rol que tendría esta estructura en la estabilización de los transcriptos correspondientes. Es así que la recurrencia de las SIs puede ser explicada tan sólo como un producto secundario de la función de la estructura de lazo y tallo (Kim & Lee, 2005), sin la necesidad de invocar explicaciones adaptativas.

3.4.1. Estabilidad de la estructura de tallo-lazo y los eventos de inversión

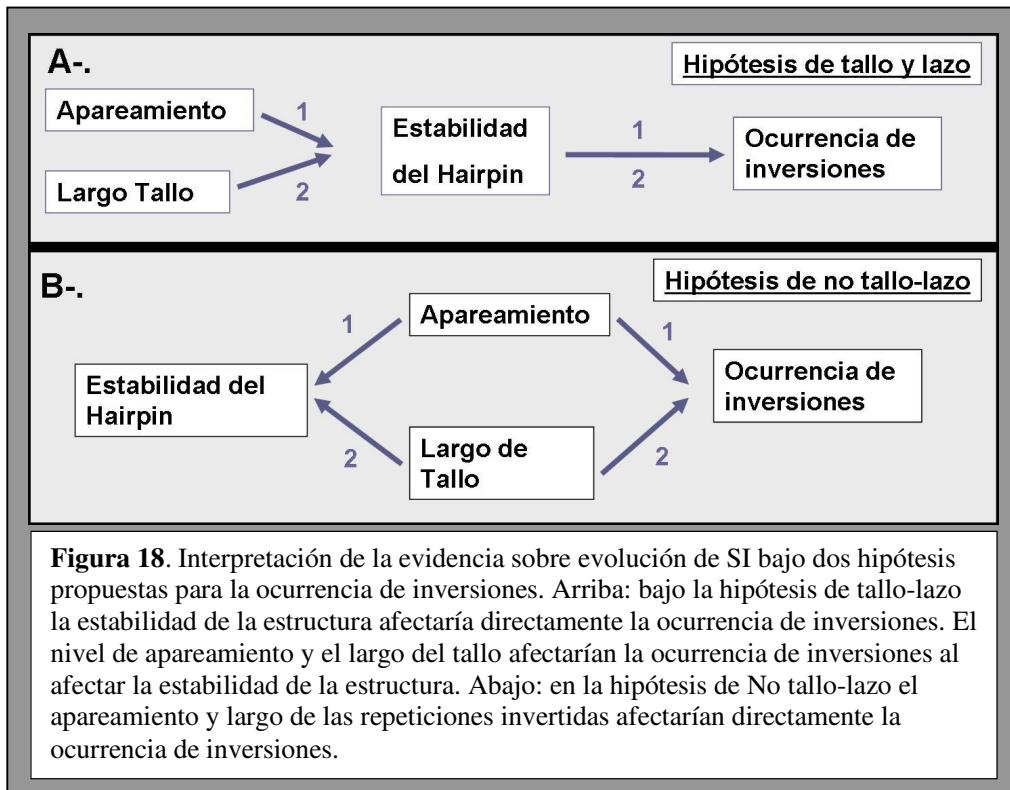
El presente estudio sugiere que la ocurrencia de las SI en el intergen *psbC-trnS* está correlacionada con una mayor estabilidad de la estructura de tallo-lazo. Este resultado está de acuerdo con el análisis efectuado por Kelchner & Wendel (1996) en otra SI. Es así que nuestro análisis plantearía una extensión de los resultados obtenidos por estos autores en dos direcciones. Primero, la observación de patrones similares en SIs presentes en diferentes regiones intergénicas sugiere que esta correlación es común a las SIs situadas a lo largo del genoma de cloroplasto. Segundo, el patrón originalmente observado en la subFam Bambusoidea se documenta ahora a nivel global dentro de angiospermas. El análisis de SI situadas en otras regiones del cloroplasto con un enfoque similar al utilizado aquí, ayudará a evaluar la generalidad de este patrón.

Kelchner & Wendel (1996) sugirieron que la estabilidad del tallo-lazo incrementaría la probabilidad de su formación que, concomitantemente, produciría un incremento en las frecuencias de las SI. Esta idea implícitamente considera que la estructura de tallo-lazo está involucrada en el proceso que finalmente conduce a la inversión (llamada a partir de aquí hipótesis de tallo-lazo). En aquel análisis, las diferencias en la estabilidad de la estructura de tallo-lazo entre taxones estaban causadas por cambios en el largo de las RI. En contraste, en esta tesis no se obtuvieron evidencias que sugirieran que el largo de las RI afecte la frecuencia de la inversión en la región intergénica *psbC-trnS*.

En el presente análisis los cambios en la estabilidad de la estructura de tallo-lazo estuvieron principalmente causados por mutaciones que produjeron apareamientos imperfectos en los tallos (quasipalíndromes). Bajo la hipótesis de tallo-lazo, ambas observaciones pueden ser explicadas en un modo muy simple: las RI cortas y el apareamiento imperfecto disminuyen las probabilidades de inversión al afectar la estabilidad del tallo-lazo (Figura 18A).

Aun si la correlación entre la estabilidad de la estructura de tallo-lazo y la ocurrencia de SI puede ser perfectamente explicada en el contexto de la hipótesis de tallo-lazo, existen igualmente explicaciones alternativas. Para alguno de los métodos propuestos para explicar la ocurrencia de SI (corrimiento de doble cadena, Leach *in* Slupska et al., 2000; excisión y re-unión de las secuencias del lazo, Kelchner & Wendel, 1996; recombinación intraestructura, Kim & Lee, 2005; invasión de hebra mediada por ADN, Musumeci et al., 2000) la formación de la estructura de tallo-lazo es un paso crucial para el proceso. Sin embargo existen modelos alternativos (modelos no-tallo-lazo) en los cuales no se requiere la estructura tallo-lazo como intermediario para la ocurrencia de inversiones (reparación de una ruptura doble en los palíndromes, Slupska et al., 2000; recombinación intramolecular en ADN de doble cadena Musumeci et al., 2000, recombinación intermolecular, Graham & Olmstead, 2000)

En todas los mecanismos propuestos para la ocurrencia de inversiones que no implican la intervención del tallo-lazo siempre se haya involucrada la acción del proceso de recombinación ya sea intra o intermolecular. Existe claras evidencias de que habría una fuerte correlación entre la frecuencia de recombinación y el largo y similitud de las secuencias interviniéntes (Rubnitz & Subramani, 1984; Shen & Huang, 1986; Bazemore et al., 1997; Sagi et al., 2006; Barnes & McCulloch, 2007).



Consecuentemente, es posible que las RI y el apareamiento imperfecto afecte la frecuencia de inversiones mediante la modificación de la eficiencia del paso de la recombinación. Bajo esta explicación alternativa, la correlación entre la ocurrencia de inversiones y la estabilidad del tallo-lazo puede ser explicada ya no como una relación causa-efecto (como es explicado por la hipótesis del tallo-lazo), sino como una consecuencia de que el largo de las RI y el apareamiento imperfecto también afectan la estabilidad del tallo-lazo. (Figura 18B).

En resumen, en este estudio tesis encontramos evidencia que apoya un patrón de correlación entre la estabilidad de la estructura de tallo-lazo y la ocurrencia de una pequeña inversión situada en la región intergénica *psbC-trnS* del genoma de cloroplasto. Este resultado está de acuerdo con la hipótesis de Kelchner & Wendel (1996) que indica que el tallo-lazo estaría involucrado en la formación de las SI. Sin embargo no puede descartarse una explicación alternativa en donde la frecuencia de

inversiones se viera afectada independientemente por el largo de las RI y el apareamiento imperfecto. La formulación de predicciones más detalladas desde esos modelos así como el análisis de nuevas SI contribuirá a una mejor comprensión de las particularidades de estos eventos moleculares.

3.4.2. Asignaciones ancestrales y homología

En el presente estudio se evaluaron diferentes métodos para asignar valores de ΔG a los nodos internos. Cuando los valores fueron obtenidos por la optimización de los valores de ΔG de los terminales, algunas de las reconstrucciones erróneamente asignaron altos valores de ΔG a algunos nodos profundos (Figura 17). Esto ocurrió porque los valores altos de ΔG en diferentes linajes fueron considerados como derivados de un ancestro común. Sin embargo, cuando las secuencias mismas fueron optimizadas usando DO o IP, las secuencias ancestrales presentaron bajos valores de ΔG para estos nodos, con altos valores de ΔG derivados independientemente en varios linajes por diferentes sustituciones en las RI. La distinción de estados fenotípicos (los valores de ΔG) que, aunque similares, no fueron derivados de un ancestro común, fue posible debido a que la relación entre genotipo y fenotipo era muy simple en este caso: una sustitución en las RI que produce un cambio en el valor de ΔG correspondiente. Esta situación contrasta fuertemente con la mayoría de los caracteres morfológicos generalmente considerados en análisis filogenéticos. Debido a la compleja red de interacciones génicas que definen estos caracteres, la evaluación de los cambios hereditarios que los afectan es imposible en la mayoría de los casos, dejando a la similitud como la evidencia más creíble de origen común.

Dada la naturaleza heurística del método “Direct Optimization”, los mismos no garantizan que las reconstrucciones sean globalmente óptimas. Además, dada la implementación actual en POY IV, sólo se puede obtener un grupo de asignaciones

ancestrales óptimas. Esto podría poner en duda las asignaciones ancestrales de la estructura de tallo-lazo y, concomitantemente, los resultados del método comparativo filogenético. Sin embargo diferentes líneas de evidencia estarían en contra de esta posibilidad. Primero, la alta conservación encontrada en las RI reduce la posibilidad de ambigüedades en la reconstrucción ancestral de secuencias. Segundo, las estructuras secundarias fueron las mismas en todos los análisis independientemente de los regímenes de costos o los algoritmos usados en la optimización. Tercero, aun cuando es posible que haya asignaciones óptimas alternativas, no implica directamente que haya habido cambios en los valores de ΔG . Finalmente no hay razón para suponer que las reconstrucciones evaluadas estuvieran sesgadas en favor de la hipótesis puesta a prueba.

3.5. Bibliografía

- Bain, J. F. & R. K. Jansen. 2006. A chloroplast DNA hairpin structure provides useful phylogenetic data within tribe Senecioneae (Asteraceae). *Canadian Journal of Botany* 84: 862-868.
- Barker, N. P., L. G. Clark, J. I. Davis, M. R. Duvall, G. F. Guala, G. F. Hsiao, E. A. Kellogg & H. P. Linder. 2001. Phylogeny and Subfamilial Classification of the Grasses (Poaceae) Grass Phylogeny Working Group. *Annals of the Missouri Botanical Garden* 88: 373-457.
- Barnes, R. L. & R. McCulloch. 2007. *Trypanosoma brucei* homologous recombination is dependent on substrate length and homology, though displays a differential dependence on mismatch repair as substrate length decreases. *Nucleic Acids Research* 35: 3478–3493.
- Bazemore, L. R., E. Folta-Stogniew, M. Takahashi & C. M. Radding. 1997. RecA tests homology at both pairing and strand exchange. *Proceedings of the National Academy of Sciences of the United States of America* 94: 11863–11868.
- Blakeley, E., K. Rennie, L. Jones, M. Elstener, Z. Chrzanowska-Lightowers, C. White, J. Shield, D. Pilz, D. Turnbull, J. Poulton & R. Taylor. 2006. Sporadic Intragenic Inversion of the Mitochondrial DNA MTND1 Gene Causing Fatal Infantile Lactic Acidosis. *Pediatric research* 59: 440-444.
- Bremer, B., K. Bremer, N. Heidari, P. Erixon, A. A. Anderberg, R. G. Olmstead, M. Källersjö & E. Barkhordarian. 2002. Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. *Molecular Phylogenetic and Evolution* 24: 274-301.
- Dumolin-Lapègue, S., M.-H. Pemonge & R. J. Petit. 1998. Association Between Chloroplast and Mitochondrial Lineages in Oaks. *Molecular Biology and Evolution* 15: 1321–1331.
- Farris, J. 1969. A Successive Approximations Approach to Character Weighting. *Systematic Zoology* 18: 374-385.
- Fisher, R. A. 1934. Statistical methods for research workers. Oliver and Boyd., Edinburgh.

- Goloboff, P. A., J. S. Farris & K. Nixon. 2003. TNT: Tree Analysis Using New Technology. Program and documentation available at <http://www.zmuc.dk/public/phylogeny>.
- Graham, S. W. & R. G. Olmstead. 2000. Evolutionary significance of an unusual chloroplast DNA inversion found in two basal angiosperm lineages. *Current Genetics* 37: 183-188.
- Harvey, P. H. & M. Pagel. 1991. *The Comparative Method in Evolutionary Biology*. Oxford University Press.
- Kelchner, S. A. & L. G. Clark. 1997. Molecular Evolution and Phylogenetic Utility of the Chloroplast rpl16 Intron in *Chusquea* and the Bambusoideae (Poaceae). *Molecular Phylogenetic and Evolution* 8: 385-397.
- Kelchner, S. A. & J. F. Wendel. 1996. Hairpins create minute inversions in non-coding regions of chloroplast DNA. *Current Genetics* 30: 259-262.
- Kim, K. J. & H.-L. Lee. 2005. Widespread occurrence of small Inversions in the chloroplast genomes of land plants. *Molecules and Cells* 19: 104-113.
- Kim, K.-J. & H.-L. Lee. 2004. Complete Chloroplast Genome Sequences from Korean Ginseng (*Panax schinseng* Nees) and Comparative Analysis of Sequence Evolution among 17 Vascular Plants. *DNA Research* 11: 247-261.
- Lavin, M., P. S. Herendeen & M. F. Wojciechowski. 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Systematic Biology* 54: 575-594.
- Maddison, W. P. 1990. A method for testing the correlated evolution of two binary characters: are gains or losses concentrated on certain branches of a phylogenetic tree? *Evolution* 44: 539-557.
- Maddison, W. P. 2000. Testing character correlation using pairwise comparisons on a phylogeny. *Journal of Theoretical Biology*. 202: 195-204.
- Markham, N. R. & M. Zuker. 2005. DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Research* 33: W577-W581.
- Mes, T., P. Kuperus, J. Kirschner, J. Stepanek, P. Oosterveld, H. Storchova & J. den Nijls. 2000. Hairpins involving both inverted and direct repeats are associated with homoplasious indels in noncoding chloroplast DNA of *Taraxacum* (Lactuceae:Asteraceae). *Genome* 43: 634-641.
- Musumeci, O., A. L. Andreu, S. Shanske, N. Bresolin, G. P. Comi, R. Rothstein, E. A. Schon & S. DiMauro. 2000. Infragenic Inversion of mtDNA: A New Type of

- Pathogenic Mutation in a Patient with Mitochondrial Myopathy. American Journal of Human Genetics 66: 1900–1904.
- Needleman, S. B. & C. D. Wunsch. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. Journal of Molecular Biology 48: 443-453.
- Olmstead, R. G., J. A. Sweere, R. E. Spangler, L. Bohs & J. D. Palmer. 1999. Phylogeny and provisional classification of the Solanaceae based on chloroplast DNA. *In* M. Nee, D. E. Symon, R. N. Lester & J. P. Jessop (eds), Solanaceae IV: Advances in biology and utilization. Royal Botanic Gardens, Kew, U.K, pp. 111-137.
- Pagel, M. & F. Lutzoni. 2001. Accounting for phylogenetic uncertainty in comparative studies of evolution and adaptation. *In* M. Laessig (ed) Biological Evolution and Statistical Physics. Springer Verlag.
- Rubnitz, J. & S. Subramani. 1984. The minimum amount of homology required for homologous recombination in mammalian cells. Mol. Cell. Biol. 4: 2253-2258.
- Sagi, D., T. Tlusty & J. Stavans. 2006. High fidelity of RecA-catalyzed recombination: a watchdog of genetic diversity. Nucleic Acids Research 34: 5021–5031.
- Sankoff, D. D. & P. Rousseau. 1975. Locating the vertices of a Steiner tree in arbitrary space. Mathematical Programming 9: 240-246.
- Shen, P. & H. V. Huang. 1986. Homologous Recombination in Escherichia coli: Dependence on Substrate Length and Homology. Genetics 112: 441-457.
- Slupska, M. M., J.-H. Chiang, W. M. Luther, J. L. Stewart, L. Amii, A. Conrad & J. H. Miller. 2000. Genes involved in the determination of the rate of inversions at short inverted repeats. Genes to Cells 5: 425-437.
- Stern, D. B. & W. Gruissem. 1987. Control of plastid gene expression: 3' inverted repeats act as mRNA processing and stabilizing elements, but do not terminate transcription. Cell 51: 1145-1157.
- Stern, D. B., H. Jones & W. Gruissem. 1989. Function of plastid mRNA 3' inverted repeats: RNA stabilization and gene-specific protein binding. Journal of Biological Chemistry 264: 18742-18750.
- Swangpol, S., H. Volkaert, R. Sotto & T. Seelanan. 2007. Utility of Selected Non-coding Chloroplast DNA Sequences for Lineage Assessment of Musa Interspecific Hybrids. Journal of Biochemistry and Molecular Biology 40: 577-587.

- Varón, A., L. Vinh, I. Bomash & W. Wheeler. 2007. POY 4.0 Beta 2398. American Museum of Natural History. <http://research.amnh.org/scicomp/projects/poy.php>.
- Wenzel, J. W. & J. M. Carpenter. 1994. Comparing methods: Adaptive traits and tests of adaptation. *In* P. Eggleton & R. I. Vane-Wright (eds), *Phylogenetics and Ecology*. Academic Press, London., pp. 79-101.
- Wheeler, W. 1996. Optimization alignment: The end of multiple sequence alignment in phylogenetics? *Cladistics* 12: 1-9.
- Wheeler, W. 2003. Iterative pass optimization of sequence data. *Cladistics* 19: 254-260.
- Wheeler, W. C. 1999. Fixed character states and the optimization of molecular sequence data. *Cladistics* 15: 379-385.

CAPÍTULO 4. TIEMPOS DE DIVERGENCIA Y DIVERSIFICACIÓN

4.1. Introducción

4.1.1. Tiempos de divergencia

La idea de estimar tiempos de divergencia a partir de datos moleculares surgió hace más de 50 años (Zuckerkandl & Pauling, 1962), cuando se observó que las tasas de sustitución aminoacídicas se mantendrían más o menos constantes a lo largo del tiempo. A partir de ese momento han existido numerosos desarrollos metodológicos para el estudio de la evolución molecular en general (Jukes & Cantor, 1969; Kimura, 1980; Felsenstein, 1981; Tavaré, 1986) y para estimación de tiempos de divergencia en particular (Felsenstein, 1988; Takezaki et al., 1995; Sanderson, 1997,2002). Los estudios de reloj molecular estuvieron desde sus comienzos acompañados de fuerte controversia. Análisis empíricos indicaron que la constancia de tasas no era una condición universal y común a todas las secuencias. Un ejemplo que suele citarse sobre esta controversia es la existencia de tasas diferentes en mamíferos: las tasas de sustitución serían mayores en roedores y menores en hominoides. (Kumar & Subramanian, 2002). Con el objeto de evaluar estadísticamente la constancia o no de tasas surgió una amplia gama de test entre los que se incluyen el test de tasas relativas (Sarich & Wilson, 1973), el de razón de verosimilitudes (Felsenstein, 1988) y los tests no-paramétricos de Gu & Li (1992) y Tajima (1993). Posteriormente, se comenzaron a desarrollar métodos de estimación de tiempos de divergencia para secuencias que no cumplen el requisito de tasas constantes de evolución. Los primeros intentos de evaluar tiempos de divergencia sin constancia de tasas eran métodos bastante “groseros” en los cuales se removían las ramas del árbol en las cuales se consideraba que la tasa había cambiado hasta que se tenía un árbol en el cual las tasas eran constantes. Estos árboles se denominaban árboles linealizados (Takezaki et al., 1995). Este método tenía diferentes problemas pero el más obvio es que podían quedar excluidas las especies de

interés. Posteriormente, Sanderson (1997) publicó el primer método que permite calcular tiempos de divergencia sin la necesidad de tener constancia en las tasas de sustitución. El método se denominó NPRS (Non-Parametric Rate Smoothing). El mismo se basaba en la idea de que aunque las tasas entre ramas posiblemente no se mantengan constantes, si es posible que las tasas de ramas vecinas se parezcan más entre ellas que entre dos ramas cualesquiera. Así el método buscaba minimizar por cuadrados las diferencias de tasas entre ramas vecinas. La lógica de este método parece estar claramente justificada, ya que es esperable que las causas que definen las tasas de sustitución se puedan heredar a lo largo del tiempo. Como consecuencia se tendría que las tasas estarán autocorrelacionadas entre ramas cercanas en el árbol. A pesar de lo promisorio del método, simulaciones realizadas mostraron que el mismo penalizaba poco los cambios entre ramas vecinas (Sanderson 2003). Con el objeto de resolver este problema, Sanderson (2002) publicó otro método denominado Penalized Likelihood (PL). La idea de este método es tomar un modelo con muchos parámetros (una tasa por rama) que estaría sobreparametrizado, y agregarle un término que penaliza cuán fuerte será la atenuación de la tasa de rama a rama, es decir el cambio de tasas aceptado entre ramas. El problema aquí es como definir cuánto se penaliza el cambio. Si la penalidad es infinita, el resultante es un modelo de tasa constante a lo largo de todas las ramas (reloj molecular) y si se le asigna una penalidad de cero se obtiene un modelo con tasas diferentes para cada rama (sobreparametrizado). Una forma de decidir el valor de atenuación es mediante un método denominado Validación Cruzada (Cross Validation, CV). La validación cruzada, es una práctica estadística separa una muestra de datos en subconjuntos de tal modo que el análisis es inicialmente realizado en uno de ellos, mientras los otros subconjuntos son retenidos para su uso posterior en la confirmación y validación del análisis inicial. En el caso particular de PL, se van removiendo ramas

terminales de a una y se elije el valor de atenuación que mejor permite predecir el número de sustituciones en la rama cortada. El método de PL ha sido extensamente utilizado en los últimos años para estudios de tiempos de divergencia (ej. Barker, 2004; Hedges, 2004; Bell, 2005) y fue elegido en el presente trabajo.

4.1.2. Diversificación en zonas áridas

Los estudios de datación molecular han aportado datos muy importantes a la hora de determinar del posible efecto de cambios ambientales en la diversificación de los grupos estudiados (ej. Richardson et al., 2001; Klak et al., 2003). El primer dato surge de correlacionar los tiempos de divergencia con el momento del aumento en la superficie de áreas con determinadas características ambientales. Así si la divergencia entre la mayoría de las especies del grupo G ocurrió en la época geológica E donde evidencias geológicas muestran un aumento en zonas áridas, entonces se concluye que tal cambio ambiental habría permitido/beneficiado/conducido a la diversificación de dicho grupo. Esta evidencia difícilmente puede ser considerada como concluyente aunque sí permite al menos decir que la evidencia está de acuerdo con las predicciones de la hipótesis. Este tipo de evidencia se utilizó por ejemplo para apoyar proponer que el desarrollo de ambientes áridos en Norteamérica habría propiciado la evolución del género *Tiquilia* (Moore & Jansen 2006). La evaluación de tiempos de divergencia permiten también refutar la hipótesis de diversificación asociada a cambios ambientales siempre que la divergencia de todas las especies (o la mayor parte de ellas) ocurra antes de la época E.

El estudio de los cambios en la tasa de diversificación es otra fuente de evidencia para evaluar la hipótesis de cambios ambientales afectando la diversificación de grupos. Dada esta hipótesis es esperable un aumento en la tasa de diversificación en los momentos posteriores al cambio ambiental. En este caso varios patrones podrían

refutar tal hipótesis: tasa constante, aumento de la tasa antes del cambio ambiental, tasa decreciente. Sólo se sostiene la hipótesis si ocurre un aumento en la tasa de diversificación y además esto ocurre en momentos posteriores al cambio ambiental. Aun si la hipótesis haya salido indemne a los esfuerzos de refutarla, es necesario además descartar hipótesis alternativas que permitan explicar el patrón observado. Supongamos que en la época E además de haber un aumento en las zonas áridas también existió una disminución en los predadores potenciales debido a una extinción masiva. De ser así existen dos posibles hipótesis y si ambas resisten los ensayos de refutarlas entonces será necesario seguir buscando fuentes de evidencias. Otra cuestión a indicar es que cuando una hipótesis se rechaza no implica que la misma sea de hecho falsa sino que también lo puede ser cualquiera de los supuestos auxiliares que se dieron como aceptados para poder hacer el análisis correspondiente (Chalmers, 1984, p. 83). Para nuestro análisis la cantidad de supuestos auxiliares son muchos pero para tener una idea de los más importantes podríamos señalar: Tiempos de divergencia: (i) la edad de los fósiles utilizados para la calibración es certera (ii) las sustituciones en el ADN de la región evaluada proceden de manera tal que pueden ser bien descriptas mediante un modelo probabilístico de transiciones entre estados (iii) las tasas de sustitución son constantes en todas las ramas del árbol, o en su defecto, es posible establecer el modo en que cambian tales tasas de sustitución de una rama a la otra. Cambios ambientales: (i) el registro geológico es lo suficientemente completo como para poder inferir las condiciones ambientales pasadas (ii) La datación realizada para los cambios en las condiciones ambientales es certera. Este conjunto de hipótesis auxiliares es tan sólo una muestra de los supuestos implícitos en los análisis indicados.

4.2. Materiales y Métodos

4.2.1. Cálculos de tiempos de divergencia en *Prosopis*

Con el objeto de estimar los tiempos de divergencia, se concatenaron las secuencias de las regiones genómicas *g3pdh* y de *trnS-psbC*. El análisis simultáneo de estos dos loci se realizó debido a que el método de Penalized Likelihood utilizado para estimar los tiempos de divergencia está sujeto a errores cuando se tratan ramas de largo cero (Sanderson, 2003). La congruencia entre ambas regiones genómicas con respecto al largo de ramas se evaluó con el test de cocientes de verosimilitud (likelihood ratio test) descrito por Lewis (1998) y Xiang et al. (2005). El estadístico de este test es:

$$-2 [\ln L - (\ln L_1 + \ln L_2)]$$

donde L_1 es la verosimilitud obtenida a partir del primer gen, L_2 es la verosimilitud obtenida a partir del segundo gen, y L es la verosimilitud obtenida a partir del análisis de ambos genes. El valor de este estadístico se comparó con una distribución χ^2 con una cantidad de grados de libertad que es igual la suma de:

$$2n + 3 - P$$

Donde n es el número de taxones del árbol y P es el número de parámetros libres usados en el cálculo de la verosimilitud.

La región genómica NIA no fue incluida en la estimación de largos de ramas debido a que no fue posible amplificar tal región para *Prosopidastrum angusticarpum* uno de los taxones que definen el único nodo fijo en el paso de la calibración. Además, *Xerocladia viridinamis* no fue incluida en este análisis debido a que en ella sólo se pudo secuenciar una parte del gen *g3pdh*. El largo de ramas fue estimado por máxima verosimilitud utilizando el programa Paup* (Swofford, 2002). El modelo evolutivo fue elegido utilizando el test de razones de verosimilitudes jerárquico implementado en Model Test 3.06 (Posada & Crandall, 1998). Para la estimación de los tiempos de

divergencia se utilizó el método Penalized Likelihood implementando en el programa r8S (Sanderson, 2003). El algoritmo utilizado para la estimación fue el de Powell y el valor de atenuación de tasas entre ramas vecinas se eligió mediante un procedimiento de validación cruzada (Sanderson, 2002).

Para derivar los tiempos absolutos de divergencia se utilizaron dos puntos de calibración combinados. El tiempo obtenido en Lavin et al. (2005) para la divergencia entre una de las especies de la sección Algarobia (*Prosopis pallida*) y una especie del género *Prosopidastrum* (*Prosopidastrum mexicanum*) se usó para fijar la edad del nodo que sostenía a *Prosopidastrum* y a las especies americanas de *Prosopis* (EAP). Con el objeto de tener en cuenta el error asociado a la estimación efectuada por Lavin (2005), el análisis se repitió considerando la media +/- un desvío estándar. Un segundo punto de calibración fue usado para asignar una edad mínima al clado de las EAP. En este caso el fósil correspondió a granos de polen obtenidos de sedimentos del Oligoceno Temprano de la región del la Columbia Británica en Canadá (Piel, 1971). Como esta época se extiende de los 28.4 a los 33.7 millones de años (MA) nosotros utilizamos la edad menor como edad mínima. Como el grupo hermano del clado EAP + *Xerocladia* no quedó claramente definido en el estudio filogenético realizado, se evaluaron dos diferentes topologías, una de ellas correspondiente a los resultados del análisis de tres marcadores y el otro correspondiente al análisis de dos marcadores. El error en la estimación de tiempos fue evaluado mediante la técnica de bootstrapping siguiendo a Sanderson (2002).

4.2.2. Análisis de diversificación

Con el objeto de estudiar los posibles cambios de la tasa de diversificación a lo largo de la historia del género se realizó un análisis de linajes a través del tiempo (c v, Nee et al., 1995). Este método se basa en graficar el logaritmo del número de linajes en

función del tiempo. Dada una constancia de tasas a lo largo del tiempo se espera una relación lineal entre estas dos variables. En cambio, la curva sería convexa si las tasas hubieran disminuido a lo largo del tiempo y cóncava en caso contrario. También es posible determinar cambios de tasas en momentos puntuales de la historia si la pendiente de la curva cambia pronunciadamente. El análisis de LTT se realizó repetidamente con los valores extremos obtenidos en el análisis de tiempos de divergencia. La falta de resolución en el clado de los Algarrobos fue considerada de dos maneras en el análisis de diversificación

(1) Como una politomía dura con todos los linajes divergiendo simultáneamente.

(2) Como si la diversificación de los mismos hubiera ocurrido a tasa constante. Para Esto se recurrió al programa PHYLOGEN versión 1.1 (Rambaut, 2002) que simula árboles dado un modelo de diversificación tipo Yule (1924). Se simuló un árbol con un número de especies finales igual al número de especies presentes en el grupo de los algarrobos (Ver Capítulo 2). Luego este subárbol fue injertado en el árbol original reemplazando la politomía del clado de los Algarrobos. Esto se repitió para 100 subárboles simulados. Como el patrón general obtenido fue el mismo, se muestra sólo el LLT para una un árbol con una de las simulaciones para el clado de los Algarrobos.

El análisis de LTT fue hecho en una primera instancia con aquellas especies incluidas en el análisis de tiempos de divergencia. Para examinar el efecto que pudieran tener análisis de LTT las especies no incluidas en el análisis de datación, se realizó un análisis en el cual se incluyeron estas especies en la posición filogenética más probable dado el conocimiento taxonómico de las mismas (Burkart, 1976). Esta estrategia fue la seguida en Barraclough & Vogler (2002) en su trabajo sobre escarabajos. Con el objeto de comparar las tasas de divergencia en *Prosopis* con aquéllas estimadas para otros géneros de plantas, se calculó la tasa neta de diversificación por linaje siguiendo a

Coyne & Orr (2004). Esta tasa indica el número de linajes sobrevivientes por unidad de tiempo y se calcula como:

$$\mathbf{NDR} = (\ln(N) - \ln(2)) / T$$

Donde N es el número de especies actuales del clado y T es el tiempo estimado para el ancestro común más reciente del clado. Como en el caso del análisis de LTT, las especies faltantes se colocaron en los sitios más probables del árbol dada la información taxonómica.

4.3. Resultados

4.3.1. Datación Molecular

El modelo óptimo elegido mediante el Test Jerárquico de razones de verosimilitud (LRT) fue el Hasegawa-Kishino-Yano plus Gamma (HKY+G). El mismo test rechazó la constancia de tasas para las regiones genómicas analizadas (Valor de verosimilitud asumiendo reloj $-\ln L = 4394.64$; no asumiendo reloj $-\ln L = 4363.50$; $df = 39$; $p < 0.05$). Así los tiempos de divergencia fueron calculados utilizando el método de Penalized Likelihood que no asume constancia de tasas. Como la existencia de ramas cortas puede afectar el resultado obtenido para todas las ramas del árbol (Sanderson, 2002) se repitió el análisis dejando solamente uno de los terminales del grupo de los algarrobos. Los resultados variaron muy poco. El análisis de bootstrapping para las edades calculadas se muestra en la Figura 19.

El patrón más relevante (Figura. 20; Tabla 18) encontrado en todos los análisis realizados fue: (1) Una divergencia antigua entre los dos principales linajes de las EAP para el Oligoceno; (2) Una divergencia más reciente, ubicada en el Mioceno tardío para el grupo-corona (Crown Group) de Strombocarpa y Algarobia s.l. (3) Un comienzo en la divergencia de las especies de algarrobos para el Plioceno.

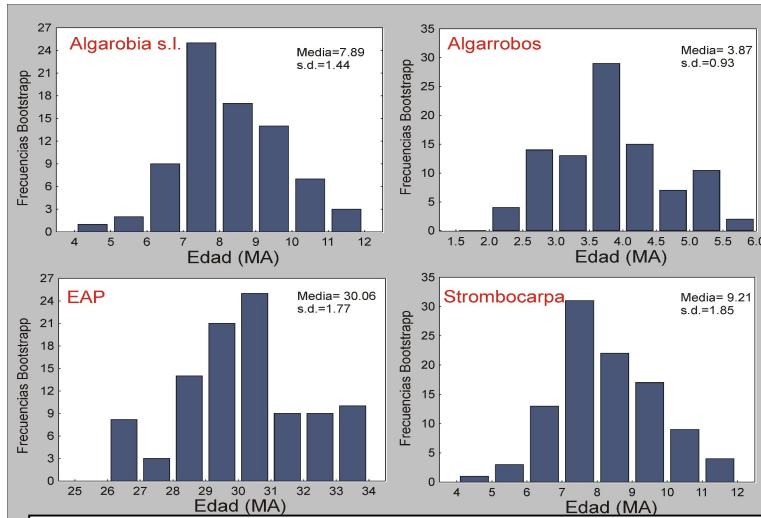


Figura 19. Análisis de bootstrapping de las edades estimadas para algunos nodos relevantes de la Figura 20

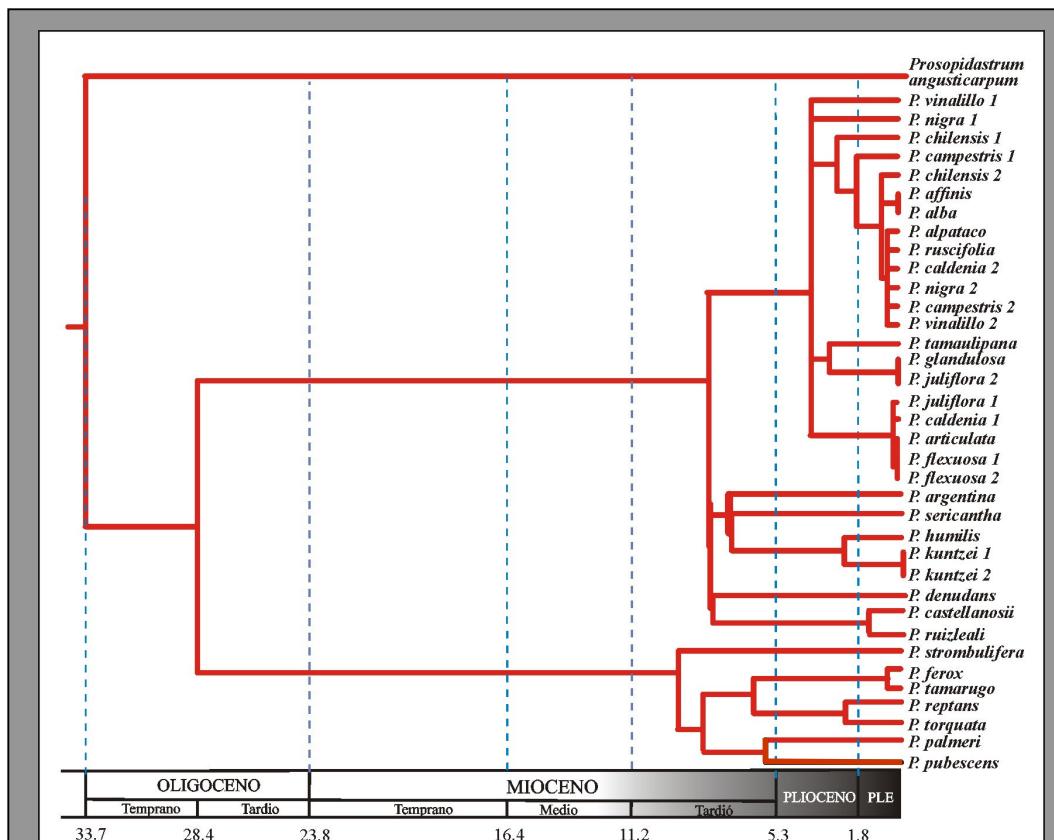


Figura 20. Cronograma resultante del análisis de Penalidad Likelihood. La topología corresponde a uno de los árboles más parsimoniosos obtenidos en el análisis de tres marcadores. El tiempo aproximado de ampliación de la distribución de zonas áridas en América se muestra sobre la escala de tiempo indicando con colores oscuros el momento de su máxima extensión. Ple: Pleistoceno

4.3.2. Análisis de diversificación

El análisis filogenético presentado en el capítulo anterior indicó que las especies de *Prosopis* del Viejo Mundo no están estrechamente emparentadas con las del Nuevo Mundo. Esto fue así para todos los análisis en el caso de *P. africana* y para *P. cineraria* en dos de los tres análisis. Tomando esto en cuenta, el estudio de diversificación y evolución de la afinidad climática se restringió al clado de las EAP + *Xerocladia*.

Tabla 18- Tiempos de divergencia (MA) obtenidos en el análisis de Penalized Likelihood. Las edades fueron obtenidas fijando el nodo por debajo del clado ASP + *Xerocladia* y *Prosopidastrum* en 33.2 MA siguiendo el valor medio obtenido por Lavin et al (2005). Las edades obtenidas considerando +/- una DE se dan entre paréntesis. En todos los análisis la edad del grupo corona de ASP + *Xerocladia* crown fue constreñido a una edad mínima de 28.4 MA siguiendo la evidencia fósil. ACMR: Ancestro Común Más Reciente.

ACMR de...	Grupo externo como en...	
	Análisis de tres marcadores	Análisis de dos marcadores
EAP + Xerocladia	28.96 (26.25 - 31.68)	29.37 (27.39 - 33.05)
Algarobia s.l.	7.89 (7.15 - 8.62)	7.96 (7.58 - 9.14)
Algarrobos	3.65 (3.31 - 3.99)	3.69 (3.38 - 4.08)
Strombocarpa	9.21 (8.35 - 10.07)	9.31 (9.18 - 11.07)

El análisis de linajes a través del tiempo (LTT) mostró claramente un incremento en la tasa de diversificación en este grupo desde el Mioceno Tardío hasta el presente (Figura 21). Dentro de este período un examen más detallado sugiere dos momentos diferentes de aceleración de la diversificación. Uno de ellos en el Mioceno Tardío y el otro en el Plioceno. Este patrón general no se vio afectado ni por la inclusión de especies que no formaban parte del estudio de tiempos de divergencia (Figura 21B) ni por las diferentes topologías utilizadas para derivar los cronogramas (Datos no mostrados). La tasa neta de diversificación estimada para diferentes clados dentro de *Prosopis* en comparación con la calculada para otros taxones se indica en la Tabla 19.

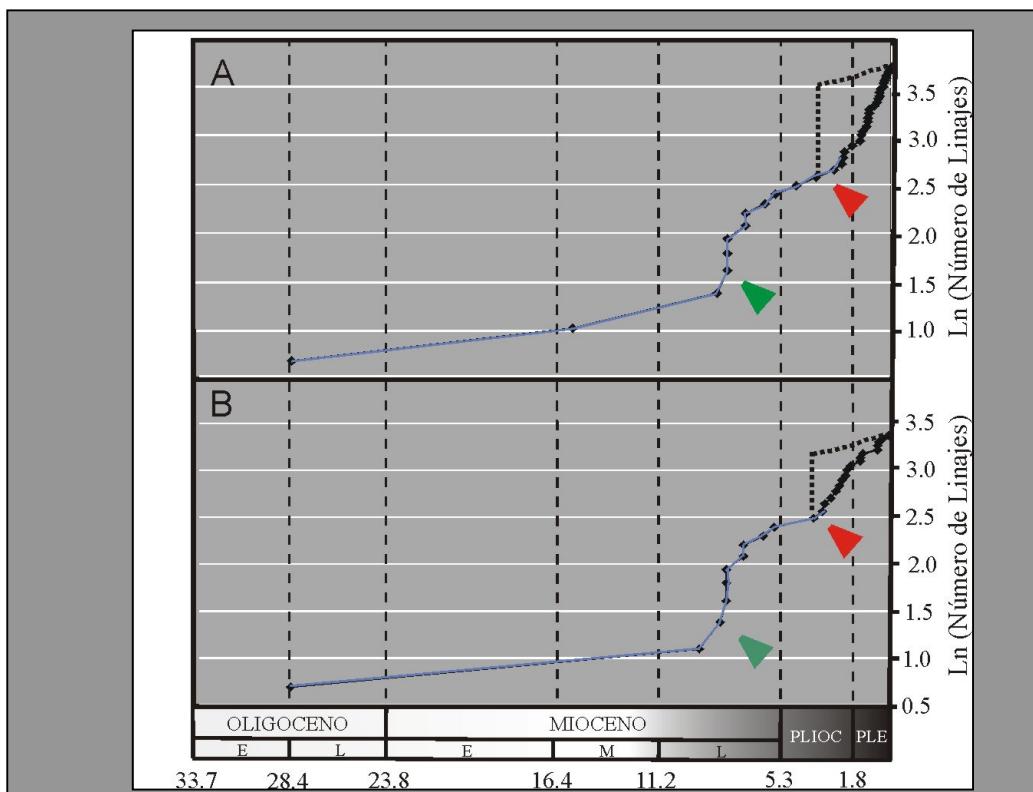


Figura 21. Análisis de linajes a través del tiempo (LTT). A-. Incluyendo especies ausentes. B-. Análisis con las especies incluidas en el análisis de tres marcadores. Las líneas punteadas representan el análisis de LTT cuando la politomía en el clado de los Algarrobos es considerada una politomía dura. Flechas verdes indican el comienzo de la diversificación dentro de Algarobia y Strombocarpa. Flechas rojas indican el comienzo de la diversificación dentro en algarrobos. E= Temprano, L= tarde, M= Medio. PLIOC= Plioceno; PLE= Pleistoceno

Tabla 19. Comparación de las tasas netas de diversificación (NDR) estimadas dentro de

	NDR	Forma de vida
<i>Prosopis</i>		
Clado Algarrobos	0.58-0.72 ^a	Árboles ^b
Clado <i>P. kuntzei-P. argentina</i>	0.16-0.21 ^a	Árboles y arbustos
Clado Strombocarpa	0.27-0.37 ^a	Árboles y arbustos
Clado EAP + <i>Xerocladia</i>	0.12-0.14 ^a	Árboles y arbustos
<i>Otros Taxa</i>		
<i>Inga</i> ¹	0.50	Árboles
<i>Lupinus</i> ²	1.92-3.84	Hierbas postradas ó arbusto
<i>Ruschioideae</i> ³	0.77-1.75	Plantas tipo piedra y arbustos
Clarkia, sect. Peripetasma ⁴	0.62	Hierbas
<i>Gentianella</i> ⁵	1.7-3.2	Hierbas y pequeños arbustos
<i>Valeriana</i> ⁶	0.8-1.3	Hierbas
<i>Yucca</i> ⁷	0.21-0.27	Plantas suculentas en roseta y arbóreas
<i>Ehrharta</i> ⁸	0.12-0.57	Hierbas
<i>Gaertnera</i> ⁹	0.71-0.83	Arbustos y pequeños árboles
Hawaiian silverwoods ¹⁰	0.55	Lianas y arbustos hasta Árboles
<i>Agave</i> ⁷	0.51	Plantas suculentas en roseta y arbóreas

Prosopis comparada con las obtenidas para otros grupos de plantas

^a Rango de valores estimados de tasas derivados de las diferentes estrategias de calibración. ^b 95 % de las especies son árboles o presentan formas arbóreas. ¹ (Richardson et al., 2001); ² (Hughes & Eastwood, 2006); ³ (Klak et al., 2003); ⁴ (Hey, 1992); ⁵ (Von Hagen & Kadereit, 2001); ⁶ (Bell & Donoghue, 2005); ⁷ (Good-Avila et al., 2006); ⁸ (Verboom et al., 2003); ⁹ (Malcomber, 2002); ¹⁰ (Baldwin & Sanderson, 1998).

4.4. Bibliografía

- Baldwin, B. G. & M. J. Sanderson. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences* 95: 9402-9406.
- Barraclough, T. G. & A. P. Vogler. 2002. Recent diversification rates in North American tiger beetles estimated from a dated mtDNA phylogenetic tree. *Molecular Biology and Evolution* 19: 1706-1716.
- Barker, F., A. Cibois, P. Schikler, J. Feinstein & J. Cracraft. 2004. Phylogeny and diversification of the largest avian radiation. *Proceedings of the National Academy of Sciences* 101: 11040-11045.
- Bell, C. D. & M. J. Donoghue. 2005. Phylogeny and biogeography of Valerianaceae (Dipsacales) with special reference to the South American valerians. *Organisms, Diversity and Evolution* 5: 147-159.
- Bell, C., D. Soltis & P. Soltis. 2005. The age of the angiosperms: a molecular timescale without a clock. 59: 1245 - 1258.
- Chalmers, A. 1984. ¿Qué es esa cosa llamada ciencia? XXI Editores, Madrid.
- Coyne, J. A. & H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland, MA.
- Felsenstein, J. 1981. Evolutionary trees from DNA sequences: A maximum-likelihood approach. *Journal of Molecular Evolution* 17: 368-76.
- Felsenstein, J. 1988. Phylogenies from molecular sequences: inference and reliability. *Annual Review of Genetics* 22: 521-565.
- Good-Avila, S. V., V. Souza, B. S. Gaut & L. E. Eguiarte. 2006. Timing and rate of speciation in *Agave* (Agavaceae). *Proceedings of the National Academy of Sciences* 103: 9124-9129.
- Gu, X. & W. Li. 1992. Higher rates of amino acid substitution in rodents than in man. *Molecular Phylogenetics and Evolution* 1: 211-214.
- Hedges, S., J. Blair, M. Venturi & J. Shoe. 2004. A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evolutionary Biology* 4.
- Hey, J. 1992. Using phylogenetic trees to study speciation and extinction. *Evolution* 46: 627-640.

- Hughes, C. E. & R. Eastwood. 2006. Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences* 103: 10334-10339.
- Jukes, T. & C. Cantor. 1969. Evolution of protein molecules. En M. HN (ed) *Mammalian Protein Metabolism*. Academic Press, New York, pp. 21-132.
- Klak, C., G. Reeves & T. Hedderson. 2003. Unmatched tempo of evolution in Southern African semi-desert ice plants. *Nature* 427: 63-65.
- Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16: 111-120.
- Kumar, S. & S. Subramanian. 2002. Mutation rates in mammalian genomes. *Proceedings of the National Academy of Sciences* 99: 803-808.
- Lavin, M., P. S. Herendeen & M. F. Wojciechowski. 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Systematic Biology* 54: 575-594.
- Lewis, P. 1998. Maximum likelihood as an alternative to parsimony for inferring phylogeny using nucleotide sequence data. In S. P. Sotis DE, Doyle JJ (ed) *Molecular systematics of plants II DNA sequencing*. Academic Publisher, Boston, MA, pp. 132-163.
- Malcomber, S. 2002. Phylogeny of *Gaertnera* Lam. (Rubiaceae) based on multiple DNA markers: evidence of a rapid radiation in a widespread, morphologically diverse genus. *Evolution* 56: 42-57.
- Nee, S., E. C. Holmes, A. Rambaut & P. H. Harvey. 1995. Inferring population history from molecular phylogenies. *Philosophical Transactions of the Royal Society of London, series B* 349: 25-31.
- Piel, K. M. 1971. Palynolgy of Oligocene sediments from British Columbia. *Canadian Journal of Botany* 49: 1885-1920.
- Posada, D. & K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14: 817-818.
- Rambaut, A. P. 2002. PhyloGen, Version 1.0. Oxford: Department of Zoology, University of Oxford.
- Richardson, J. E., R. T. Pennington, T. D. Pennington & P. M. Hollingsworth. 2001. Rapid diversification of a species-rich genus of Neotropical rain forest trees. *Science* 293: 2242 - 2245.

- Sanderson, M. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* 14: 1218–1231.
- Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* 19: 101-109.
- Sanderson, M. J. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19: 301-302.
- Sarich, V. & A. Wilson. 1973. Generation time and genomic evolution in primates. *Science* 179: 1200-1203.
- Swofford, D. L. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer, Sunderland, Mass.
- Tajima, F. 1993. Simple methods for testing molecular clock hypothesis. *Genetics and Molecular Biology* 135: 599-607.
- Takezaki, N., A. Rzhetsky & M. Nei. 1995. Phylogenetic test of the molecular clock and linearized trees. *Molecular Biology and Evolution* 12: 823-833.
- Tavaré, S. 1986. Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences. 17 American Mathematical Society: Lectures on Mathematics in the Life Sciences: 57–86.
- Verboom, G. A., H. P. Linder & W. D. Stock. 2003. Phylogenetics of the grass genus *Ehrharta*: Evidence for radiation in the summer-arid zone of the South African cape. *Evolution* 57: 1008–1021.
- Von Hagen, K. B. & J. W. Kadereit. 2001. The phylogeny of *Gentianella* (Gentianaceae) and its colonization of the southern hemisphere as revealed by nuclear and chloroplast DNA sequence variation. *Organisms Diversity & Evolution* 1: 61-79.
- Xiang, Q., S. Manchester, D. Thomas, W. Zhang & C. Fan. 2005. Phylogeny, biogeography, and molecular dating of cornelian cherries (*Cornus*, Cornaceae): tracking tertiary plant migration. *Evolution* 59: 1658–1700.
- Yule, G. U. 1924. A mathematical theory of evolution based on the conclusions of Dr. J. C. Willis. *Philosophical Transactions of the Royal Society of London, series B* 213.

Zuckerkandl, E. & L. Pauling. 1962. Molecular disease, evolution, and genetic heterogeneity. *In* M. K. a. B. Pullman (ed) *Horizons in Biochemistry*. Academic Press, New York., pp. 189–225.

CAPÍTULO 5 - OPTIMIZACIÓN DE CARACTERES AMBIENTALES

El presente capítulo tiene como objetivo inferir el ambiente en el cual habrían vivido los ancestros de *Prosopis*. En particular se buscó analizar la relación histórica de este género con los ambientes áridos. En la primera parte del capítulo se incluye una evaluación crítica de los métodos propuestos hasta el momento para inferir las condiciones ambientales ancestrales de un cierto grupo de especies a partir de la optimización sobre su filogenia de caracteres que describan el ambiente. Posteriormente, y tomando en cuenta las cuestiones discutidas en la primera parte del capítulo, se realizó la estimación condiciones ambientales ancestrales dentro de género *Prosopis*.

5.1. Optimización de condiciones ambientales: análisis crítico de la teoría subyacente a dichos estudios

5.1.1. Introducción

En los últimos años ha habido un creciente interés en inferir las condiciones ambientales en las cuales habrían vivido los ancestros de especies actuales (Graham et al., 2004; Hardy, 2006; Hardy & Linder, 2005; Hoffmann, 2005; Linder et al., 2006; Stephens & Wiens, 2003; Yesson & Culham, 2006a,b). Los estudios que analizan variables climáticas han sido bautizados “Modelaje Filoclimático” (Yesson & Culham, 2006a). Estos estudios parten de un “modelo climático” para especies actuales en el cual se definen los rangos de diferentes variables en los cuales habitan las especies estudiadas. Posteriormente estos valores son optimizados sobre la filogenia del grupo y se establecen las condiciones en las cuales habrían vivido los ancestros hipotéticos. En los últimos años diferentes estudios han evaluado ciertos puntos metodológicos sobre la reconstrucción de caracteres ecológicos (Hardy & Linder, 2005; Hardy, 2006), sin embargo en ningún caso se ha evaluado la lógica subyacente a tales estudios ¿Es posible obtener información de las condiciones ambientales en que vivían los ancestros del

grupo de estudio a partir de la optimización de los caracteres ambientales? ¿Cuáles son los supuestos que subyacen a estos estudios?

5.1.2. Carácteres ecológicos, nicho y herencia

Los estudios de estimación de estados para los nodos internos de una filogenia están totalmente basados en la noción de que los caracteres analizados pueden ser transmitidos, al menos parcialmente, entre ancestros y descendientes, es decir los caracteres son heredables. Sin este supuesto, no habría base teórica para realizar dichos análisis. La evaluación sobre la factibilidad o no de estimar los valores ancestrales de caracteres ecológicos debe incluir necesariamente la evaluación de los factores que determinan la heredabilidad de estos caracteres.

Dentro de los estudios a nivel macroevolutivos, la terminología asociada a la noción de heredabilidad es bastante deficiente. Los problemas surgen, al menos en parte, al querer realizar una analogía entre genética de poblaciones y macroevolución. En el marco de genética de poblaciones un carácter se considera heredable si presenta las siguientes condiciones:

- 1-. El mismo puede ser transmitido de generación en generación (esta genéticamente definido)
- 2-. Si presenta variación genética en dicho carácter.

Concomitantemente un carácter puede no presentar heredabilidad si: a-. el carácter no está genéticamente definido b-. el carácter está genéticamente definido pero no presenta variación. Dado esto, la heredabilidad de un carácter en genética de poblaciones no es la propiedad del mismo en ser transmitido de generación en generación sino la capacidad que tiene de poder cambiar (evolucionar) de una generación a la otra. Esta noción de genética de poblaciones ha sido utilizada en diferentes estudios macroevolutivos y puede ser de utilidad al intentar estudiar posibles

procesos de selección a nivel de especies (Jablonski, 1987; Savolainen et al., 2002). Sin embargo, en estudios de optimización de caracteres lo que nosotros debemos analizar es solamente la primera de las condiciones indicadas: si dicho carácter puede ser transmitido de ancestros a descendientes. Así, a lo largo del trabajo se considerará esta definición restringida de heredabilidad.

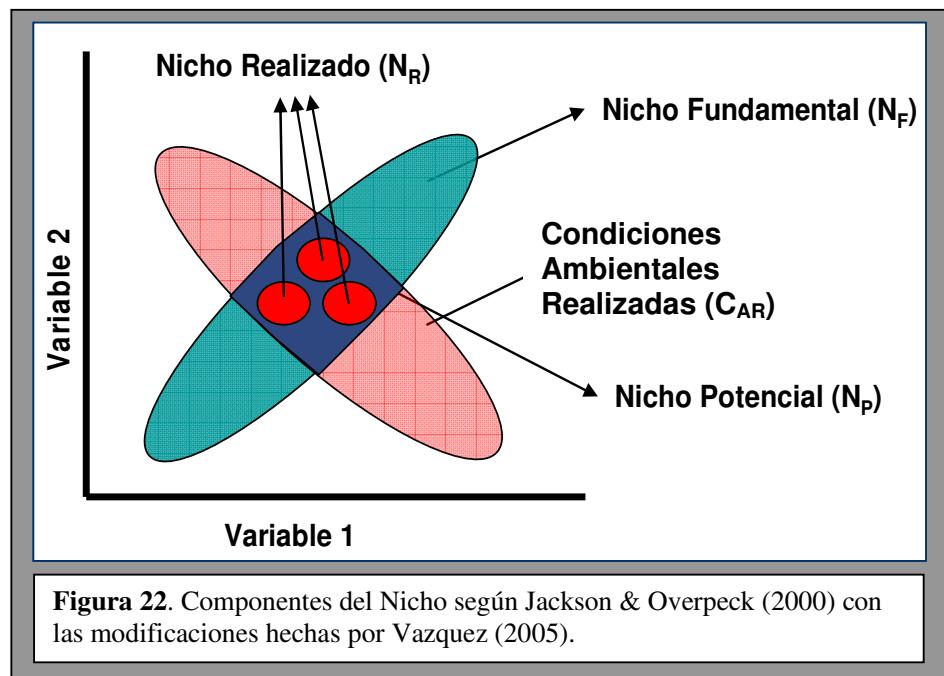
Otra cuestión asociada al término heredabilidad reside en la definición utilizada en estudios macroevolutivos: “*offspring resembling (significantly) their parents*” es decir “descendencia que se parece significativamente a sus padres” haciendo nuevamente una analogía con genética de poblaciones (Jablonski, 1987; Savolainen et al., 2002). Esta definición confunde causas con consecuencias. El proceso es la herencia del carácter (su transmisión), mientras que la similitud es la resultante de tal proceso. En el momento de evaluar si un carácter es heredable, es posible hacerlo mediante el estudio de la similitud entre padres e hijos, pero heredabilidad no *es* similitud. Aquello que define que un carácter sea heredable no es la similitud sino la transmisión de información entre padres e hijos; un legado que produce similitud. Especies hermanas pueden parecerse en determinadas características, por otras causas que no impliquen herencia. Esta distinción adquiere mucha relevancia al estudiar caracteres ecológicos, donde la similitud no sólo surge de la herencia sino también de la interacción entre la especie y su ambiente.

La discusión previamente planteada sobre la definición de heredabilidad tuvo como fin buscar una definición que simplifique la presentación de los problemas asociados a la optimización de caracteres ambientales. Las conclusiones de este capítulo no cambiarían al considerar una definición alternativa de heredabilidad, sólo harían más complicada la presentación de las ideas.

5.1.3. Teoría de nicho y herencia

En la mayoría de los trabajos previos los caracteres considerados para describir el ambiente (temperatura, humedad, tipo de suelo, etc.) eran denominados caracteres “ecológicos” con el objeto de diferenciarlos de los caracteres utilizados generalmente en estudios filogenéticos (morfológicos o moleculares). Estas características del ambiente en el que habitan las especies definirían el nicho de las mismas, siendo la estimación de valores ancestrales para estos caracteres la estimación del nicho ancestral. Desde el trabajo seminal de Hutchinson (1957), la definición de nicho y la teoría subyacente han estado en continua revisión y modificación. En algunos casos la diferencia en las definiciones han estado relacionadas con la preponderancia que cada autor le ha dado a diferentes facetas de tal teoría, siendo posible que diferentes definiciones puedan ser útiles en diferentes contextos. En nuestro caso necesitamos una definición que nos permita comprender si es posible establecer conclusiones sobre las condiciones ambientales de los ancestros a partir de los caracteres de las especies actuales. Considero que la mejor definición para lograr este objetivo es la propuesta por Jackson & Overpeck (2000) con las modificaciones hechas por Vazquez (2005). Desde este punto de vista (Figura 22), el nicho estaría compuesto por:

- 1-. **Condiciones Ambientales Realizadas (CA_R)**: son las condiciones ambientales que existen en la naturaleza en un momento histórico dado.
- 2-. **Nicho Fundamental (N_F)** es la combinación de condiciones donde una especie lograría sobrevivir, existan o no tales condiciones en un momento histórico dado.



3-. **Nicho Potencial (N_P):** surge de la intersección de los dos primeros. Es la porción de las C_{AR} en las cuales la especie puede potencialmente sobrevivir.

4-. **Nicho Realizado (N_R):** son aquellas condiciones ocupadas realmente por la especie en un momento dado, siendo un subconjunto del N_P . La diferencia entre el N_P y el N_R se debe entonces a la contingencia histórica que hace que la especie no esté presente en todos aquellos sitios que podría ocupar.

La diferencia entre las visiones de Vazquez (2005) y de Jackson & Overpeck (2000) proviene de que el primero incluye a las interacciones biológicas como parte que delimita el N_F y no el N_R . La utilización de esta definición permite considerar que el N_F se mantendrá constante si la especie no cambia su composición genética.

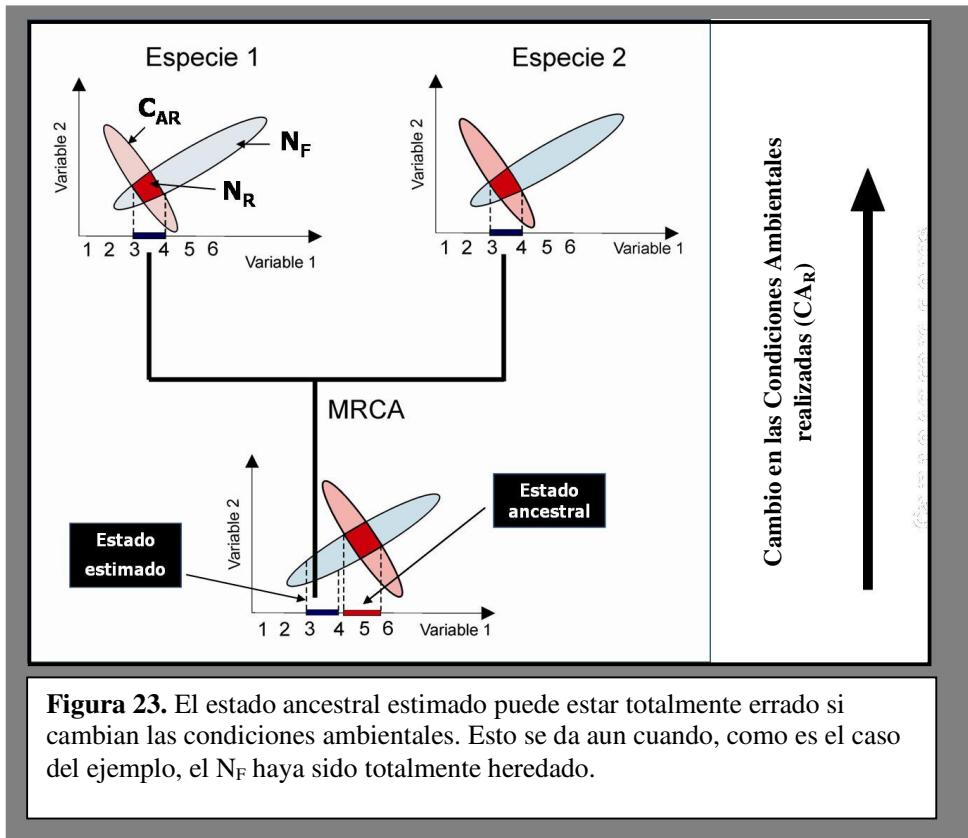
Durante esta exposición supondremos que el N_F es heredado perfectamente. Esto no se debe a que considere que siempre sea así sino porque intentaré discutir aquí aquellos problemas asociados a optimizar caracteres ambientales. La optimización del

N_F puede tener los mismos problemas que la optimización de cualquier carácter morfológico con lo cual los problemas se sumarían.

La heredabilidad del N_R implica que los componentes del mismo sean heredables. Para simplificar la exposición del problema, primero supondré que el N_R es igual al N_P es decir la especie está presente en todo el rango de condiciones ambientales presentes en un momento dado en las cuales dicha especie puede sobrevivir. Dada esa condición el N_R surge de la intersección entre el N_F y las CA_R . Siguiendo la definición adoptada, el N_F está determinado completamente por las características biológicas propias de la especie, y consecuentemente podrá ser perfectamente heredado siempre y cuando el acervo genético de la especie sea heredado (recordar que consideramos esto como supuesto). Si las características biológicas de la especie fueran las únicas que determinaran el N_R de la especie, sería factible estimar directamente el N_R ancestral a partir del N_R actual. Sin embargo el N_R está también determinado por las CA_R y tales variables (e.g. rango de O_2 , humedad, etc.) son por definición externas a las especies y consecuentemente imposible de ser heredadas por las mismas. La idea de que el N_R es heredable supondría que las especies tuvieran un mecanismo para transmitir las condiciones ambientales *per se*, algo que para la extensísima mayoría de los casos es imposible. Así, estrictamente hablando, en N_R no puede ser heredado.

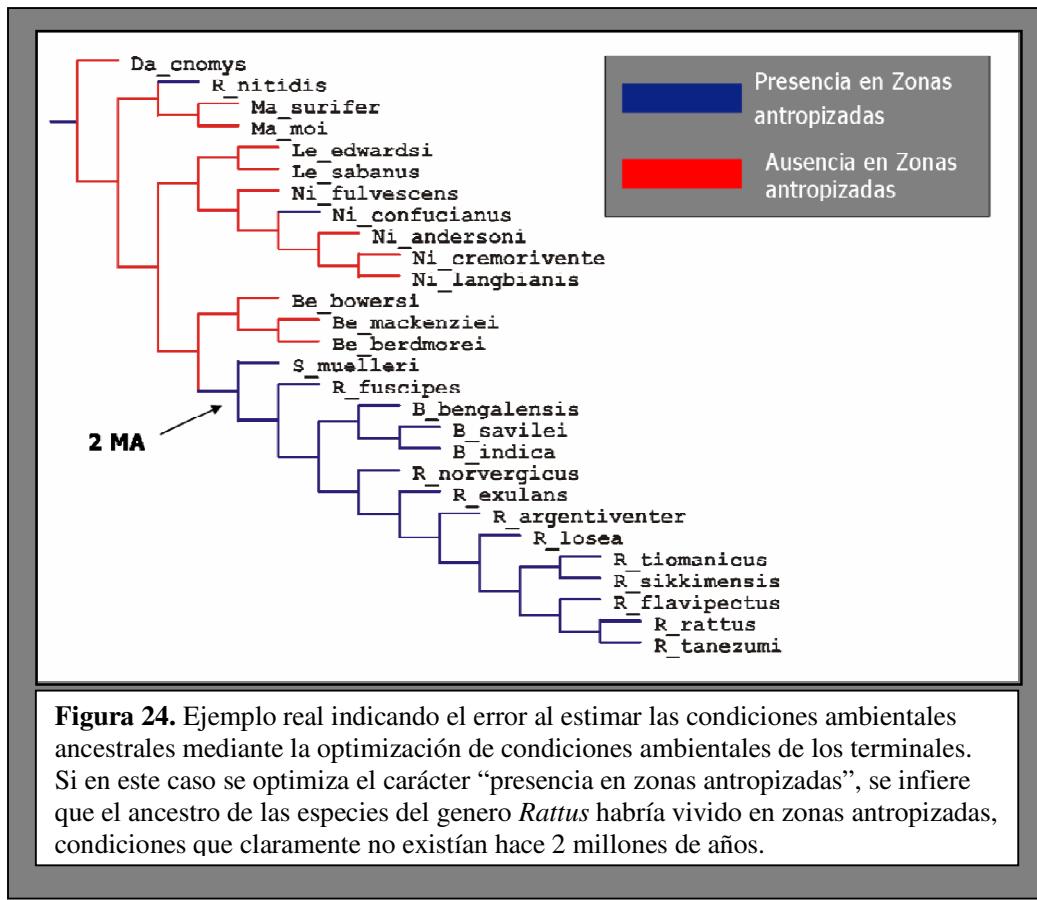
Que el N_R no pueda ser heredado no implica que las especies cercanamente emparentadas no puedan parecerse unas a otras con respecto al ambiente en que habitan. De hecho son innumerables los ejemplos de especies cercanas ocupando ambientes muy similares. Sin embargo esta similitud no es debida a que el N_R sea heredable ya que las especies pueden sólo heredar el N_F y no el N_R . Como fue anteriormente indicado similitud y herencia no son lo mismo.

Además de heredar el N_F , la otra cuestión necesaria para que dos especies hermanas sean similares en su N_R es que las CA_R no cambien al menos en el límite definido por la intersección con el N_F . Si un cambio en la CA_R modifica la intersección entre CA_R y N_F , ambas especies podrán tener el mismo N_R pero diferente al de su ancestro (Figura 23).



Este punto puede ilustrarse con un ejemplo real. Dentro de género *Rattus* existen una serie de especies que hoy en día se hallan en ambientes antropizados: edificios, ciudades, casas y terrenos cultivados (Figura 24). Si nosotros intentásemos estimar el ambiente ancestral a partir de la optimización de un carácter que defina el ambiente en que viven encontraríamos que la asignación al nodo ancestral (datado en aproximadamente 2 MA, (Verneau et al., 1998) incluiría ambientes antropizados. Sin embargo el conocimiento sobre la historia de la humanidad indica claramente que el

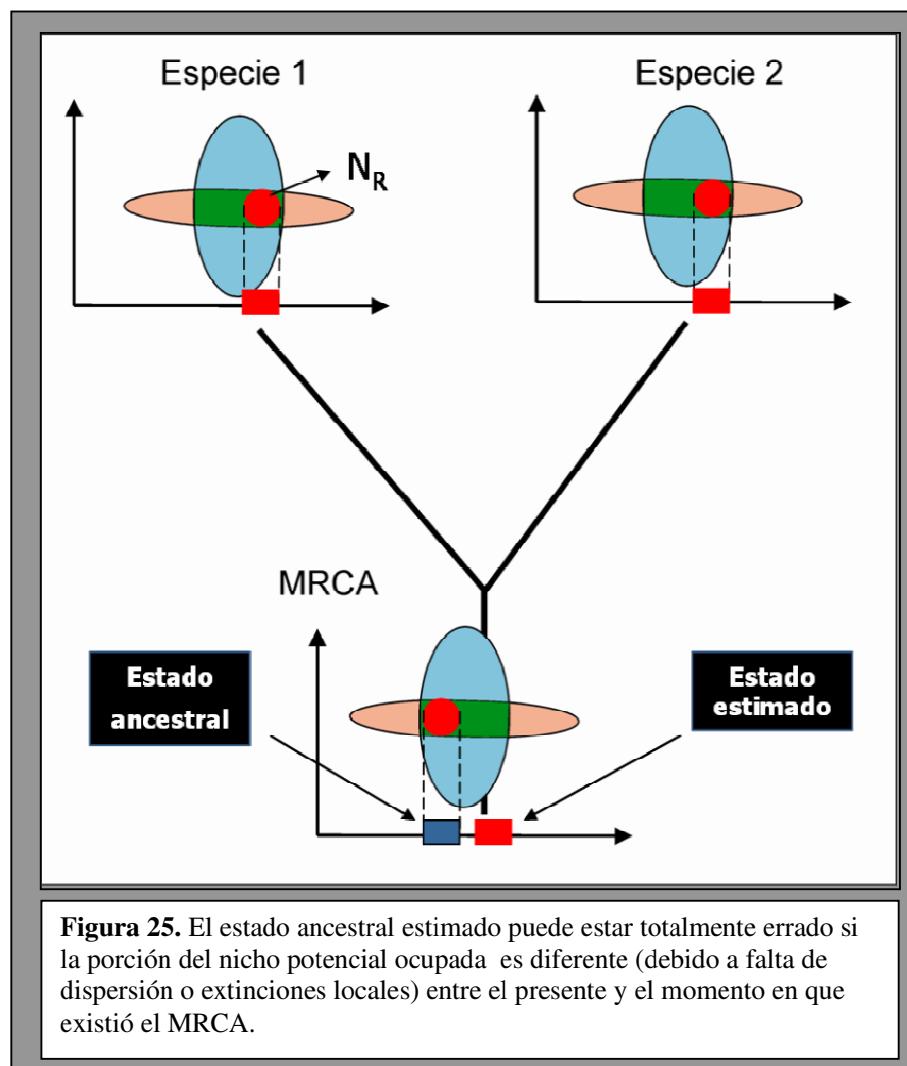
hombre comenzó a generar tales ambientes mucho más tarde, siendo imposible que tal ancestro hubiera ocupado ambientes antrópicos. El marco de la Teoría de Nicho nos permite reconocer cual fue la falla en nuestra evaluación: todas las especies de este grupo habría *heredado la capacidad* de ocupar estos ambientes de su ancestro (es decir la capacidad ya formaba parte de su N_F). Sin embargo el origen de los ambientes antrópicos fue mucho posterior, haciendo que el N_R haya cambiado en paralelo en todas estas especies. La optimización inicial entonces sólo indicó que el N_F había sido heredado del ancestro común más reciente (ACMR) y nada más. Sin otra evidencia (las CA_R ancestrales), sólo esto puede ser inferido.



5.1.4. Cuando el Nicho Realizado es diferente al Nicho Potencial

En la sección anterior se indicó un supuesto no explicitado en estudios empíricos previamente publicados sobre optimización de este tipo de caracteres: para estimar

estados ancestrales de caracteres ambientales las CA_R no deben cambiar (en el límite del N_F) o su cambio debe ser conocido. Para llegar a esta conclusión hicimos una simplificación: N_P=N_R. Como se indicó previamente, la diferencia entre N_P y N_R estaría dada, según Vázquez (2005), por contingencia histórica: extinciones locales y/o dispersión limitada determinarían que las especies no estén en todo su N_P (Es importante recordar que dada la definición utilizada la interacción con otros organismos definirían el N_F y no el N_R). Esta diferencia que puede existir entre el N_R y el N_P es un problema adicional a la hora de determinar el N_R ancestral. De hecho, este problema puede producir una estimación completamente errónea de las condiciones ancestrales si el ancestro ocupaba una porción diferente del N_P que sus descendientes (Figura 25).



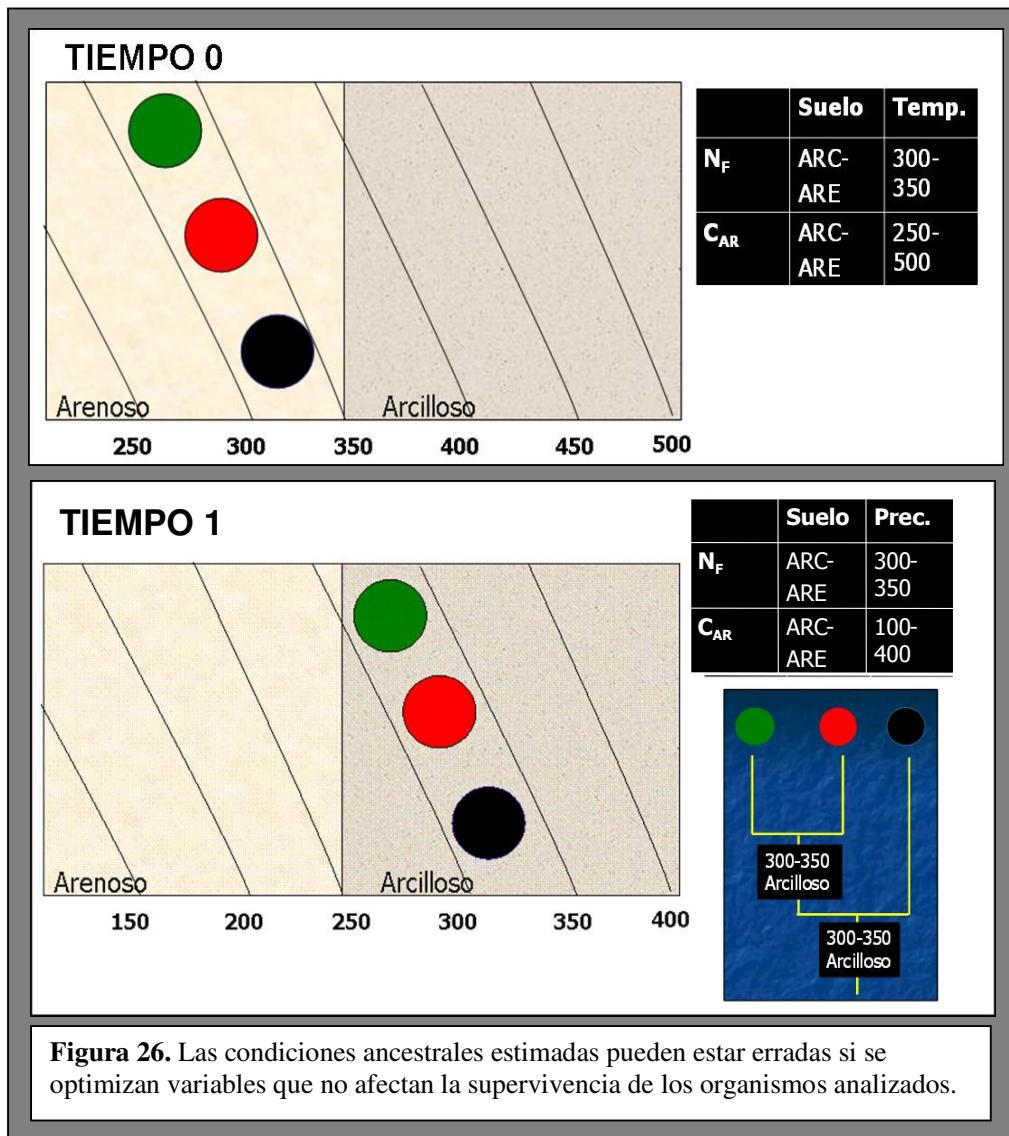
Las causas de las diferencias entre el N_P y el N_R pueden ser muchas: extinción local; falta de dispersión; especies recién formadas. Cuanto menor sea la dispersión, mayor la tasa de extinción local y menor el tiempo desde la formación de las especies es de esperar que el N_R sea más diferente que el N_P .

5.1.5. Variables: más y más problemas en la estimación del nicho ancestral

Hasta aquí hemos supuesto que el nicho en todos sus componentes está formado por un grupo de variables independientes, en el cual si tomamos por ejemplo los valores del nicho para dos variables, el nicho final sería un rectángulo que queda definido por los límites de ambas variables. Sin embargo, como otros autores han indicado (Jackson & Overpeck, 2000; Ackery, 2003) muchas veces el N_R no tendrá esta forma. Una de las causas de tal fenómeno se debe a que es posible que para un momento y lugar dado no estén presentes todas las combinaciones de valores para las variables ambientales analizadas. Por ejemplo, la temperatura media anual puede variar entre 10 y 20°C y la precipitación entre 100 y 200 mm. Sin embargo es posible que no existan regiones con 100 mm y 17°C. Esto tiene como consecuencia que si modelamos a la distribución de las especies a lo largo de diferentes variables como rectángulos, no podremos definir el ambiente (entendiendo ambiente como combinación de los valores de las diferentes condiciones ambientales).

Una cuestión asociada a la previamente indicada está relacionada ya no con la ausencia de determinadas combinaciones de valores para las variables ambientales sino con la forma en la que las especies perciben tales variables. Uno de estos casos ocurre cuando la especie presenta mucha tolerancia a cambios en una de las variables o directamente no se ve afectada de manera alguna por la misma. En el ejemplo de la Figura 26 las tres especies se originaron y diversificaron en un ambiente con terreno arenoso y con valores de precipitación de entre 300 a 350 mm. Sin embargo su N_F

también incluye condiciones arcillosas, es decir está biológicamente preparada para vivir con suelos arcillosos. Estas especies no ocupan tal ambiente solo porque la precipitación del lugar no se lo permite. Ahora supongamos que existe un cambio en la precipitación de manera tal que haya una nueva combinación de CA_R (Tiempo 1, Figura 26).



Dado nuestro supuesto que el N_F es perfectamente heredado, estas especies tienen 2 posibles caminos: (i) se extinguen (ii) migran “siguiendo” el ambiente, en este caso en particular siguiendo los valores de precipitación, ya que el tipo de suelo les es indiferente. En el caso que sobrevivan, estas especies seguirán ocupando un ambiente con precipitaciones entre 300 y 350 mm, pero ahora ocupando un terreno arcilloso. Si quisiésemos establecer las condiciones del ancestro de estas especies a partir de la distribución del tiempo 1, la determinación del estado ancestral sería la correcta para la precipitación pero incorrecta para el tipo de suelo.

5.1.6. Problemas en los Test propuestos para evaluar la heredabilidad de nicho

La idea que el N_R debe ser heredable para poder establecer conclusiones a partir de los valores ancestrales inferidos no es una idea original de este trabajo. De hecho, diferentes trabajos han intentado evaluar tal hipótesis auxiliar antes de optimizar este tipo de caracteres (Yesson & Culham, 2006a,b). Sin embargo a diferencia del presente análisis, estos estudios empíricos evaluaron esta hipótesis mediante una aproximación probabilística (Yesson & Culham, 2006a,b) utilizando el índice de QVI (Ackerly & Donoghue, 1998) un índice que evalúa el nivel de convergencia de un carácter dado una determinada optimización. Según estos autores si el QVI es significativamente menor que el esperado por azar entonces los resultados de la optimización serían válidos. El problema principal de esta aproximación es que puede dar resultados altamente significativos aun cuando el estado de muchos de sus nodos o inclusive todos ellos fueran diferentes a los que realmente presentaron los ancestros. La causa de esto ya fue explicada en el apartado 5.1.3. y ejemplificado en la Figura 24. Si se calcula la significación de QVI para el ejemplo de la Figura 24, la probabilidad que el nivel de QVI observado (0.308) sea menor que el esperado por azar es de 0.0017 (Ver Anexo II). Es decir el test de QVI indicaría que no existe evidencia para desconfiar de la

optimización. Sin embargo, como ya dijimos, la mayoría de los nodos de esa filogenia tienen asignaciones que son incorrectas. Este resultado no es debido al índice utilizado. Cualquier índice que intente analizar de esta manera la heredabilidad del N_R caerá en el mismo problema. Una pregunta lógica que surge entonces es porqué el valor del test da significativo. La respuesta es que dicho test da significativo porque el N_F sí ha sido heredado. Cómo las CA_R para todas las especies actuales son las mismas y considerando que el N_R puede ser similar al N_P de cada especie entonces si el N_F se hereda el test da un resultado significativo.

Otro problema asociado al uso de test como el QVI para evaluar la validez de las optimizaciones es que se refieren a patrones observados a lo largo de todo el árbol. Aun si dieran altamente significativos y no existieran los problemas indicados en el párrafo anterior, no podríamos saber a ciencia cierta si para un nodo de interés, la asignación es la correcta o no.

5.1.7. Posibles correcciones o mejoras a la metodología de estimación de condiciones ambientales ancestrales

Las críticas hechas a la metodología utilizada para la estimación de condiciones ambientales ancestrales son un punto de partida necesario para la incorporación de correcciones y/o mejoras a tales estudios. Una de las principales cuestiones a tener en cuenta es la necesidad de contar con información sobre las CA_R para el marco temporal en que se busca estimar el N_R de los ancestros hipotéticos. Es así que se requiere del conocimiento del tiempo en el cual habrían divergido las especies estudiadas así como también de las CA_R en este período. La fuente de evidencia del primer punto podría ser dado por datos paleontológicos sobre el grupo o por datación molecular, mientras que el segundo de los puntos debe ser definido a partir de datos paleoclimáticos y paleoecológicos. Cuanto más sustancial sea la evidencia que apoya las reconstrucciones

paleoclimáticas y mayor también la precisión en el marco temporal de diversificación, mayor será la verosimilitud de las reconstrucciones de N_R ancestrales.

Otra cuestión importante a considerar es si en las especies estudiadas el N_R puede ser considerado como similar al N_P (Inciso 5.1.4.). Así si las especies tiene alta tasa de dispersión, baja tasa de extinción local (por ejemplo debido a grandes tamaños poblacionales) y son especies que se han formado hace ya largo tiempo, es posible N_R y N_P sean similares y las estimaciones de condiciones ambientales ancestrales presenten mayor verosimilitud.

En último término se requiere de tener información sobre la posible interacción entre las variables ambientales así como también de su interacción con las especies analizadas. Como fue indicado en el apartado 4.1.5. sólo tiene sentido optimizar variables que se considera que realmente afectan de algún modo la supervivencia de las especies estudiadas.

5.1.8. Consideraciones Finales

Este capítulo no intenta ser una proclama en contra de la optimización de caracteres ambientales, ni tampoco pretende indicar que la mayoría de los resultados que se obtienen mediante esta metodología son incorrectos. La motivación de este trabajo ha sido poner en relieve algunos puntos que considero no han sido discutidos de manera profunda en la bibliografía y deberían ser tenidos en cuenta a la hora de evaluar condiciones ambientales de ancestros dada una filogenia. Si los problemas planteados en este trabajo afectan a la mayoría de los estudios o no es una cuestión sin respuesta clara. Uno de los puntos que pude ser considerado en contra de las críticas planteadas en este trabajo es que mucho trabajos empíricos (Peterson et al., 1999; Prinzing et al., 2001; Webb et al., 2002) mostrarían que el existiría una clara conservación del nicho. Sin embargo muchos de ellos se basaron en test estadísticos como los indicados

previamente, y que como fue dicho los mismos no evalúan la conservación del N_R. Inclusive uno de los trabajos más citados sobre conservación de nicho evaluando especies hermanas (Peterson et al., 1999) concluye que el mismo patrón se podría haber obtenido si las dos especies hubieran cambiado en paralelo sus nichos. Resultaría entonces más útil la evaluación de cada problema en particular, teniendo en cuenta la biología e historia del grupo y no mediante de Test “mágico-estadísticos”.

5.2. Análisis de condiciones ambientales ancestrales en *Prosopis*

5.2.1 Introducción

Uno de los objetivos de la presente tesis es la evaluación del marco ambiental en el cual se habría desarrollado la diversificación del género. Para ello hemos establecido primero las relaciones históricas entre las especies (Capítulo 2) y luego hemos visto que el marco temporal de su diversificación coincide con el desarrollo de ambientes áridos en América (Capítulo 4). Además también hemos visto que existía un aumento en la tasa de diversificación en momentos de mayor extensión de ambientes áridos en Sudamérica (Capítulo 4). Podríamos decir que nuestros datos sugieren una relación entre la evolución de *Prosopis* y los ambientes áridos. Con el objeto complementar estos análisis realizamos un estudio de optimización de condiciones ambientales sobre la filogenia de *Prosopis* buscando evaluar si los ancestros de *Prosopis* ya habitaban condiciones áridas o si las especies de este género solo colonizaron tales ambientes posteriormente.

Para estudiar la relación histórica entre *Prosopis* con ambientes áridos es necesario elegir la/las variables que sean las mejores descriptoras de este ambiente, teniendo en cuenta además que cumpla con las consideraciones teóricas remarcadas en la primera parte de este capítulo. Entre las posibles variables a considerar para analizar nuestra problemática (asociación histórica de *Prosopis* con ambientes áridos) se

encuentran la precipitación y la evapo-transpiración. La correlación entre los valores de estas variables con ambientes áridos es en líneas generales proporcional en el caso de la evapo-transpiración e inversamente proporcional en el caso de la precipitación. Sin embargo cada una de las variables por separado no permiten determinar de manera correcta si un ambiente es árido o no. Así, es posible encontrar ambientes áridos con valores más altos de precipitación que en ambientes semiáridos si el nivel de evapo-transpiración también es alto. Es entonces la evaluación conjunta de los valores de estas variables las que permiten una mejor caracterización del ambiente. Una variable más adecuada para considerar tal aspecto del ambiente es el índice de humedad (IH). Dicho índice se define como la razón entre la precipitación y la evapo-transpiración potencial. (UNEP, 1992). Esta será la variable que consideraremos en el presente análisis.

Otra cuestión a evaluar en estudios de optimización de caracteres ambientales es la codificación de los caracteres. Existen una gama de modos de codificar tales variables: codificación binaria, continua, de polimorfismos, etc. (Hardy & Linder, 2005). Una de estas formas de codificar los caracteres es la denominada Min-Max (Graham et al., 2004) en la cual los valores máximos y mínimos son optimizados independientemente y recién a posteriori se define el intervalo de valores para los nodos ancestrales. Esta codificación es particularmente adecuada para aquellos casos en que, como el nuestro, la tolerancia a altos valores de IH puede estar desacoplada de la tolerancia a valores altos de IH. Es así que a lo largo de este trabajo utilizamos tal codificación.

Con el objeto de estudiar la relación histórica de *Prosopis* con ambientes áridos, se calcularon los valores máximos y mínimos de IH (Tabla 20) para la distribución natural de cada una de las especies de *Prosopis* incluidas en la filogenia presentada en la Figura 8 del Capítulo 2. La distribución de cada especie fue definida a partir de la

bibliografía (Ross, 1975; Burkart, 1976; Rzedowski, 1988; Roig, 1993) y de especímenes de herbario (CTES, LIL, SI, BAFC). Los datos climáticos para la distribución de cada una de las especies estudiadas fueron obtenidas de la base de datos de clima mundial de la FAO usando el programa New LocClim (FAO, 2005). Las regiones fueron clasificadas en función de los valores de IH siguiendo a UNEP (1992). 0.05 < Hiperárido; 0.05-0.20 Árido; 0.20-0.50 Semiárido; 0.50-0.65 Seco-subhúmedo; > 0.65 Húmedo.

Tabla 20. Valor Máximo y Mínimo para el índice de humedad para las especies incluidas en el análisis de optimización de variables ambientales.

Especies	Índice de Humedad	
	Minima	Maxima
<i>X. viridinamis</i>	1,44	2,39
<i>P. affinis</i>	5,05	15,44
<i>P. alba</i>	2,52	11,75
<i>P. alpataco</i>	0,65	3,13
<i>P. argentina</i>	0,65	2,65
<i>P. articulata</i>	1,26	1,52
<i>P. caldenia</i>	2,61	7,11
<i>P. campestris</i>	4,17	4,86
<i>P. castellanosii</i>	1,16	1,16
<i>P. chilensis</i>	0,09	5,86
<i>P. denudans</i>	1,02	2,46
<i>P. ferox</i>	1,96	2,77
<i>P. flexuosa</i>	0,65	6,43
<i>P. glandulosa</i>	0,23	5,09
<i>P. humilis</i>	2,22	15,02
<i>P. juliflora</i>	4,25	13,33
<i>P. kuntzei</i>	2,53	11,75
<i>P. nigra</i>	3,62	15,44
<i>P. palmeri</i>	0,28	0,84
<i>P. pubescens</i>	0,42	1,19
<i>P. reptans</i>	2,43	8,87
<i>P. ruizleali</i>	1,9	1,9
<i>P. ruscifolia</i>	4,11	8,02
<i>P. sericantha</i>	2,22	8,18
<i>P. strombulifera</i>	0,02	2,75
<i>P. tamarugo</i>	0,02	0,09
<i>P. tamaulipana</i>	1,44	5,3
<i>P. torquata</i>	1,03	4,86
<i>P. vinalillo</i>	4,44	6,48

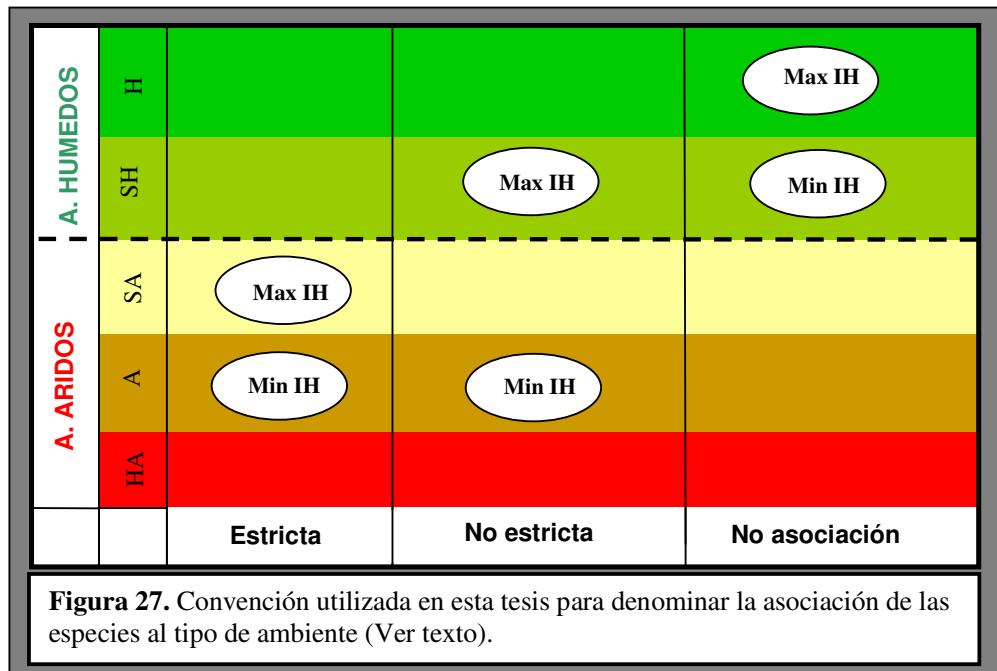
Los valores máximos y mínimos de aridez fueron optimizados como caracteres continuos siguiendo la implementación en TNT (Goloboff et al., 2003) del algoritmo Wagner (Farris, 1970). Con el objeto de evaluar de que manera la incertidumbre en el estado de la raíz afectaba los resultados, todos los análisis se repitieron considerando dos valores extremos de IH: húmedo e hiper-árido. Además, con el objeto de evaluar si incertidumbre en las relaciones dentro del clado de los algarrobos afectaba los resultados, el análisis se repitió considerando 10.000 resoluciones diferentes de este clado. Este estudio se realizó mediante scripts de TNT incluidos en el Apéndice II.

A la hora de evaluar las asignaciones a cada nodo obtenidas mediante la optimización se siguió la siguiente convención (Figura 27):

Asociación estricta con un ambiente árido: Todo el rango de valores de IH está en ambientes semiáridos, áridos o hiper-áridos. Es decir el valor de la variable IH mínimo (IH Min) y el valor de la variable IH máximo (IH Max) tienen valores menores al límite entre las clases subhúmedo y semiárido.

Asociación no-estricta con un ambiente árido: Parte del rango de valores de IH esta en ambiente semiáridos-áridos o hiper-áridos. Es decir el valor de la variable IH mínimo es menor que el valor límite entre las clases subhúmedo y semiárido mientras que el valor de IH máximo es mayor que tal límite.

Asociación no-estricta con un ambiente árido: Todo el rango de valores de IH esta en ambiente subhúmedos o húmedos. Es decir el valor de la variable IH mínimo y el valor de la variable IH máximo tienen valores mayores al límite entre las clases subhúmedo y semiárido.

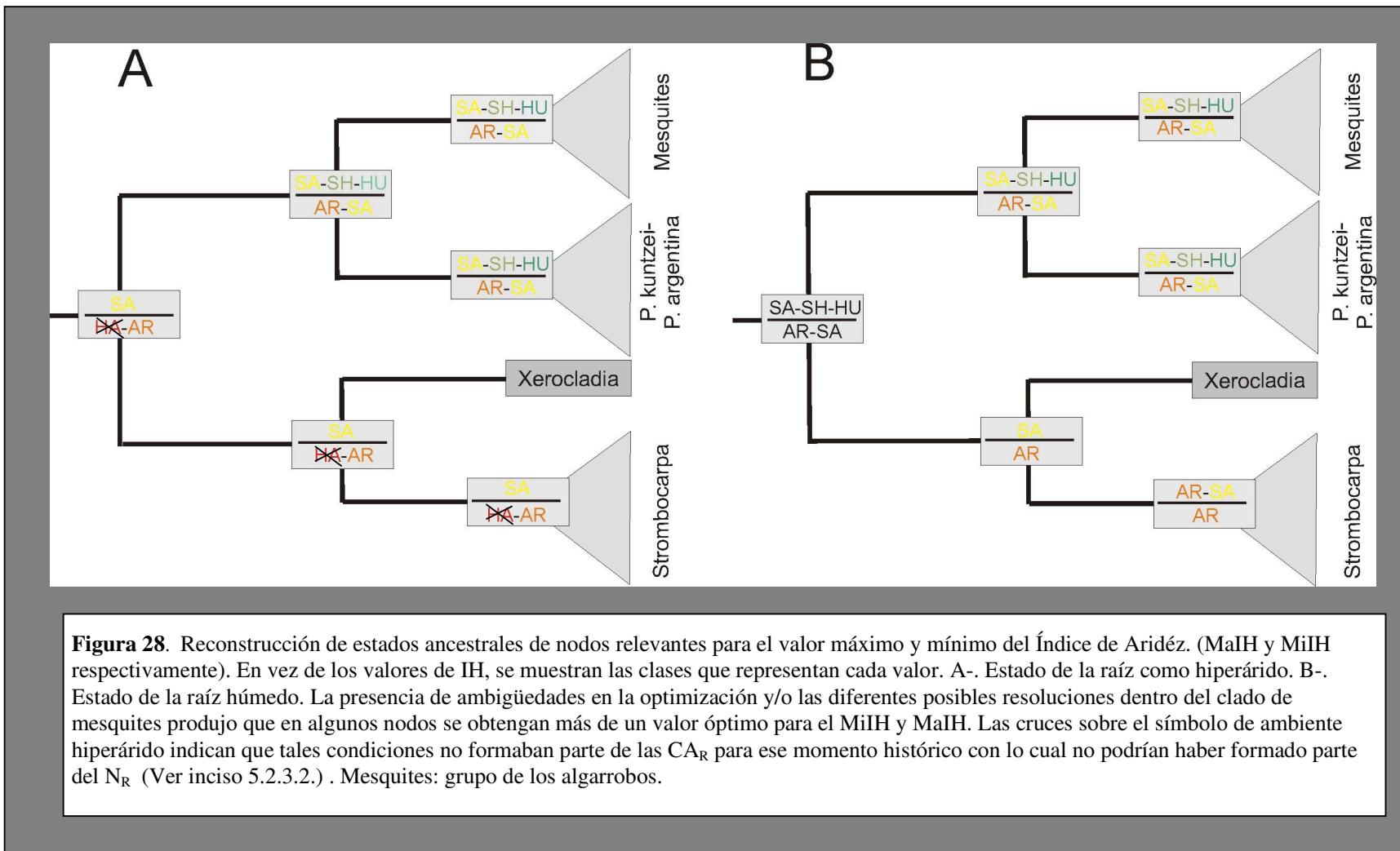


5.2.2. Resultados

Los resultados de la optimización de los valores de las variables IH Min y IH Max sobre la filogenia de *Prosopis* se resumen en la Figura 28. Con el objeto de facilitar la presentación de los resultados en la misma sólo se muestran los valores obtenidos para algunos nodos relevantes. La presencia de más de un valor de cada variable se debe a dos factores: (a) más de un estado óptimo para una topología determinada (b) diferentes valores obtenidos en las diferentes resoluciones de la politomía presente en el clado de los Algarrobos. A pesar de esta ambigüedad, fue posible extraer conclusiones válidas de este estudio.

La optimización de los valores del índice de humedad indicó que el rango de condiciones para el ancestro común más reciente (ACMR) del clado formado por las especies americanas de *Prosopis* más *X. viridiramis* incluyó valores representativos de ambientes áridos (Figura 28a). Al considerar el estado de la raíz como hiper-árido, la asignación al nodo indicado previamente implicó una asociación estricta con ambientes

áridos en todas las reconstrucciones más parsimoniosas. Cuando el estado asignado a la raíz fue húmedo, tal asociación estricta sólo se obtuvo en algunas de las reconstrucciones, mientras que en el resto de las reconstrucciones la asociación con ambientes áridos fue no estricta. Para el caso del clado *Strombocarpa*, las asignaciones obtenidas fueron muy similares independientemente del estado considerado para la raíz: una asociación estricta del nodo ancestral con condiciones áridas. La única diferencia fue que al considerar al estado ancestral como hiper-árido, este estado también fue incluido entre los óptimos del nodo ACMR de *Strombocarpa*. El mismo resultado se obtuvo para el ACMR de *X. viridiramis* + clado *Strombocarpa*. Tanto para el ACMR de Algarobia s.l., como para el de *P. kuntzei* - *P. argentina* y el del clado de los algarrobos, algunas de las reconstrucciones óptimas indicaron una asociación estricta con ambientes áridos, mientras que en otros casos las reconstrucciones indicaron una asociación no estricta con ambientes áridos. En ninguno de los casos se obtuvieron reconstrucciones con una asociación estricta a ambientes húmedos.



5.2.3. Evaluación de los resultados en función de las críticas hechas en el Inciso 5.1

En la primera parte de este capítulo se ha evaluado el marco teórico en el cual debería circunscribirse un análisis de optimización de condiciones ambientales. A partir de ese desarrollo se derivaron algunas consideraciones metodológicas que deberían ser atendidas a la hora de realizar dichos análisis. Una de estas consideraciones estuvo relacionada con la elección de las variables. Una conclusión general del estudio fue que carecía de sentido una evaluación de un gran número de variables ambientales siendo preferible centrarse en algunas pocas variables que sean las mejores descriptoras de las características del ambiente que se quiera analizar. Esto permite hacer un estudio mucho más profundo sobre cada variable y considerar cuestiones tales como la respuesta de las especies a cada variable ambiental analizada, posibles interacciones entre variables, cambios a lo largo del tiempo en los valores y distribuciones geográficas de las variables ambientales, etc. A continuación se discuten estas cuestiones para el análisis realizado:

5.2.3.1. Variable analizada e interacción con otras variables.

Como fue previamente indicado, la variable considerada (índice de humedad) permite una muy buena caracterización del ambiente en función de la disponibilidad de agua para estas plantas, aun cuando existen otros factores microambientales que afectan la disponibilidad de agua (profundidad de la capa freática, pendiente, tipo de suelo, etc.). La necesaria influencia de la disponibilidad de agua en la supervivencia de cualquier vegetal permite cumplir con uno de los requisitos necesarios para optimizar una variable ambiental (ver inciso 5.1.5.). Con respecto a la interacción de las variables analizadas, hemos considerado directamente una variable que es la combinación de la precipitación y la evapotranspiración. Esto permite que se pueda caracterizar directamente el ambiente sin necesidad de evaluar si todas las combinaciones de

precipitación y evapotranspiración se encuentran en las condiciones ambientales realizadas (CA_R) o el nicho realizado (N_R) (ver inciso 5.1.5.).

5.2.3.2. *Condiciones Ambientales Ancestrales.*

Una de las conclusiones principales de la primera parte del capítulo fue que la optimización de los valores de variables ambientales donde viven las especies actuales (optimización del N_R) debe ser considerada tan sólo una etapa en la evaluación de las condiciones del hábitat ancestral y que de no cumplirse algunas hipótesis auxiliares, poco se puede decir de los resultados obtenidos.

Una cuestión importante es considerar si las condiciones ambientales realizadas (CA_R) del pasado serían las mismas que en el presente (ver inciso 5.1.3.). De no ser así es posible que la inferencia del N_R hecha para cada nodo estuviera errada. La existencia de ambientes húmedos durante el cenozoico es indudable, siendo más difícil de evaluar en qué momento se habrían desarrollado ambientes con condiciones áridas. La evidencia paleoambiental apoya sin embargo la existencia de este tipo de ambientes en América desde hace al menos 28 MA. Aun cuando el desarrollo a gran escala de estos ambientes habría comenzado en el Mioceno tardío, habría habido igualmente condiciones áridas de manera local tanto en Norteamérica (Axelrod, 1979a) como en Sudamérica. (Jordan et al., 2001; Volkheimer, 1971) durante la mayor parte del Terciario. Es así que las CA_R en el marco temporal de la diversificación de *Prosopis* serían aproximadamente las mismas que las actuales, difiriendo sólo en el grado de distribución relativa de cada una de ellas (con una clara tendencia al aumento de regiones áridas). Un ambiente particular que se habría desarrollado mucho después en América son los ambientes hiperáridos. Por ejemplo el desierto de Atacama habría surgido recién en el Plioceno Tardío, es decir hace 3 MA. Algunas de las especies actuales de *Prosopis* habitan tales ambientes (ej. *P. tamarugo*, *P. strombulifera*, *P.*

argentina) y debido a esto algunas de las reconstrucciones indicaron que el N_R ancestral incluiría tales ambientes. Estos valores deben ser excluidos del N_R ya que las CA_R no incluía tales ambientes (Ver Figura 28)

5.2.3.3. Nicho Realizado vs. Nicho Potencial

En el inciso 5.1.4. se indicó que cuando el N_R difiere del N_P , la estimación de condiciones ancestrales podría dar resultados erróneos. También se indicó que existirían diferentes cuestiones asociadas al grupo de estudio que pueden producir una diferencia significativa entre el N_P y el N_R : alta extinción local y falta de dispersión. Las especies de *Prosopis* presentan en general distribuciones geográficas muy grandes, muchas veces superando los miles de Km^2 . Esto sugiere que la dispersión en estas especies no sería limitada. Por otra parte el comportamiento invasivo documentado en muchas de las especies del género indicaría que estas especies con gran capacidad de dispersión (Shiferaw et al. 2004). Tampoco es esperable que la extinción local sea muy común debido a los grandes tamaños poblacionales que presentan estas especies. Esta cuestión podría ser un factor relevante para algunas especies tales como *P. castellanosii* cuya distribución si es mucho más acotada y discontinua ya que esta asociada a la distribución de las condiciones edáficas propicias para su desarrollo (suelos arenosos).

Lo anteriormente indicado sugeriría que el N_P y el N_R en las especies de *Prosopis* serían bastante similares. Una evidencia adicional proviene comparar la distribución de las especies de *Prosopis* con la distribución de los ambientes. Una de las primeras conclusiones que se pueden sacar es que los ambientes húmedos ó subhúmedos son contiguos a los ambientes semiáridos en diferentes partes de la distribución de *Prosopis*. Ejemplos de esto puede verse en el Oeste de la región chaqueña donde habitan muchas especies de *Prosopis* pero que se encuentran totalmente ausentes en una región mucho más húmeda como es la región de las Yungas.

También pudo servir como ejemplo la presencia de *Prosopis* en el Noreste Argentino pero ausente en la selva Misionera. Parece ser entonces que la ausencia de especies de *Prosopis* en ambientes húmedos se debería a que no son capaces de sobrevivir en ese lugar y no porque tuvieran una imposibilidad de dispersarse a tales ambientes. Esto indicaría que el N_R y el N_P podrían ser muy similares al menos para la variable analizada.

La evaluación hecha en este apartado (5.2.3.) sobre los posibles errores en los valores obtenidos en el análisis de optimización en *Prosopis* sugiere que los estados estimados en la optimización podrían ser un buen reflejo de los estados que realmente presentaban los ancestros. La única excepción es la asignación de ambientes hiperáridos a los ACMR de los clados *Strombocarpa*, *Strombocarpa + Xerocladia* y de las EAP.

5.3. Bibliografía

- Axelrod DI. 1979a. Age and origin of Sonoran desert vegetation. *Occasional Papers of the California Academy of Sciences* 132: 1–74.
- Ackerly, D. & M. Donoghue. 1998. Leaf size, sapling allometry, and Corner's rules: Phylogeny and correlated evolution in maples (*Acer*). *American Naturalist* 152: 767-791.
- Ackerly, D. 2004. Adaptation, Niche Conservatism, and Convergence: Comparative Studies of Leaf Evolution in the California Chaparral. *American Naturalist* 163: 654-671.
- Burkart, A. 1976. A monograph of the genus *Prosopis* (Leguminosae subFam Mimosoideae). *Journal of the Arnold Arboretum* 57: 219-249.
- Farris, S. 1970. Methods for computing Wagner trees. *Systematic Zoology* 34: 21-24.
- FAO. 2005. New LocClim 1.06 Local Climate Estimator. Agrometeorology Group. FAO/SDRN, Rome, Italy.
- Goloboff, P. A., J. S. Farris & K. Nixon. 2003. TNT: Tree Analysis Using New Technology. Program and documentation available at <http://www.zmuc.dk/public/phylogeny>.
- Graham, C. H., S. R. Ron, J. C. Santos, C. J. Schneider & C. Moritz. 2004. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in Dendrobatid frogs. *Evolution* 58: 1781-1793.
- Hardy, C. R. 2006. Reconstructing ancestral ecologies: challenges and possible solutions. *Diversity and Distributions* 12: 7–19.
- Hardy, C. R. & H. P. Linder. 2005. Intraspecific variability and timing in ancestral ecology reconstruction: A test case from the Cape flora. *Systematic Biology* 54: 299-316.
- Hoffmann, M. 2005. Evolution of the realized climatic niche in the genus *Arabidopsis* (brassicaceae). *Evolution* 59: 1425-1436.
- Hutchinson, G. E. 1957. Concluding remarks *In* Cold Spring Harbor Symposia on Quantitative Biology.
- Jablonski, D. 1987. Heritability at the Species Level: Analysis of Geographic Ranges of Cretaceous Mollusks. 238: 360 - 363.
- Jackson, S. T. & J. T. Overpeck. 2000. Responses of plant populations and communities to environmental changes of the Late Quaternary. *Paleobiology* 26: 194-220.

- Jordan, T. E., F. Schlunegger & N. Cardozo. 2001. Unsteady and spatially variable evolution of the Neogene Andean Bermejo foreland basin, Argentina. *Journal South American Earth Sciences* 14: 775-798.
- Linder, H., T. Dlamini, J. Henning & G. Verboom. 2006. The evolutionary history of *Melianthus* (Melianthaceae). *American Journal of Botany* 93: 1052–1064.
- Peterson, A., J. Soberon & V. Sanchez-Cordero. 1999. Conservatism of ecological niches in evolutionary time. *Science* 285: 1265–1267.
- Prinzing, A., W. Durka, S. Klotz & R. Brandl. 2001. The niche of higher plants: evidence for phylogenetic conservatism. *Proceedings of the Royal Society B: Biological Sciences* 268.
- Roig, F. A. 1993. Informe nacional para la selección de germoplasma en especies de *Prosopis* en la República Argentina. IADIZA-CRICYT.
- Ross, J. H. 1975. Flora of Southern Africa. Vol. 16, Part I. Mimosoideae, Fabaceae. Department of Agriculture, Pretoria, South Africa.
- Rzedowski, J. 1988. Análisis de la distribución geográfica del complejo *Prosopis* (Leguminosae, Mimosoideae) en Norteamérica. *Acta Botanica Mexicana* 3: 7-19.
- Savolainen, V., S. B. Heard, Powell M.P, Davies T.J & M. A.Ø. 2002. Is cladogenesis heritable? *Systematic Biology* 51: 835-843.
- Shiferaw, H., D. Teketay, S. Nemomissa & F. Assefa. 2004. Some biological characteristics that foster the invasion of *Prosopis juliflora* (Sw.) DC. at Middle Awash Rift Valley Area, north-eastern Ethiopia. *Journal of Arid Environments* 58: 135–154.
- Stephens, P. & J. Wiens. 2003. Ecological diversification and phylogeny of emydid turtles. *Biological Journal of the Linnean Society* 79: 577–610.
- UNEP. 1992. World Atlas of Desertification, London (UK).
- Vazquez, D. P. 2005. Reconsiderando el nicho Hutchinsoniano. *Ecología Austral* 2005.
- Verneau, O., F. Catzeflis & A. Furano. 1998. Determining and dating recent rodent speciation events by using L1 (LINE-1) retrotransposons. *Proceedings of the National Academy of Sciences* 95: 11284–11289.
- Volkheimer, W. 1971. Aspectos paleoclimáticos del Terciario argentino. *Revista del Museo Argentino de Ciencias Naturales Bernardino Rivadavia* 1: 243-262.
- Yesson, C. & A. Culham. 2006a. Phyloclimatic Modeling: Combining Phylogenetics and Bioclimatic Modeling. *Systematic Biology* 55: 785–802.

- Yesson, C. & A. Culham. 2006b. A phyloclimatic study of *Cyclamen*. *BMC Evolutionary Biology* 6: 72.
- Webb, C., D. Ackerly, M. McPeek & M. Donoghue. 2002. Phylogenies and community ecology. *Annual Review of Ecology and Systematics* 33: 475–505.

CAPÍTULO 6 - DISCUSIÓN Y CONCLUSIONES

6.1. Relaciones filogenéticas

Uno de los resultados más relevantes del estudio de diversificación en el género *Prosopis* obtenidos en esta tesis fue la no-monofilia del grupo. El género no apareció formando un grupo natural en ninguno de los análisis, y las búsquedas con constreñimientos indicaron que los árboles más cortos que presentaban a *Prosopis* como monofilético tenían entre 6 y 11 pasos más que los árboles óptimos, aportando una ostensible evidencia en contra de esta hipótesis. Las especies americanas del género estarían más relacionadas con la especie africana *Xerocladia viridinamis*, única especie de su género, que con los representantes de *Prosopis* del viejo mundo (*P. cineraria* y *P. africana*). Esta relación entre *X. viridinamis* y especies de *Prosopis* ya había sido propuesta por Luckow et al. (2004) en base a secuencias de ADN de cloroplasto (regiones *trnL-trnF*, *trnK-matK* y *ndhF*), pero dicho trabajo sólo incluía especies de la sección Algarobia (*P. nigra*, *P. elata*, *P. rojasiana*, *P. pallida*). Por este motivo, la relación de *X. viridinamis* con las especies americanas de *Prosopis* (EAP) es un resultado novedoso. Aún si la posición filogenética de *X. viridinamis* fuera incorrecta, la monofilia del género *Prosopis* estuvo fuertemente contradicha por la posición de *P. africana* en todos los análisis y por la posición de *P. cineraria* en algunos de ellos. El único caso en que *P. cineraria* aparece como grupo hermano de las EAP es en el análisis de dos marcadores pero con bajo soporte. La no-monofilia de *Prosopis* no había sido adecuadamente puesta a prueba en los trabajos previos. En el trabajo de Luckow et al. (2004) el análisis incluía un grupo importante de géneros afines a *Prosopis* pero dentro de éste sólo se incluyeron unas pocas especies todas pertenecientes a la sección Algarobia. Por su parte, en el trabajo de Bessegå et al. (2006a) los grupos externos estaban lejanamente emparentados con el género. Finalmente, el trabajo de Burghardt & Espert (2007) incluye sólo las especies americanas y un único grupo externo. En esta

tesis el diseño aplicado permitió obtener fuerte evidencia para rechazar la hipótesis de monofilia de *Prosopis*.

Sin descartar la posibilidad de que nuestros resultados puedan ser incorrectos y que nuevos estudios puedan apoyar la monofilia del género, la discrepancia entre nuestros resultados filogenéticos y la delimitación genérica de Burkart (1976) proviene posiblemente de que la taxonomía realizada por este autor no fue confeccionada siguiendo los principios de la sistemática filogenética. La delimitación de grupos en la sistemática moderna esta basada en agrupamientos definidos por caracteres derivados compartidos, es decir por sinapomorfías, en desmedro de aquellos caracteres que, aunque compartidos, sean primitivos (simplesiomorfías). Los caracteres que según Burkart (1976) definen al género son: flores actinomorfas, pentámeras decandras de tipo hermafrodita con anteras con extremo glandular, hojas parabipinnadas, y fruto de tipo “lomento drupáceo”. Todos estos caracteres (con excepción del tipo de fruto) según Burkart (1976) serían primitivos dentro de toda la subfamilia, es decir serían simplesiomorfías para *Prosopis* y por lo tanto no serían validos para la definición del grupo. El carácter restante, tipo de fruto, podría ser considerado la única sinapomorfía del género. Sin embargo, este rasgo aporta una evidencia morfológica pobre en apoyo a la monofilia de *Prosopis* puesto que no puede descartarse que el fruto de *Xerocladia* sea un estado derivado dentro del género (una autapomorfía de *X. viridinamis*) como parecen apoyar algunas evidencias (ver más abajo).

X. viridiramis Burch (Taub.) es la única especie de su género, cuya distribución se halla restringida a los desiertos de Namibia (África). Es un arbusto enano cuyos tallos fotosintéticos presentan carenas y valéculas al igual que las especies del género *Prosopidastrum*. El fruto de *X. viridinamis* es algo único dentro de las leguminosas: está formado por un único artejo chato, membranoso y alado (Lewis & Elias en Polhill

& Raven, 1981). Palacios & Hoc (2005) en un trabajo sobre *Prosopidastrum* señalan una cuestión muy interesante sobre el fruto de *Xerocladia*. Estos autores indican que en Spegazzini (1896) describe a *Xerocladia pampeana*, basándose en un ejemplar que constaba de ramas con flores y frutos que tenían agallas y que eran monospermios. Este ejemplar era muy similar a *Prosopidastrum striata*, excepto por el fruto monospermio que sería la causa por la cual Spegazzini lo incluyó en *Xerocladia*. Debido a las agallas que presentaba, consideramos que es muy posible que el ejemplar descrito por Spegazzini haya sido un ejemplar teratomorfo perteneciente a *Prosopidastrum striata*. Este caso excepcional demostraría que cambios en la regulación génica (similar a los producidos durante la formación de las agallas) pueden modificar de manera muy rápida y rotunda la morfología del fruto en mimosoideas, haciendo factible que un cambio similar pudiera haber ocurrido en la historia del linaje que dio origen a *Xerocladia*. Este caso podría también ser considerado como una evidencia en favor de una relación cercana entre *Prosopidastrum* y *Xerocladia*. Nuestros datos no parecen apoyar tal hipótesis: en ningún análisis estos géneros formaron un grupo monofilético y en los casos de los análisis de 5 y de 3 marcadores, los datos contradicen fuertemente este agrupamiento (Capítulo II, Tabla 10). Sin embargo, si se considera al grupo formado por *EAP* + *Prosopidastrum* + *Xerocladia* el mismo es solo un paso más largo en los análisis de 2 y 3 marcadores y 3 pasos en el caso de análisis de 5 marcadores. Estos resultados no descartan entonces una relación cercana entre *Xerocladia*, *Prosopidastrum* y las EAP.

Una relación más próxima entre especies de las secciones americanas de *Prosopis* (Strombocarpa, Algarobia y Monilicarpa) que con las secciones del viejo mundo (Anonychium y *Prosopis*) está de acuerdo con la subdivisión del género propuesta por Guinet & Bessedik (1984). La división subgenérica se basó

principalmente en caracteres polínicos: el subgénero *Neoprosopis* incluye a todas las especies americanas del género cuyos granos de polen se caracterizan por las siguientes características (Guinet & Bessedik 1984): ausencia de costae; endexina progresivamente engrosada alrededor de los poros; exina lisa a escabrosa o rugosa a finamente rugosa; tectum atravesado por muy numerosas perforaciones de diámetro variable en general repartidos regularmente dando a la superficie tectal una aspecto reticulado. Los otros dos subgéneros coinciden con las dos secciones del viejo mundo: el Sub-Género *Anonychium* coincide con la sección *Anonichium* Bentham y su polen se caracteriza por caracteriza por: presencia de costae, exina lisa; tectum atravesado por raras perforaciones repartidas irregularmente y de pequeño diámetro. El Subgénero *Prosopis* coincide con la sección *Prosopis* (Burkart 1976) y a nivel polínico se caracterizaría por: ausencia de costae, endexina progresivamente engrosada alrededor de los poros; exina areolada verrugosa; numerosas perforaciones tectales con un diámetro de entre 0.8 y 1.2 μm .

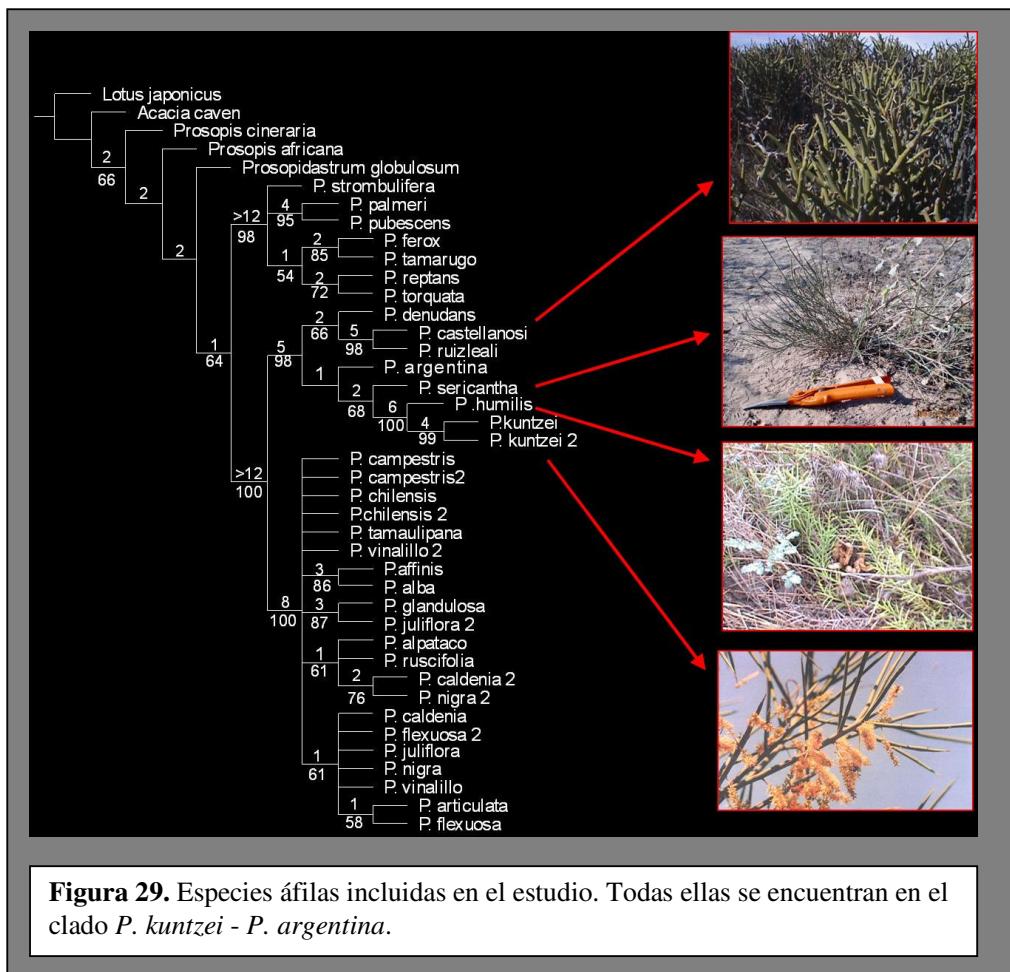
Esta relación cercana entre las secciones americanas de *Prosopis* indicada por Guinet & Bessedik (1984) contrasta con la idea de Burkart (1976) de que la sección *Strombocarpa* estaría cercanamente emparentada con la sección *Prosopis* (ya que ambas presentan legumbres comparativamente más pequeñas que el resto de las especies del género y comparten la capacidad de dispersarse por rizomas). En ninguno de nuestros análisis las especies de la sección *Strombocarpa* aparecieron más relacionadas con *P. cineraria* que con el resto de las especies americanas. Además, las búsquedas realizadas forzando la monofilia del grupo formado por *P. cineraria* más la sección *Strombocarpa* mostraron que los árboles eran hasta 63 pasos más largos que los árboles óptimos (Tabla 10, Capítulo 2), indicando que, al menos para el conjunto de datos analizado, esta hipótesis esta fuertemente contradicha.

La sección *Strombocarpa* fue monofilética en nuestro análisis así como en Ramírez et al. (1999) y Burghardt & Espert (2007). En contraste, la monofilia de esta sección no fue corroborada en el trabajo de Bessegá et al. (2006) en el cual las especies de *Strombocarpa* (junto a *P. cineraria*) aparecieron formando un grado basal al resto de las especies de *Prosopis* incluidas en dicho análisis. Desde el punto de vista morfológico esta sección presentaría dos sinapomorfías: presencia de una corola gamopétala y presencia de espinas foliares (Burghardt & Espert, 2007). Las relaciones entre especies dentro de la sección *Strombocarpa* no quedaron claramente establecidas en nuestros análisis. Burkart (1976) dividió la sección en dos series: *Cavenicarpae* que comprende a *P. ferox* y *P. tamarugo* y *Strombocarpae* que incluye a las 7 especies restantes de la sección. Este autor basó la separación entre estas series en un sólo carácter morfológico: la posición de los segmentos del endocarpo. En las especies de *Strombocarpae* estos segmentos son longitudinales mientras que en las especies de *Cavenicarpae* están en posición transversal. Como estos dos estados son los únicos encontradas en el género, es imposible que este carácter pueda simultáneamente delimitar como grupos monofiléticos a ambas series: si la posición longitudinal es una sinapomorfía de la serie *Cavenicarpae*, entonces la posición transversal tiene que ser un carácter primitivo, y consecuentemente no pude delimitar grupo alguno. Contrariamente si la posición transversal es una sinapomorfía de la serie *Strombocarpae*, entonces la posición longitudinal es un carácter primitivo. Como consecuencia, no existe evidencia morfológica para sustentar que ambas series sean monofiléticas. En nuestro estudio sólo dos grupos aparecieron fuertemente apoyados por la evidencia. Uno de ellos es el clado formado por *P. ferox* y *P. tamarugo*, resultado congruentes con la evidencia morfológica indicando que la sección *Cavenicarpae* es monofilética y que el carácter derivado en esta sección es la posición de trasversal.

El otro clado dentro de la sección *Strombocarpae* que fue recobrado en todos los análisis es el formado por *P. palmeri* y *P. pubescens*. Estas dos especies son norteamericanas, la primera de ellas endémica de la región de Baja California, México mientras que *P. pubescens* está presente en el sur de California, norte de Chihuahua y suroeste de Texas. La posición filogenética de *P. palmeri*, asociada al resto de las especies de la sección *Strombocarpa*, merece una consideración especial puesto que su afiliación taxonómica ha sido un problema desde su descripción original a finales del siglo XIX. Watson (1889) le dio el nombre que actualmente presenta y consideró que pertenecería a la sección *Algarobia sensu* De Candolle (Simpson et al., 1975). A finales de la década del 20 del siglo pasado, Britton & Rose (1928) consideraron que esta especie debería ser incluida en un nuevo género monotípico (*Soprosopis*), basándose en la estructura del fruto. Más tarde (Burkart, 1940), asignó provisionalmente esta especie a la sección *Anonychium*, junto con *P. africana*. Un estudio más detallado realizado por Simpson et al. (1975) asignó esta especie a la sección *Strombocarpa*. Estos últimos autores consideran que, excepto por la falta de enrollamiento del fruto, la mayoría de los caracteres morfológicos (florales, palinológicos, estipulares y flavonoides) indicarían que esta especie pertenecería al género *Prosopis* y no a un género diferente. Por otra parte, la presencia de estípulas pareadas indicarían que *P. palmeri* pertenecería a la sección *Strombocarpa*. En el trabajo de Bessegå et al. (2006), esta especie aparece dentro del grado basal formado por especies de la sección *Strombocarpa* + *P. cineraria*. En el trabajo de Burghardt & Espert (2007), *P. palmeri* aparece formando un grupo monofilético con el resto de las especies de *Strombocarpa* en dos de las tres condiciones analíticas consideradas. Parece que luego años de incertidumbre sobre su posición taxonómica, la ubicación de esta especie junto al resto de las especies de la sección *Strombocarpa* estaría corroborada.

Las especies de la sección Algarobia junto con *P. argentina* formaron un clado altamente apoyado (llamado aquí Algarobia s.l.). El mismo presentó dos grupos principales también fuertemente apoyados por la evidencia; uno de ellos correspondió a las especies de las series Pallidae, Chilenses y Ruscifoliae (nombrado en esta tesis “grupo de los algarrobos”) mientras que el otro correspondió a las especies de las series Sericanthae, Humiles y Denudantes más *P. argentina* de la sección Monilicarpa (nombrado como grupo *P. kuntzei* – *P. argentina*). Esta separación en dos grupos de la sección Algarobia fue uno de los resultados más importantes de la tesis y coincide con los resultados obtenido bajo una de las tres condiciones analíticas ensayadas por Burghardt & Espert (2007). Mientras que el grupo de los algarrobos apareció en todos los análisis del trabajo de Burghardt & Espert (2007), el grupo *P. kuntzei* - *P. argentina* apareció solamente en uno de ellos. El grupo de los algarrobos no fue monofilético en el trabajo de Bessega et al. (2006) ya que en éste *P. argentina* aparece anidada dentro de este grupo. Burghardt & Espert (2007) basándose en sus resultados, propusieron reconsiderar el estatus taxonómico del grupo de los Algarrobos y llevarlo al nivel de sección quizás con el nombre de Algarobia *sensu stricto*, una idea con la que nosotros acordamos.

El clado *P. kuntzei* - *P. argentina* incluyó varias especies (*P. castellanosii*, *P. humilis*, *P. denudans* y *P. ruizleali*) que no habían sido estudiadas previamente a nivel genético. Una característica interesante de este clado es que aquí se encuentran todas las especies áfilas o subáfilas incluidas en este análisis (Figura 29).



Las especies de la serie Denudantes aparecieron formando un grupo monofilético en nuestro estudio al igual que bajo algunas condiciones analíticas en Burghardt & Espert (2007). Según estas autoras esta serie debería ser considerada una nueva sección ya que presentaría diferencias con el resto de las especies de la sección Algarobia que justificarían este nuevo estatus taxonómico. Nuestros resultados son compatibles con tal cambio taxonómico. Según Burkart (1976) *P. ruizleali* y *P. denudans* podrían ser una sola especie, correspondiéndoles un estatus de subespecie a ambas debido a la existencia de una variación morfológica casi continua entre las mismas. Burghardt & Espert (2007) encontraron que estas dos especies formaron un grupo monofilético apoyando una estrecha relación filogenética. Sin embargo, en

nuestro estudio *P. ruizleali* apareció formando un grupo monofilético con *P. castellanosii*, mientras que *P. denudans* apareció como grupo hermano del anterior. Este es uno de los pocos grupos en que el que nuestro análisis difiere de un resultado que fue apoyado por los trabajos de Burkart (1976) y Burghardt & Espert (2007). A pesar de que son muchas las posibles causas de la discrepancia, una de las explicaciones posibles es que haya existido un polimorfismo genético compartido entre las especies y que debido a ello la filogenia de genes no refleje la relación histórica entre estos taxones (Ver inciso 6.2.). Una explicación similar podría esgrimirse para explicar por qué la serie Sericanthae no fue monofilética. En este caso *P. kuntzei* apareció formado un clado con *P. humilis* de la serie Humiles, mientras que la otra especie de la serie Sericanthae (*P. sericantha*) se ubicó como grupo hermano de estas dos especies.

La comparación entre los resultados obtenidos en la presente tesis y los obtenidos en estudios filogenéticos previos indica que nuestros resultados serían más coincidentes con la filogenia basada en caracteres morfológicos y bioquímicos de Burghardt & Espert (2007) que con aquella basada en caracteres de secuencias (Bessegá et al., 2006). Consideramos que existen diferentes líneas de evidencia para suponer que el análisis presentado en esa tesis podría reflejar más certeramente las relaciones históricas entre especies del género *Prosopis* que el trabajo de molecular de Bessegá et al. (2006). (1) El muestreo del grupo interno fue mayor y más representativo de la diversidad taxonómica del género (2) La filogenia obtenida se basó en el análisis de cinco regiones genómicas diferentes mientras que en el trabajo de Bessegá et al. (2006) sólo se utilizó solo ITS, una secuencia que ha sido ampliamente criticada para su uso como marcador filogenético fidedigno (Alvarez & Wendel, 2003) (3) Los resultados obtenidos en nuestro trabajo coinciden en mayor medida con la taxonomía tradicional (Burkart, 1976; Guinet & Bessedik, 1984), análisis fenéticos de frecuencias alélicas

(Ramírez et al., 1999; Bessegå et al., 2005) y análisis cladísticos caracteres morfológicos y bioquímicos (Burghardt & Espert, 2007).

Como punto final a destacar sobre cuestiones filogenéticas en *Prosopis* considero que la solución de las cuestiones no irresueltas requerirán de un enfoque que abarque no sólo caracteres moleculares sino también caracteres morfológicos. Aunque el trabajo de Burghardt & Espert (2007) ha hecho un avance considerable en el estudio de morfología en las especies Americanas *Prosopis*, es imprescindible no sólo su extensión a las especies de *Prosopis* del Viejo Mundo sino también a especies de géneros afines. Estudios descriptivos morfológicos tales como los publicados sobre anatomía foliar (Caccavari & Dome, 2000; Luckow, 2002; Flores-Cruz et al., 2004) o palinología (Cacaccavari & Dome, 2000) son una etapa necesaria para lograr este objetivo.

6.2. Genes utilizados, ventajas y desventajas.

La mayoría de los estudios de filogenia moleculares en plantas han utilizado marcadores de ADN de cloroplasto y/o secuencias de ITS. Las primeras han mostrado ser muy útiles en la resolución de las relaciones filogenéticas supragenéricas pero en general carecen de un nivel de variación suficiente como para resolver las relaciones entre especies cercanas. Por otra parte, las secuencias de ITS son altamente variables pero presentan serias desventajas entre las que se encuentran la existencia de genes parálogos y la falta de una evolución concertada completa. En los últimos años comenzaron a publicarse filogenias moleculares en plantas basadas en datos de genes nucleares (Small & Wendel, 2000; Tank & Sang, 2001; Simmons et al., 2001; Cronn et al., 2002; Malcomber, 2002; Small et al., 2004), en particular regiones no codificantes de los mismos, que suelen presentar niveles de variación adecuados para la resolución de las relaciones filogenéticas a niveles infragenéricos. En el presente estudio

utilizamos dos genes nucleares (*nia* y *g3pdh*) para resolver las relaciones filogenéticas entre las especies de *Prosopis*. La concordancia general de nuestros resultados con la mayoría de los estudios previos indicaría que los marcadores nucleares utilizados en este análisis serían una fuente muy útil para resolver las relaciones filogenéticas entre especies cercanamente emparentadas. Sin embargo su uso también presentó algunas desventajas. Una de las mismas fue la dificultad encontrada para amplificar estos genes en algunas de las especies. Debido a la baja eficiencia de las reacciones de amplificación, fue necesario realizar múltiples reacciones en paralelo para lograr la cantidad de producto amplificado necesario para su posterior secuenciación. Otra desventaja propia de la utilización de los genes nucleares es la existencia de polimorfismos compartidos entre especies. Es sabido que para poder inferir las relaciones filogenéticas entre especies a partir de genes es necesario que dichos genes tengan la misma historia que las especies que los portan (Nichols, 2001). Debido a que están presentes en doble copia y que se heredan por vía materna y paterna, el tamaño poblacional efectivo de los genes nucleares es 4 veces mayor que el de los genes de cloroplaso. Esto hace que la probabilidad de presentar polimorfismos compartidos y/o de sufrir “lineage sorting” sea mucho mayor, produciéndose consecuentemente un desacople entre las relaciones establecidas a partir del análisis de una región genómica con las relaciones históricas entre las especies. Es esta una de las posibles causas por las cuales no fue posible resolver las relaciones filogenéticas entre el grupo de especies de los algarrobos y por la cual la serie Sericanthae no fue monofilética. La probada existencia de polimorfismos isoenzimáticos compartidos entre especies de la sección Algarobia (Saidman 1986; Bessega et al. 2000a,b; Ferreyra et al. 2007) es una evidencia que sugiere que lo mismo podría estar ocurriendo para los genes secuenciados.

Small et al. (2004) han criticado la utilización de genes de un número bajo de copias en estudios filogenéticos en plantas por la posible existencia de secuencias parálogas. Diferentes líneas de evidencias indicarían que no hemos tenido ese problema en nuestro estudio. Por un lado, las relaciones del análisis filogenético teniendo en cuenta los genes nucleares estuvieron, en línea generales, de acuerdo tanto con la taxonomía tradicional y con otros estudios previos. Por otra parte, cuando se clonaron varias copias de estos genes para el mismo individuo, en ningún caso obtuvimos más de 2 variantes diferentes de cada individuo. Esto sugiere que estas dos variantes se deberían a heterocigosis y no a paralogía.

La inclusión de genes de cloroplasto fue muy útil, principalmente para los análisis de dos y de cinco marcadores aunque presentó también desventajas. Una de ellas es que los mismos no presentaron variación entre especies cercanamente emparentadas. Por ejemplo todas las especies del grupo de los algarrobos presentaban secuencias idénticas para el gen *psbC-trnS*. Algo similar se encontró para las regiones *trnK-matK* y *trnL-trnF*. Al igual que estudios previos (ej. Kelcher, 2000; Graham et al., 2000), en los genes de cloroplasto analizados encontramos una alta frecuencia de rearreglos genómicos en relación con el número de sustituciones nucleotídicas observadas. En el caso de la región *trnL-trnF* se encontró una delección de 300 pb. en *P. cineraria* mientras que para la región *trnS-psbC* encontramos una inversión de 10 pares de bases en varias de las especies de este género (Capítulo 3). Una característica común a ambos rearreglos fue su ocurrencia independiente en diferentes linajes. La misma delección fue también observada en géneros de otras tribus de mimosoideas (ej. *Samanea*). La inversión por su parte, como se indicó en el Capítulo 3 de esta tesis, habría ocurrido al menos tres veces en la historia de *Prosopis* y 10 veces en Angiospermas.

Existen dos cuestiones importantes a ser consideradas a la hora de establecer relaciones filogenéticas a partir de regiones que presentan rearreglos genómicos. La primera es la necesidad de reconocer la ocurrencia de tales eventos. No reconocerlos puede llevar a errores: si una inversión involucra 200 pares de bases y no es detectada, puede ser considerada como 200 sustituciones que independientemente apoyan un determinado agrupamiento cuando en realidad es un solo evento. Una segunda cuestión está relacionada la posibilidad de incluir la información filogenética de tales eventos en los análisis. Métodos como los propuestos en el Capítulo III para analizar inversiones pueden perfectamente ser usados en la inferencia de filogenias. Varios autores han indicado que las pequeñas inversiones no serían útiles para establecer relaciones filogenéticas debido a que las mismas serían muy homoplásicas (Kelchner & Wendel, 1996; Kelchner, 2000). Esto contrastaría con lo que ocurre con inversiones que involucran un número mucho mayor de nucleótidos que aparecerían como sinapomorfías de grupos superiores (Jansen & Palmer, 1987; Raubeson & Jansen, 1992). Consideramos que la exclusión de las pequeñas inversiones de los análisis filogenéticos no tiene sustento alguno ya que la existencia de homoplasia en un carácter no implica que el mismo carezca de información filogenética. De hecho en la inversión estudiada, uno de los sentidos de la misma puede ser considerado una sinapomorfía de la serie Denudantes (Figura 15, Capítulo 3)

6.3. Diversificación, tiempos de divergencia y restos fósiles

Ciertas cuestiones sobre la estimación de tiempos de divergencia a partir de datos moleculares hacen que los tiempos estimados deban ser considerados con cierta cautela. Entre estas se encuentran la incertidumbre del punto usado como calibración, tanto con respecto a la determinación taxonómica del mismo como también respecto a su edad. Otro punto es el método de cálculo propiamente dicho. Aunque el método de Penalizad

likelihood no requiere de tasas constantes a lo largo del árbol, igualmente considera una función para el cambio entre ramas de las tasas de sustitución que bien puede no ajustarse a los cambios que realmente ocurrieron en la historia del grupo. Un tercer punto es que los tiempos se calculan teniendo en cuenta una pequeñísima fracción de todo el genoma siendo posible la existencia de errores de estimación asociados a este pequeño tamaño de muestreo.

A pesar de las limitaciones previamente indicadas realizamos este análisis con el objeto de aproximar un marco temporal para la divergencia del grupo ya que el único dato publicado sobre este tema indicaba que *Prosopis* se habría originado hace 120 millones de años (Ramírez et al., 1999), una edad que se contradice de manera total con toda la evidencia existente sobre evolución de la familia de las leguminosas, cuya edad no sería mayor a los 60 MA (Lavin et al., 2005).

El análisis de datación molecular sugirió que la divergencia entre los dos grupos principales dentro del clado de las *EAP + Xerocladia* es notablemente antiguo (26.2 - 33 MA) en comparación con el comienzo de la diversificación de los miembros actuales de toda la subfamilia Mimosoideae que, según Lavin et al. (2005), habría ocurrido hace 40 MA. Sin embargo la mayor parte del proceso de diversificación dentro de cada uno de los principales clados de *Prosopis* habría ocurrido más recientemente, en el Mioceno Tardío. Independientemente del marco temporal absoluto, este contraste entre una divergencia temprana entre los principales linajes americanos y una divergencia más tardía dentro de cada uno de estos grupos ya había sido postulada en base a evidencia morfológica (Burkart & Simpson, 1977; Pasiecznik et al., 2001).

El número de fósiles asignados al género es alto (Figura 30). Sin embargo debido a que ninguno de ellos a sido incorporado a un análisis filogenético, la asignación de los mismos al género y /o a sus subgrupos es tentativa.

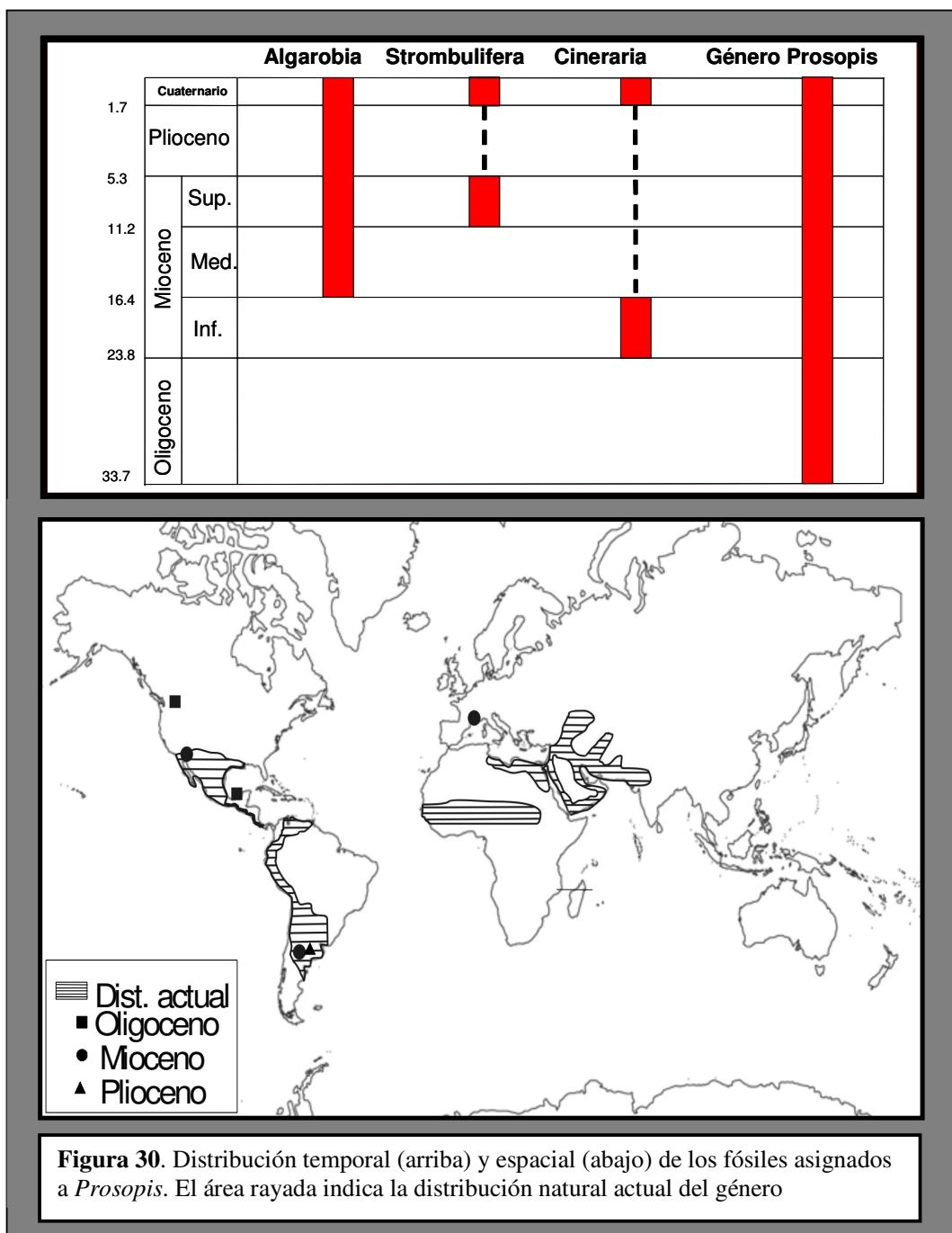


Figura 30. Distribución temporal (arriba) y espacial (abajo) de los fósiles asignados a *Prosopis*. El área rayada indica la distribución natural actual del género

El fósil más antiguo que puede ser atribuido al grupo de las EAP y que fue usado como uno de los puntos de calibración en el análisis de datación molecular, pertenece a granos de polen (*P. quesneli*) del Oligoceno temprano de la Columbia Británica, Canadá (Piel, 1971). Una edad similar fue estimada para un resto de un fruto (*P. lazarii*) encontrado en la paleoflora de Puebla, México (Magallon-Puebla &

Ceballos-Ferriz, 1994). Sin embargo la afinidad de este fósil con los grupos actuales dentro del género es incierta. Recientemente (Anzótegui & Herbst, 2004) describieron, en base a restos foliares, una nueva especie de *Prosopis*, para el Mioceno Medio de Argentina que estaría relacionado con miembros actuales de la sección *Strombocarpa*. La edad de este fósil es compatible con los tiempos estimados en esta tesis, ya que ambos indicarían una divergencia entre los grupos principales dentro de *Prosopis* antes del Mioceno Medio. Una comparación más detallada de los tiempos de divergencia con los tiempos derivados de los restos fósiles requiere de especies actuales y fósiles en un mismo análisis filogenético.

6.4. Diversificación del grupo de los algarrobos

Los resultados obtenidos en este estudio apoyan fuertemente la monofilia del clado formado por especies de las series Chilenses, Pallidae y Ruscifoliae (grupo de los algarrobos). Es interesante considerar que muchas de estas especies fueron en un momento consideradas parte de un Singameon (Palacios & Bravo, 1981) debido a su capacidad de hibridar naturalmente, su alta similitud morfológica y las distribuciones geográficas parcialmente solapadas. Sin embargo en las ultimas dos décadas, los estudios moleculares realizados en este grupo han indicado que a pesar de la alta similitud genética, las poblaciones se agrupaban coincidentemente con su especie (Palacios & Bravo, 1981; Saidman & Vilardi, 1987; Burghardt, 1995; Bessega et al., 2000a,b) aun en el caso de poblaciones de diferentes especies que habitaban en simpatría. Además (Bessega et al., 2000a) mostró que las especies simpátricas no son genéticamente más similares entre sí que con respecto a especies alopátricas. Un dato adicional está dado por el análisis de híbridos: el único trabajo publicado hasta el momento sobre viabilidad de híbridos (Naranjo et al. 1984) sugiere que al menos para algunas combinaciones de especies, los híbridos formados serían parcialmente o

totalmente estériles. Toda esta evidencia apoya la idea de que la formación de híbridos no produciría (o al menos no ha producido hasta el momento) un flujo génico significativo entre estas especies. Los tiempos derivados de dataciones moleculares sugieren que la alta similitud genética entre especies puede ser explicada por su origen reciente (< 3.5 MA) sin necesidad de tener que invocar la existencia de introgresión entre ellos.

La tasa de diversificación estimada para el grupo de los algarrobos es comparable a los valores más altos obtenidos para radiaciones en plantas. Sólo el clado Andino del género *Lupinus* (Hughes & Eastwood, 2006) y la subfamilia *Ruschoidea* (Klak et al., 2003) tienen una tasa mayor a aquélla calculada para los algarrobos.

En el presente estudio no fue posible establecer las relaciones filogenéticas dentro del grupo de los algarrobos debido a que los individuos de la misma especie no formaban un grupo monofilético y además el consenso estuvo altamente politomizado. Esta dificultad de establecer las relaciones entre especies de algarrobos ya había sido observado en el trabajo de Burghardt & Espert (2007), excepto para el clado formado por las especies *P. alba*, *P. nigra*, *P. ruscifolia*, *P. hasleri* que fue monofilético en todos los análisis presentados en ese trabajo. Una posible explicación a este resultado está asociada a la rápida y reciente radiación de los algarrobos indicada por los análisis realizados en la presente tesis. Cuando los intervalos entre eventos de especiación son cortos, existe una mayor probabilidad que los polimorfismos ancestrales sean retenidos por los diferentes linajes divergentes (Pamilo & Nei, 1988). La presencia de este polimorfismo ancestral compartido además de la existencia posterior de “sorting” afecta derivación de relaciones filogenéticas entre especies a partir de los árboles de genes (Nichols, 2001). Una explicación similar ha sido invocada para la falta de resolución en

diversos géneros de plantas (Hodges & Arnold, 1995; Baldwin & Sanderson, 1998; Richardson et al., 2001; Hughes & Eastwood, 2006).

6.5. Hibridación y Evolución de mecanismos de aislamiento en Algarobia

Los resultados obtenidos en el análisis filogenético así como en la datación molecular nos dan una oportunidad de comprender el modo en que han evolucionado los mecanismos de aislamiento reproductivo en *Prosopis*. Una de las cuestiones más interesantes a analizar esta relacionada con el origen de los mecanismos de aislamiento en la sección Algarobia. A partir de la filogenia obtenida es posible inferir que la capacidad para hibridar no sería una característica extendida a lo largo de todo el grupo Algarobia s.l. De hecho los dos grupos principales en los cuales se divide este clado, presentan un comportamiento claramente diferenciado (Figura 30). Mientras que en el clado de los Algarrobos hay documentadas al menos 11 combinaciones híbridas (Palacios & Bravo, 1981; Naranjo et al., 1984; Hunziker et al., 1986), no hay cita alguna para la formación de híbridos entre especies del clado *P. kunzzei* – *P. argentina*. Aun cuando las especies de este último clado pueden haber desarrollado mecanismos de aislamiento que evitan la formación de híbridos, es posible que el reducido nivel de simpatría entre la mayoría de estas especies sea una explicación suficiente para la falta de híbridos entre las mismas. Por ejemplo *P. kunzzei* esta presente en el Chaco, *P. humilis* en la región central del país, *P. denudans* en Patagonia y *P. castellanosii* restringido al norte de Neuquén y sur de Mendoza, haciendo de hecho imposible la formación de híbridos. Una de las excepciones serían las especies *P. kunzzei* y *P. sericantha* que si aparecen en simpatría.

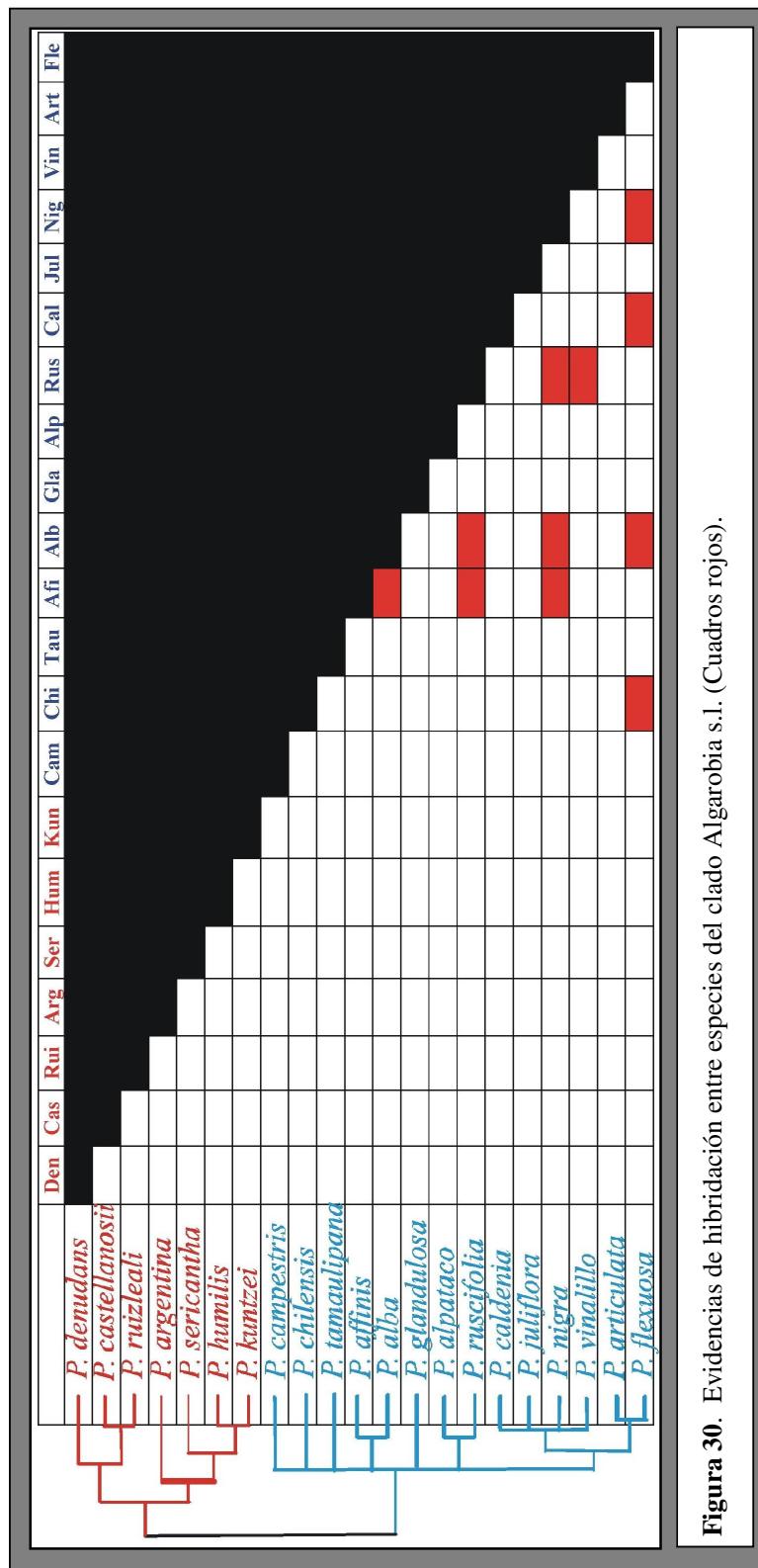


Figura 30. Evidencias de hibridación entre especies del clado Algarobia s.l. (Cuadros rojos).

Tampoco existen evidencias documentadas de la existencia de híbridos entre alguna especie del grupo de los Algarrobos con las del clado *P. kuntzei* - *P. argentina* aun cuando especies de ambos grupos sí aparecen comúnmente en simpatría. Esto sugiere que otra barrera, intrínseca o extrínseca podría estar actuando previniendo la formación de híbridos. Independientemente de cual podría ser tal mecanismo el mismo debería haberse desarrollado, teniendo en cuenta los datos aportados por la datación molecular, en menos de 8 millones de años (MA) que sería el momento en que estos dos clados habrían divergido.

Como ya fue previamente indicado, en el clado de los Algarrobos el flujo génico efectivo entre las especies habría estado restringido aun cuando en algunos casos las especies son simpátricas en alto grado. Esto sugiere la existencia de algún(os) mecanismo(s) de aislamiento que evitaría(n) la introgresión entre especies que hibridan. Palacios & Bravo (1981) discutieron cuales podrían ser los mecanismos de aislamiento que estarían actuando para mantener la identidad de las especies de este grupo y basaron su discusión en la clasificación de mecanismos de aislamiento reproductivo de Dobzhansky (1970) y Dobzhansky et al. (1977). Nosotros utilizaremos la clasificación realizada por Coyne & Orr (2004) que se sustenta en el nuevo conocimiento tanto empírico como teórico obtenido en los últimos 30 años sobre mecanismos de aislamiento y especiación.

6.5.1. Barreras de Aislamientos Precigóticos:

Como fue previamente indicado, la existencia de híbridos naturales entre especies de algarrobos está ampliamente documentada. Este dato es una de las principales evidencia en contra de la eficiencia de las barreras precigóticas entre las especies estudiadas. Sin embargo, también existen evidencias independientes que permiten descartar la posibilidad de que alguna de las barreras precigóticas juegue un

rol importante entre estas especies. Muchas de las especies de *Prosopis* que habitan el Chaco son simpátricas, pudiendo descartarse en este caso un aislamiento ecológico por Hábitat. Los polinizadores tampoco serían una barrera (Palacios & Bravo, 1981; Genisse et al., 1990), ya que la información disponible indica que las especies de *Prosopis* son visitadas por las mismas especies de insectos. Con respecto a la barrera temporal, Palacios & Bravo (1981) indicaron que al menos para las especies de la región chaqueña que ellos estudiaron, la misma no sería importante ya que todas las especies estudiadas tuvieron una floración simultánea. Tampoco existen evidencias de existen evidencias de barreras mecánicas. Sin embargo, anteriormente se pensaba que *Prosopis* presentaba protoginia, algo que ayudaría a la fecundación cruzada y por ende la hibridación (Palacios & Bravo, 1981). Sin embargo esto ha sido descartado para al menos 4 de las especies de *Prosopis* (Hoc et al., 1994).

6.5.2. Barreras de Aislamientos Postcigóticos:

Existen claras evidencias de que la principal barrera entre la mayoría de estas especies sería la Inviabilidad Ecológica, es decir que los individuos híbridos son viables siempre y cuando existan las condiciones ecológicas adecuadas para su desarrollo. Esto ya fue indicado por Palacios & Bravo (1981). La principal evidencia es que los híbridos aparecen principalmente en lugares disturbados, producidos en general por la acción humana. Como casos ejemplares de tales modificaciones Palacios & Bravo (1981) citan el sobre-pastoreo de las sabanas del centro y norte de Formosa y de los bañados y esteros de Formosa y Chaco; la desaparición de la vegetación natural por efecto de las crecientes del Pilcomayo; la aparición de ambientes abiertos en la proximidad a las poblaciones debido al sobrepastoreo; y la tala indiscriminada tanto en Santiago del Estero como en el Chaco. Esta asociación entre ambientes disturbados y formación de híbridos fue también indicada por Verga (1995) para las especies de algarrobo típicas

del monte (*P. flexuosa* y *P. chilensis*). Estas especies en condiciones “normales” ocupan diferentes ambientes: *P. flexuosa* crece principalmente en sitios donde existe un aporte de agua extra ya sea superficial o subterránea, mientras que *P. chilensis* aparece en valles y llanuras donde no hay aporte de agua extra. Las especies entran en contacto en el pedemonte de las sierras del Monte donde la formación de híbridos es evidente en zonas altamente disturbadas por la ganadería y la agricultura. Estas observaciones hechas por de Palacios & Bravo (1981) y Verga (1995) indicarían que los híbridos no son intrínsecamente inviables pero que su desarrollo está limitado por la falta de un ambiente adecuado. Así en el marco conceptual de Coyne & Orr (2004) el mecanismo de aislamiento reproductivo sería el de inviabilidad ecológica. Esta idea es compatible con el plantea propuesto por Saidman et al. (1995) en el cual se indica que las especies del grupo de los algarrobos podrían ser consideradas especies cohesivas (sensu Templeton 1989) manteniéndose la integridad específica más por mecanismos cohesivos ecológicas que genéticos.

Evidencias de una segunda barrera postcigótica han sido obtenida por Naranjo et al. (1984), quienes encontraron una fertilidad disminuida en la combinación híbrida *P. affinis* X *P. nigra*. Lamentablemente no han existido trabajos posteriores que hayan permitido continuar con la evaluación de este fenómeno. Independientemente de qué mecanismo postcigótico en particular haya actuado, el mismo se tuvo que haber formado en menos de 3.1-3.5 millones de años, el lapso total de la diversificación de este grupo.

6.6. Un escenario posible para la evolución de *Prosopis* en América

6.6.1. Evidencias obtenidas en esta tesis

Los resultados obtenidos en esta tesis nos permiten especular un posible escenario histórico para la diversificación de *Prosopis* en América que debería

someterse a prueba en el futuro con la incorporación de nuevos datos. Consideramos que cualquier ensayo para un análisis más profundo sobre la historia evolutiva de *Prosopis* debe tomar como punto de partida una hipótesis que sea lo más explicativa de toda la evidencia relevante existente hasta el momento. La discusión que daremos en esta sección tiene ese objetivo.

Una de las predicciones derivadas de la hipótesis de una diversificación de *Prosopis* impulsada por el desarrollo de ambientes áridos es la existencia de una correlación temporal entre el aumento de la extensión geográfica de los ambientes áridos en América y la diversificación del género (ver Capítulo 4). El desarrollo de ambientes áridos en América del Sur y América del Norte comenzó en el Mioceno Superior y continuó en el Plioceno y Pleistoceno. En Norte América, el desarrollo de áreas áridas a escala regional habría comenzado en el Mioceno Superior o Plioceno (Graham, 1999; Riddle & Hafner, 2006) asociado con el levantamiento de mesetas en el Oeste de Norteamérica. Este proceso de aridificación continuó durante el Pleistoceno, alcanzando una considerable distribución para los períodos interglaciares (Axelrod 1979a,b). En Sudamérica, la formación de grandes áreas con condiciones semiáridas y áridas comenzó posiblemente en el Mioceno Superior en concomitancia con la elevación de los Andes en la fase distrófica Quechua, momento en el cual la cordillera alcanzó la mitad de la elevación que presenta hoy (Gregory-Wodzicki, 2000) y comenzó a actuar efectivamente como barrera para los vientos cargados de humedad. Durante la fase distrófica ‘Diaguita’ el levantamiento final de las sierras Pampeanas y el Andes Central generó un efecto de “sobra de lluvia” que dio origen a las condiciones extremadamente áridas que existen hoy en día en el área presente entre estas estribaciones (Pascual Ortiz-Jaureguizar & Prado, 1996; Alberdi, Bonadonna & Ortiz-Jaureguizar, 1997). La datación molecular realizada en esta tesis sugirió que, excepto

para los dos grupos principales, la divergencia entre los linajes de las EAP habría ocurrido en el Mioceno Superior o en tiempos posteriores, que como quedó anteriormente expuesto, coincide con la formación de zonas áridas en América. Esto está de acuerdo con lo predicho por la hipótesis de una diversificación de *Prosopis* asociada a un aumento de la distribución de ambientes áridos en América.

Los cambios climáticos previamente señalados permitieron la conformación de los biomas en los cuales se encuentran hoy presentes las especies de *Prosopis*. Tanto el Monte (Axelrod et al., 1991) como la Patagonia (Pascual et al., 1996) se habrían desarrollado concomitantemente con el surgimiento de los Andes desde el Mioceno Superior. Algo similar habría ocurrido en la Puna, donde desde su levantamiento se habría instalado las condiciones extremas y áridas que prevalecen hoy en día. El área actualmente ocupada por el desierto de Atacama, tuvo condiciones semidesérticas desde el Mioceno, 8 millones de años antes del presente (MA ap) hasta el plioceno tardío (3 MA ap) puntuado por una fase de mayor aridez hace 6 millones de años. El desierto de Sonora, donde actualmente habitan algunas especies del género, se habría originado entre los 15 y los 8 MA ap (van Devender, 2000). En conclusión, estos biomas conformados por especies con neta afiliación a ambientes áridos y en los cuales se hayan hoy presente las especies de *Prosopis* se habrían desarrollado en los últimos 10 MA en un marco temporal consistente con los valores de diversificación obtenidos para la mayoría de los linajes de este género.

Una diversificación del género asociada a la expansión de ambientes áridos se ve apoyada también por la evidencia obtenida en el análisis de tasas de diversificación. El análisis de linajes a través del tiempo, LTT (Capítulo 4, Figura 21) indicó que existiría con un claro incremento en la tasa de diversificación en el género en momentos en los cuales se habrían desarrollado a gran escala los ambientes áridos en América, es

decir desde el Mioceno Superior al Plioceno. Si este patrón se analiza en detalle, se pueden distinguir dos cambios hacia tasas más altas. El primero de ellos ocurrió en el Mioceno tardío y es coincidente con el comienzo de la divergencia dentro de cada uno de los dos principales linajes de las EAP (Figura 20, Capítulo 4). El segundo aumento en la tasa estaría asociado con el comienzo de la divergencia de los algarrobos durante el Plioceno y está de acuerdo con la alta tasa de diversificación estimada para este grupo.

El análisis de condiciones ancestrales indicó una asociación antigua de las especies de *Prosopis* con ambientes áridos posiblemente desde el comienzo de la diversificación dentro de los dos clados principales (Strombocarpa + *X. viridinamis* y Algarobia SL) en el Mioceno Superior. Para el caso del ancestro de la sección Strombocarpa, el análisis indicó que el mismo habría estado presente sólo en ambientes áridos. En el caso del ancestro de la sección Algarobia SL, los resultados difirieron entre reconstrucciones pero todas coinciden en que al menos parte de la distribución habría estado en ambientes áridos. La diversificación habría ocurrido dentro de ambientes áridos y no por múltiples ocupaciones independientes de este ambiente, indicando una larga historia de relaciones entre las especies del género y estos ambientes. Estos resultados no pueden ser interpretados como una evidencia a favor de una diversificación de *Prosopis* asociada al desarrollo de ambientes áridos ya que una única ocupación o múltiples ocupaciones de ambientes áridos son compatibles con tal hipótesis.

6.6.2. Diversificación de taxones en zonas áridas: evidencias para un patrón común de diversificación

La idea que los cambios ambientales son un factor clave en la diversificación de la mayoría de los grupos tiene un sustento tanto lógico como empírico. Dada la estrecha relación que necesariamente existen entre los organismos y su entorno, y que los

procesos de especiación y extinción (únicos procesos responsables de determinar el nivel de diversificación de los organismos) están altamente afectados por el ambiente, se deduce que la relación entre cambios en el ambiente y cambios en los patrones de diversificación de los organismos es directa. Existen innumerables trabajos en los cuales, con diferentes niveles de corroboración, se ha mostrado esta relación entre cambios ambientales y cambios en la tasa de diversificación. Entre los posibles ejemplos se pueden indicar trabajos relacionados con la diversificación en mamíferos (Delsuc et al., 2004; Gingerich, 2006), en Angiospermas para el Cretácico Inferior (Heimhofer et al., 2005), en comunidades vegetales completas como es el caso de las Guayanas (Rull, 2005), en peces cíclidos en África (Sturmbauer, 2005). Para el caso de particular de las leguminosas, los cambios climáticos han sido considerados por Axelrod (1992) como claves en su evolución. Según este autor la radiación de las tres subfamilias habría estado asociada a grandes cambios climáticos que en algunos casos estuvieron representados por modificaciones hacia climas más secos y calurosos y en otros casos hacia climas más fríos y secos. También los cambios cíclicos ocurridos durante el cuaternario habrían dado las condiciones para un aumento de la diversificación.

Una de las herramientas más comúnmente usadas hoy en día para analizar asociaciones entre cambios ambientales y diversificación es la datación molecular. Esto permite tener un marco temporal para aquellos grupos sin registro fósil. Entre estos análisis existe un número importante referido a taxones de zonas áridas (ej. Verboom et al., 2003; Moore & Hansen 2006; Good-Avila et al. 2006). En la mayoría de los casos se ha encontrado evidencia para plantear que la diversificación de las especies habría estado asociada al aumento de condiciones áridas en los últimos 20 MA. Estos estudios han estado enfocados generalmente a nivel genérico, como es el caso *Tiquilla*

(Boraginaceae). Dicho estudio (Moore & Hansen 2006) indicó que los principales linajes dentro de ese género habrían aumentado su diversificación concomitantemente con el aumento en la distribución en zonas áridas a finales del Mioceno. Good-Avila et al. (2006) encontraron algo similar en el género *Agave*, otro integrante de ambientes áridos de América. En este trabajo los autores siguieren que el punto máximo de diversificación habría ocurrido entre los 8 y 6 MA ap coincidiendo con el incremento de las condiciones áridas en México. La evidencia geológica sugeriría que la actividad volcánica habría cambiado el clima entre 30 y 15 MA ap, llevando a condiciones más áridas en el centro del país después de este período y con la formación del desierto de Sonora entre 15 y 8 MA ap. Entre los 5 y 1.8 MA ap hubo una reversión hacia climas más tropicales y húmedos, después de lo cual el clima se volvió mucho más frío y seco durante los períodos glaciares. En este trabajo se ha planteado además que *Agave* tendría una diversificación coincidente con otros grupos de plantas de ambientes áridos durante los últimos 10 MA (Good- Avila et al., 2006). Nuestro estudio sería el primero en encontrar un patrón similar para un grupo con distribución principalmente sudamericana como es *Prosopis*.

Esta asociación entre la diversificación de taxones vegetales con aumento de la aridez no se restringe a América, también ha sido indicada por ejemplo para géneros australianos como es el caso de *Tratheca* (Tremandraceae) en el cual la mayoría de sus especies habrían divergido en los últimos 6-7 MA coincidiendo con el período en el cual Australia sufrió un proceso de aridificación. La región para la que, sin dudas, hay una mayor cantidad de trabajos sobre diversificación en ambientes áridos es la región Capense (Richardson et al., 2001; Klak et al., 2003; Verboom et al., 2003; Mummenhoff et al., 2005; Bakker et al., 2005), situada al sur de África y caracterizada por un altísimo grado de endemismo (Kreft & Jetz 2007). La mayoría de estos estudios

han indicado que los grupos estudiados presentarían un aumento en la tasa de diversificación en un período que va desde el Mioceno tardío al Plioceno (Cowling et al, en prensa). La interpretación dada para este patrón ha sido siempre un aumento de la tasa de diversificación debido al cambio de las condiciones climáticas en el Mioceno tardío-Plioceno que habrían ampliado las áreas con condiciones ambientales áridas o semiáridas. Esta idea fue sostenida principalmente por el grupo de trabajo del Dr. Peter Linder. En uno de sus trabajos (Verboom et al., 2003), este grupo analizó al género *Enharta* (Poaceae). Partiendo de la estimación de tiempos de divergencia a partir de los datos moleculares realizaron un análisis de linajes a través del tiempo (LTT). En el mismo encontraron un cambio en la tasa de diversificación hacia el Mioceno Tardío. En otro de los estudios realizados analizaron la diversificación de un grupo de la familia Aizoaceae (Klak et al., 2003), y encontraron que la diversificación de la subfamilia Ruschoideae habría ocurrido en los últimos 8 millones de años, más posiblemente en los últimos 4 MA. Esto implica una tasa altísima de diversificación considerando que la subfamilia comprende cerca de 1000 especies. En este estudio también se plantea al aumento de zonas de clima semiárido como principal control de la diversificación de este grupo.

Recientemente se ha planteado una visión alternativa a la previamente indicada sobre la dinámica de la diversificación en la región Capense (Cowling et al. En prensa). La misma sostiene que en realidad el principal determinante del aumento de la tasa de diversificación habría sido un aumento en la diversidad de ambientes que habría propiciado un aumento en la tasa de especiación de estos grupos de plantas. Tal idea se basa en estudios geológicos e indican que habría habido una exposición de suelos arcillosos ricos en nutrientes, la incisión de valles fluviales en las zonas bajas, la exposición de cuarcitas frescas en las montañas, y el amplio desarrollo de extensas áreas

de un substrato calcáreo a lo largo de la costa. En particular Cowling et al. (En prensa) plantean que la causa de tal diversificación habría sido una relación sinergética entre cambio climático y geomorfológico. El cambio climático según estos autores no sería suficiente para explicar el fenómeno. Es así que nuevos datos, no previamente involucrados en la discusión plantean al menos una visión alternativa válida para el aumento de la diversificación en el caso de la región Capense. La resolución de esta disputa esta totalmente abierta.

Una explicación similar a esta visión alternativa se ha dado para explicar la diversificación de un grupo de especies del género *Rheum* (Polygonaceae) que ocupan zonas áridas/semiáridas de la Meseta Tibetana Qinghai (China), al Norte del cordón del Himalaya (Wang et al., 2005). El estudio mostró que la diversificación habría ocurrido en los últimos 7 MA, señalando diferentes factores como posibles causas de su diversificación. Entre los mismos los autores indican (i) Una oscilación entre ambientes secos y húmedos durante el período Plioceno-Holoceno habría promovido la diversificación al producir un aislamiento de habitats que había permitido la divergencia en pequeñas poblaciones y su posterior diversificación; (ii) una amplia diversidad geológica y ecológica en la meseta y zonas cercanas que también habría permitido una fuerte diversificación. Sin embargo, uno solo de los posibles factores habría sido una condición necesaria para la ocurrencia de la diversificación: el levantamiento de la Meseta Tibetana Qinghai que habría permitido la formación de los ambientes en los cuales hoy en día viven las especies del género: praderas alpinas secas y frías, estepas desérticas y laderas secas.

También se ha ensayado una explicación similar al estudiar el género *Encelia*, cuyas especies habitan el desierto de la Península de California. Según Fehlberg & Ranker (2007) la diversificación del género en esta región se habría visto favorecida por

la diversidad de habitats representados por variaciones en elevación, precipitación, temperatura estacionalidad y sustrato.

Es así que existirían dos hipótesis sobre el factor principal de la diversificación de géneros con afinidad a ambientes áridos: el aumento de la distribución espacial de condiciones propicias y una mayor diversidad de ambientes. Es posible que estos factores hayan tenido importancias relativas diferentes en cada uno de los grupos analizados. Nuestros análisis están de acuerdo con que el aumento de zonas áridas habría colaborado a la diversificación de *Prosopis* en América. Sin embargo la posibilidad de que también haya jugado un rol el origen de mayor variación de ambientes es algo que no puede descartarse. Es posible que una interacción entre ambas causas haya sido importante para la diversificación del género. Al desarrollarse los ambientes áridos en condiciones geomorfológicos y físicas diversas, habrían aparecido subambientes que habrían también propiciado la diversificación. Así dada esta hipótesis, el desarrollo de ambientes áridos habría sido una condición necesaria para la diversificación aunque no suficiente. El segundo componente habría sido un escenario geomorfológico diverso.

Este posible control externo asociado al clima, no niega la posible existencia de caracteres claves de estas especies (Miller, 1949; Simpson, 1953) que hayan promovido la diversificación de *Prosopis*. De hecho se han sugerido diferentes cambios fisiológicos y morfológicos que habrían promovido la diversificación de *Prosopis* en zonas áridas (Roig, 1993): cambio en hábito de árboles a arbustos, desarrollo de raíces profundas, reducción en el tamaño y cantidad de hojas, lignificación del fruto y adaptaciones fisiológicas para maximizar la absorción de agua y minimizar su pérdida. Un análisis de la evolución de estos caracteres junto con estudios de la base molecular de la tolerancia de las especies de *Prosopis* a factores de estrés ayudarán a tener una

mejor imagen de los procesos de adaptación y diversificación de *Prosopis* en ambientes áridos

6.7. Conclusiones

- El análisis de filogenia molecular indicó que el género *Prosopis* no sería monofilético debido a que las especies del viejo mundo no formaron un clado monofilético con las especies americanas y además porque *Xerocladia viridiramis* apareció cercanamente relacionada con estas últimas
- Las especies Americanas de *Prosopis* aparecieron formando dos grupos principales, el primero incluyendo todas las especies de la sección *Strombocarpa* y el segundo incluyendo las especies de las secciones *Algarobia* y *Monilicarpa*. Dentro de este último grupo aparecieron 2 clados altamente apoyados. Uno de ellos formados por las especies de las series *Chilenses*, *Pallidae* y *Ruscifoliae* (grupo de los algarrobos) y el otro formados por especies de las series *Sericanthae*, *Denudantes* de la sección *Algarobia* y *P. argentina* de la sección *Monilicarpa*.
- La divergencia entre los linajes americanos de *Prosopis* habría ocurrido en tres momentos claramente diferenciados en el tiempo: una temprana divergencia entre las secciones *Strombocarpa* y *Algarobia* para el Mioceno medio; una posterior divergencia dentro de cada una de ellas para el Mioceno tardío; y una diversificación del grupo de los algarrobos para el Plioceno.
- La diversificación de *Prosopis* en América se habría acelerado desde el Mioceno tardío hasta el presente. En particular la tasa de diversificación alcanzada por el grupo de los algarrobos es comparable con las más altas encontradas en plantas.
- La diversificación del género *Prosopis* en América habría estado asociada con el paulatino aumento del área ocupada por ambientes áridos las zonas áridas a partir de fines del Mioceno y el Pleistoceno.
- La intención de evaluar las condiciones ambientales en las cuales habrían vivido los ancestros de este género requirió de una previa evaluación crítica de los métodos empleados hasta el momento para cumplir tal tarea. Dicha evaluación indicó que los mismos partían de hipótesis auxiliares nunca explicitadas y que

además presentaban varios problemas metodológicos. En esta tesis se plantearon soluciones para algunos de estos problemas, que fueron consideradas en el estudio empírico desarrollado en *Prosopis*.

- La optimización de caracteres ambientales indicó que las especies de *Prosopis* habrían ocupado ambientes áridos en América desde el comienzo de su diversificación y que posiblemente estuvieran ocupando tales ambientes desde el Mioceno.
- Durante el proceso de secuenciación de la región *trnS-psbC* del genoma de cloroplasto se descubrió una pequeña inversión de 10 pares de bases. Un análisis detallado de la misma indicó que habría ocurrido al menos 10 veces en diferentes linajes de angiospermas. Por otra parte la ocurrencia de la inversión estaría fuertemente condicionada a la estabilidad de la estructura secundaria de la región genómica adyacente a la inversión.

6.8. Bibliografía

- Alberdi, M. T., F. P. Bonadonna & E. Ortiz-Jaureguizar. 1997. Chronological correlation, paleoecology, and paleobiogeography of the Late Cenozoic South American “Rionegrán” land-mammal fauna: a review. *Revista Española de Paleontología* 12: 249-255.
- Alvarez, I. & J. F. Wendel. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29: 417-434.
- Anzótegui, L. M. & R. Herbst. 2004. Megaflora (hojas y frutos) de la Formación San José (Mioceno Medio) en río Seco, departamento Santa María, provincia de Catamarca, Argentina. *Ameghiniana* 41: 423-436.
- Axelrod DI. 1979a. Age and origin of Sonoran desert vegetation. *Occasional Papers of the California Academy of Sciences* 132: 1-74.
- Axelrod DI. 1979b. Desert vegetation, its age and origin. In: Goodin JR, Northington DK, ed. *Arid land plant resources*. Lubbock, TX: International Center for Arid and Semiarid Land Studies, 1-72.
- Axelrod, D., M. Kalin Arroyo & P. Raven. 1991. Historical development of temperate vegetation in the Americas. *Revista Chilena de Historia Natural* 64: 413-446
- Axelrod, D. I. 1992. Climatic pulses a major factor in legume evolution. In P. S. Hendereen and D. L. Dilcher (eds), *Climatic pulses a major factor in legume evolution*. The royal botanical garden, kew, pp. 259-279.
- Baldwin, B. G. & M. J. Sanderson. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences* 95: 9402-9406.
- Bakker, F.T., Culham, A., Marais, E.M. & Gibby, M., 2005. Nested radiation in Cape *Pelargonium*. In: Bakker, F.T., Chartrou, L.W., Gravendeel, B. & Pelser, P.B. (Ed.), *Plant species-level systematics. New perspectives on pattern and process*. ARG Ganter Verlag, Ruygel, Lichtenstein, pp. 76-100.
- Bessegá, C., L. I. Ferreyra, J. C. Vilardi & B. O. Saidman. 2000a. Unexpected low genetic differentiation among allopatric species of Section Algarobia of *Prosopis* (Leguminosae). *Genetica* 109: 255-266.
- Bessegá, C., B. Saidman & J. Vilardi. 2000b. Isozyme and RAPD studies in *Prosopis glandulosa* and *P. Velutina* (Leguminosae, Mimosoideae). *Genetics and Molecular Biology* 23: 1-5.

- Bessegá, C., J. C. Vilardi & B. O. Saidman. 2006. Genetic relationships among American species of the genus *Prosopis* (Mimosoideae, Leguminosae) inferred from ITS sequences: evidence for long-distance dispersal. *Journal of Biogeography* 33: 1905-1915.
- Britton, N. & J. Rose. 1928. Mimosaceae. *North American Flora*. 23: 1 -194.
- Burghardt, A. & S. Espert. 2007. Phylogeny of *Prosopis* (Leguminosae) as shown by morphological and biochemical evidence. *Australian Journal of Botany* 20: 332-339.
- Burghardt, A. D. 1995. La identidad de cuatro especies de *Prosopis* L. expresada a través de sus patrones electroforeticos. *Mendeliana* 12: 38-50.
- Burkart, A. 1940. Materiales para una monografía del género *Prosopis*. *Darwiniana* 4: 57-128.
- Burkart, A. 1976. A monograph of the genus *Prosopis* (Leguminosae subFam Mimosoideae). *Journal of the Arnold Arboretum* 57: 219-249.
- Burkart, A. & B. Simpson. 1977. The genus *Prosopis* and annotated key to the species of the world. In B. Simpson (ed) *Mesquite: Its Biology in Two Desert Ecosystems*. Dowden, Hutchinson and Ross, Stroudsburg, Pennsylvania, USA., pp. 201-215.
- Caccavari, M. & E. Dome. 2000. An account of morphological and structural characterization of american mimosoideae pollen. part I: Tribe Acacieae. *Palynology* 24: 231-248.
- Catalano, S. A., C. Bessegá, B. Saidman & J. Vilardi. 2001. Análisis de Marcadores RAPD en *Prosopis kuntzei* (Leguminosae). Comparación con otras especies de la sección Algarobia In. XXX Congreso Argentino de Genética y IV Jornadas Argentino-Uruguayas de Genética. Mar del Plata-Buenos Aires.
- Coyne, J. A. & H. A. Orr. 2004. *Speciation*. Sinauer Associates, Sunderland, MA.
- Cowling R. M., Ş. Proches & T.C. Partridge. In press. Explaining the uniqueness of the Cape flora: Incorporating geomorphic evolution as a factor for explaining its diversification. *Molecular Phylogenetic and Evolution*.
- Cronn, R., R. Small, T. Haselkorn & J. Wendel. 2002. Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. *American Journal of Botany* 89: 707-725.
- Delsuc, F., S. Vizcaíno & E. Douzery. 2004. Influence of Tertiary paleoenvironmental changes on the diversification of South American mammals: a relaxed molecular clock study within xenarthrans. *BMC Evolutionary Biology* 4.

- Dobzhansky, T., F. Ayala, G. Sttebins & J. Valentine. 1977. *Evolution*. W.H. Freeman Co., San Francisco, EEUU.
- Dobzhansky, T. 1970. *Genetics of the Evolutionary Process*. Columbia University Press, New York.
- Fehlberg, S. & T. Ranker. 2007. Phylogeny and Biogeography of *Encelia* (Asteraceae) in the Sonoran and Peninsular Deserts Based on Multiple DNA Sequences. *Systematic Botany* 32: 692–699.
- Ferreira, L., C. Bessega, J. Vilardi & B. Saidman. 2007. Consistency of population genetics parameters estimated from isozyme and RAPDs dataset in species of genus *Prosopis* (Leguminosae, Mimosoideae). *Genetica* 131: 217-230.
- Flores-Cruz, M., H. Santana-Lira, S. Koch & R. Grether. 2004. Taxonomic Significance of Leaflet Anatomy in *Mimosa* series Quadrivalves (Leguminosae, Mimosoideae). *Systematic Botany* 29: 892–902.
- Gingerich, P. 2006. Environment and evolution through the Paleocene–Eocene thermal maximum. *Trends in Ecology & Evolution* 21: 246-253.
- Good-Avila, S. V., V. Souza, B. S. Gaut & L. E. Eguiarte. 2006. Timing and rate of speciation in *Agave* (Agavaceae). *Proceedings of the National Academy of Sciences* 103: 9124–9129.
- Graham, A. 1999. Late Cretaceous and Cenozoic History of North American Vegetation, New York, NY, USA
- Graham, S., P. Reeves, A. Burns & R. Olmstead. 2000. Microstructural Changes in Noncoding Chloroplast DNA: Interpretation, Evolution, and Utility of Indels and Inversions in Basal Angiosperm Phylogenetic Inference. *International Journal of Plant Sciences* 161: S83–S96.
- Gregory-Wodzicki, K. M. 2000. Uplift history of the central and northern Andes: a review. *Geological Society of America Bulletin* 112: 1091-1105
- Guinet, P. H. & M. Bessedik. 1984. Presence de Genre *Prosopis* (Leguminosae-Mimosoideae) a l'Aquitainien Basal dans l'aude (Languedoc-france). *Pollen et spores* 26: 101-108.
- Heimhofer, U., P. Hochuli, S. Burla, J. Dinis & H. Weissert. 2005. Timing of Early Cretaceous angiosperm diversification and possible links to major paleoenvironmental change. *Geology* 33: 141-144.

- Hoc, P. S., M. A. Agullo, & R. A. Palacios. 1994. Stylar trimorphism in four functionally andromonoecious *Prosopis* species (Mimosaceae). *Plant Systematics and Evolution* 190: 143–156
- Hodges, S. A. & M. L. Arnold. 1995. Spurring plant diversification: are floral nectar spurs a key innovation? *Proceedings of the Royal Society of London serie B*: 343-348.
- Hughes, C. E. & R. Eastwood. 2006. Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences* 103: 10334-10339.
- Jansen, R. K. & J. D. Palmer. 1987. A chloroplast DNA inversion marks an ancient evolutionary split in the sunflower family (Asteraceae). *Proceedings of the National Academy of Sciences of the United States of America* 84: 5818–5822.
- Kelchner, S. A. 2000. The Evolution of Non-Coding Chloroplast DNA and Its Application in Plant Systematics. *Annals of the Missouri Botanical Garden* 87: 482-498.
- Kelchner, S. A. & J. F. Wendel. 1996. Hairpins create minute inversions in non-coding regions of chloroplast DNA. *Current Genetics* 30: 259-262.
- Klak, C., G. Reeves & T. Hedderson. 2003. Unmatched tempo of evolution in Southern African semi-desert ice plants. *Nature* 427: 63-65.
- Kreft, H., Jetz, W., 2007. Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences* 104, 5925–5930.
- Lavin, M., P. S. Herendeen & M. F. Wojciechowski. 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Systematic Biology* 54: 575-594.
- Luckow, M. 2002. Anatomical Features of the Leaves in the *Dichrostachys* Group (Leguminosae: Mimosoideae) and their Utility for Phylogenetic Studies. *Systematic Botany* 27: 29–40.
- Luckow, M., J. Miller, R. Jobson & D. Murphy. 2004. Generic relationships in the Mimosoideae (Fabaceae) Botany *In* Alpine Diversity: Adapted to the Peaks. Session 19 (Colloquium) Rupert Barneby and his legume legacy, Salt Lake City, Utah, USA.
- Magallon-Puebla, S. & S. R. S. Ceballos-Ferroz. 1994. Fossil legume fruits from tertiary strata of Puebla, Mexico. *Canadian Journal of Botany* 72: 1027-1038.

- Malcomber, S. 2002. Phylogeny of *Gaertnera* Lam. (Rubiaceae) based on multiple DNA markers: evidence of a rapid radiation in a widespread, morphologically diverse genus. *Evolution* 56: 42-57.
- Miller, A.H. 1949. Some ecological and morphologic considerations in the evolution of major higher taxonomic categories. 84-88. in E. Mayr & E Schuz (ed) Winter Heidelberg, pp. 84- 88.
- Moore, M. J. & R. K. Jansen. 2006. Molecular evidence for the age, origin, and evolutionary history of the American desert plant genus *Tiquilia* (Boraginaceae). *Molecular Phylogenetics and Evolution* 39: 668-687.
- Mummenhoff, K., I. Al-Shehbaz, F. Bakker, H. Linder & A. Muhlhausen. 2005. Phylogeny, morphological evolution, and speciation of endemic Brassicaceae genera in the Cape flora of southern Africa. *Annals of the Missouri Botanical Garden* 92: 400-424.
- Naranjo, C. A., L. Poggio & S. Enus Zeiger. 1984. Phenol chromatography, morphology and cytogenetics in three species and natural hybrids of *Prosopis* (Leguminosae-Mimosoideae). *Plant Systematic and Evolution* 144: 257-276.
- Nichols, R. 2001. Gene trees and species trees are not the same. *Trends in Ecology & Evolution* 16: 358-364.
- Palacios, R. & L. Bravo. 1981. Hibridación natural en *Prosopis* (Leguminosae) en la región chaqueña argentina. Evidencias morfológicas y cromatográficas. *Darwiniana* 23: 3-35.
- Palacios, R. & P. Hoc. 2005. Revisión del género *Prosopidastrum* (Leguminosae) para la Argentina. *Boletín de la Sociedad Argentina de Botánica* 40: 113-128.
- Pamilo, P. & M. Nei. 1988. Relationships between gene trees and species trees. *Molecular Biology and Evolution* 5: 568-583.
- Pascual, R., E. Ortiz-Jaureguizar & J. L. Prado. 1996. Land mammals: paradigm of Cenozoic South American geobiotic evolution. In G. Arratia (ed) Contribution of Southern South America to Vertebrate Paleontology. *Müncher Geowissenschaftliche Abhandlungen* (A).
- Pasiecznik, N. M., P. Felker, P. J. C. Harris, L. N. Harsh, G. Cruz, J. C. Tewari, K. Cadoret & L. J. Maldonado. 2001. The *Prosopis juliflora*-*Prosopis pallida* Complex: A Monograph. HDRA, Coventry UK.
- Piel, K. M. 1971. Palynolgy of Oligocene sediments from British Columbia. *Canadian Journal of Botany* 49: 1885-1920.

- Ramírez, L., A. De La Vega, N. Razkin, V. Luna & P. J. C. Harris. 1999. Analysis of the relationships between species of the genus *Prosopis* revealed by the use of molecular markers. *Agronomie* 19: 31-43.
- Raubeson, L. & R. Jansen. 1992. Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. *Science* 255: 1697-1699.
- Richardson, J. E., R. T. Pennington, T. D. Pennington & P. M. Hollingsworth. 2001. Rapid diversification of a species-rich genus of Neotropical rain forest trees. *Science* 293: 2242 - 2245.
- Richardson, J., F. Weitz, M. Fay, Q. Cronk, H. Linder, G. Reeves & M. Chase. 2001. Rapid and recent origin of species richness in the Cape Flora of South Africa. *Nature* 412: 181-183.
- Riddle, B. R. & D. J. Hafner. 2006. A step-wise approach to integrating phylogeographic and phylogenetic biogeographic perspectives on the history of a core North American warm deserts biota. *Journal of Arid Environments* 66: 435-461
- Rull, V. 2005. Biotic diversification in the Guayana Highlands: a proposal. *Journal of Biogeography* 32: 921 - 927.
- Saidman, B. & J. Vilardi. 1987. Analysis of the genetic similarities among seven species of *Prosopis* (Leguminosae: Mimosoideae). *Theoretical and Applied Genetics* 75.
- Simmons, M., V. Savolainen, C. Clevinger, R. Archer & J. Davis. 2001. Phylogeny of the Celastraceae inferred from phytochrome B gene sequence and morphology. 19: 353-366.
- Simpson, G. G. 1953. The major features of evolution. Columbia University Press, New York.
- Simpson, B., A. Burkart & N. Carman. 1975. *Prosopis palmeri*: a relict of an ancient North American colonization. *Madroño* 23: 220-227.
- Small, R., R. Cronn & J. Wendel. 2004. L.A.S. Johnson Review no. 2. Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* 17: 145-170.
- Small, R. & J. Wendel. 2000. Phylogeny, duplication, and intraspecific variation of *Adh* sequences in new world diploid cottons. *Molecular phylogenetics and Evolution* 16: 73-84.

- Sturmbauer, C. 2005. Explosive speciation in cichlid fishes of the African Great Lakes: a dynamic model of adaptive radiation. *Journal of Fish Biology* 53: 18 - 36.
- Tank, D. & T. Sang. 2001. Phylogenetic utility of the glycerol-3-phosphate acyltransferase gene: Evolution and implications in *Paeonia* (Paeoniaceae). *Molecular Phylogenetics and Evolution* 19: 421–429.
- Templeton, A.R. 1989. The meaning of species and speciation: a genetic perspective. En D. Otte & J.A. Endler (Eds.): *Speciation and its Consequences* :3-27. Sinauer, Sunderland, MA.
- van Devender, T. R. 2000. The Deep history of the Sonoran Desert. In S. J. Phillips and P. W. Comus (eds), *The Deep history of the Sonoran Desert*. Arizona-Sonora Desert Museum Press, Tucson, and University of California Press, Berkeley., pp. 61-69
- Verboom, G. A., H. P. Linder & W. D. Stock. 2003. Phylogenetics of the grass genus *Ehrharta*: Evidence for radiation in the summer-arid zone of the South African cape. *Evolution* 57: 1008–1021.
- Wang, A., M. Yang & J. Liu. 2005. Molecular phylogeny, recent radiation and evolution of gross morphology of the rhubarb genus *Rheum* (Polygonaceae) inferred from chloroplast DNA trnL-F sequences. *Annals of Botany* 96: 489–498.
- Watson, S. 1889. Proc. Am. Acad. Arts & Sci 24: 36-82.

ANEXOS

ANEXO I. SCRIPTS UTILIZADOS PARA EL ANÁLISIS DE INVERSIONES

Anexo 1. Script 1

Este *script* genera la matriz de costos para realizar la optimización de Fixed States con Inversiones (FSI) en TNT. En el cuerpo del *script* se incluye una matriz en la cual cada “taxon” representa a cada uno de los motivos del espaciador que fueron observados (con un máximo de 31). Los primeros caracteres representan la secuencia de cada uno de los espaciadores seguidos por el complemento reverso de cada uno de ellos. Si algún espaciador es más corto que el resto, entonces se le agregan el símbolo “7” luego de la secuencia original y luego del complemento reverso.

Los argumentos que se le deben dar al correr es script son la penalidad de inversión y la penalidad de indels. Para este script todas las sustituciones cuestan uno.

Los archivos de salida son:

1- La matriz de costos de Sankoff que debe ser añadida al set de datos.

Un proto-script que, luego de algunas modificaciones, debe ser añadido al script que será el que optimice la inversión. Para optimizar la inversión se debe construir una nueva matriz donde cada Terminal tendrá un símbolo que represente el motivo del espaciador correspondiente. El motivo que fue el numero 0 en el script **FSI_COSTS.run** tendrá el simbolo “0”, el segundo tendrá el estado 1 mientras que el 11vo tendrá el símbolo “A”, etc.

```

nstates 31
xread
20 31
0      AGGGGAA777TTCCCCT777
1      AGTGGAA777TTCCACT777
2      AGTGGGA777TCCCCT777
3      AGTGGTA777TACCACT777
4      AGTGGTGATATATCACCACT
5      AGTGGTGGTATACCACT777
6      GATAGAA777TTCTATC777
7      GATGGAA777TTCCATC777
8      GGGAGGA777TCCTCCC777
9      GGGGGA777TCCCCC7777
10     GGGGGGA777TCCCCCCC777
11     GGGTGGA777TCCACCC777
12     GGTAGCA777TGCTACC777
13     GGTAGGA777TCCTACC777
14     GGTGCAA777TTCGACC777
15     GGTGAGA777TCTCACC777
16     GGTGGGA777TCCCACC777
17     GGTGTGA777TCACACC777
18     GTTGGGA777TCCCAAC777

```

```

19      TACCACTACTGGTGGTA
20      TATCACCCTAGTGGTGATA
21      TCATACC777GGTATGA777
22      TCCCACC777GGTGGGA777
23      TCCCCT777AGTGGGA777
24      TCCCCCCC777GGGGGAA777
25      TCCCGCC777GGCGGGAA777
26      TCTCACCC777GGTGAGA777
27      TCTCGCC777GGCGAGA777
28      TTCCACT777AGTGGAA777
29      TTCCCCC777GGGGGAA777
30      TTTCACC777GGTGAAA777
;
ra 1 ;
;
macro-;
macro *700 100200 ;
macro [99999 ;
macro=;
report-;
macreport-;
sil=all;
var:
costo cola colas cccc fifa fifas ffff invcost coo stpmatrix [(ntax+1) (ntax+1)]
bases seqs largof largol largofi largola luu suma match [24 24] matcho [24 24] wave [24 24] indel
inv_pen
;
set indel %1 /* First argument: Indel Penalty */
set inv_pen %2 /*Second argument: Inversion Penalty" */
log Sankoff.txt ;
log/;
log batch.txt ;
log/;
set bases (((nchar+1)/2)) ;
set seqs ntax ;
loop=fi 0 'seqs'
set largof 0 ;
loop=c 0 ('bases'-1)
if ((states [#c #fi 0]) != 128)
    set largof ++ ;
end
stop
log +batch.txt; sil-file; quote if (variable_Y [nod] ==#fi ) ; sil=file ; log/;
loop=las 0 'seqs'
loop=w 0 23
loop=r 0 23
    set match [#w #r] 0 ;
stop
stop
loop=d 0 23
loop=e 0 23
    set matcho [#d #e] 0 ;
stop
stop
set largol 0 ;
loop=l 0 ('bases'-1)
if ((states [#l #las 0]) != 128)
    set largol ++ ;
end
stop

```

```

loop=co 0 ('largof'-1)
  loop=lu 0 ('largol'-1)
    set coo #co+1 ;
    set luu #lu+1 ;
    if ((states [#co #fi 0] != (states [#lu #las 0]))
      set match ['coo' 'luu'] 1 ;
    end
    stop
  stop
  set coo 0 ;
  set luu 0 ;
  set ffff 1 ;
  set cccc 1 ;
  loop=colu 'bases' ('largof'+ 'bases'-1)
    loop=mnas 0 ('largol'-1)
      if ((states [#colu #fi 0] != (states [#mnas #las 0]))
        set matcho [ 'ffff' 'cccc' ] 1 ;
      end
      set cccc++ ;
      stop
      set cccc 1 ;
      set ffff++ ;
      stop
      set ffff 0 ;
    loop=uco 0 'largof'
      set wave [#uco 0] 'suma' ;
      set suma 'suma' + 'indel' ;
    stop
    set suma 0 ;
    loop=ufil 0 'largol'
      set wave [0 #ufil] 'suma' ;
      set suma 'suma' + 'indel' ;
    stop
    set suma 0 ;
    loop=coa 1 ('largof')
      loop=fif 1 ('largol') /* aca hago la directa */
      set fifa (#fif-1) ;
      set cola (#coa-1) ;
      set wave [#coa #fif] ('match [#coa #fif] + 'wave ['cola' 'fifa']) ;
      if ('wave [#coa #fif] > ('wave [#coa 'fifa'] + 'indel'))
        set wave [#coa #fif] ('wave [#coa 'fifa'] + 'indel') ;
      else
        if ('wave [#coa #fif] > ('wave ['cola' #fif] + 'indel'))
          set wave [#coa #fif] ('wave ['cola' #fif] + 'indel') ;
      end
    end
    set costo 'wave [#coa #fif]' ;
  stop
  stop
  loop=cc 0 23
  loop=ff 0 23
    set wave [#cc #ff] 0 ;
  stop
  stop
  loop=uco 0 'largof'
    set wave [#uco 0] 'suma' ;
    set suma 'suma' + 'indel' ;
  stop
  set suma 0 ;
  loop=ufi 0 'largol'

```

```

set wave [0 #ufi] 'suma' ;
set suma 'suma' + 'indel' ;
stop
set suma 0 ;
loop=oac 1 ('largo1') /* acá hago la inversa */
loop=ifa 1 ('largo1')
set fifas (#ifa-1) ;
set colas (#oac-1) ;
set wave [#oac #ifa] ('matcho [#oac #ifa] + 'wave ['colas' 'fifas']) ;
if ('wave [#oac #ifa] > ('wave [#oac 'fifas] + 'indel'))
    set wave [#oac #ifa] ('wave [#oac 'fifas] + 'indel') ;
else
    if ('wave [#oac #ifa] > ('wave ['colas' #ifa] + 'indel'))
        set wave [#oac #ifa] ('wave ['colas' #ifa] + 'indel') ;
    end
end
set invcost 'wave [#oac #ifa]' ;
stop
stop
loop=ccc 0 23
loop=fff 0 23
set wave [#ccc #fff] 0 ;
stop
stop
if (('invcost' + 'inv_pen') < 'costo')
set stpmatrix [#fi #las] ('invcost' + 'inv_pen') ;
log +batch.txt; sil-file; quote if (variable_Y [myanc] == #las) ; sil=file ; log/;
log +batch.txt; sil-file; quote set trans [pasos 0] nod .. ; sil=file ; log/;
log +batch.txt; sil-file; quote set trans [pasos 1] myanc .. ; sil=file ; log/;
log +batch.txt; sil-file; quote set pasos++ .. ; sil=file ; log/;
log +batch.txt; sil-file; quote log +invertido silfile., qote arbol arbol ,recy
rcy ,rcx, ancestro myanc, descendiente, anc .. ; sil=file ; log/;
log +batch.txt; sil-file; quote log +invertido silfile., qote
***** .. ; sil=file ; log/;
log +batch.txt; sil-file; quote end; sil=file ; log/;
else
set stpmatrix [#fi #las] 'costo' ;
end
stop
log +batch.txt; sil-file; quote end; sil=file ; log/;
stop
log/;
log Sankoff.txt;
loop=ma 0 'seqs'
loop=triz 0 'seqs'
    lquote = ;
    sil-file; quote #ma/#triz 'stpmatrix [#ma #triz]' ; sil=file ;
    lquote -;
stop
stop
log/;
proc/;
```

Anexo I. Script 2

El siguiente *script* calcula la probabilidad de un patrón observado pero usando los valores de energía libre estimada en el análisis de estructura secundaria ancestral. Estos valores aparecen en el *script* como variable_x [4] 12.5 ; donde el numero entre corchetes es el número del nodo de acuerdo a TNT y el último número es la enegía asignada al mismo (estado forzado a ese nodo).

```

report=;
macro- ;
macro *700 100020 ;
macro [200000 ;
macro= ;
macreport-;
report-;
echo-;
var:
xtrue ytrue yre xre ar arbol arbtot alfa atrec azar grupos [10] enramas entrando itera limpasos lim limite
myanc nodo nod numero nodtot nodazarA [(2*ntax+1)] numrecs pasos perejil [1000] probabilidad
rama ramas_obs recs realsimul recy simul trans [(2*ntax+1) 2] raiz ulnode variable_Y [(2*ntax+1)]
variable_X [(2*ntax+1)] visitado valor yahecho [50] za rex rey chance gilun [1000] chancey proba
entrar target_total no_target_total totales no_enramas no_target_total chgtot
;
log sheet_f.txt;
sil-file; quote BRANCH_TOT, BRANCH_TAR, BRANCH_NO_TAR, CHG_TOT, CHG_TAR,
CHG_NO_TAR ; sil=file; log/;
taxname=;
sil=all ;
set rey 0 ;
set rex 0 ;
set limite 10.0 ; /* LIMITE ENTRE LAS CLASES ALTA Y BAJA ENERGIA LIBRE */
set recy 0 ;
set itera 0;
set pasos 0 ;
set simul 1000000 ;
set recs 0 ;
set numrecs 0 ;
set atrec 0 ;
set nodtot (2*ntax+1) ;
set azar 0 ;
set ulnode (2*ntax+1) ;
set arbtot 0 ;
set xre 1 ;
set raiz root ;
set yre 1 ;
set variable_X [91] 12.57 ; set variable_X [92] 12.09 ; set variable_X [93] 12.09 ; set
variable_X [94] 12.09 ; set variable_X [95] 13.08 ; set variable_X [96] 12.98 ; set variable_X
[97] 12.98 ; set variable_X [98] 12.98 ; set variable_X [99] 12.98 ; set variable_X [100] 12.68 ;
set variable_X [101] 12.68 ; set variable_X [102] 12.98 ; set variable_X [103] 12.98 ; set
variable_X [104] 12.98 ; set variable_X [105] 12.98 ; set variable_X [106] 12.98 ; set variable_X
[107] 12.97 ; set variable_X [108] 13.09 ; set variable_X [109] 13.09 ; set variable_X [110] 13.09
; set variable_X [111] 13.09 ; set variable_X [112] 13.09 ; set variable_X [113] 13.09 ; set
variable_X [114] 13.09 ; set variable_X [115] 13.09 ; set variable_X [116] 12.98 ; set variable_X
[117] 12.09 ; set variable_X [118] 6.19 ; set variable_X [119] 6.65 ; set variable_X [120] 6.65 ;
set variable_X [121] 7.13 ; set variable_X [122] 7.13 ; set variable_X [123] 7.13 ; set variable_X
[124] 12.98 ; set variable_X [125] 12.98 ; set variable_X [126] 12.98 ; set variable_X [127] 12.98
; set variable_X [128] 12.98 ; set variable_X [129] 12.98 ; set variable_X [130] 7.13 ; set
variable_X [131] 7.13 ; set variable_X [132] 7.13 ; set variable_X [133] 7.13 ; set variable_X

```

```

[134] 12.98 ; set variable_X [135] 12.98 ; set variable_X [136] 12.68 ; set variable_X [137] 13.09
; set variable_X [138] 13.09 ; set variable_X [139] 13.09 ; set variable_X [140] 13.09 ; set
variable_X [141] 12.98 ; set variable_X [142] 8.49 ; set variable_X [143] 7.13 ; set variable_X
[144] 12.98 ; set variable_X [145] 12.98 ; set variable_X [146] 6.19 ; set variable_X [147] 21.66
; set variable_X [148] 21.66 ; set variable_X [149] 13.09 ; set variable_X [150] 13.09 ; set
variable_X [151] 13.09 ; set variable_X [152] 13.09 ; set variable_X [153] 6.65 ; set variable_X
[154] 6.65 ; set variable_X [155] 12.98 ; set variable_X [156] 12.98 ; set variable_X [157] 12.98
; set variable_X [158] 12.98 ; set variable_X [159] 12.09 ; set variable_X [160] 6.65 ; set
variable_X [161] 8.98 ; set variable_X [162] 8.98 ; set variable_X [163] 8.98 ; set variable_X
[164] 8.98 ; set variable_X [165] 13.08 ; set variable_X [166] 13.08 ; set variable_X [167] 13.08
; set variable_X [168] 13.09 ; set variable_X [169] 13.09 ; set variable_X [170] 7.13 ; set
variable_X [171] 21.14 ; set variable_X [172] 7.13 ; set variable_X [173] 8.98 ; set variable_X
[174] 13.08 ; set variable_X [175] 7.13 ; set variable_X [176] 7.62 ; set variable_X [177] 12.98
; set variable_X [178] 8.98 ; set variable_X [179] 8.98 ; set variable_X [180] 7.13 ; goto = %0 ;
collapse none ;
outgroup Amborella ; /* INDICAR EL OUTGROUP AQUÍ */
collapse none ;
loop=tre 0 'arbtot'
  set xre 0 ;
  set yre 0 ;
  k 0;
  rseed*;
tread
(Amborella (Nymphaea (Illicium ((Nandina (Platanus (Buxus ((Beta Spinacia )(((Citrus (((Coccina
(Benincasa (Citrillus ((Paecytrulus (Dyplocyclos (Achanthosatacys (Lagenaria Trichosanthes )))))(Luffa_a
Luffa_c ))))))(Ecballiumelaterium (Sicyosangulatus ((Bryonia dioica (Cucumis (Sechiumedule (Cucurbita
Marahoreganus ))))))(Cyclantherapedata Echinocystislobata )))))(Castanea (Carpinus_a Carpinus_b
))))((((Affinis (Ferox Xerocladia ))(Acacia Prosopidastrum )))(Lotus (Cicer (Medicago Pisumsativum
))))((Morus (Humuluslupulusdemas HumuluslupulusHaplotypeP ))(Photinia Prunus )))(Eumynus
(Elaeocarpus (Sapiumjaponicum (Populus Populustrichocarpa ))))))(Eucalyptus Pelargonium
))(Schizophagma (((((Ajania (Chrysanthemum_a Chrysanthemum_b ))(Helianthus Lactuca ))(Viburnum
(Daucus Panax )))(Jasminum ((Atropa (Lycopersicon (Solanbulbocastanum Solantuberosum
))))(Nicotislyvestris Nicotitabacum )))))(Arcterica ((Camelia_a Camelia_b )(Pierisfloribunda
Pieristaiwanensis )))))))))(Chloranthus ((Dioscorea ((Phalaenopsis (Allium_cepa Allium_th
))))(Bactrisgasipae1 Bactrissetulosa2 )((Phyllostachis ((Hordeum Triticum )(Oryzanivara Oryzasativa
))))(Zeamays (Saccharum Sorghum )))))(Calycanthus (Drimys Liriodendron ))))))));
/* EL ARBOL DEBE INSERTARSE AQUÍ */
loop=gr 0 ntrees
  set grupos [#gr] tnodes [#gr] ;
  stop
  set arbol #tre ;
  iterrecs 0 1 perejil
    set yre ++ ;
  endrecs ;
  iterrecs 0 0 gilun
    set xre ++ ;
  endrecs ;
  loop=tt 0 (2*ntax)
    set trans [#tt 0] 0 ;
    set trans [#tt 1] 0 ;
  stop
  loop=y 0 (2*ntax)
    set variable_Y [#y] 0 ;
  stop
  macreport-;
  report-;
  iterrecs 0 0 &variable_X
    set chance getrandom [0 100] ;
    set xtrue++;
    if (('chance'/100 <= ('rex'/'xre')) || ('xre' <=1000) )

```

```

iterrecs 0 1 variable_Y
    set ytrue++ ;
    set chancey getrandom [0 100] ;
    if (('chancey'/100 <= ('rey'/'yre')) || ('yre' <=1000) )
        set pasos 0 ;
        loop=nod 0 (2*ntax)
        if (#nod != 'raiz')
            set myanc anc [0 #nod] ;
            set nod #nod ;
            if ('variable_Y ['nod']' != ('variable_Y ['myanc']'))
                goto INVERSION ;
            end
            if (('.2variable_X ['nod']' < 'limite') && ('.2variable_X ['myanc']'
                < 'limite'))
                set target_total++ ;
            end
            if ('(.2variable_X ['nod']'>='limite') && ('.2variable_X
                ['myanc']>='limite'))
                set no_target_total++ ;
            end
        end
        stop
    loop=trapa 0 ('pasos'-1)
    if (('.2variable_X ['trans [#trapa 0 ]]'<'limite') && ('.2variable_X ['trans [
        #trapa 1 ]]'<'limite'))
        set enramas ++ ;
        log +cambios_en_ramas_target4.txt ; sil-file; quote Cambio en rama
        Nodo = 'trans [#trapa 0]' / Ancestro = 'trans [#trapa 1]' ; sil=file; log/
    end
    if ((('.2variable_X ['trans [#trapa 0 ]]'>='limite') && ('.2variable_X ['trans [
        #trapa 1 ]]'>='limite')))
        set no_enramas++ ;
    end
    stop
    set totales 'target_total' + 'no_target_total' ;
    set chgtot 'enramas' + 'no_enramas' ;
    log +sheet_f.txt ; sil-file ; quote 'totales','target_total','no_target_total','chgtot',
    'enramas','no_enramas' ; sil=file ;log/
    set enramas 0 ;
    set pasos 0 ;
    set recy ++ ;
    set ramas_obs 'enramas' ;
    set target_total 0 ;
    set totales 0 ;
    set no_enramas 0 ;
    set no_target_total 0 ;
    set chgtot 0 ;
    end
endrecs
    set recy 0 ;
    set itera ++ ;
    loop=vx 0 (2*ntax)
        set variable_X [#vx] 0 ;
    stop
end
endrecs ;
    set itera 0 ;
    set limpasos 0 ;
    set arbol++ ;
stop

```

```

proc/;
label INVERSION /* ESTO PROVIENE DEL OUTPUT DEL SCRIPT FSI_COSTS.run */;

if ('variable_Y ['nod']' ==0 )
  if ('variable_Y ['myanc']' == 19)
    set trans ['pasos' 0] 'nod';
    set trans ['pasos' 1] 'myanc';
    set pasos++;
  end
if ('variable_Y ['myanc']' == 20)
  set trans ['pasos' 0] 'nod';
  set trans ['pasos' 1] 'myanc';
  set pasos++;
end
if ('variable_Y ['myanc']' == 21)
  set trans ['pasos' 0] 'nod';
  set trans ['pasos' 1] 'myanc';
  set pasos++;
end
if ('variable_Y ['myanc']' == 22)
  set trans ['pasos' 0] 'nod';
  set trans ['pasos' 1] 'myanc';
  set pasos++;
end
if ('variable_Y ['myanc']' == 23)
  set trans ['pasos' 0] 'nod';
  set trans ['pasos' 1] 'myanc';
  set pasos++;
end
if ('variable_Y ['myanc']' == 24)
  set trans ['pasos' 0] 'nod';
  set trans ['pasos' 1] 'myanc';
  set pasos++;
end
if ('variable_Y ['myanc']' == 25)
  set trans ['pasos' 0] 'nod';
  set trans ['pasos' 1] 'myanc';
  set pasos++;
end
if ('variable_Y ['myanc']' == 26)
  set trans ['pasos' 0] 'nod';
  set trans ['pasos' 1] 'myanc';
  set pasos++;
end
if ('variable_Y ['myanc']' == 27)
  set trans ['pasos' 0] 'nod';
  set trans ['pasos' 1] 'myanc';
  set pasos++;
end
if ('variable_Y ['myanc']' == 28)
  set trans ['pasos' 0] 'nod';
  set trans ['pasos' 1] 'myanc';
  set pasos++;
end
if ('variable_Y ['myanc']' == 29)
  set trans ['pasos' 0] 'nod';
  set trans ['pasos' 1] 'myanc';
  set pasos++;
end
if ('variable_Y ['myanc']' == 30)

```

```
set trans ['pasos' 0] 'nod';
set trans ['pasos' 1] 'myanc';
set pasos++;
end
end
proc /;
```

Anexo I. Script 3

El script ***Fisher_FSI.run*** produce como output los valores que se necesitan para calcular la probabilidad hipergeométrica de los patrones de correlación observado. En este script en particular las inversiones se analizan con el método FSI.

Estos valores son:

BRANCH_TOT= Número total de ramas

BRANCH_LowFE = Numero de ramas con valores bajos de energía libre

BRANCH_HighFE = Numero de ramas con valores altos de energía libre

CHG_TOTAL= Número total de inversiones

CHG_HigFE = Numero de inversiones en ramas con alto valores de energía libre

CHG_LowFE = Numero de inversiones en ramas con bajos valores de energía libre

Estos valores se dan para cada una de hasta 50 reconstrucciones de cada carácter y hasta 50 resoluciones de las posibles politomías presentes en la topología original. La topología debe pegarse en el espacio indicado dentro del cuerpo del script. Este script también da como salida un archivo en el que se encuentran todas las resoluciones de la topología original que fueron evaluadas en el análisis. El archivo de salida esta en formato compacto TNT.

```

report=;
macro- ;
macro *700 100020 ;
macro [200000 ;
macro= ;
macreport-;
report-;
echo-;
var:
xtrue ytrue yre xre matriz [ (2*ntax+1) (2*ntax+1) ] arbol arbtot alfa azar arbmax cuenta grupos [10]
enramas itera limpasos limite myanc nodo nod nodito numero nnn nodB nodtot nodazarA [(2*ntax+1)]
numrecs pasos perejil [1000] probabilidad raiz rama ramas_obs recs recy simul trans [ (2*ntax+1) 2 ]
ulnode variable_Y [ (2*ntax+1) ] variable_X [ (2*ntax+1) ] za rex rey chance gilun [1000] chancey proba
entrar target_total no_target_total totales no_enramas no_target_total chgtot
;
log sheet.txt; sil-file; quote TREE, REC_X,REC_Y,BRANCH_TOT, BRANCH_HighFE,
BRANCH_LowFE, CHG_TOT, CHG_HigFE, CHG_LowFE ; sil=file; log/;

taxname=;
sil-all ;
set rey 50 /* NUMERO DE RECONSTRUCCIONES ANALIZADAS PARA LA VARIABLE
RESPUESTA (CARACTER 1) */
set rex 50 /* NUMERO DE RECONSTRUCCIONES ANALIZADAS PARA LA VARIABLE
PREDICTORA (CARACTER 0) */
set limite 10.0 /* LIMITE ENTRE LAS CLASES ALTA Y BAJA ENERGIA LIBRE */
set recy 0 ;
set itera 0;

```

```

set arbmax 0;
set pasos 0 ;
set raiz root ;
set simul 1000000 ; /*INDICAR AQUI EL NUMERO DE REPLICAS DE LA SIMULACION*/
set recs 0 ;
set numrecs 0 ;
set nodtot (2*ntax+1) ;
set azar 0 ;
set ulnode (2*ntax+1) ;
set arbtot 50 ; /* NUMERO DE RESOLUCIONES DIFERENTES DE LA TOPOLOGIA A
EVALUAR (SOLO SI HAY POLITOMIAS) */
set xre 0 ;
set yre 0 ;
goto =%0 ;
collapse none ;
outgroup Amborella ; /* INDICAR EL OUTGROUP AQUÍ */

collapse none ;

loop=tre 0 'arbtot'

    set xre 0 ;
    set yre 0 ;
    k 0;
    rseed*;
    tread
    (Amborella (Nymphaea (Illicium ((Nandina (Platanus (Buxus ((Schizophagma ((Arcterica
    ((Camelia_a Camelia_b) (Pierisfloribunda Pieristaiwanensis )))(Viburnum ((Ajania (Chrysanthemum_a
    Chrysanthemum_b ))(Helianthus Lactuca ))(Daucus Panax ))(Jasminum ((Atropa (Lycopersicon
    (Solanbulbocastanum Solantuberosum ))(Nicotisylvestris Nicotitabacum ))))))(Eucalyptus Pelargonium
    (Citrus (((Castanea (Carpinus_a Carpinus_b ))((Coccina (Benincasa (Citrillus ((Luffa_a Luffa_c
    )(Paecytrulus (Dyplocyclos (Achanthosatacys (Trichosanthes Lagenaria )))))))))(Ecballiumelaterium
    (Sicyosangulatus ((Bryoniadioica (Cucumis (Sechiumedule (Cucurbita Marahoreganus
    )))))(Cyclantherapedata Echinocystislobata ))))))(Eumynus (Elaeocarpus (Sapiumjaponicum (Populus
    Populustrichocarpa ))))((Prunus Photinia )(Morus (Humuluslupulusdemas HumuluslupulusHapl0EPyP
    ))))((Acacia Prosopidastrum (Affinis (Ferox Xerocladia )))(Lotus (Cicer (Medicago Pisumsativum
    ))))))((Beta Spinacia ))))(Chloranthus (Dioscorea (((Phylostachis ((Hordeum Triticum )(Oryzanivara
    Oryzasativa )))(Zeamays (Saccharum Sorghum )))(Bactrisgasipaes1 Bactrissetulosa2 )))(Phalaenopsis
    (Allium_cepa Allium_th )))(Calycanthus (Drimys Liriodendron )))))));

    /* EL ARBOL DEBE INSERTARSE AQUÍ */
randtrees*;
loop=gr 0 ntrees
    set grupos [#gr] tnodes [#gr] ;
stop
set arbol #tre ;
iterrecs 0 1 perejil
    set yre ++ ;
endrecs ;
iterrecs 0 0 gilun
    set xre ++ ;
endrecs ;
loop=tt 0 (2*ntax)
    set trans [#tt 0] 0 ;
    set trans [#tt 1] 0 ;
stop
loop=y 0 (2*ntax)
    set variable_Y [#y] 0 ;
stop
iterrecs 0 0 variable_X
    set chance getrandom [0 100] ;

```

```

set xtrue++;
if (('chance'/100 <= ('rex'/'xre')) || ('xre' <=4) )
  iterrecs 0 1 variable_Y
    set ytrue++;
    set chancey getrandom [0 100];
    if (('chancey'/100 <= ('rey'/'yre')) || ('yre' <=4) )
      set pasos 0;
      loop=nod 0 (2*ntax)
        if (#nod != 'raiz')
          set myanc anc [0 #nod];
          set nod #nod;
          if ('variable_Y ['nod']' != ('variable_Y ['myanc']'))
            goto INVERSION;
        end
        if ((/.2variable_X ['nod'] < 'limite') && (/.2variable_X ['myanc'] < 'limite'))
          set target_total++;
        end
        if ((/.2variable_X ['nod']>='limite') && (/.2variable_X ['myanc']>='limite'))
          set no_target_total++;
        end
      end
    end
  stop
loop=trapa 0 ('pasos'-1)
  if ((/.2variable_X ['trans [#trapa 0 ]']<'limite') && (/.2variable_X ['trans [#trapa 1 ]']<'limite'))
    set enramas++;
  end
  if (((/.2variable_X ['trans [#trapa 0 ]']>='limite') && (/.2variable_X ['trans [#trapa 1 ]']>='limite')))
    set no_enramas++;
  end
  stop
  set totales 'target_total' + 'no_target_total';
  set chgtot 'enramas' + 'no_enramas';
  log +sheet.txt ; sil-file ; quote 'arbol','itera','recy','totales','target_total','no_target_total','chgtot','enramas','no_enramas'; sil=file ;log/ ;
  set enramas 0;
  set pasos 0;
  set recy++;
  set ramas_obs 'enramas';
  set target_total 0;
  set totales 0;
  set no_enramas 0;
  set no_target_total 0;
  set chgtot 0;
end
endrecs;
set recy 0;
set itera++;
loop=vx 0 (2*ntax)
  set variable_X [#vx] 0;
stop
end
endrecs;
set itera 0;
set limpaso 0;
set arbol++;
tsave arboles.ctf +; save; tsave/;

```

```

stop
proc/;

label INVERSION
/* ESTO PROVIENE DEL OUTPUT DEL SCRIPT FSI_COSTS.run */
if ('variable_Y ['nod']= ==0 )
  if ('variable_Y ['myanc']= == 19)
    set trans ['pasos' 0] 'nod';
    set trans ['pasos' 1] 'myanc';
    set pasos++;
  end
  if ('variable_Y ['myanc']= == 20)
    set trans ['pasos' 0] 'nod';
    set trans ['pasos' 1] 'myanc';
    set pasos++;
  end
  if ('variable_Y ['myanc']= == 21)
    set trans ['pasos' 0] 'nod';
    set trans ['pasos' 1] 'myanc';
    set pasos++;
  end
  if ('variable_Y ['myanc']= == 22)
    set trans ['pasos' 0] 'nod';
    set trans ['pasos' 1] 'myanc';
    set pasos++;
  end
  if ('variable_Y ['myanc']= == 23)
    set trans ['pasos' 0] 'nod';
    set trans ['pasos' 1] 'myanc';
    set pasos++;
  end
  if ('variable_Y ['myanc']= == 24)
    set trans ['pasos' 0] 'nod';
    set trans ['pasos' 1] 'myanc';
    set pasos++;
  end
  if ('variable_Y ['myanc']= == 25)
    set trans ['pasos' 0] 'nod';
    set trans ['pasos' 1] 'myanc';
    set pasos++;
  end
  if ('variable_Y ['myanc']= == 26)
    set trans ['pasos' 0] 'nod';
    set trans ['pasos' 1] 'myanc';
    set pasos++;
  end
  if ('variable_Y ['myanc']= == 27)
    set trans ['pasos' 0] 'nod';
    set trans ['pasos' 1] 'myanc';
    set pasos++;
  end
  if ('variable_Y ['myanc']= == 28)
    set trans ['pasos' 0] 'nod';
    set trans ['pasos' 1] 'myanc';
    set pasos++;
  end
  if ('variable_Y ['myanc']= == 29)
    set trans ['pasos' 0] 'nod';
    set trans ['pasos' 1] 'myanc';
    set pasos++;
  end

```

```
end
if ('variable_Y ['myanc']' == 30)
  set trans ['pasos' 0] 'nod';
  set trans ['pasos' 1] 'myanc';
  set pasos++;
end
end
proc/;
```

Anexo I. Script 4.

El script ***simulation_.run*** calcula las probabilidades asociadas a la combinación de topología-reconstrucción que presentó el valor máximo o mínimo de probabilidad. Dentro del script se deben indicar:

Número total de inversions

- Número de inversions en ramas con bajo valor de energía libre
- Número de ramas con valores bajos de energía libre
- Número de ramas con valores bajos de energía libre

La simulación funciona de la siguiente manera:

Todos los nodos presentan un estado inicial de 0. Primero, uno de los nodos se elige al azar y su estado y el de los nodos descendientes son cambiados (como ahora todos están en 0, pasan a 1). Luego se elige un segundo nodo cuidando que no haya sido elegido previamente. El mecanismo para este segundo nodo es el mismo que para el primero. Si los descendientes del nodo en cuestión tuvieran una asignación de 1 entonces pasan a 0. Se sigue así tantas veces como cambios (inversiones) haya. Las probabilidades calculadas de esta manera deberían ser las mismas que las calculadas utilizando el test exacto de Fisher (1935). Análisis de prueba indicaron que al considerar 15.000 réplicas, la diferencia entre las probabilidades calculadas mediante el test exacto de Fisher y las simulaciones fueron en promedio del 2%.

Cuando se han elegido una cantidad de nodos igual al numero observado de inversiones y los estados de los descendientes se hayan modificado, los nodos terminales pueden haber cambiado su estado con respecto al original. Si estos estados son optimizados en el árbol, se puede calcular el número de cambios inferidos por parsimonia. Así se pueden aceptar a la hora de calcular las probabilidades, solo aquellas reconstrucciones que, dada la optimización de parsimonia, tengan el mismo número de cambios que los originalmente inferidos. Las simulaciones hechas con este script se han hecho de esta manera.

macro- ;

```

macro *700 100020 ;
macro [200000 ;
macro= ;
macreport-;
report-;
echo-;
var:
yre xre matriz [ (2*ntax+10) (2*ntax+10) ] distri [ (2*ntax+1) ]
ar arbol arbtotal aza azar nodaz arbmax larg cuenta columna
grupos [10] enramas final [ (2*ntax+1) ] istr itera limpitos lim
limite maxprob myanc nodo nod numero nodtot nodazarA [ (2*ntax+1) ]
numrecs pasos perejil [1000] probabilidad raiz rama ramas_obs
recomax realsimul recy simul target trans [ (2*ntax+1) 2 ]
ulnode variable_Y [ (2*ntax+1) ] variable_X [ (2*ntax+1) ] visitado valor
za rex rexxy rey chance gilun [1000] chancey proba target_total no_target_total totales
no_enramas no_target_total chgtot pasos_pars original [(ntax+1)] simulacion_Y [ (2*ntax+1) ] champion
[ (2*ntax+1) ] rasa [ (2*ntax+1) ] abajo arriba enra pas no_enra entra
;

macseed 0;
rseed 0 ;
taxname=;
sil=all ;
set rey 1 ;
set rex 1 ;
set limite 10.0 ;/* LIMITE ENTRE LAS CLASES CON VALORES ALTOS Y BAJOS DE ENERGÍA LIBRE/LARGO DE IR */
set recy 0 ;
set itera 0;
set arbmax 0;
set recomax 0;
set maxprob 0;
set pasos 0 ;
set raiz root ;
set simul 1000000 ;
set realsimul %1; /* PRIMER ARGUMENTO: NUMERO DE REPLICAS */
set reccs 0 ;
set numrecs 0 ;
set nodtot (2*ntax+1) ;
set azar 0 ;
set ulnode (2*ntax+1) ;
set arbtotal 0 ;
set xre 0 ;
set yre 0 ;
goto =%0 ;
set no_target_total 0 ;
set target_total 0 ;
collapse none ;
outgroup Amborella ;/* INDICAR ACA EL GRUPO EXTERNO */
collapse none ;

loop=tre 0 'arbtotal'
  set xre 0 ;
  set yre 0 ;
  k 0;
  rseed*;
  tread
    (Amborella (Nymphaea (Illicium ((Nandina (Platanus ((Eucalyptus      (Pelargonium
    (Citrus (((Acacia (Prosopidastrum (Affinis (Ferox Xerocladia )))))(Lotus (Cicer (Medicago
    Pisum sativum )))))(Morus (Humulus lupulus demas Humulus lupulus HaplotypeP ))(Photinia

```

```

Prunus ))))))((Coccina (Benincasa (Citrillus ((Paecytrulus (Dyplocyclos (Achanthosatacys
(Lagenaria Trichosanthes ))))((Luffa_a Luffa_c ))))((Ecballiumelaterium (Sicyosangulatus
((Bryoniadioica (Cucumis (Sechiumedule (Cucurbita Marahoreganus ))))((Cyclantherapedata
Echinocystislobata ))))((Castanea (Carpinus_a Carpinus_b ))((Eumynus (Elaeocarpus
(Sapiumjaponicum (Populus Populustrichocarpa ))))))))((Schizophragma (((((Ajania
(Chrysanthemum_a Chrysanthemum_b ))((Helianthus Lactuca ))((Viburnum (Daucus Panax
))))((Jasminum ((Atropa (Lycopersicon (Solanbulbocastanum Solantuberosum ))))((Nicotisylvestris
Nicotitabacum ))))((Arcteria ((Camelia_a Camelia_b ))((Pierisfloribunda Pieristaiwanensis
))))))((Beta Spinacia ))))((Chloranthus ((Dioscorea ((Phalaenopsis (Allium_cepa Allium_th
))))((Bactrisgasipaes1 Bactrissetulosa2 ))((Phylostachis ((Hordeum Triticum )(Oryzanivara
Oryzasativa ))((Zeamays (Saccharum Sorghum ))))))))((Calycanthus (Drimys Liriodendron ))))))));
; /* PASTE THE CORRESPONDING TOPOLOGY HERE */

loop=gr 0 ntrees
  set grupos [#gr] tnodes [#gr];
  stop
  set arbol #tre;
  iterrecs 0 0 variable_X

  set no_target_total 0;
  set target_total 0;
  loop=nd 0 (2*ntax) /* aca se fija en que nodos ocurrieron los cambios y pone
    ambos nodos en trans */
    if (#nd != 'raiz')
      set myanc anc [0 #nd];
      set nod #nd;
    I if ((/.2variable_X ['nod'] < 'limite') && ('/.2variable_X ['myanc'] <
      'limite'))
      set target_total++;
    end
    if (( ('/.2variable_X ['nod'] >='limite') && ('/.2variable_X
      ['myanc'] >='limite')))
      set no_target_total++;
  end
  end
  stop
  set pasos 13;
  set enramas 0; /* NUMBER OF CHANGES IN BRANCHES WITH IHGH ΔE */
  set chgtot 13; /* NUMBER OF TOTAL OBSERVED CHANGES*/
  if ( ('no_target_total'==111 ) && ('target_total'==59) ) /* NUMBER OF OBSERVED
TARGET AND NON-TARGET BRANCHES*/
    goto SIMULACION;
    if ('enramas'==0)
      loop=f 0 'enramas'
        set probabilidad 'probabilidad' + ('distri [#f]' / 'numero');
        stop
        log +probability.txt ; sil-file ; quote* Probability=
        'probabilidad' ; sil=file ; log/;
        set proba 'probabilidad';
        set probabilidad 0;
    else
      loop=ram 0 ('enramas'-1)
        set rama #ram;
        loop=e 0 'enramas'
          set probabilidad 'probabilidad' + ('distri [#e]' / 'numero');
    stop
    log +probability.txt ; sil-file ; quote* Probability=
    'probabilidad' ; sil=file ; log/;
    set proba 'probabilidad';
    set probabilidad 0;
  stop

```

```

        end
        killrecs
    end
    set rama 0 ;
    loop=ri 0 (2*ntax)
        set distri [#ri] 0 ;
        stop
    set enramas 0 ;
    set pasos 0 ;
    set recy ++ ;
    set ramas_obs 'enramas' ;
    set target_total 0 ;
    set totales 0 ;
    set no_enramas 0 ;
    set no_target_total 0 ;
    set chgtot 0 ;
    set itera ++ ;
    loop=vx 0 (2*ntax)
        set variable_X [#vx] 0 ;
        stop
    endrecs ;
    set itera 0 ;
    set limpasos 0 ;
    set arbol++ ;
stop
proc/;
label SIMULACION
set no_target_total 0 ;
set target_total 0 ;
loop=simu 0 'simul'
    if ( 'numero' == 'realsimul')
        endloop
    end
    rseed*;
    loop=node 0 (2*ntax)
        set simulacion_Y [ #node ] 0 ;
        stop
    loop=aza 0 ('chgtot'-1)
        rseed*;
        set azar #aza ;
        set nodazarA ['azar'] getrandom[0 (2*ntax)] ;
        if ('azar'!=0)
            loop=zar 0 ('azar'-1)
                set za #zar ;
                if ('nodazarA ['azar']' == 'nodazarA ['za']')
                    set nodazarA ['azar'] getrandom[0 (2*ntax)] ;
                if ('nodazarA ['azar']' == 'nodazarA ['za']')
                    set nodazarA ['azar'] getrandom[0 (2*ntax)] ;
                    if ('nodazarA ['azar']' == 'nodazarA ['za']')
                        set nodazarA ['azar'] getrandom[0 (2*ntax)] ;
                end
            end
        end
    stop
    end
stop
loop=az 0 ('chgtot'-1)
    set az a #az ;
    travtree up terms 0 'nodazarA ['aza']' visitado
    set matriz [ 'aza' 'visitado' ] 1 ;

```

```

endtrav
set visitado 0 ;
stop
loop=col 0 (2*ntax)
  loop=fili 0 ('chgtot'-1)
    set valor 'matriz [#fili #col]' ;
    set columna ('columna' + 'valor') ;
    stop
    set final [#col] 'columna' ;
    set cuenta 'final [#col]' / 2 ;
    if ('/.10cuenta' ) == ('/.0cuenta')) ;
      set simulacion_Y [#col] 0 ;
    else
      set simulacion_Y [#col] 1 ;
    end
    set cuenta 0 ;
    set columna 0 ;
  stop
loop=forz 0 ntax
  xread=1 #forz 'simulacion_Y [#forz]' ;
  stop
  set pasos_pars length [0 1] ;
  set aza 'chgtot';
  loop=lsas 0 10
    if ('chgtot'!= 'pasos_pars'))
      loop=loc 0 (2*ntax)
        set simulacion_Y [#loc] 0 ;
        stop
        set nodaz getrandom[0 (2*ntax)] ;
        travtree up terms 0 'nodaz' visitado
          set matriz [ 'aza' 'visitado' ] 1 ;
        endtrav
        set visitado 0 ;
      loop=col 0 (2*ntax)
        loop=fl 0 ('aza')
          set valor 'matriz [#fl #col]' ;
          set columna ('columna' + 'valor') ;
          stop
          set final [#col] 'columna' ;
          set cuenta 'final [#col]' / 2 ;
          if ('/.10cuenta' ) == ('/.0cuenta')) ;
            set simulacion_Y [#col] 0 ;
          else
            set simulacion_Y [#col] 1 ;
          end
          set cuenta 0 ;
          set columna 0 ;
        stop
      loop=forz 0 ntax
        xread=1 #forz 'simulacion_Y [#forz]' ;
        stop
        set pasos_pars length [0 1] ;
        iterrecs 0 1 rasa
          set rexx++ ;
        endrecs
        set chance getrandom [0 ('rexx'-1)] ;
        iterrecs 0 1 champion
          if ('chance'== 'entra')
            loop=orz 0 (2*ntax)
              set simulacion_Y [#orz] 'champion [#orz]' ;

```

```

        set larg recsteps ;
        stop
    end
        set entra++ ;
        endrecs
        set rex 0 ;
        set entra 0 ;
    else
        endloop
    end
        set aza++;
    stop
    set aza 0;
loop=nod 0 (2*ntax)
if (#nod != 'raiz')
    set myanc anc [0 #nod] ;
    set nod #nod ;
    if ('simulacion_Y ['nod']' != 'simulacion_Y ['myanc']')
        set trans ['pas' 0] 'nod' ;
        set trans ['pas' 1] 'myanc' ;
        set pas++ ;
    end
    if (('.2variable_X ['nod']' < 'limite') && ('.2variable_X ['myanc']' < 'limite'))
        set target_total++ ;
    end
    if ( ( ('.2variable_X ['nod']'>='limite') && ('.2variable_X
        ['myanc']'>='limite')) )
        set no_target_total++ ;
    end
end
stop
if ('chgtot' == 'pasos_pars')
    loop=trap 0 ('pas'-1)
        if (('.2variable_X ['trans [ #trap 0 ]]'<'limite') && ('.2variable_X ['trans
            [ #trap 1 ]]'<'limite'))
            set enra ++ ;
        end
        if (('.2variable_X ['trans [ #trap 0 ]]'>'limite') && ('.2variable_X ['trans [
            #trap 1 ]]'>'limite'))
            set no_enra ++ ;
        end
        if (('.2variable_X ['trans [ #trap 0 ]]'>'limite') && ('.2variable_X ['trans [
            #trap 1 ]]'<'limite'))
            set abajo ++ ;
        end
        if (('.2variable_X ['trans [ #trap 0 ]]'<'limite') && ('.2variable_X ['trans [
            #trap 1 ]]'>'limite))
            set arriba ++ ;
        end
    stop
    if ((!'arriba') && (!'abajo'))
        set numero ++ ;
        set distri ['enra'] ++ ;
    end
end
set no_enra 0 ;
set pas 0 ;
set enra 0 ;
set pasos_pars 0 ;
set target_total 0 ;

```

```
set no_target_total 0 ;
set abajo 0 ;
set arriba 0 ;
loop=mat 0 (2*ntax+2)
    loop=iz 0 (2*ntax+2)
        set matriz [#iz #mat] 0 ;
        stop
    stop
stop ;
proc/;
```

ANEXO II. Lista de especies descargadas de GenBank e incluidas en el análisis de inversiones. La columna *Análisis* indica si la secuencia fue incluida en ambos análisis (análisis de dos especies por género y el análisis con el muestreo completo: II) o solamente en el análisis de 2 especies por género (I). Además se indica aquellas secuencias que fueron excluidas del análisis por no presentar la estructura de tallo – lazo (N1); por tener una estructura fuertemente modificada (N2) o tener diferentes límites entre el tallo-lazo (N3). IN la columna ΔG se indica la energía libre asociada con la estructura de tallo–lazo correspondiente. Motivo: estado asignado a cada especie en la codificación binaria de la inversión. RIs: repeticiones invertidas

Espece	Número de accesion	Análisis	ΔG (kcal mol⁻¹)	Largo RIs	Motivo
<i>Acacia caven</i>	EF165294	II	11,79	10	0
<i>Acanthosicyos naudinianus</i>	AY396176.1	II	12,5	11	1
<i>Acer micranthum</i>	AB237387.1	N3	-	-	-
<i>Acorus calamus</i>	AJ879453.1	*	12,98	11	1
<i>Aethionema grandiflorum</i>	AP009367	N3	-	-	-
<i>Aethionema cordifolium</i>	AP009366	N3	-	-	-
<i>Agrostis stolonifera</i>	NC_008591.1	*	8,98	11	1
<i>Ajania shiwogiku</i>	AB234823.1	II	6,65	11	0
<i>Allium cepa</i>	AB292296.1	II	12,17	11	0
<i>Allium fistulosum</i>	AB292295.1	II	12,17	11	1
<i>Amborella trichopoda</i>	AJ506156.2	II	7,21	10	1
<i>Arabidopsis thaliana</i>	NC_000932.1	N3	-	-	-
<i>Arabis hirsuta</i>	AP009369.1	N3	-	-	-
<i>Arcterica nana</i>	AB206620	II	7,13	11	0
<i>Atropa belladonna</i>	NC_004561	II	7,5	11	0
<i>Bactris gasipaes</i>	DQ645572.1	II	12,68	11	0
<i>Bactris riparia</i>	DQ645578.1	I	12,98	11	1
<i>Bactris setulosa</i>	DQ645580.1	II	12,98	11	1
<i>Barbarea verna</i>	NC_009269	N3	-	-	-
<i>Benincasa hispida</i>	AY396184.1	II	12,5	11	1
<i>Beta vulgaris</i>	EF534108.1	II	12,2	11	0
<i>Brassica napus</i>	AJ578474.1	N3	-	-	-
<i>Bryonia dioica</i>	AY396175	II	12,79	11	0
<i>Buxus microphylla</i>	EF380351.1	II	7,43	11	1
<i>Calycanthus fertilis</i>	AJ428413.1	II	12,68	11	0
<i>Camellia furfuracea</i>	AY741465.1	II	8,59	12	0
<i>Camellia tenuifolia</i>	AY741466.1	II	8,59	12	0
<i>Capsella bursa-pastoris</i>	NC_009270	N3	-	-	-
<i>Carpinus tschonoskii</i>	AB237219.1	II	7,14	11	
<i>Carpinus japonica</i>	AB237206.1	II	12,5	11	1
<i>Castanea sativa</i>	AY497341.1	II	13,28	11	1
<i>Chloranthus spicatus</i>	NC_009598	II	12,98	11	1
<i>Chrysanthemum boreale</i>	AB234795.1	II	6,65	11	0
<i>Chrysanthemum zawadskii</i>	AB234786	II	6,65	11	0
<i>Cicer arietinum</i>	AC161102	II	17,46	18	1
<i>Citrullus lanatus</i>	AY396178.1	II	12,5	11	1
<i>Citrus sinensis</i>	DQ864733.1	II	12,98	11	1
<i>Coccinia palmata</i>	AY396174.1	II	12,5	11	1
<i>Coffea arabica</i>	NC_008535.1	N2	-	-	-
<i>Crucihimalaya wallichii</i>	NC_009271	N3	-	-	-
<i>Cyclanthera pedata</i>	AY396182.1	II	13,09	11	1
<i>Cucumis sativus</i>	DQ119058.1	II	13,09	11	1
<i>Cucurbita pepo</i>	AY396185.1	II	13,09	11	1
<i>Cuscuta sandwichiana</i>	AY936348	N3	-	-	-

<i>Daucus carota</i>	DQ898156.1	I	6,65	11	0
<i>Dioscorea elephantipes</i>	EF380353	II	8,63	8	1
<i>Draba nemorosa</i>	AP009373	N3	-	-	-
<i>Drimys granadensis</i>	DQ887676.1	I	8,11	11	1
<i>Diplocyclos palmatus</i>	AY396173.1	II	12,5	11	1
<i>Elaeocarpus sylvestris</i>	AB111778.1	II	15,3	12	1
<i>Ecballium elaterium</i>	AY396172.1	II	13,09	11	1
<i>Echinocystis lobata</i>	AY396181.1	II	13,09	11	1
<i>Epifagus virginiana</i>	NC_001568	N1	-	-	-
<i>Eucalyptus globulus</i>	AY780259.1	II	13,08	11	1
<i>Euonymus oxyphyllus</i>	AB237408.1	II	12,5	11	1
<i>Glycine max</i>	NC_007942	N1	-	-	-
<i>Gossypium barbadense</i>	AP009123.1	N3	-	-	-
<i>Gossypium hirsutum</i>	DQ345959.1	N3	-	-	-
<i>Helianthus annuus</i>	DQ383815.1	I	6,65	11	0
<i>Humulus japonicus</i>	AB112635.1	II	13,08	11	1
<i>Humulus lupulus</i>	AB112627.1	II	12,78	11	0
<i>Hordeum vulgare</i>	EF115541.1	II	8,98	11	1
<i>Illicium oligandrum</i>	EF380354.1	II	10,12	9	1
<i>Jasminum nudiflorum</i>	DQ673255.1	II	7,23	11	0
<i>Lactuca sativa</i>	AP007232.1	II	6,65	11	0
<i>Lagenaria siceraria</i>	AY396179.1	II	12,5	11	1
<i>Lepidium virginicum</i>	NC_009273	N3	12,3	-	-
<i>Liriodendron tulipifera</i>	DQ899947.1	II	8,8	11	1
<i>Lycopersicon esculentum</i>	DQ347959	II	7,5	11	0
<i>Lobularia maritima</i>	AP009375	N3	-	-	-
<i>Lotus japonicus</i>	AP002983.1	II	20,19	18	0
<i>Luffa echinata</i>	AY396186.1	II	12,5	11	1
<i>Luffa graveolens</i>	AY396187.1	I	12,5	11	1
<i>Luffa quinquefida</i>	AY396180.1	II	12,5	11	1
<i>Marah oreganus</i>	AY396183.1	II	13,09	11	1
<i>Medicago truncatula</i>	AC093544.8	II	15,78	18	0
<i>Morus indica</i>	DQ226511.1	II	13,08	11	1
<i>Nandina domestica</i>	DQ923117.1	II	15,24	12	1
<i>Nasturtium officinale</i>	NC_009275	N3	-	-	-
<i>Nicotiana sylvestris</i>	AB237912.1	II	7,5	11	0
<i>Nicotiana tomentosiformis</i>	AB240139.1	I	7,5	11	0
<i>Nicotiana tabacum</i>	NC_001879	II	7,5	11	0
<i>Nuphar advena</i>	NC_008788	N3	-	-	-
<i>Nymphaea alba</i>	AJ627251.1	II	13,27	11	1
<i>Oenothera elata</i>	NC_002693	N2	-	-	-
<i>Olimarabidopsis pumila</i>	NC_009267	N3	-	-	-
<i>Oryza nivara</i>	AP006728.1	II	8,98	11	1
<i>Oryza sativa</i>	X15901.1	II	8,98	11	1
<i>Panax ginseng</i>	NC_006290	II	7,5	11	0
<i>Pelargonium x hortorum</i>	DQ897681	II	8,17	10	1
<i>Phalaenopsis aphrodite</i>	AY916449.1	II	12,68	11	0
<i>Phyllostachys pubescens</i>	DQ288240.1	II	8,98	11	1
<i>Photinia glabra</i>	AB235112.1	II	12,6	11	1
<i>Pieris floribunda</i>	AB206619.1	II	7,5	12	0
<i>Pieris phillyreifolia</i>	AB206618.1	I	7,13	11	0
<i>Pieris formosa</i>	AB206617.1	I	7,13	11	0
<i>Pieris taiwanensis</i>	AB206615.1	II	7,5	11	0
<i>Pieris koidzumiana</i>	AB206614.1	I	7,13	11	0
<i>Pieris japonica</i>	AB206612.1	I	7,5	11	0
<i>Piper cenocladum</i>	DQ887677.1	I*	10,12	9	1
<i>Pisum sativum</i>	M27309	II	20,21	17	0
<i>Platanus occidentalis</i>	DQ923116.1	II	7,62	11	1
<i>Populus alba</i>	AP008956.1	II	8,98	12	1
<i>Populus trichocarpa</i>	NC_009143	II	8,98	12	1

<i>Praecitrullus fistulosus</i>	AY396191.1	II	12,5	11	1
<i>Prosopis africana</i>	EF165297	I	12,09	10	1
<i>Prosopis affinis</i>	EF165296	II	12,09	10	1
<i>Prosopis alba</i>	EF165298	I	12,09	10	1
<i>Prosopis alpataco</i>	EF165299	I	12,09	10	1
<i>Prosopis argentina</i>	EF165300	I	12,09	10	1
<i>Prosopis articulata</i>	EF165301	I	12,09	10	1
<i>Prosopis caldenia</i>	EF165302	I	12,09	10	1
<i>Prosopis campestris</i>	EF165304	I	12,09	10	1
<i>Prosopis castellanosii</i>	EF165306	I	11,79	10	0
<i>Prosopis chilensis</i>	EF165307	I	12,09	10	1
<i>Prosopis cineraria</i>	EF165309	I	11,79	10	0
<i>Prosopis denudans</i>	EF165310	I	11,79	10	0
<i>Prosopis ferox</i>	EF165311	II	12,09	10	1
<i>Prosopis flexuosa</i>	EF165312	I	12,09	10	1
<i>Prosopis glandulosa</i>	EF165314	I	12,09	10	1
<i>Prosopis humilis</i>	EF165315	I	12,09	10	1
<i>Prosopis juliflora</i>	EF165316	I	12,09	10	1
<i>Prosopis kuntzei</i>	EF165318	I	12,09	10	1
<i>Prosopis nigra</i>	EF165320	I	12,09	10	1
<i>Prosopis palmeri</i>	EF165322	I	11,79	10	0
<i>Prosopis pubescens</i>	EF165323	I	11,79	10	0
<i>Prosopis reptans</i>	EF165333	I	12,09	10	1
<i>Prosopis ruizleali</i>	EF165324	I	11,79	10	0
<i>Prosopis ruscifolia</i>	EF165325	I	12,09	10	1
<i>Prosopis sericantha</i>	EF165326	I	12,09	10	1
<i>Prosopis strombulifera</i>	EF165327	I	12,09	10	1
<i>Prosopis tamarugo</i>	EF165328	I	12,09	10	1
<i>Prosopis tamaulipana</i>	EF165329	I	12,09	10	1
<i>Prosopis torquata</i>	EF165330	I	12,09	10	1
<i>Prosopis vinalillo</i>	EF165331	I	12,09	10	1
<i>Prosopidastrum angusticarpum</i>	EF165295	II	12,09	10	1
<i>Prunus grayana</i>	AB237333.1	II	12,6	11	1
<i>Ranunculus macranthus</i>	NC_008796	N2	-	-	-
<i>Sapium japonicum</i>	AB237358.1	II	13,09	11	1
<i>Saccharum officinarum</i>	AP006714.1	II	8,98	11	1
<i>Sechium edule</i>	AY396190.1	II	13,09	11	1
<i>Schizophragma hydrangeoides</i>	AB237309	II	7,24	11	0
<i>Sicyos angulatus</i>	AY396189.1	II	13,09	11	1
<i>Spinacia oleracea</i>	AJ400848.1	II	12,68	11	0
<i>Solanum bulbocastanum</i>	DQ347958.1	II	7,5	11	0
<i>Solanum tuberosum</i>	NC_008096	II	7,5	11	0
<i>Sorghum bicolor</i>	NC_008602.1	II	8,98	11	1
<i>Trichosanthes cucumerina</i>	AY396188.1	II	12,5	11	1
<i>Triticum aestivum</i>	AB042240.3	II	8,98	11	1
<i>Viburnum furcatum</i>	AB237448.1	II	6,47	10	0
<i>Vigna angularis</i>	AB035901	N1	-	-	-
<i>Vitis vinifera</i>	DQ424856.1	N2	-	-	-
<i>Zea mays</i>	X86563.2	II	8,98	11	1
<i>Xerocladia viridiramis</i>	EF585493	II	11,79	10	0

ANEXO III. Scripts usados en el análisis de caracteres ambientales

Script 1. Script utilizado para evaluar el estado ancestral del ACMR de Algarobia considerando diferentes resoluciones de la politomía del clado de los Algarrobos.

```

macro=;
var:
  caracter
  favo
  target
  Xerocl
  nodo
  totarbol
  valmaximo
  valminimo
  ;
set caracter %1 /* primer argumento indica el carácter a analizar */
set Xerocl 2 ;
set totarbol %2 /* segundo argumento es la cantidad de resoluciones de la politomía del clado de los
algarrobos a ser evaluadas */
log %3.txt; /* tercer argumento es el nombre del archivo de salida */
sil-all; quote arbol,caracter,nodo, estado ;sil=all;
ccode +.:
loop=tre 0 'totarbol'
rseed*;
force
=
(out1 (out2 ((Xerocladia (((Ferox Tamarugo )(Reptans Torquata ))(Strombulifera (Palmeri Pubescens
))))((Affinis Alba Alpataco Articulata Caldenia Campestris Chilensis Flexuosa Glandulosa Juliflora Nigra
Ruscifolia Tamaulipana Vinalillo )((Argentina (Sericantha (Humilis Kuntzei )))(Denudans (Castellanosii
Ruizleali ))))));;
copytree * ;
randtrees*;

loop=nod 31 59
set nodo #nod ;
if (( (isanc[0 'nodo' 2]) && (isanc[0 'nodo' 3]) && (isanc[0 'nodo' 4]) && (isanc[0 'nodo' 5])
&& (isanc[0 'nodo' 6]) && (isanc[0 'nodo' 7]) && (isanc[0 'nodo' 8]) && (isanc[0 'nodo' 9]) &&
(isanc[0 'nodo' 10]) && (isanc[0 'nodo' 11]) && (isanc[0 'nodo' 12]) && (isanc[0 'nodo' 13]) &&
(isanc[0 'nodo' 14]) && (isanc[0 'nodo' 15]) && (isanc[0 'nodo' 16]) && (isanc[0 'nodo' 17]) &&
(isanc[0 'nodo' 18]) && (isanc[0 'nodo' 19]) && (isanc[0 'nodo' 20]) && (isanc[0 'nodo' 20]) &&
(isanc[0 'nodo' 21]) && (isanc[0 'nodo' 22]) && (isanc[0 'nodo' 23]) && (isanc[0 'nodo' 24]) &&
(isanc[0 'nodo' 25]) && (isanc[0 'nodo' 26]) && (isanc[0 'nodo' 27]) && (isanc[0 'nodo' 28]) &&
(isanc[0 'nodo' 29])) && (isanc[0 'nodo' 30]) && (!isanc[0 'nodo' 1]))
  set target 'nodo' ;  set valminimo contmins [1 'nodo' 0] ;
  set valmaximo contmaxs [1 'nodo' 0] ;
  sil-all; quote #tre,'caracter', 'target', 'valminimo','valmaximo' ;sil=all;
end
stop
k 0 ;
stop
;
log/
proc/;
```