

Tesis Doctoral

Genética del paisaje en poblaciones de *Prosopis alba* de la provincia de Santiago del Estero

Roser, Leandro Gabriel

2015-12-17

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en digital.bl.fcen.uba.ar. Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the doctoral theses collection of the Central Library Dr. Luis Federico Leloir, available in digital.bl.fcen.uba.ar. It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Roser, Leandro Gabriel. (2015-12-17). Genética del paisaje en poblaciones de *Prosopis alba* de la provincia de Santiago del Estero. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

Cita tipo Chicago:

Roser, Leandro Gabriel. "Genética del paisaje en poblaciones de *Prosopis alba* de la provincia de Santiago del Estero". Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2015-12-17.

EXACTAS UBA

Facultad de Ciencias Exactas y Naturales



UBA

Universidad de Buenos Aires



UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales
Departamento de Ecología, Genética y Evolución

**Genética del paisaje en poblaciones de *Prosopis alba* de la
provincia de Santiago del Estero**

Tesis presentada para optar al título de Doctor de la Universidad de Buenos
Aires en el área **CIENCIAS BIOLÓGICAS**

Lic. Leandro Gabriel Roser

Directora de tesis: Dra. Beatriz Ofelia Saidman
Directora Asistente: Dra. Laura Inés Ferreyra
Consejera de Estudios: Dra. Beatriz Ofelia Saidman

Buenos Aires, 17 de diciembre de 2015

ÍNDICE

AGRADECIMIENTOS	5
RESUMEN	7
ABSTRACT	8
I. INTRODUCCIÓN	9
1. INTRODUCCIÓN AL GÉNERO <i>PROSOPIS</i>	9
1.1. El género <i>Prosopis</i>	9
1.2. Estudios poblacionales previos en <i>Prosopis</i>	11
1.3. <i>Prosopis alba</i> (sección Algarobia)	12
1.4. Estudios poblacionales previos en <i>P. alba</i>	14
2. INTRODUCCIÓN AL ÁREA DE ESTUDIO: DESCRIPCIÓN AMBIENTAL DE LA PROVINCIA DE SANTIAGO DEL ESTERO	15
2.1. Descripción general de la provincia	15
2.2. Clima	16
2.3. Geomorfología	17
2.4. Suelos	20
3. INTRODUCCIÓN AL ANÁLISIS GENÉTICO-POBLACIONAL: COANCESTRÍA Y ENDOGAMIA - LOS ÍNDICES F y ϕ - AMOVA	26
3.1. Definiciones teóricas elementales	26
3.2. Coancestría y endogamia	27
3.3. Los índices F jerárquicos	28
3.4. Análisis Molecular de la Varianza (AMOVA)	31
4. LA GENÉTICA DEL PAISAJE Y LA ESTRUCTURACIÓN DE LAS POBLACIONES A ESCALA FINA: EL MODELO DE AISLAMIENTO POR DISTANCIA Y LOS PATRONES DE AUTOCORRELACIÓN ESPACIAL	33
4.1. Introducción	33
4.2. Aislamiento por distancia	34
4.3. Autocorrelación espacial	35
5. ALGUNAS HERRAMIENTAS PARA LA OBTENCIÓN DE DATOS EN GENÉTICA DEL PAISAJE: MARCADORES GENÉTICOS, TELEDETECCIÓN Y LENGUAJE R	39
5.1. Marcadores Genéticos	39
5.1.1. Marcadores moleculares ISSR	40
5.1.2. Marcadores moleculares SSR	42
5.2. Sensores remotos	42
5.2.1. El índice de vegetación de diferencia normalizada (NDVI)	45
5.2.2. El satélite Landsat	47
5.3. El lenguaje R	48
5.3.1. Propiedades de R como lenguaje orientado a objetos	49
II. OBJETIVOS E HIPÓTESIS	52
1. OBJETIVO GENERAL	52
2. OBJETIVOS ESPECÍFICOS	52
3. HIPÓTESIS	54

III. ESTUDIOS GENÉTICOS Y MORFOMÉTRICOS EN UN HUERTO EXPERIMENTAL CON CLONES DE *PROSOPIS ALBA*, SELECCIONADOS POR TOLERANCIA A LA SALINIDAD

1. RESUMEN.....	55
2. INTRODUCCIÓN.....	56
3. OBJETIVOS.....	58
4. MATERIALES Y MÉTODOS.....	58
4.1. Obtención de individuos tolerantes a la salinidad.....	58
4.2. Obtención de individuos sensibles a la salinidad.....	62
4.3. Métodos morfométricos.....	64
4.3.1. Colección de datos morfométricos.....	64
4.3.2. Análisis estadístico de los datos morfométricos.....	64
4.4. Métodos de análisis molecular.....	65
4.4.1. Obtención de datos moleculares.....	65
4.4.2. Análisis estadístico de los datos moleculares.....	65
A. Análisis de los datos de ISSR.....	65
A.1. Análisis de escalado multidimensional (MDS).....	66
A.2. Análisis entre grupos (BGA).....	66
A.3. Análisis de árboles de clasificación.....	66
B. Análisis de los datos de SSR.....	67
B.1. Escalado multidimensional y análisis entre grupos de los marcadores SSR.....	67
B.2. Análisis de coancestría.....	67
4.4.3. Consistencia entre los marcadores ISSR y SSR.....	67
5. RESULTADOS.....	68
5.1. Análisis de datos morfométricos.....	68
5.2. Análisis de ISSR.....	70
5.2.1. Análisis de escalado multidimensional.....	72
5.2.2. Análisis entre grupos.....	73
5.2.3. Análisis de árboles de clasificación.....	74
5.3. Análisis de datos SSR.....	74
5.3.1. Análisis de escalado multidimensional.....	75
5.3.2. Análisis entre grupos.....	76
5.3.3. Análisis de coancestría.....	76
5.4. Congruencia entre los conjuntos de datos ISSR y SSR.....	77
6. DISCUSIÓN.....	79

IV. DESARROLLO DE UN PAQUETE ESTADÍSTICO DE R PARA EL ANÁLISIS DE DATOS FENOTÍPICOS, GENOTÍPICOS Y AMBIENTALES EN BIOLOGÍA DE POBLACIONES

1. RESUMEN.....	83
2. INTRODUCCIÓN.....	84
3. OBJETIVOS.....	85
4. MÉTODOS.....	85
4.1. Estructura básica del paquete.....	85
4.2. Algoritmos de análisis espacial.....	87
A. Módulo espacial general.....	87
B. Módulo de lag espacial.....	88
4.3. Estructura de los objetos “ecogen”.....	90

5. APLICACIONES.....	91
----------------------	----

V. ESTUDIOS CON UNA POBLACIÓN NATURAL FRAGMENTADA DE *PROSOPIS ALBA*

1. DESCRIPCIÓN DEL ÁREA DE ESTUDIO.....	100
2. RECOLECCIÓN DE MUESTRAS.....	101

PRIMERA PARTE: ESTUDIO DE LA ESTRUCTURA GENÉTICA ESPACIAL A ESCALA FINA

1. RESUMEN.....	105
2. INTRODUCCIÓN.....	106
3. OBJETIVOS	109
4. MATERIALES Y MÉTODOS.....	109
4.1. Aislamiento de ADN y genotipado. Análisis de SSR.....	109
4.2. Análisis de datos.....	110
4.2.1. Estimación de la diversidad genética.....	110
4.2.2. Análisis de estructura genética.....	111
4.2.3. Análisis de coancestría.....	111
4.2.4. Análisis demográfico.....	113
5. RESULTADOS.....	114
5.1. Diversidad genética.....	114
5.2. Análisis de estructura genética.....	116
5.3. Análisis de coancestría.....	117
5.4. Análisis demográfico.....	117
6. DISCUSIÓN.....	121

SEGUNDA PARTE: ESTUDIO DE LA ESTRUCTURA ESPACIAL DE CARACTERES CUANTITATIVOS A ESCALA FINA.....

1. RESUMEN.....	125
2. INTRODUCCIÓN.....	126
2.1. Elementos conceptuales para el estudio a escala fina.....	126
2.2. Elementos para el estudio a escala fina en <i>Prosopis alba</i>	128
3. OBJETIVOS.....	129
4. MATERIALES Y MÉTODOS.....	130
4.1. Colección de datos morfométricos y de crecimiento.....	130
4.2. Análisis de NDVI.....	131
4.3. Asociación entre variables cuantitativas y ambientales.....	132
4.3.1. Obtención de datos.....	132
4.3.2. Análisis de árboles de inferencia condicional.....	133
4.3.3. Evaluación de la robustez del análisis con árboles de inferencia condicional ante la presencia de autocorrelación espacial en las variables cuantitativas.....	134
4.4. Análisis de superficies de tendencia.....	135
4.4.1. Modelado de los datos.....	135
4.4.2. Evaluación de la capacidad del análisis de superficies de tendencias para remover las tendencias globales conocidas de las variables cuantitativas.....	137
4.5. Análisis espacial a escala fina.....	137
4.5.1. Análisis espacial global.....	137
4.5.2. Análisis espacial local.....	139

5. RESULTADOS.....	140
5.1. Estadísticas básicas.....	140
5.2. Análisis de NDVI.....	141
5.3. Asociación entre los caracteres cuantitativos y las variables ambientales.....	142
5.3.1. Resultados del modelo.....	142
5.3.2. Evaluación de la robustez del análisis con árboles de inferencia condicional ante la presencia de autocorrelación espacial en las variables cuantitativas.....	144
5.4. Evaluación de la capacidad el análisis de superficies de tendencias para remover las tendencias globales conocidas de las variables cuantitativas.....	144
5.5. Análisis espacial a escala fina.....	145
5.5.1. Análisis espacial global.....	145
5.5.2. Análisis espacial local.....	146
6. DISCUSIÓN.....	148
VI. DISCUSIÓN GENERAL Y CONCLUSIONES.....	151
APÉNDICES.....	155
<i>APÉNDICE 1: PROTOCOLOS UTILIZADOS PARA LA OBTENCIÓN DE DATOS MORFOMÉTRICOS, MOLECULARES Y AMBIENTALES.....</i>	
Protocolo 1: obtención de datos de morfología foliar.....	155
Protocolo 2: obtención de datos moleculares.....	157
A. Análisis de ISSR.....	157
B. Análisis de SSR.....	158
Protocolo 3: Análisis de NDVI.....	159
<i>APÉNDICE 2: FÓRMULAS DE RELEVANCIA PARA EL DESARROLLO DE LA TESIS.....</i>	
1. Cálculo de la heterocigosis observada (H_o) y esperada (H_E).....	162
2. Indicadores de autocorrelación espacial.....	163
2.1. Indicadores globales.....	163
A. I de Moran (Moran, 1950).....	163
B. I_{xy} de Moran (Reich y col., 1994).....	163
2.2 Indicadores locales: G_i^* de Getis y Ord (Ord y Getis, 1995).....	164
<i>APÉNDICE 3: UN MÉTODO PARA LA SELECCIÓN DE PESOS ESPACIALES PARA G_i^*.....</i>	
1. Descripción del método.....	165
2. Aplicación del método al modelo de estudio.....	167
BIBLIOGRAFÍA.....	168

AGRADECIMIENTOS

Esta tesis ha sido posible gracias a numerosas personas e instituciones, a las que deseo agradecer:

A mi directora, la Dra. Beatriz Saidman, por permitirme ser parte de su equipo de trabajo; por el estímulo, confianza, observaciones y sugerencias en cada una de las instancias que cristalizaron en este texto.

A mi directora asistente, la Dra. Laura Ferreyra, quien ha sido una guía imprescindible; por todo lo que me ha enseñado y por su apoyo incondicional en todos estos años.

Al Dr. Juan Vilardi, por sus valiosos consejos a lo largo de esta investigación; en particular por los conocimientos transmitidos y el estímulo permanente al aprendizaje en materia de estadística y lenguaje R.

A mi madre, por el impulso emocional que me empujó siempre hacia adelante.

A mis familiares y amigos.

A mis compañeras de grupo y de espacio: Cecilia, Carolina, Paula, Andrea, Laura, Rosario y Rocío.

Al Ing. Mauricio Ewens, por brindarme los materiales de *Prosopis alba* estudiados en el capítulo III, y la asistencia para la colección de los ejemplares estudiados en el capítulo V.

A todos aquellos con los que compartí estos años de una u otra manera, que son muchos y la limitación de estas páginas no me permite nombrar, pero que están presentes.

Al Estado Nacional, por garantizar el acceso a la educación superior pública, gratuita y de calidad en la que me formé; a la Universidad de Buenos Aires y al CONICET, que me otorgaron las becas que posibilitaron la presente investigación; a la Facultad de Ciencias Exactas y Naturales, que

me permitió mediante la beca Manuel Sadosky concluir mis estudios de grado en una instancia previa, y al Departamento de Ecología, Genética y Evolución, donde desarrollé mis tareas.

Esta investigación fue financiada por la Universidad de Buenos Aires (UBACYT 20020130100043BA y UBACYT 20020100100008), la Agencia Nacional de Promoción Científica y Tecnológica (PICTO OTNA 00081 y PICT-2013 0478) y el CONICET (PIP 11220090100147).

Genética del paisaje en poblaciones de *Prosopis alba* de la provincia de Santiago del Estero

RESUMEN

La sección Algarobia del género *Prosopis* comprende especies arbóreas, la mayoría de las cuales constituyen importantes recursos naturales en las regiones áridas y semiáridas de América. Una de las especies más apreciada dentro de esta sección por la calidad de su madera y el sabor dulce de sus frutos es *Prosopis alba*. En la presente tesis se combinaron herramientas informáticas, moleculares y morfométricas para analizar la variación espacial a escala fina en esta especie. Se estudió en primer lugar un huerto experimental, continuando la investigación iniciada en la tesis de licenciatura de Roser (2010). Este capítulo aportó un mayor conocimiento general sobre la base genética de 11 rasgos de morfología foliar, determinando que la variación mostrada por 5 de ellos (largo, ancho y área del foliólulo, número de pares de foliólulos por pinna y número de pinnas) tiene un importante componente genético en la muestra analizada. Se estudió en segundo lugar una población natural de individuos adultos a escala espacial fina. Para poder realizar esta investigación, se elaboró un paquete estadístico en R denominado “EcoGenetics” (disponible en <https://cran.r-project.org/web/packages/EcoGenetics>), útil para el manejo y análisis de información espacial en poblaciones, con particular énfasis en enfoques a escala fina. El estudio de la población natural consistió en *i*) el análisis de la estructura genética espacial usando 7 loci microsatélites, evaluando el ajuste de los datos a un modelo de aislamiento por distancia y la presencia de señales de cambios demográficos recientes y *ii*) el análisis de la estructura espacial para los 11 caracteres de morfología foliar investigados en el huerto y el diámetro del tronco a la altura del pecho (DAP), examinando la presencia de patrones de autocorrelación espacial y la relación entre estos caracteres cuantitativos y variables ambientales. Los patrones moleculares mostraron la ausencia de discontinuidades genéticas en la población, ajuste a un modelo de aislamiento por distancia e indicios de una posible expansión o evento de inmigración en uno de los sitios de estudio. El análisis de los caracteres cuantitativos mostró estructuración espacial (autocorrelación positiva) para varios de ellos, lo que podría asociarse con flujo génico restringido y efectos de micrositio. En esta tesis se evaluaron de modo sistemático y por primera vez los patrones espaciales a escala fina en *Prosopis alba* mediante técnicas de autocorrelación espacial, resaltando la complejidad de la dinámica mostrada por poblaciones de distribución continua en paisajes modelados por el hombre.

Palabras clave: *Prosopis alba*, clones, genética del paisaje, estructura espacial, autocorrelación espacial, aislamiento por distancia, coancestría, morfometría, SSR, ISSR, lenguaje R, software libre.

Landscape genetics in populations of *Prosopis alba* of the province of Santiago del Estero

ABSTRACT

The section Algarobia of the genus *Prosopis* includes tree species that constitute important natural resources in arid and semiarid regions of America. One of the most appreciated species of this section for its wood quality and the sweet flavour of its pods is *Prosopis alba*. In the present thesis, informatics, molecular and morphometric tools were combined in order to analyze the patterns of spatial variation at fine scale in this species. In the first place, an experimental orchard was studied, continuing the research started in the thesis by Roser (2010). This chapter provided a deeper understanding of the genetic basis of 11 morphological leaf traits, determining that the variation showed by five of them (leaflet length, width and area, number of leaflets pairs per pinna and number of pinnae) has an important genetic component in the analyzed sample. In second place, a natural population of adults trees was studied at fine spatial scale. To carry out this research, a statistical R package called "EcoGenetics" (available in <https://cran.r-project.org/web/packages/EcoGenetics>), useful for managing and analyzing spatial information in populations, with particular emphasis on fine-scale approaches, was developed. The study of the natural population consisted in *i*) the analysis of the spatial genetic structure using 7 microsatellite loci in order to determine whether the data fit the isolation by distance model and to detect signals of recent demographical changes, and *ii*) the analysis of the spatial structure in the 11 leaf traits investigated in the orchard and the trunk diameter at breast height (DAP), assessing the presence of spatial autocorrelation patterns, and the relationship between these quantitative traits and environmental variables. The molecular patterns showed the absence of genetic discontinuities in the population. The microsatellite data fit an isolation by distance model, and signals of a possible expansion or immigration event were observed in one of the study sites. The analysis of the quantitative traits showed spatial structure (positive autocorrelation) for several of them, which could be associated with restricted gene flow and microsite effects. In this thesis, the fine-scale spatial patterns in *Prosopis alba* were evaluated systematically and for the first time using spatial autocorrelation techniques, highlighting the complexity of the dynamics showed by continuously distributed populations in anthropized landscapes.

Keywords: *Prosopis alba*, clones, landscape genetics, spatial autocorrelation, spatial structure, isolation by distance, kinship, morphometrics, SSR, ISSR, R programming language, open source software.

Introducción



I. INTRODUCCIÓN

1. INTRODUCCIÓN AL GÉNERO *PROSOPIS*

1.1. El género *Prosopis*

El género *Prosopis* Linnaeus emend. Burkart (Fabaceae, Mimosoidae) comprende alrededor de 44 especies leñosas, en su mayoría formas arbóreas y arbustivas, ampliamente distribuidas en ecosistemas áridos y semiáridos del planeta (Figura 1.1). Como género perteneciente a la familia Fabaceae, se caracteriza por producir semillas en legumbres; como perteneciente a la subfamilia Mimosoideae, por poseer hojas pinnadas compuestas con pequeñas flores agrupadas en inflorescencias. *Prosopis* se encuentra presente principalmente en América, ya que del total de especies que lo integran, sólo tres son nativas de Asia y una de África (Burkart, 1976). De las 40 restantes, 34 son nativas de Sudamérica (Solbrig, 1977). En Argentina se hallan 27 especies y 13 de ellas son endémicas (Burkart, 1976). La distribución geográfica se da a lo largo y ancho del territorio nacional, exceptuando al sur de la Patagonia y la provincia de Misiones (Burkart, 1976; Hunziker y col., 1986). La mayor diversidad específica está concentrada en las Provincias Biogeográficas del Monte, Espinal y Chaco (Cabrera y Willink, 1973).



FIGURA 1.1 Distribución del género *Prosopis*. Los números indican la cantidad de especies nativas en cada región. Imagen extraída de Catalano y col. (2008).

La historia taxonómica del género es compleja y extensa. En 1976 Arturo Burkart lo dividió en cinco secciones, basándose principalmente en el tipo de espinas y la estructura del tallo: Algarobia (30 especies americanas), Strombocarpa (9 especies americanas), Monilicarpa (1 especie endémica de Argentina), Anonychium (1 especie africana) y Prosopis (3 especies afro-asiáticas); a su vez, las primeras dos fueron subdivididas en series. La sección Algarobia, ampliamente representada en Argentina, está compuesta por las series Ruscifoliae, Sericanthae, Denudantes, Humiles, Pallidae y Chilenses. Es estrictamente americana, con especies presentes tanto en el sur como en el norte del continente, que se caracterizan por poseer espinas axilares y ramas diferenciadas en braquiblastos y macroblastos, aunque algunas de ellas carecen de espinas (Catalano y col., 2008). Dentro de Algarobia la distinción entre especies es confusa, ya que en zonas de simpatria los individuos clasificados dentro de diferentes grupos son capaces de cruzarse entre sí, dejando descendencia fértil (Hunziker y col., 1986). Las especies de esta sección son en su mayoría autoincompatibles y de fecundación cruzada (Solbrig y Cantino, 1975; Simpson, 1977); en algunas de ellas sin embargo se ha observado un pequeño porcentaje de autofecundación (Bessegga y col., 2000). Diversos estudios (Bessegga, 2005, 2006; Catalano y col., 2008) sugieren que el género *Prosopis* y la sección Algarobia no serían grupos monofiléticos.

Desde el punto de vista ecológico, las especies de *Prosopis* cumplen importantes funciones, que incluyen la fijación de nitrógeno, el control de la erosión del suelo y la estabilización de médanos. Las especies de zonas áridas y semiáridas poseen adaptaciones que les permiten soportar las condiciones ambientales extremas en las que habitan. De hecho, algunas de ellas son capaces de crecer sin problemas dentro de un amplio intervalo de temperaturas que va desde -20° C a 44° C (Medina y Cardemil, 1993). También logran crecer en suelos arenosos y/o salinos, así como en zonas donde la precipitación pluvial es muy escasa (250 mm a 70 mm). Las especies con sistema radicular dimórfico poseen una raíz leñosa vertical de crecimiento rápido y una corona de raíces laterales, que les permiten utilizar agua en niveles superficiales y profundos (Galera y col., 2000; Villagra y col., 2011).

Las especies representantes del género con valor económico (por ejemplo, *P. alba*, *P. flexuosa*, *P. tamarugo*, *P. juliflora*, *P. chilensis*, *P. pallida* y *P. glandulosa*, entre otras), han despertado gran interés en los últimos años, ya que son recursos promisorios para forestación (Dutton, 1989; Geesing y col., 2000), debido a la alta tolerancia a la salinidad que presentan y a los productos que

se pueden obtener de ellas (Pasiecznik y col., 2001; Felker y col., 2008). Exhiben una excelente adaptación a sistemas agroforestales y sus usos son de los más diversos. Presentan una madera de buena calidad, dura, densa y durable, la cual es empleada en la construcción de pilares, soportes y aberturas, así como en la fabricación de muebles; también es utilizada como combustible, leña o en la elaboración de carbón vegetal (Hunziker y col., 1986; Pasiecznik y col., 2001). Las vainas, ricas en hidratos de carbono, azúcares y proteínas, son una importante fuente de alimento tanto para el hombre¹, como para la fauna silvestre y el ganado; en este último caso, los frutos son recolectados y secados, utilizándolos como forraje en invierno (Galera y col., 1992).

1.2. Estudios poblacionales previos en *Prosopis*

El género *Prosopis* ha sido objeto de gran número de estudios citogenéticos, genético-poblacionales y morfológicos. Los primeros estudios citogenéticos mostraron que sus especies tendrían un número cromosómico somático $2n = 28$. *P. juliflora* sería la única especie descrita como tetraploide ($2n = 56$) hasta el momento (Hunziker y col., 1975; Burkart, 1976; Bukhari, 1997; Saidman y col., 1997; Harris y col., 2003; Trenchard y col., 2008). Los estudios genético-poblacionales utilizaron múltiples técnicas que en general mostraron resultados concordantes. Estos incluyeron análisis basados en cromatografía (Naranjo y col., 1984; Palacios y Bravo, 1984; Zallocki y col., 1990), proteínas seminales (Burghardt, 1996; Burghardt y Palacios, 1997), isoenzimas (Solbrig y Bawa, 1975; Saidman, 1985; Saidman y Vilardi, 1987; Verga, 1995; Bessega y col., 2000; Julio, 2000), polimorfismos para el largo de fragmentos de restricción (RFLP) (Saidman y col., 1998; Ferreyra y col., 2004), polimorfismos de loci amplificados al azar (RADP) (Juárez-Muñoz y col., 2002; Ferreyra y col., 2004, 2007, 2010), microsatélites (SSR) (Bessega y col., 2012; Teich y col., 2014; Bessega y col. 2015b), y caracteres cuantitativos (Bessega y col., 2009, 2011; Darquier y col., 2013; Ferreyra y col., 2013; Bessega y col., 2015b). 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 llegó a través del uso de técnicas bioquímicas y moleculares es que las especies de estas series presentan una gran similitud genética. Por otra parte, se observó que la mayor parte de la diversidad genética se encuentra dentro de las poblaciones y que la divergencia entre poblaciones de una misma especie es baja. A pesar de la alta similitud interespecífica, la simpatria de las especies y la presencia de

¹ Por ejemplo, mediante la elaboración de harina para la preparación de “patay” o torta de algarrobo y bebidas como la “chicha”, la “añapa” y el “arope”.

hibridación (Hunziker y col., 1975, 1977; Palacios y Bravo, 1981; Hunziker y col., 1986), las poblaciones de una misma especie se agrupan generalmente de acuerdo a su origen específico. Los estudios con caracteres cuantitativos mostraron alta heredabilidad en rasgos de crecimiento y foliares, variación no neutral en el tamaño y forma de la hoja, importante variación geográfica, asociación con variables ambientales y concordancia con las descripciones dadas por marcadores moleculares.

Estudios realizados en siete especies de la sección Algarobia (Bessega y col., 2000) mostraron una preponderancia de reproducción exogámica. El polen de *Prosopis* es dispersado por insectos (Genise y col. 1990) y las semillas de modo endozoico por mamíferos (Reynolds, 1954; Mares y col., 1977), mecanismos que se consideran asociados con dispersión limitada. Las semillas duras germinan luego de atravesar el tracto digestivo de los mamíferos que ingieren los frutos, siendo depositadas junto a sus heces, lo que provoca que árboles altamente emparentados crezcan juntos en áreas limitadas (Steibel y col., 1999; Bessega y col., 2012).

1.3. *Prosopis alba* (sección Algarobia)

Prosopis alba Griseb. (“algarrobo blanco”, Figura 1.2) es una importante especie nativa que compone el bosque Chaqueño, distribuida en Argentina, Uruguay, Paraguay, sur de Bolivia y Perú. Es un árbol de porte mediano, de hasta 12 m de altura y con un tronco de hasta 150 cm de diámetro (Tortorelli, 2009).

Los árboles de *P. alba* se caracterizan por su gran adaptación a las regiones semiáridas del NE argentino. El algarrobo blanco puede ayudar a mejorar las condiciones del suelo por medio de la prevención de la erosión y la formación de agregados de materia orgánica. Diversos estudios indicaron además que esta especie presenta una gran tolerancia a la salinidad (Rhodes y Felker, 1988; Velarde y col., 2003; Felker y col., 2008), lo que permitiría utilizarla en planes de forestación para la recuperación de suelos salinos. Desde el punto de vista económico es una especie multipropósito, pudiéndose obtener a partir de ella productos tales como leña, carbón, forraje, alimentos para consumo humano y una madera de muy alta calidad (Pometti y col., 2009, 2010; Pizzo y col., 2011). Actualmente, el algarrobo blanco es la principal especie nativa promocionada para realizar plantaciones comerciales en la región del Parque Chaqueño, en el marco de la Ley

Nacional 25.080/99 (y su reforma, la Ley 26.432/08), que incentiva y subsidia la forestación con especies nativas y exóticas en todo el territorio nacional (Coronel de Renolfi y Gomez, 2014).

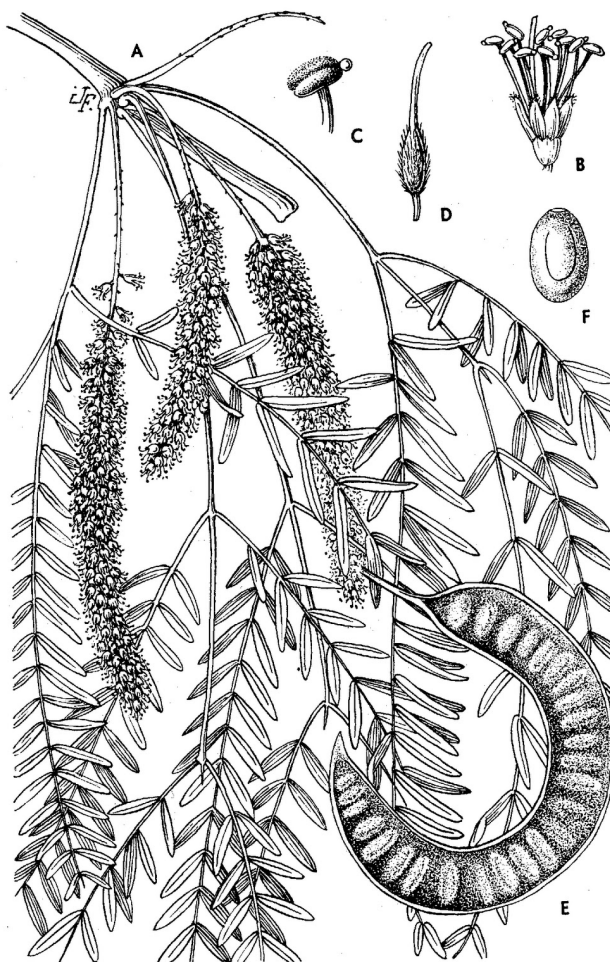


FIGURA 1.2 *Prosopis alba* (A) rama florífera; (B) flor; (C) detalle del estambre; (D) gineceo; (E) fruto; (F) semilla. Imagen extraída de Rapoport, y col. (1998).

Los montes naturales de *P. alba* se han visto reducidos de manera drástica a causa de la sobreexplotación y el avance de la frontera agrícola. El panorama actual de la distribución de esta especie combina bosques relativamente prístinos, reducidos a algunas zonas de Chaco y Formosa; parches remanentes de la amplia distribución original, especialmente en Santiago del Estero, usualmente rodeados de áreas con uso productivo (agroganadero o forestal); y zonas de colonización relativamente reciente, como consecuencia de modificaciones antrópicas asociadas con cambios en los cursos de agua por la construcción de diques y aparición de hábitats abiertos luego del desmonte. En Santiago del Estero, las poblaciones naturales de algarrobo blanco están prácticamente reducidas a pequeñas áreas protegidas entre ríos, a los costados de los caminos o son

mantenidas en pequeños parches en fincas donde proveen alimento y sombra al ganado.

1.4. Estudios poblacionales previos en *P. alba*

La información disponible sobre la estructura poblacional de *P. alba* (Ferreyra y col., 2004, 2007; Bessega y col., 2012; Ferreyra y col., 2013) indica que la mayor parte de la variabilidad genética se encuentra dentro de las poblaciones, con una limitada diferenciación regional. La selección natural podría ser un factor importante en la generación de patrones morfológicos a escalas gruesas (Bessega y col., 2015a). A partir de ensayos, se pudo estimar una alta heredabilidad para rasgos relacionados con la forma y el tamaño de hojas y espinas, y con rasgo de crecimiento (altura y diámetro del tronco), lo que sugiere la posibilidad de obtener una respuesta significativa a la selección de caracteres de importancia adaptativa y económica (Bessega y col. 2009, 2011; Ewens y col., 2012).

Estudios sobre el sistema de fecundación (Bessega y col., 2000) y el análisis de la estructura fina a nivel familiar (Bessega y col., 2012), mostraron que esta especie es principalmente de fecundación cruzada, que la dispersión del polen sería limitada (entre 5 y 30 m) y que el parentesco entre árboles sería nulo a distancias mayores de 100 m. Al momento de esta revisión bibliográfica, la última publicación realizada con *P. alba* (Bessega y col., 2015b) mostró que el uso de estimas de parentesco mejoraría el proceso de selección de semillas para formar plantaciones.

2. INTRODUCCIÓN AL ÁREA DE ESTUDIO: DESCRIPCIÓN AMBIENTAL DE LA PROVINCIA DE SANTIAGO DEL ESTERO

2.1. Descripción general de la provincia

La provincia de Santiago del Estero se encuentra situada casi en su totalidad en la región fitogeográfica Chaqueña (Figura 1.3) (Cabrera, 1976), una vasta planicie que posee características semiáridas en el área central, y subhúmedas hacia el este (Pérez Carrera y col., 2008).

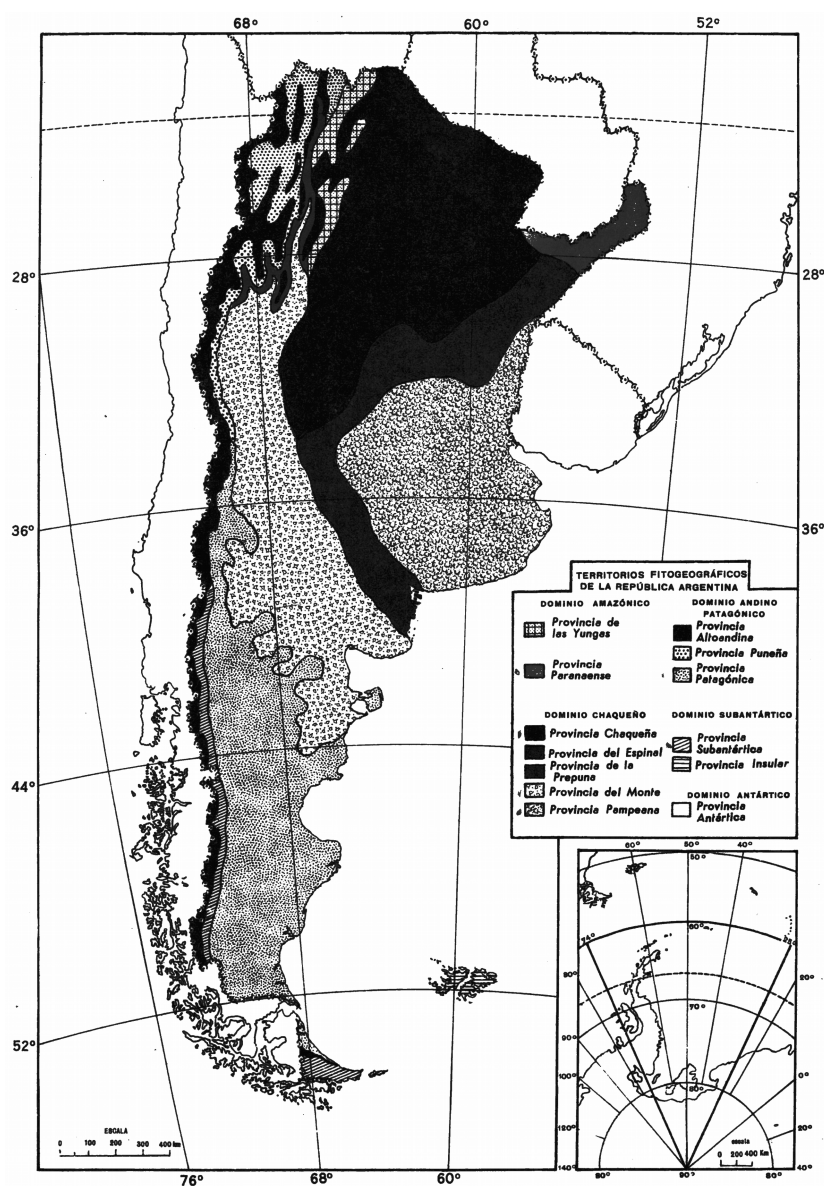


FIGURA 1.3 Regiones fitogeográficas de la Argentina según Cabrera (1971). Imagen extraída del trabajo original.

I. Introducción

La vegetación de Santiago del Estero consiste en bosques xerófitos, estepas halófitas y sabanas edáficas o inducidas por fuego o deforestación (Cabrera, 1976; SAyDS, 2004). El área boscosa, compuesta de bosques naturales y zonas forestadas para producción, ocupa un 65% de la provincia. Las especies más explotadas son el quebracho colorado (*Schinopsis lorentzii*) y el quebracho blanco (*Aspidosperma quebracho blanco*). Además, se explotan especies como el algarrobo (*Prosopis* sp.), el mistol (*Ziziphus mistol*) y el chañar (*Geoffroea decorticans*) (Pérez-Carrera y col., 2008).

La agricultura de la región se realiza en condiciones de secano y bajo riego. Las zonas de agricultura de secano se encuentran comprendidas en el sudeste de la provincia (límite con Santa Fe). Las zonas de agricultura de riego abarcan áreas relacionadas con los ríos Salado, Horcones y Urueña, una subzona de riego con aguas surgentes, y el sistema del río Dulce. La base productiva agrícola está compuesta por soja, sorgo granífero, girasol, maíz, cultivos forrajeros y frutihortícolas (Pérez-Carrera y col., 2008).

Santiago del Estero es la provincia con el mayor desarrollo de la ganadería en el NOA, contando con un 50% del stock ganadero regional. En zonas de secano, se realiza explotación de cría y recría sobre campos naturales con monte, llegando a haber hasta un 10% de zonas con pasturas cultivadas y rastrojos de cultivo, en las que también se realiza engorde. En zonas de riego, la actividad ganadera que se realiza es la de engorde. La ganadería caprina se distribuye en forma uniforme y dispersa en toda la provincia (Pérez-Carrera y col., 2008).

La marcada degradación del suelo es consecuencia de la intensa utilización agrícola, el pastoreo excesivo, y por sobre todo, la deforestación: la explotación irracional de los bosques naturales ha destruido millones de hectáreas desde comienzos del Siglo XX. El proceso de expansión de la frontera agrícola "ha generado un fuerte impacto en el área, removiendo cubierta boscosa de grandes superficies y afectando a la población rural" (Aristide, 2010).

2.2. Clima

El clima de Santiago del Estero es caluroso y equivalente al de las regiones subtropicales. La temperatura anual promedio es de 21°C (Figura 1.4A), con valores máximos que exceden los 45 °C

en verano y mínimos entre 4.5 y 5 °C en invierno (Pérez Carrera y col., 2008).

Las precipitaciones presentan una distribución marcadamente estacional, distinguiéndose una temporada húmeda, de octubre a marzo, y una seca, de abril a septiembre. La precipitación varía desde 550 a 950 mm, en sentido sudeste-noroeste (Figura 1.4B). Un 85% de la precipitación anual se concentra en el período estival, con 30 días de lluvia y con una media de 380 mm (Moscuza y col., 2003). La evapotranspiración varía entre 1250 y 1550 mm (Figura 1.4C), lo que determina un balance hídrico con déficit a lo largo de todo el año.

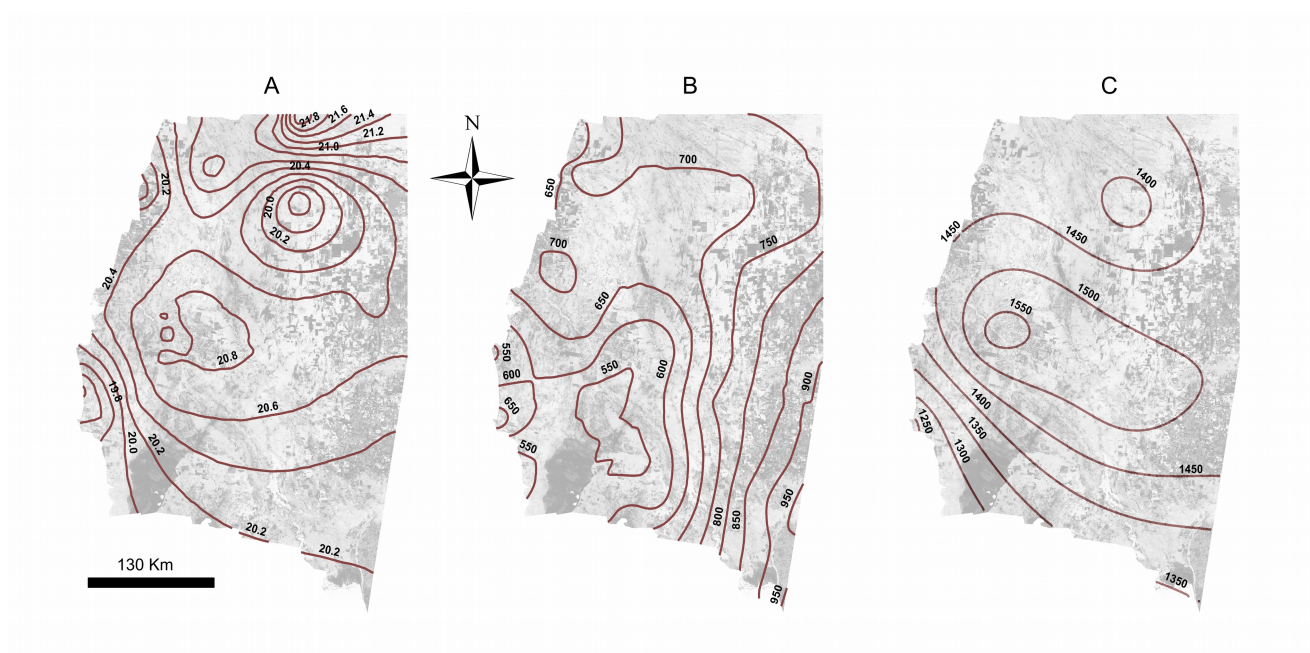


FIGURA 1.4 Mapas de variables climáticas de Santiago del Estero. (A) Mapa de temperatura media anual (°C); (B) Mapa de precipitación media anual (mm); (C) Mapa de evapotranspiración potencial (mm). Adaptado de Angueira y col. (2007).

2.3. Geomorfología

La planicie del Chaco, de la que forma parte la provincia, tiene un carácter prácticamente llano, con una pendiente media de oeste a este de 0.04% (Matteucci y Camino, 2012). Es una gran planicie sedimentaria descrita como un bloque hundido relleno con los materiales de los ríos que la atraviesan en sentido noroeste-sudeste (Aristide, 2010). El territorio de Santiago del Estero se encuentra interrumpido por elevaciones en las que se registran las mayores altitudes provinciales (Figura 1.5) y por dos ríos principales y ríos menores. Cinco ríos lo cruzan: Dulce, Salado,

Horcones, Urueña y Albigasta (Figura 1.6), siendo los principales, de mayor caudal y transporte permanente de agua, los ríos Salado y Dulce.

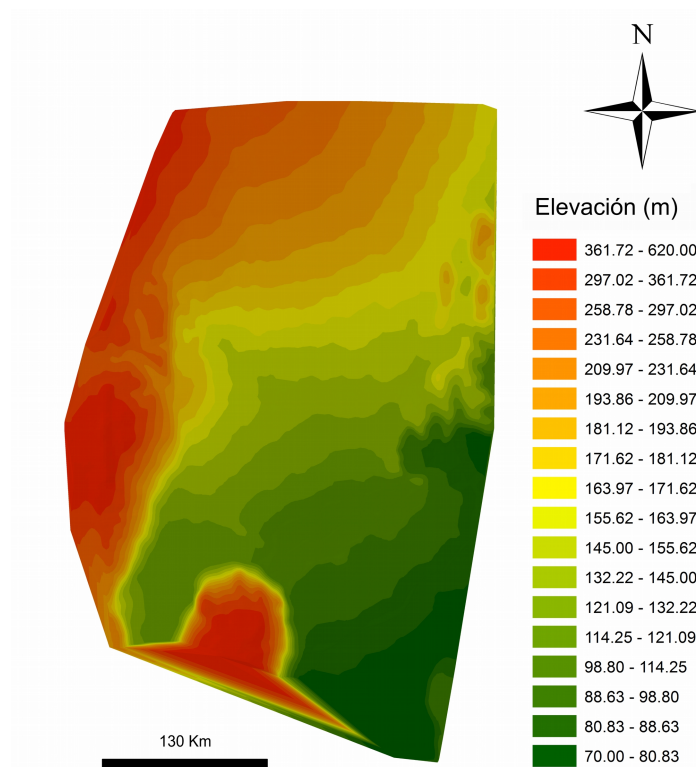


FIGURA 1.5 Mapa de elevación de Santiago del Estero. Superficie construida con ArcGIS 10.1 a partir de curvas de nivel disponibles en Angueira y col., 2007.

El río Salado nace en Salta-Catamarca, en las sierras occidentales del borde de la Puna y desemboca en Santa Fe. Este río se seca casi todos los inviernos. El río Dulce, el más importante desde el punto de vista socioeconómico, pasa en su recorrido por Salta, Tucumán y Santiago del Estero, desembocando en la provincia de Córdoba. En Santiago del Estero inunda el Dique Frontal de Río Hondo y se transforma en la línea divisoria de los departamentos Capital y La Banda. En este recorrido, se encuentra el dique Los Quiroga, base del Sistema de Riego del Área del Río Dulce, que provee de agua a 110 mil hectáreas en los departamentos Capital, La Banda y Robles.

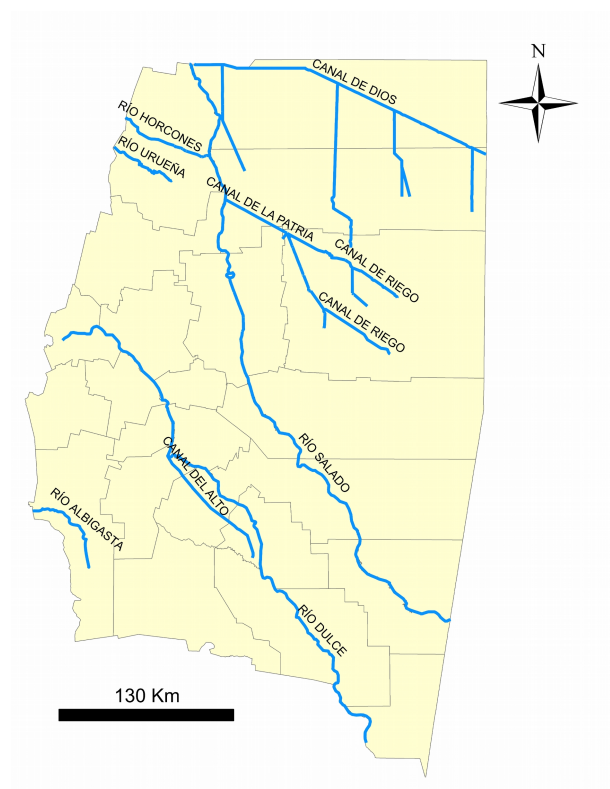


FIGURA 1.6 Mapa hidrográfico de Santiago del Estero, mostrando los ríos y canales principales. Adaptado de Angueira y col. (2007).

La acción de la orogenia y el clima, dos factores modeladores fundamentales del relieve, ha desarrollado una amplia llanura aluvial sobre el río Dulce (Figura 1.7). La importante dinámica hídrica que afectó al lugar trajo como consecuencia una manifiesta heterogeneidad de sedimentos depositados y suelos asociados. Los sitios más elevados están formados por suelos de texturas medias a gruesas. Los antiguos cauces del río, actualmente secos y desactivados, poseen suelos de texturas arenosas y presentan alveolos bien definidos con albardones de alturas variables. En las depresiones, el tamaño de las partículas minerales de los suelos es de menor tamaño y manifiestan deficiencias de drenaje con significativa afectación salina (UESRRD, 2015).

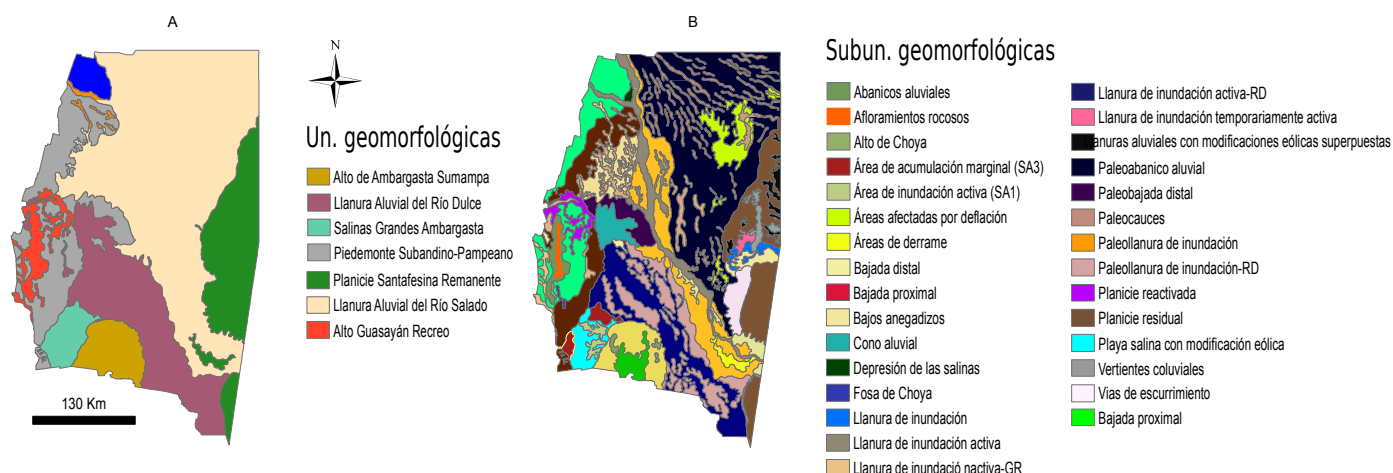


FIGURA 1.7 Mapas de geomorfología de Santiago del Estero. A. Mapa de unidades geomorfológicas; B. Mapa de subunidades geomorfológicas. Adaptado de Angueira y col. (2007).

2.4. Suelos

Los suelos son el resultado de la interacción del material originario, el clima y la biósfera, modificados por la topografía y el tiempo. En su mayoría, los suelos de Santiago del Estero se han desarrollado sobre materiales fluvio-eólicos en un relieve casi llano sin formas sobresalientes significativas, exceptuando las Sierras Pampeanas (Angueira y col., 2007). Los suelos de la parte oriental de la planicie santiagueña, originados a partir de materiales loésicos y aluviales, presentan moderado desarrollo, con horizontes orgánicos relativamente espesos. En el sudoeste de la provincia, zona de salinas, los suelos no permiten el crecimiento y el desarrollo de ningún tipo de vegetación. Hacia el oeste, con la disminución de las precipitaciones, los suelos presentan menor desarrollo, con muy bajos tenores de materia orgánica. El suelo de aspecto rojizo cubre casi la totalidad de Santiago del Estero. Dentro de los sistemas fluviales de los ríos Salado y Dulce pueden observarse una importante variedad de suelos.

El sistema Soil Taxonomy (Soil Survey Staff, 1999), que es el sistema de referencia a nivel mundial para la clasificación de suelos, establece 6 órdenes jerárquicos (Orden, Suborden, Gran grupo, Subgrupo, Familia y Serie, Tabla 1.5). La nomenclatura utilizada para nombrar los suelos es informativa, en el sentido que se van incorporando una serie de prefijos y sufijos que refieren a propiedades de los mismos, formándose de este modo grupos taxonómicos.

I. Introducción

TABLA 1.5 Características principales de las taxones de suelos establecidos en el sistema Soil Taxonomy (fuente: SMIN, <http://www.mineria.gob.ar/estudios/irn/cordoba/X-6suec.asp>).

Categoría	Naturaleza de las propiedades diferenciadoras
Orden	Procesos de formación de suelos, indicados por la presencia o ausencia de importantes horizontes diagnósticos
Suborden	Subdivisión de los órdenes en función de una homogeneidad genética, definida por propiedades tales como régimen de humedad, materiales parentales y estado de descomposición de fibras orgánicas
Gran Grupo	Subdivisión de los subórdenes de acuerdo a la clase, ordenamiento y grado de expresión de los horizontes, presencia o ausencia de determinados rasgos diagnósticos o clima edáfico
Subgrupo	Concepto central del taxón perteneciente al gran grupo y propiedades indicando intergradación a otros grandes grupos, subórdenes y órdenes o extragradación a "no suelo"
Familia	Propiedades importantes para el crecimiento de las plantas que se relacionan con amplias clases texturales y de temperatura del suelo y mineralogía predominante del solum
Serie	Tipo y ordenamiento de horizontes, color, textura, estructura, consistencia y reacción de los horizontes; propiedades químicas y mineralógicas de los horizontes

Los órdenes de suelos que pueden encontrarse en Argentina, se muestran en la Figura 1.8. En Santiago del Estero en particular, se hallan presentes los órdenes **Molisol**, **Aridisol**, **Entisol** y **Alfisol** (Angueira y col., 2007). Las características principales de estos órdenes se muestran en la Tabla 1.6.

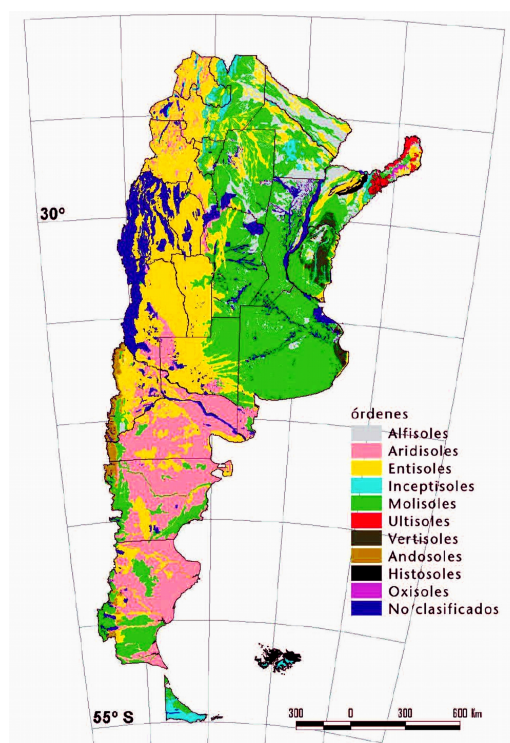


FIGURA 1.8 Órdenes de suelos presentes en Argentina, de acuerdo al sistema Soil Taxonomy (fuente: INTA).

I. Introducción

TABLA 1.6 Tipo de suelos que pueden encontrarse en Santiago del Estero (fuente: SMIN, <http://www.mineria.gob.ar/estudios/irn/cordoba/X-6suec.asp>).

Denominación	Características principales
Molisol	Suelos ricos en materia orgánica, horizonte superficial bien estructurado, ricos en bases
Aridisol	Falta de agua por aridez del clima
Entisol	Suelos jóvenes o escasamente desarrollados
Alfisol	Suelos con iluviación de arcillas, riqueza en bases de media a alta

Los **molisoles** (Figura 1.9) de las planicies chaco-pampeanas constituyen los suelos de tipo dominante entre aquellos con la mejor aptitud para agricultura (Moscatelli, 1990).



FIGURA 1.9 Perfil de un Molisol. Estos son suelos con un horizonte superficial oscuro y ricos en bases (fuente: Soil Survey Staff, <http://www.nrcs.usda.gov>).

Los suelos de este orden **se caracterizan por poseer un horizonte A enriquecido en materia orgánica**, debido a la incorporación sistemática de los residuos vegetales y su mezcla con la parte mineral, produciéndose un oscurecimiento que se refleja más profundamente en la parte superficial. Poseen una alta saturación de bases y de moderada a alta capacidad de intercambio catiónico. Son suelos más o menos desarrollados, húmedos, franco-limosos a arcillolimosos, con moderada a fuerte estructuración, bien a moderadamente drenados, sin problemas de encharcamiento, con cierta acumulación de sales o carbonatos (Angueira y col., 2007). Se disponen en las áreas que reciben mayor aporte hídrico por precipitaciones o escurrimiento superficial. Se pueden encontrar en la bajada proximal del Piedemonte Subandino-Pampeano, la Llanura Aluvial del Río Salado en su Paleobanico Aluvial, la Llanura de Inundación Activa en la Depresión de las Salinas, y en la mayor parte de la Planicie Santafecina Remanente (Figura 1.7). El **suborden Ustol**

puede presentar períodos sub-húmedos con sequías que pueden ser prolongadas durante algunos períodos. Dentro de este a su vez, de particular interés en la presente tesis es el **gran grupo Haplustol**. Representantes de este tipo de suelos se encuentran ubicados en relieves positivos, en las Bajadas del Piedemonte Subandino Pampeano, en la Planicie Santaferina Remanente y en el Paleoabánico Aluvial del Río Salado (Figura 1.7). Son suelos con escaso desarrollo, moderadamente profundos, presentan un horizonte con poca meteorización, con un incremento de arcilla sin llegar a ser una capa arcillosa, bien drenados y sin problemas de encharcamiento, con cierta acumulación de sales o carbonatos y perfil franco-arcillo-limoso (Angueira y col., 2007).

Los **aridisoles** (Figura 1.10) son suelos secos por períodos prolongados del año y/o salinos con napa freática alta, que **no disponen de agua suficiente para el crecimiento de cultivos o pasturas polifíticas durante largos períodos** (Moscatelli, 1990).



FIGURA 1.10 Perfil de un Aridisol. Estos se caracterizan por ser suelos muy secos para el crecimiento de plantas mesófitas (fuente: Soil Survey Staff, <http://www.nrcs.usda.gov>).

Presentan bajos contenidos de materia orgánica, color claro, textura franca y débil estructuración. Cuando no se encuentran **bajo riego**, los aridisoles se asocian con una vegetación xerófila, en general escasa y que no cubre completamente la superficie del suelo. Se encuentran en amplias zonas de la provincia: en el Piedemonte Pampeano, la Llanura Aluvial del Río Salado, el Paleoabánico Aluvial, la Depresión de las Salinas y la Llanura Aluvial del Río Dulce en sus componentes salinos (Angueira y col., 2007) (Figura 1.7).

Los **entisoles** (Figura 1.11) son suelos ampliamente distribuidos en zonas áridas y semiáridas.

Son suelos **recientes, poco evolucionados, con poca o ninguna evidencia de desarrollo de horizontes pedogenéticos, siendo en su mayoría el material parental inalterado**. Este escaso desarrollo es resultado del material originario inerte, pendientes pronunciadas activamente erosionadas, escasas precipitaciones o falta de tiempo para su maduración (Angueira y col., 2007).



FIGURA 1.11 Perfil de un Entisol. Estos se caracterizan por ser suelos con poca o ninguna evidencia de desarrollo de horizontes podogenéticos (fuente: Soil Survey Staff, <http://www.nrcs.usda.gov>).

Este tipo de suelo se da mayormente a lo largo de los ríos principales y en zonas bajo irrigación (Moscatelli, 1990), apareciendo en relieves recientemente erosionados. Pueden encontrarse en la Llanura Aluvial del Río Dulce, en la Planicie de Inundación de los Ríos Salado y Dulce, en el Alto Ambargasta-Sumampa y la Elevación Guasayán (Figura 1.7). Poseen materia orgánica distribuida regularmente, un contacto lítico a poca profundidad y textura franca o más fina. Presentan una buena aptitud agrícola en caso de no estar afectados por salinización o sodificación. De particular interés en esta tesis es el tipo de suelos del **suborden Fluventes**. Estos suelos se localizan en los pantanos con drenaje pobre de las Planicies de Inundación de los Ríos Salado y Dulce (Figura 1.7). Son **suelos inundados frecuentemente, pero que no permanecen saturados todo el año**. Poseen textura arenosa-franca o más fina. El material es estratificado. Los sedimentos aluviales son ricos en carbono orgánico, generalmente asociados a la fracción arcilla. El porcentaje de materia orgánica varía irregularmente con la profundidad, dependiendo de la fracción granulométrica presente. En la zona se encuentran los **grandes grupos Torripsamente**, con régimen de humedad **tórrido** (árido) y **Ustifluvente**, con régimen de humedad **ústico**, donde el período de lluvias coincide con la estación cálida, con una importante pérdida del agua por evapotranspiración, común de regiones semiáridas (Angueira y col., 2007).

Los **alfisoles** (Figura 1.12) se caracterizan por presentar un **horizonte subsuperficial de enriquecimiento secundario de arcillas** desarrollado en condiciones de acidez o de alcalinidad sódica, y **asociado con un horizonte superficial claro**, generalmente pobre en materia orgánica o de poco espesor.



FIGURA 1.12 Perfil de un Alfisol. Estos se caracterizan por poseer un horizonte subsuperficial con abundante arcilla y saturación de bases de 35% o mayor (fuente: Soil Survey Staff, <http://www.nrcs.usda.gov>).

Los suelos que pertenecen a este orden presentan una alta saturación con bases en todo el perfil. Se suelen utilizar para el pastoreo del ganado, para tierras de pastura natural o para pasturas tolerantes al exceso de agua y al sodio (Moscatelli, 1990). Son suelos de escasa extensión en la provincia, presentes en las Llanuras de Inundación Activa de los Ríos Salado y Dulce, en las Áreas de Derrame del Río Salado y en la Depresión de Mar Chiquita (Angueira y col., 2007) (Figura 1.7). De particular interés en esta tesis es el tipo de suelos del **suborden Acualfes**. Estos suelos se caracterizan por estar **estacionalmente saturados con agua por períodos prolongados** (régimen de humedad ácuico) y asociados generalmente con una capa de agua freática fluctuante y cercana a la superficie. A causa de esto último se generan condiciones reductoras y de falta de oxigenación en el suelo. Normalmente los acualfes se ubican en áreas planas o cóncavas expuestas a encharcamiento durante las épocas de lluvias dado el escaso escurrimiento superficial y al aporte de aguas de escorrentía de sectores vecinos más elevados. Dentro de los acualfes, los **natracualfes** presentan un horizonte enriquecido en sales y en sodio de intercambio, procedentes de capas de agua salinas.

3. INTRODUCCIÓN AL ANÁLISIS GENÉTICO-POBLACIONAL: COANCESTRÍA Y ENDOGAMIA - LOS ÍNDICES F y ϕ - AMOVA

3.1. Definiciones teóricas elementales

- Se denomina **panmixia** al apareamiento aleatorio de los individuos que componen una población.
- En una población panmíctica, suficientemente grande y no sometida a migración, mutación, deriva genética o selección, las frecuencias génicas y genotípicas se mantienen constantes de generación en generación. Cuando se cumplen estas condiciones, se dice que la población está en **equilibrio de Hardy-Weinberg** (EHW). Si el equilibrio en una población es alterado, pero volvieran a restablecerse las condiciones de EHW , el mismo se alcanzaría nuevamente en la siguiente generación (tras un evento de apareamiento aleatorio).
- Dos alelos que descienden de un mismo alelo ancestral, se dicen idénticos por descendencia (**IBD**). Dos individuos **emparentados** (*related* en inglés) poseerán alelos **IBD** derivados de un antepasado en común (Weir y col., 2006).
- La probabilidad de que dos alelos tomados al azar, cada uno de individuos distintos X e Y, sean **IBD**, es el **coeficiente de coancestría** (*coancestry* o *kinship*) entre X e Y (θ_{XY}).
- Sea Z un descendiente de los individuos X e Y. La probabilidad de que dos alelos en un locus de Z sean **IBD**, se denomina **coeficiente de endogamia** (simbolizado con la letra **F**). El coeficiente de endogamia (F_Z) de un descendiente Z, es igual a la coancestría (θ_{XY}) (*coancestry* o *kinship*) entre sus padres X e Y (Figura 1.13).
- Cuando dos alelos de un locus son iguales, independientemente de si descienden de un alelo ancestral, se dice que son idénticos en estado (**IBS**). Dos alelos pueden ser **IBS**, pero eso no implica que sean necesariamente **IBD**, por ejemplo, la misma base en dos SNP o el mismo número de repeticiones en microsatélites (Weir y col., 2006), cuando derivan de eventos mutacionales distintos. En cambio alelos **IBD** son siempre **IBS**.

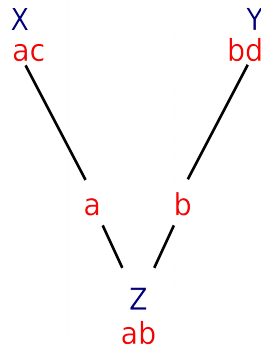


FIGURA 1.13 El coeficiente de coancestría entre X y Y es la probabilidad de que dos alelos que segregan a sus gametas sean **IBD**. El coeficiente de endogamia es la probabilidad de que dos gametas que se unen para formar al individuo Z, lleven alelos **IBD**. $F_Z = \Theta(X, Y)$.

3.2. Coancestría y endogamia

La coancestría fue inicialmente concebida como una medida de ancestría reciente dada por relaciones en un pedigree (ver por ejemplo Wright, 1921; Malécot, 1948; Jacquard, 1970 y Cockerham, 1971). En el método desarrollado por Wright (1921), cada trayectoria posible en un pedigree que lleva desde un individuo a su ancestro y regresa nuevamente al individuo (sin tener en cuenta trayectorias que impliquen pasar dos veces por un mismo individuo), contribuye con una proporción $(1/2)^{n+1} (1+f_a)$ al coeficiente de endogamia F , donde n es el número de pasos en la trayectoria y f_a es la endogamia del ancestro común (método del “loop”). De este modo, $F = \sum (1/2)^{n+1} (1+f_a)$.

Otra forma de evaluar la coancestría, es a través de la medida del grado de relación entre individuos **en una población**, a partir de la identidad en estado (**IBS**). En poblaciones naturales, esto puede estimarse mediante la correlación entre loci homólogos, que puede definirse como:

$$f_{ij} \equiv \frac{Q_{ij} - \bar{Q}}{1 - \bar{Q}} \quad (\text{Expresión 1.1})$$

Donde $Q_{ij} = 1$ si los loci i y j son **IBS** y 0 en caso contrario, y $\bar{Q} = \sum_l p_l^2$, siendo p_l la frecuencia del alelo l en la población (Cockerham y Weir, 1987; Hardy y Vekemans, 1999). En esta definición,

los coeficientes Q representan probabilidades de *IBS*. El símbolo " $\equiv$ " se utiliza para diferenciar la definición de los parámetros de sus valores bajo modelos particulares (empleando en estos casos el símbolo "=").

Alternativamente, f_{ij} puede definirse en términos de *IBD*:

$$f_{ij} \equiv \frac{\theta_{ij} - \bar{\theta}}{1 - \bar{\theta}} \quad \text{(Expresión 1.2)}$$

Donde θ_{ij} es el coeficiente de coancestría *a priori* entre los loci i y j , y $\bar{\theta}$ es el coeficiente de coancestría entre dos locus elegidos al azar de la población (Hardy y Vekemans, 1999). El término *a priori* hace referencia a que el coeficiente de coancestría representa probabilidades de *IBD*, entre loci muestreados adecuadamente (Hardy y Vekemans, 1999). Las expresiones 1.1 y 1.2 son equivalentes bajo el modelo de alelos infinitos (Hardy y Vekemans, 1999).

Se han propuesto varios estimadores de f_{ij} , tales como los descritos en Thompson (1975), Lynch (1988), Queller y Goodnight (1989), Loiselle y col. (1995), y Ritland (1996).

3.3. Los índices F jerárquicos

Los índices de fijación o índices F jerárquicos (Wright, 1921; Malécot, 1948; Wright, 1951; Cockerham y Weir 1968; Cockerham, 1969, 1973; Weir y Cockerham, 1984) se utilizan para cuantificar el desvío que presentan las frecuencias alélicas de los valores esperados bajo apareamiento aleatorio (panmixia) en una población. Fueron definidos conceptualmente por Sewall Wright (1921) como la **correlación entre las gametas que se unen**. En aquel trabajo, Wright había definido el coeficiente de relación (*relationship*) entre los individuos X e Y, como la correlación (una correlación de Pearson) para algún rasgo cuantitativo, asumiendo efectos aditivos entre diferentes loci y entre alelos de un locus. Para los índices F jerárquicos, la correlación corresponde a las frecuencias de alelos provenientes de gametas elegidas al azar, en un nivel poblacional de referencia. Siguiendo el trabajo de Wright, una población puede dividirse en tres niveles de complejidad: individuos (I), subpoblaciones (S ; también llamadas “demos” o “poblaciones locales”) y población total (T); los índices de fijación jerárquicos se definen de esta manera como:

- F_{IT} : correlación entre gametas dentro de individuos (I) respecto a la población total (T)
- F_{ST} (llamado también *índice de fijación*): correlación entre gametas elegidas al azar dentro de una subpoblación (S) respecto a la población total (T).
- F_{IS} (llamado también *coeficiente de endogamia*): correlación entre gametas dentro de individuos (I) respecto a la subpoblación a la que pertenecen (S)

Aunque por razones históricas se toman a los niveles I , S y T como referencia, la jerarquía puede definirse de distinto modo en un diseño anidado, por ejemplo considerando los niveles “demo”, “región”, o cualquier otro (Holsinger y Weir, 2009).

Estos tres índices **son coeficientes de coancestría entre sus respectivos niveles jerárquicos**, y se relacionan entre sí de la siguiente manera (Wright, 1951; Cockerham, 1969):

$$F_{IS} \equiv \frac{F_{IT} - F_{ST}}{1 - F_{ST}} \quad \text{(Expresión 1.3)}$$

Puede verse en este caso la conexión con la Expresión 1.1. De hecho, los tres coeficientes F pueden ser tomados como coeficientes de coancestría (Wang, 1997a,b; Rousset, 2002). Se puede definir F de modo genérico (Rousset, 1996, 2002) como:

$$F \equiv \frac{Q_d - Q_e}{1 - Q_e} \quad \text{(Expresión 1.4)}$$

Donde Q_d es la probabilidad de identidad en estado para un locus "dentro" de una unidad de estructura poblacional determinada (población, demo, etc.) y Q_e es la probabilidad de identidad en estado "entre" unidades de estructura (Rousset, 2002). Los coeficientes Q en esta expresión son parámetros que no dependen de un modelo en particular. El "1" del denominador puede verse en esta definición como la probabilidad de identidad de un locus consigo mismo (Rousset, 2002).

Weir y Cockerham (1984) desarrollaron los conceptos originales de Wright (1943, 1951, 1952, 1965), que habían sido pensados en un modelo para subpoblaciones discretas, donde cada subpoblación era del mismo tamaño y recibía individuos de todas las demás con la misma tasa de

migración. Hay que tener en cuenta al respecto que los datos disponibles para estudios genéticos (caracteres fenotípicos) en la primera mitad del siglo XX, convertían a los conceptos que se desprenden de los trabajos de Wright y Malécot en desarrollos más bien teóricos, pudiendo ser reelaborados posteriormente a partir de los datos genéticos provenientes de la secuenciación y del estudio de proteínas en la segunda mitad del mismo siglo. Bajo el concepto original de "correlación entre gametas que se unen", Weir y Cockerham formalizaron la estimación de los índices de fijación por medio de Análisis de la Varianza (ANOVA) de las frecuencias alélicas. Weir y Cockerham (1984) asumen que la población total hace referencia a la población ancestral más reciente (cada subpoblación es una réplica de una historia evolutiva común, compartiendo el mismo tiempo de divergencia desde un ancestro común). Holsinger y Weir (2009) sugieren que al usar a F_{ST} (según la definición de Weir y Cockerham) como un descriptor de la partición de la diversidad entre "subpoblaciones", no es necesario hacer supuestos respecto a la naturaleza discreta o no de las mismas, o al proceso evolutivo que generó la diferenciación.

Con el estudio de componentes de varianzas (Cockerham y Weir 1968; Cockerham, 1969, 1973; Weir y Cockerham, 1984) se llega a los siguientes estimadores de los coeficientes (para un locus):

$$\hat{F}_{ST} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2 + \sigma_w^2} \quad \text{(Expresión 1.5)}$$

$$\hat{F}_{IT} = \frac{\sigma_a^2 + \sigma_b^2}{\sigma_a^2 + \sigma_b^2 + \sigma_w^2} \quad \text{(Expresión 1.6)}$$

$$\hat{F}_{IS} = \frac{\sigma_b^2}{\sigma_w^2 + \sigma_b^2} \quad \text{(Expresión 1.7)}$$

Donde $\sigma_a, \sigma_b, \sigma_w$ son los componentes de varianza entre poblaciones, entre individuos dentro de poblaciones y dentro de individuos, respectivamente. Las estimas multilocus pueden obtenerse haciendo un promedio ponderado (relación de la suma de los numeradores y denominadores a través de loci) (Reynolds y col., 1983; Weir y Cockerham, 1984).

F_{ST} toma valores entre 0 y 1; un valor alto indica que las las frecuencias alélicas entre subpoblaciones son diferentes, mientras que si es bajo estas no se diferencian. F_{IS} y F_{IT} varían entre -1 y 1; valores positivos indican exceso de homocigotas, mientras que valores negativos indican exceso de heterocigotas respecto a los valores esperados bajo equilibrio de Hardy-Weinberg. Para F_{IS} , el valor -1 representa la máxima exogamia, 0 la presencia de apareamiento aleatorio y 1 la máxima endogamia.

3.4. Análisis Molecular de la Varianza (AMOVA)

Excoffier y col. (1992) introdujeron el Análisis Molecular de la Varianza (AMOVA) para analizar la variación genética en sistemas haploides (haplotipos). Posteriormente, Michalakis y Excoffier (1996) extendieron este método a sistemas diploides (no siendo necesario conocer la fase gamética). La técnica está basada en un Análisis de la Varianza (ANOVA), similar al desarrollado por Weir y Cockerham (1984) pero usando una matriz de distancias, que permite el análisis de diferentes estructuras poblacionales y niveles jerárquicos. La entrada para un análisis de AMOVA consiste en una matriz de distancias euclidianas entre haplotipos y las frecuencias de cada haplotipo en las subpoblaciones.

Dependiendo de la forma en que se calcule la matriz euclidiana, se pueden obtener distintos estadísticos φ (los propios del AMOVA). Cuando se asume un modelo de alelos infinitos, se obtiene un índice equivalente al F definido por Weir y Cockerham (1984). Cuando se usa un modelo de mutación por pasos, el estadístico calculado equivale al R de Slatkin (1995).

De modo similar al tratado por Weir y Cockerham (1984), se pueden calcular para un locus los siguientes estadísticos φ :

$$\varphi_{ST} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2 + \sigma_w^2} \quad \text{(Expresión 1.8)}$$

$$\varphi_{IT} = \frac{\sigma_a^2 + \sigma_b^2}{\sigma_a^2 + \sigma_b^2 + \sigma_w^2} \quad \text{(Expresión 1.9)}$$

$$\phi_{IS} = \frac{\sigma_b^2}{\sigma_w^2 + \sigma_b^2} \quad \text{(Expresión 1.10)}$$

Aquí se observan las mismas fórmulas de las Expresiones 1.5, 1.6 y 1.7, siendo $\sigma_a, \sigma_b, \sigma_w$ los componentes de varianza entre poblaciones, entre individuos dentro de poblaciones y dentro de individuos, respectivamente. La diferencia entre ambos métodos (AMOVA y F jerárquicos de Weir y Cockerman) radica en la forma de calcular estos componentes.

4. LA GENÉTICA DEL PAISAJE Y LA ESTRUCTURACIÓN DE LAS POBLACIONES A ESCALA FINA: EL MODELO DE AISLAMIENTO POR DISTANCIA Y LOS PATRONES DE AUTOCORRELACIÓN ESPACIAL

4.1. Introducción

El paisaje, en términos generales, comprende las características visibles de un área, incluyendo elementos físicos de la morfología del terreno (montañas, cuerpos de agua, etc.), elementos vivientes (cobertura vegetal nativa), elementos de diferente naturaleza incorporados por el hombre (distintas formas de uso de la tierra, construcciones, etc.) y elementos transitorios (iluminación y condiciones climáticas). La genética del paisaje amalgama la genética de poblaciones, la estadística espacial y la ecología del paisaje² (Manel y col., 2003; Storfer y col., 2007). El propósito de esta ciencia es proveer información acerca de la forma en que las características del ambiente y el paisaje influyen la estructura genética poblacional y la variabilidad genética, a través de sus efectos sobre procesos microevolutivos, tales como el flujo génico, la deriva genética y la selección. Su surgimiento ha sido posible en gran parte gracias a los avances tecnológicos recientes, tales como las herramientas moleculares hipervariables, la información de calidad brindada por satélites y radares, y los desarrollos computacionales (hardware y software) y estadísticos (geoestadística y aplicaciones basadas en máxima verosimilitud y métodos bayesianos). Gracias al creciente número de marcadores moleculares polimórficos disponibles es posible examinar la biogeografía a escala espacial y temporal fina, lo que permite detectar discontinuidades en la distribución de las subpoblaciones y grupos de individuos. El avance en materia de teledetección y de geolocalización permite obtener información de alta calidad sobre el paisaje, tal como imágenes a distintas escalas, modelos de elevación digital e información geográfica precisa, que puede ser incorporada en Sistemas de Información Geográfica (SIG). Los avances en geoestadística de los últimos 50 años y la teoría en general disponible (tal como la estadística bayesiana), posible de ser utilizados hoy día gracias a la alta capacidad de procesamiento de las computadoras corrientes, son elementos que también, en conjunto, han facilitado la convergencia disciplinaria.

² La ecología del paisaje es el estudio de la interacción entre los patrones espaciales y los procesos ecológicos; concierne a la extensión y configuración espacial dentro de las cuales ocurren dichos patrones y procesos (Manel y col., 2003). Esta disciplina incluye las causas y consecuencias biofísicas y sociales de la heterogeneidad del paisaje.

4.2. Aislamiento por distancia

Tres modelos básicos de **estructura poblacional** fueron desarrollados por diversos autores. En el modelo de **aislamiento por distancia** (Wright, 1943) el área que ocupa una población continua es grande en comparación a la distancia de dispersión de los individuos que la componen (dispersión geográfica restringida), hecho que previene la formación de una única unidad panmíctica (Kimura y Weiss, 1964) y produce la acumulación de diferencias genéticas locales (Slatkin, 1993); dicho de otra manera, una población está distribuida uniformemente a lo largo de un gran territorio, pero los padres de cualquier individuo provienen de un pequeño área alrededor del mismo; individuos lejanos están aislados por distancia, en el sentido que no se cruzan entre sí (Ishida, 2009). Este modelo fue elaborado por Wright (1943, 1946), Malécot (1948), Kimura y Weiss (1964), Maruyama (1971), y Nagylaky (1976), entre otros. En contraposición al mismo, en el denominado **modelo de islas** (Wright, 1943) una población se divide en subpoblaciones, interconectadas por migración (Figura 1.13A). En el **modelo de piedras de paso** (*stepping stone*) (Kimura y Weiss, 1964) (Figura 1.13B, C), los individuos también están distribuidos en subpoblaciones pero la migración ocurre con más alta probabilidad entre las cercanas (en el caso general, tal como está discutido en Kimura y Weiss, 1964).

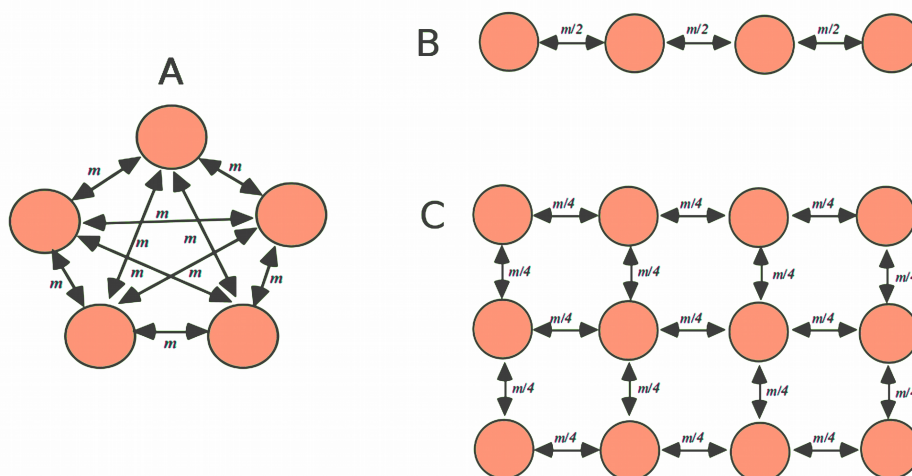


FIGURA 1.13 Modelos de estructura poblacional. (A) modelo de islas; (B) modelo de piedras de paso en una dimensión; (C) modelo de piedras de paso en dos dimensiones. En los tres modelos los círculos representan subpoblaciones y m es la tasa de migración entre las mismas. En los modelos B y C, cada individuo puede migrar entre subpoblaciones dando un paso por generación. Kimura y Weiss (1964), además del modelo con un solo paso, presentaron una generalización para múltiples pasos.

El modelo de aislamiento por distancia podría concebirse como un caso límite del de piedras de paso, donde la distancia entre las subpoblaciones tiende a cero (Kimura y Weiss, 1964). Como mencionan estos autores, se pueden tomar dos enfoques complementarios para el estudio de la *diferenciación local*: el utilizado por Wright (1940, 1943, 1946, 1951), con los coeficientes F jerárquicos, o el enfoque desarrollado por Malécot (1948, 1955, 1959), con un cambio de correlación en función de la distancia, donde los individuos más cercanos se parecen más entre sí que los distantes. Ishida (2009) sugiere que la teoría de Wright concierne a lo que denomina **aislamiento por distancia ecológico** mientras que el trabajo de Malécot apunta hacia lo que denomina **aislamiento por distancia genético**. El primero se refiere a la interacción local entre los individuos: para Wright el aislamiento por distancia sería un proceso donde **el apareamiento se da entre los individuos cercanos, debido a la presencia de una distancia de dispersión corta**, proceso que puede producir diferenciación genética y cambio evolutivo; para Malécot, sería un proceso probabilístico, donde **la coancestría entre pares de individuos disminuye al aumentar la distancia geográfica que los separa**, reduciéndose la probabilidad de compartir ancestros comunes recientes. En la teoría de Wright, el aislamiento por distancia es una *causa* de diferenciación genética local. En la de Malécot, dado que la probabilidad de migración depende sólo de la distancia entre los lugares de nacimientos de los ancestros y sus descendientes, el aislamiento por distancia sería una *consecuencia*, un estado al que llega una población en el equilibrio. En esta situación (en el equilibrio), la población mostraría un patrón (el patrón de aislamiento por distancia) en el que la coancestría disminuye cuando la distancia entre individuos aumenta. Según Epperson (1999), el enfoque de Wright descansa en el *estudio estadístico* de muestras, mientras que el de Malécot se enfoca en el problema como un *proceso estocástico*.

4.3. Autocorrelación espacial

La **autocorrelación espacial (AE)** es una **propiedad de los objetos en pares de sitios de ser más similares** (autocorrelación positiva) **o menos similares** (autocorrelación negativa) **entre ellos que lo esperado por azar**. El tópico comenzó a ser explorado en los años 50, en el área de la geografía analítica, pero su estudio se volvió riguroso tras la publicación de la monografía *spatial autocorrelation* de Cliff y Ord (1973). Con el uso masivo de las computadoras de escritorio, esta temática alcanza una difusión plena (Getis, 2008). Sokal fue pionero en el estudio de la AE en biología de poblaciones. En el primer trabajo que publica sobre esta temática (Sokal y Oden, 1978a)

(Figura 1.14), detalla la metodología general, que actualmente se ha vuelto más compleja, aunque puede considerarse a esta investigación como la base de los desarrollos posteriores. El trabajo que sigue en la serie (Sokal y Oden, 1978b), describe las implicancias biológicas del análisis de AE. Podemos considerar a los métodos descritos en el primer trabajo como medidas de **autocorrelación global**. Los estudios de AE global pueden realizarse con datos cuantitativos o cualitativos. Para datos cuantitativos, las medidas comúnmente usadas son los estadísticos I de Moran (Moran, 1950) y C de Geary (Geary, 1954). Para los datos nominales, se utiliza el estadístico *Join-count* (Cliff y Ord, 1973; Sokal y Oden, 1978a). Un método común para estudiar el patrón de autocorrelación espacial de una variable, es realizar un **correlograma** (Figura 1.14), donde se grafica el valor de un coeficiente de AE en función de la distancia geográfica entre pares de individuos, agrupada en clases. Esto puede realizarse sin discriminar la dirección (correlograma omnidireccional), o pueden calcularse correlogramas direccionales (Oden y Sokal, 1986).

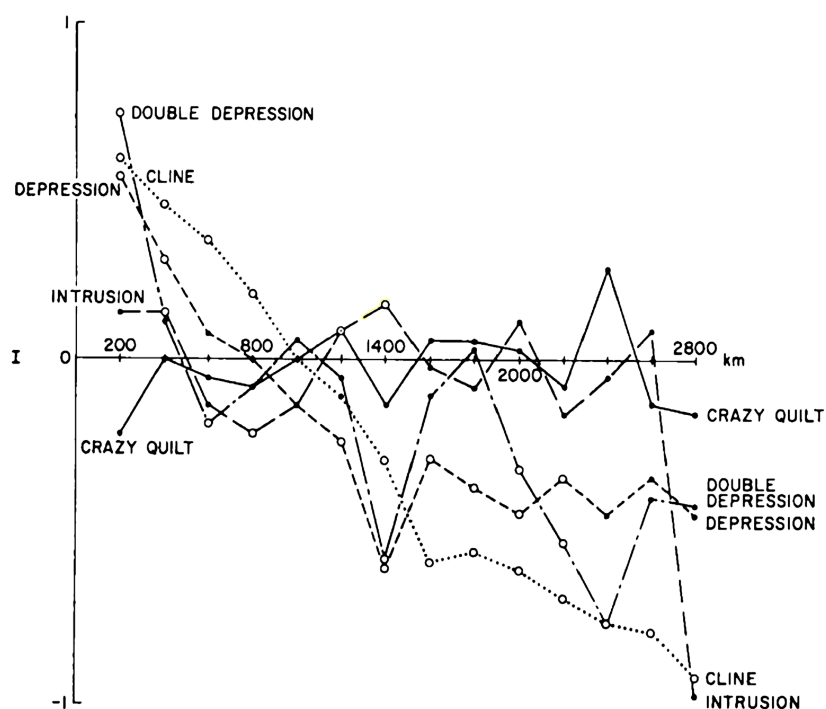


FIGURA 1.14 Correlogramas basados en el coeficiente I de Moran, para 5 variables simuladas (imagen extraída de Sokal y Oden, 1978a). El eje de las abscisas representa la distancia en km. El patrón de “clina” (*cline*) representa una transición gradual de valores, el patrón de “depresión” (*depression*) es un núcleo central de valores bajos que incrementan gradualmente hacia la periferia, el patrón de “doble depresión” (*double depression*) representa dos áreas con valores bajos separadas por áreas con valores altos. El patrón de “intrusión” (*cline intrusion*) es un área de valores altos con una porción bordeando sitios con valores bajos. Finalmente, el patrón de “mosaico” (*crazy quilt*), representa zonas con valores altos rodeadas por zonas con valores bajos (como en un tablero de ajedrez). Para más detalles ver el trabajo de Sokal y Oden (1978a).

A partir de la década del 90, el estudio de la AE se ha enriquecido con el desarrollo de medidas de **autocorrelación local**³. Estas medidas sirven para estudiar la **autocorrelación en torno a sitios individuales**, para descubrir "paquetes" locales (*hot spots*; Anselin, 1995) de puntos autocorrelacionados, y para estudiar la presencia de outliers (Sokal y col., 1998). Hay dos tipos de indicadores locales disponibles: los estadísticos LISA y los estadísticos G_i y G_i^* de Getis y Ord (Getis y Ord, 1992; Ord y Getis, 1995). Los estadísticos LISA (del inglés *Local Indicators of Spatial Association*) fueron definidos por Anselin (1995). Para que un indicador espacial sea considerado LISA, debe satisfacer dos requisitos: a) el valor de LISA para cada observación es indicador del grado de agrupamiento significativo de valores similares alrededor de la misma, y b) la suma de los valores LISA de todas las observaciones es proporcional a un indicador global de asociación espacial. Anselin (1995), describió dos indicadores que satisfacen estos requerimientos: los coeficientes I de Moran y C de Geary, en versiones locales. Los estadísticos locales de Getis y Ord, G_i y G_i^* no son de tipo LISA ya que no cumplen la condición (b). Estos últimos describen el agrupamiento espacial de valores altos (*hot spots*) y bajos (*cold spots*) de una variable, en torno a un pivote (Sokal y col., 1998) y sólo detectan autocorrelación espacial positiva⁴. Ord y Getis (1995) redefinieron estos estadísticos de modo estandarizado. Nuevamente, Sokal fue pionero en sugerir la aplicación de estos métodos a nivel biológico, y en el trabajo escrito por Sokal, Oden y Thomson (1998), se estudian este tipo de indicadores, donde puede encontrarse su estudio técnico.

En el campo de la genética, el análisis de AE ha sido también llevado a cabo por autores tales como Epperson (1993), Hardy y Vekemans (1999), Rousset (1997) y Smouse y Peakall (1999). El potencial del análisis de AE para examinar los patrones de estructura genética espacial es bien reconocido entre los botánicos (Epperson y Alvarez-Buylla 1997; Epperson y Li 1997; Sokal y col., 1997; Smouse y Peakall 1999). Es fácilmente observable la relación entre los patrones de AE y el modelo de aislamiento por distancia, tal como fue desarrollado por Malécot. Este modelo implica la disminución del parentesco con la distancia geográfica, y el resultado esperado de este proceso es la ocurrencia de estructura genética espacial, con una disminución lineal en el coeficiente de parentesco promedio en función del logaritmo de la distancia entre individuos. Además de los métodos más tradicionales, tales como los usados por Sokal en muchos de sus trabajos, donde se estudia el comportamiento de las frecuencias alélicas con la distancia geográfica, el análisis global

3 En este caso, el desarrollo proviene de la econometría espacial, en contraste con los indicadores globales, que provienen de la geografía analítica.

4 Biológicamente la más interesante, porque se asocia con la detección de grupos locales con valores similares de una variable.

I. Introducción

de parentesco permite una aproximación multilocus con correlogramas, en la que se calcula y se grafica un valor de parentesco en función de grupos discretos de distancia geográfica creciente. Un método usual en genética del paisaje es el uso del coeficiente de parentesco de Nason (f_{ij}), implementado por Loiselle (1995). Recientemente, Double y col. (2005) usaron el método de Smouse y Peakall (1999) para realizar un análisis de autocorrelación espacial bidimensional (2D-LSA), generando así una aplicación a nivel local del análisis de AE basado en una estima del parentesco.

5. ALGUNAS HERRAMIENTAS PARA LA OBTENCIÓN DE DATOS EN GENÉTICA DEL PAISAJE: MARCADORES GENÉTICOS, TELEDETECCIÓN Y LENGUAJE R

5.1. Marcadores Genéticos

Un marcador genético puede ser definido como cualquier variación estable y de herencia mendeliana con base genética, capaz de ser medida o detectada con métodos adecuados, y que puede ser utilizada para identificar la presencia de un determinado genotipo o fenotipo, que de otro modo sería inconmensurable o difícil de detectar (Mitra y col., 1994). Esta variación puede darse en diferentes niveles biológicos, desde la morfología hasta la variación presente en una sola base del ADN. Un marcador ideal debe ser: 1) altamente polimórfico, 2) de herencia mendeliana no epistática, 3) insensible a los efectos ambientales, 4) codominante, 5) de rápida identificación y análisis simple y 6) de posible detección en los estadios tempranos del desarrollo. Los marcadores más antiguos y clásicos utilizados en genética de poblaciones son los **marcadores fenotípicos**, de fácil identificación visual, que permiten examinar los genotipos subyacentes mediante cruzamientos controlados o análisis de pedigríes (D'Surney y col., 2001). Los **marcadores bioquímicos** comprenden compuestos orgánicos tales como terpenos, alcaloides, etc., o bien, proteínas (isoenzimas y proteínas de reserva). Las isoenzimas fueron los primeros marcadores no fenotípicos ampliamente utilizados en estudios poblacionales (a partir de los años 60), tras el desarrollo de la técnica de electroforesis. Los **marcadores moleculares**, basados en la detección de la variación a nivel de ADN, se comenzaron a utilizar a partir de mediados de la década del 70, con el desarrollo de la técnica RFLP (polimorfismo de longitud de los fragmentos de restricción). Los marcadores moleculares poseen numerosas ventajas respecto a sus predecesores: 1) son específicos para cada individuo, grupo de individuos, especies o aún grupos sistemáticos mayores, 2) son independientes de la influencia del medio ambiente y los efectos pleiotrópicos, 3) No están bajo control poligénico, 4) se dispone de un número potencialmente ilimitado de marcadores, y 5) muchos marcadores moleculares se suponen neutros, es decir, no sujetos a una presión de selección.

Los marcadores moleculares pueden ser clasificados en dos categorías principales, dependiendo de los fundamentos técnicos utilizados para su detección: basados en hibridación (RFLP) o basados en PCR (RAPD, ISSR, AFLP, SSR, SNP, etc.). La técnica de RFLP implica el corte del ADN en fragmentos de restricción, que se separan mediante electroforesis; las otras, la

I. Introducción

amplificación del ADN con la técnica de PCR (desarrollada a mediados de los 80) y su análisis posterior mediante electroforesis o secuenciación.

La Tabla 1.7 describe las principales diferencias entre los marcadores comúnmente utilizados: RAPD (amplificación al azar del polimorfismo del ADN), SSR (secuencias simples repetidas o microsatélites), ISSR (inter-secuencias simples repetidas), AFLP (polimorfismos en la longitud de los fragmentos amplificados) y SNP (polimorfismo de nucleótido simple). **En esta tesis se utilizaron marcadores ISSR y SSR.**

TABLA 1.7 Características de los marcadores moleculares más utilizados. Adaptado de Kesawat y col. (2009) y Miah y col. (2013).

Característica	RAPD	SSR	ISSR	AFLP	SNP
abundancia en el genoma	alta	alta	media-alta	alta	alta
nivel de polimorfismo	medio	alto	medio	medio	alto
locus-especificidad	no	si	no	no	si
codominancia de alelos	no	si	no	no	si
reproducibilidad	baja	alta	media-alta	media-alta	alta
tiempo para su obtención	bajo	bajo	bajo	medio	bajo
demanda técnica	baja	baja-media	baja-media	media	media
costo de operación	bajo-medio	bajo	bajo-medio	medio	medio
costo de desarrollo	bajo-medio	alto	bajo	bajo	medio
cantidad de ADN requerida	baja	baja	baja	media	baja
calidad de ADN requerida	crudo	relativamente puro	relativamente puro	relativamente puro	puro

5.1.1. Marcadores moleculares ISSR

Los ISSR (Zietkiewicz y col., 1994) son un tipo de marcador genético que permite la cuantificación de los niveles de variación en las regiones microsatélites del genoma nuclear. La técnica se basa en el uso de cebadores diseñados a partir de regiones microsatélites, que consisten en repeticiones en tándem de motivos simples como (CT)_n o (CA)_n, ubicadas entre secuencias no repetitivas (Figura 1.15). Los motivos repetidos, denominados también SSR (secuencias simples repetidas) pueden ser penta, tetra, tri y dinucleótidos. La longitud de las secuencias de microsatélites tiende a ser altamente variable entre individuos debido a las altas tasas de mutación que experimentan, ya que la ADN polimerasa puede eliminar o agregar unidades a la cadena de ADN durante la duplicación en meiosis (proceso denominado *slippage*), resultando en secuencias con menos o más unidades de repetición (o pares de bases) que las parentales (Zietkiewicz y col.,

1994; Wolfe, 2000). Los ISSR son marcadores semiarbitrarios debido a que en esta técnica se utiliza como cebador un oligonucleótido complementario a las secuencias microsatélites (motivo repetido de tri o dinucleótidos). Además, en ocasiones es posible agregar a esta secuencia un par de nucleótidos extra arbitrarios en el extremo 3' o en el 5' que juegan el papel de “anclas”, asegurando que la amplificación se inicie siempre en el extremo correspondiente del microsatélite (Zietkiewicz y col., 1994; Bornet y Branchard, 2001; Reddy y col., 2002).

Cuando dos secuencias repetidas se presentan a una distancia amplificable y con una orientación invertida, el cebador complementario a ellas puede inducir la amplificación del segmento de ADN intermedio (Figura 1.15). La molécula generada, con un tamaño particular (peso molecular), se considera un “locus”, que representa el segmento de ADN entre los microsatélites (el marcador propiamente dicho).

Con esta técnica se pueden obtener entre 25 a 50 bandas en una sola reacción y este patrón característico de productos de PCR se considera la “huella genética” de cada uno de los individuos analizados. Las bandas de ISSR son marcadores dominantes que siguen herencia mendeliana simple, por lo que la presencia de la banda representa el genotipo dominante (homocigota o heterocigota) y su ausencia el genotipo homocigota recesivo. La falta de amplificación del locus, que se manifiesta como la ausencia de banda, se produce si se presenta divergencia entre los sitios de unión y los cebadores, si uno de los sitios de unión no existe y/o si existen rearrreglos estructurales genómicos.

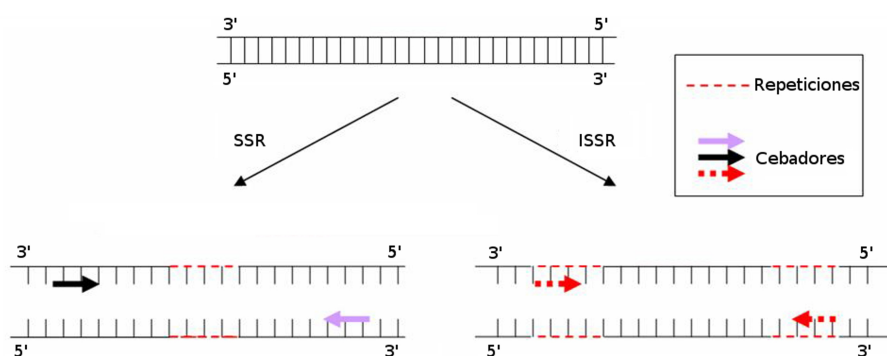


FIGURA 1.15 Proceso de amplificación con marcadores SSR e ISSR. Se muestra de modo esquemático un fragmento de ADN. La técnica SSR (esquema inferior izquierdo) usa cebadores complementarios (en negro y en violeta) que flanquean una región repetida; la técnica ISSR (esquema inferior derecho) usa un único cebador (en rojo) que amplifica el ADN entre dos regiones repetidas. Adaptado de Yip y col. (2007).

5.1.2. Marcadores moleculares SSR

Los microsatélites o SSR son motivos repetitivos compuestos de 1 a 6 pares de bases, que pueden encontrarse tanto en regiones codificantes como no codificantes del genoma (Miah y col., 2013). Debido a que las secuencias flanqueantes de cada uno de estos sitios es única, una vez que los loci SSR son clonados y secuenciados se pueden diseñar cebadores específicos para poder amplificarlos (Figura 1.15). El resultado final identifica usualmente un locus simple, el cual es con frecuencia multialélico debido a la alta tasa de mutación de los SSR (Jones y col., 1997). El polimorfismo que se observa en un locus microsatélite particular es debido a la variación en la longitud del fragmento amplificado⁵, por la presencia de más o menos repeticiones de un motivo.

En estos marcadores es común la existencia de alelos nulos, que se refieren a alelos donde no hay productos específicos de PCR. La existencia de alelos nulos puede explicarse por *a*) la delección del microsatélite en un locus específico y *b*) mutaciones (inserciones, delecciones o sustituciones en el sitio de unión del cebador) (Varshney y col., 2005). La presencia de alelos nulos puede conducir a una subestimación de la heterocigosis en una población, debido a que en el heterocigota la presencia de un alelo nulo no puede ser reconocida, observándose el producto de un solo alelo como en los homocigotas. Los productos amplificados son analizados en geles corrientes o utilizando secuenciadores automáticos.

5.2. Sensores remotos

El estudio con sensores remotos comprende el análisis y la interpretación de mediciones de la radiación electromagnética reflejada o emitida por un objeto y observada o registrada por un instrumento que no está en contacto con el mismo (Mather y Koch, 2011). Una imagen obtenida a partir de un sensor remoto está compuesta por una matriz rectangular de mediciones de **flujo** de radiación electromagnética (esto es, la cantidad de radiación incidente o transmitida por una superficie de área unitaria), en forma de puntos o **píxeles** (Figura 1.16), de modo que cada valor registrado representa la magnitud de la radiación proveniente de un área pequeña. La radiación emitida contiene información sobre *a*) la naturaleza del material presente en el área que compone el

⁵ La variación en la longitud de los microsatélites se produce por el proceso anteriormente descrito (*slippage*). En este proceso, la ADN polimerasa elimina o agrega unidades a la cadena de ADN durante la duplicación en meiosis.

píxel, *b*) la posición topográfica del área (si es un sitio horizontal, o una pendiente iluminada directamente por el Sol o con sombra) y *c*) el estado de la atmósfera a través de la cual viaja la radiación. La información se registra como matrices de números, donde la magnitud de la radiación reflejada o emitida por el píxel está representada usualmente por un número entero graduado en el rango de 0 (0% reflexión) a 255 (100% de reflexión) (Figura 1.16).

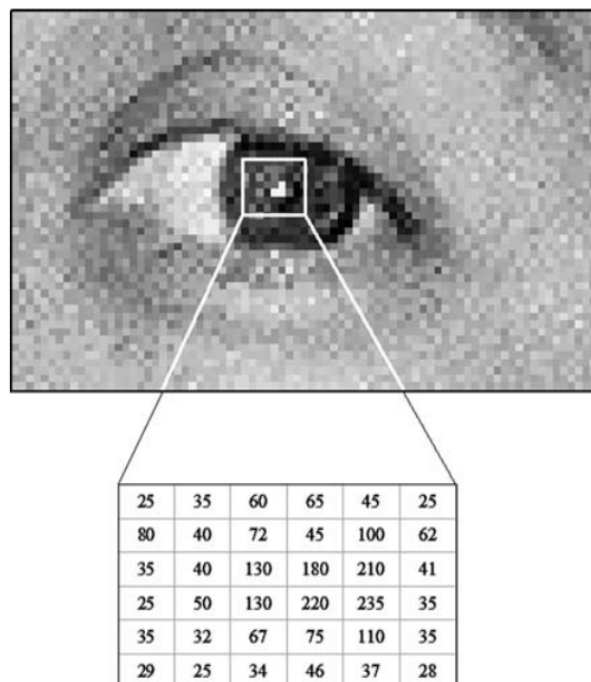


FIGURA 1.16 Imagen digital de un ojo humano (Mather y Koch, 2011), mostrando la correspondencia entre los niveles de gris de los píxeles y la representación numérica del nivel de gris en la computadora (valores entre 0 y 255 [8 bits]).

La **radiancia** es el flujo transmitido por una unidad de área de la superficie terrestre a través de una unidad de ángulo sólido (tridimensional). La **reflectancia** es la relación entre la radiación emitida por un objeto, y la que incide sobre el mismo. Como es una relación independiente de la cantidad de luz incidente, se suele utilizar para comparar imágenes tomadas a distintos tiempos. Radiancia y reflectancia pueden utilizarse en referencia a longitudes de onda particulares.

La radiación solar reflejada en la Tierra y detectada por un satélite tiene que pasar a través de la atmósfera dos veces, interactuando con partículas en su trayecto. Esta interacción suele describirse en términos de **dispersión** y **absorción**, procesos que pueden tener una mayor o menor

importancia en distintas regiones del espectro electromagnético y que pueden variar espacial y/o temporalmente. Por un lado, una cierta cantidad de la radiación reflejada y dispersada desde los alrededores de cada objeto es registrada por el sensor, volviendo a la imagen menos nítida; por otro, la absorción a lo largo del trayecto por el que la radiación viaja reduce el brillo en la imagen. Estos dos fenómenos en conjunto generan un efecto de “niebla” (en inglés **haze**) sobre la imagen captada por el sensor.

Diferentes objetos en la superficie terrestre tienen una manera individual y característica de interactuar con la radiación, y estas diferencias permiten identificarlos a través de una **firma espectral** (un patrón cuantitativo de la respuesta que un objeto tiene para distintas regiones del espectro electromagnético en condiciones determinadas) (Figura 1.17). La respuesta espectral de un objeto puede variar temporalmente, por ejemplo, en el caso de la vegetación a través de sus distintos estados de fisiológicos, de crecimiento, etc. A su vez, depende de la orientación del Sol, del ángulo que este forma con el cenit y la dirección del sensor respecto del nadir (ángulo de visión).

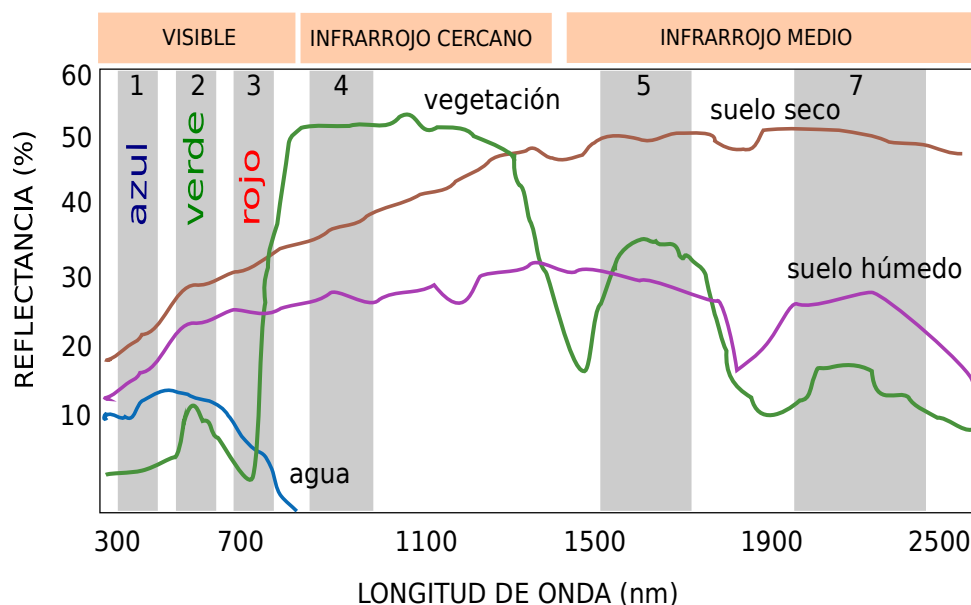


FIGURA 1.17 Representación de firmas espectrales de distintos materiales captadas por el satélite Landsat 5. Los números en la parte superior de cada columna gris indican el número de banda. Los rectángulos encima de la imagen indican la parte cualitativa del espectro a la que corresponde cada longitud de onda.

5.2.1. El índice de vegetación de diferencia normalizada (NDVI)

Distintos **índices de vegetación** (IV) han sido desarrollados para el estudio de la superficie terrestre. Un índice de vegetación es una medida cuantitativa, indicadora de vigor (Campbell, 2002) y de biomasa vegetal (Asrar y col., 1984) en un píxel. Usualmente es el resultado de la combinación de operaciones aritméticas (suma, resta, multiplicación, división) de un conjunto de valores espectrales. Estos índices han sido ampliamente utilizados para derivar atributos funcionales y biofísicos de las coberturas vegetales (Tucker, 1979), debido a su alta correlación con distintas variables, tales como la biomasa verde o el índice de área foliar, y por ello con la tasa de fotosíntesis, lo que permite utilizarlos como una herramienta para la discriminación de la vegetación (Tucker y col., 1985; Zerda y Tiedemann, 2010).

Los sensores con bandas espectrales en el rojo e infrarrojo cercano (IRc) permiten monitorear bien la vegetación, ya que la diferencia entre estas bandas ha mostrado ser un fuerte indicador de la cantidad de biomasa verde fotosintéticamente activa. Esto se debe a que la clorofila absorbe las longitudes de onda en el rojo visible (630-690 nm), mientras que la estructura interna de la hoja refleja en el infrarrojo cercano (760-900 nm) (Lay, 2005). El **NDVI** (Tucker, 1979) es un índice que relaciona la diferencia entre las porciones del rojo visible y del infrarrojo cercano del espectro electromagnético (Figura 1.18), normalizada por su suma, hecho que compensa los cambios de condiciones lumínicas, pendiente, aspecto y otros factores:

$$NDVI = \frac{IRc - ROJO}{IRc + ROJO} \quad (\text{Expresión 1.10})$$

El NDVI toma valores entre -1 y 1 y permite tener una medida comparable (Tucker, 1979). Nubes, agua y nieve, con valores importantes de reflectancia en el IRc, toman por lo general valores negativos de NDVI. La roca y el suelo toman valores cercanos a cero, y la vegetación por lo general valores positivos, acercándose a 1 con el incremento de clorofila en las plantas o con la biomasa verde (Ramachandra y col., 2012). La relación entre los valores de reflectancia del infrarrojo respecto al rojo carece de un límite superior, factor que puede resultar en valores excepcionalmente elevados cuando la reflectancia del rojo alcanza el valor de cero en áreas densamente pobladas con vegetación (Jiang y Huete, 2010). La forma en que está construido el NDVI permite superar esta dificultad, acotando el rango de valores.

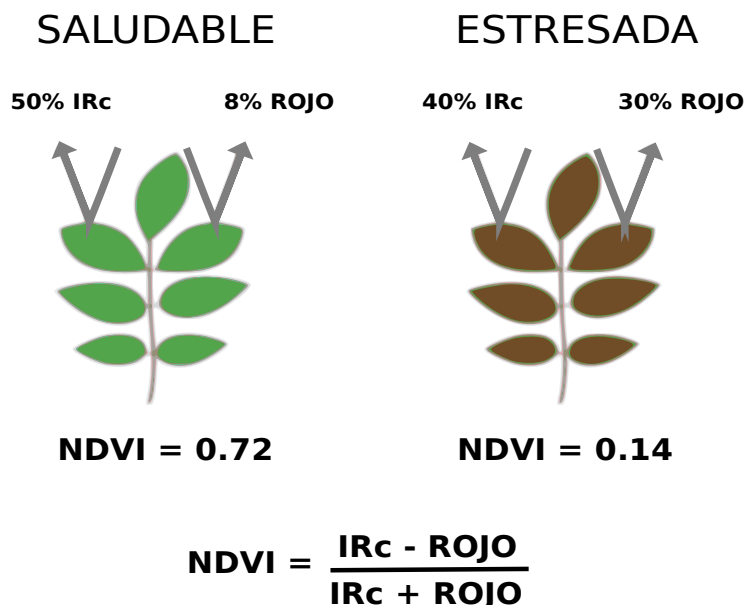


FIGURA 1.18 Valores de NDVI para una vegetación saludable y una estresada. La saludable refleja en alta proporción el infrarrojo cercano (IRc) y absorbe el rojo, lo que produce un alto valor de NDVI. La vegetación estresada refleja menos IRc y absorbe menos el rojo, resultando en un valor bajo de NDVI. Mayores valores de biomasa también generan valores elevados de NDVI, ya que se refleja más IRc debido a fenómenos de dispersión y se absorbe más rojo por efecto de la clorofila.

El NDVI se ha mostrado ampliamente correlacionado con distintos parámetros de la vegetación, tales como densidad espacial, cobertura, biomasa foliar, productividad primaria neta, fenología, índice de área foliar (*LAI*), fracción de radiación fotosintéticamente activa interceptada (*fPAR*), contenido de clorofila y situación de los cultivos (Lenney y col, 1996; Wilson y Sader, 2002; Kabthimer, 2012). El NDVI ha sido utilizado, por ejemplo, para monitorear la vegetación y estudiar su comportamiento ante las fluctuaciones climáticas a escalas globales, continentales y regionales, en la detección de sequías, y en estudios sobre procesos de desertificación y degradación del paisaje (Propastin y Kappas, 2008; Zerda y Tiedemann, 2010).

5.2.2. El satélite Landsat

A partir de mediados de los 60, la NASA se embarcó en el desarrollo del primer satélite para monitorear la Tierra. En julio de 1972 se lanza el primero de los satélites de la serie Landsat (más adelante llamado Landsat 1) (Williams y col., 2006). En años siguientes (1975, 1978, 1982 y 1984) se lanzan los satélites Landsat 2, 3, 4, y 5 (Lauer y col., 1997) (Figura 1.19). Los satélites Landsat 1 a 3 llevaban incorporados dos sensores: el **sensor RBV** (Return Beam Vidicon), con cobertura radiométrica en el verde, rojo e infrarrojo cercano, y el **sensor MSS** (MultiSpectral Scanner), con una resolución aproximada de 80 m y una cobertura radiométrica en las longitudes de onda del rojo y del verde, y dos bandas IR. El sensor MSS de Landsat 3 poseía además una quinta banda en el IR-térmico. El sensor RBV era esencialmente una cámara de televisión. Los satélites Landsat 4 y 5 también llevaban el sensor MSS e incluían otro, el **sensor TM** (Thematic Mapper). El nuevo sensor TM constaba de siete bandas espectrales, que abarcaban desde el azul hasta el IR-térmico. Poseía una resolución de 120 m para la banda IR-térmica y de 30 m para las seis restantes. Landsat 5 dejó de recabar información en enero de 2013, permitiendo disponer de un registro satelital de más de 20 años.

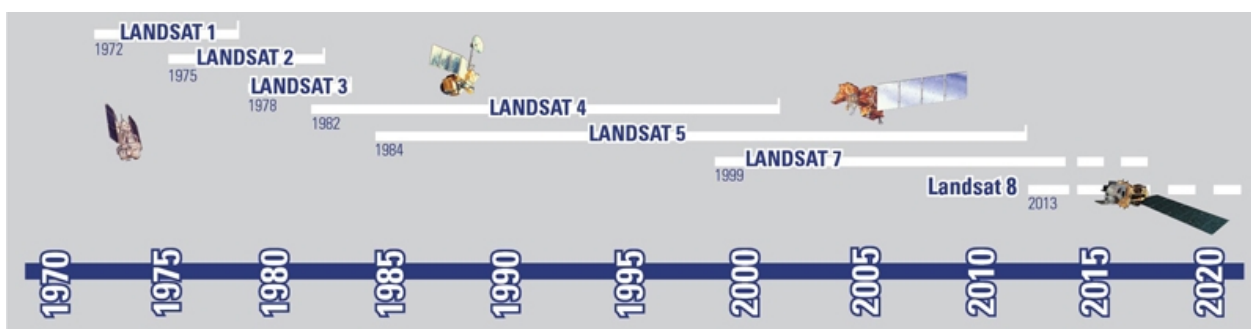


FIGURA 1.19 Línea de tiempo de la serie de satélites Landsat (fuente: USGS, http://landsat.usgs.gov/about_mission_history.php). Las líneas blancas indican el período en actividad de los satélites. Las líneas punteadas para Landsat 7 y 8 indican el tiempo de actividad restante proyectado para estos satélites antes de ser puestos fuera de funcionamiento.

Landsat 6 se lanzó en 1993, pero falló en alcanzar la órbita. Incluía un sensor TM mejorado, el **sensor ETM** (Enhanced Thematic Mapper). Landsat 7, aún es funcional, se lanzó en 1998 y posee el **sensor ETM+** (Enhanced Thematic Mapper +). El sensor ETM+ tiene el diseño básico de los TM pero con algunas mejoras (Obenschain y col., 1996). En 2003, sufrió una falla en el

dispositivo Scan Line Corrector (SLC), generando agujeros en las imágenes, que pueden rellenarse con métodos dados por el USGS (U.S. Geological Survey). El satélite Landsat 8 se lanzó en 2013. Lleva el **sensor OLI** (Operational Land Imager) y el **sensor TIRS** (Thermal Infrared Sensor). Al igual que Landsat 4, 5 y 7, sigue una órbita que regresa a un punto determinado de la Tierra cada 16 días. Actualmente la NASA está trabajando en Landsat 9, que se planea lanzar en 2023 (NASA, 2015).

Las imágenes Landsat registradas durante tres décadas han mostrado ser útiles para un sinnúmero de usos, tales como el estudio de cambios en la biosfera (por ejemplo deforestación) y para el monitoreo agrícola.

5.3. El lenguaje R

Durante los años 70, empieza a tener lugar el desarrollo del **lenguaje S** en Laboratorios Bell (AT&T) (Krause y Olson, 2005). Este lenguaje es el predecesor de R. En un principio consistía de una serie de rutinas escritas en Fortran y una interfaz, que podría pensarse como el primer lenguaje S, cuya función era manejar los programas basados en Fortran. S se pensó íntegramente en base a funciones y es de los primeros lenguajes escritos para el análisis de datos, teniendo una programación de alto nivel de abstracción, con funciones para regresión, agrupamiento, etc. (AT&T Staff, 2013). A principios de los 80, se decidió reescribir las rutinas en C y migrar hacia UNIX, estando S a partir de entonces disponible fuera de Laboratorios Bell. En la versión 4 de S, el lenguaje pasó a estar estructurado bajo un paradigma orientado a objetos, donde “todo es un objeto” (ver sección 5.3.1). Las funciones de S podían ahora ser pasadas a otras funciones, y ser la salida de otros cálculos. Hacia los 90, Laboratorios Bell le da a StatSci (ahora Insightful Corporation) una licencia exclusiva para desarrollar y vender el lenguaje S, y posteriormente la empresa decide comprar el lenguaje, llamándolo S-PLUS (convirtiéndose así en un programa pago). A la vez, en la Universidad de Auckland durante la década del 90 nace el **lenguaje R**. Este es una versión mejorada de S, que tiene **licencia libre** y es de **código abierto**, es decir, cualquiera puede usar el código y modificarlo, lo que le otorga amplias ventajas respecto a un software privativo como S-PLUS. Si la característica de S era ser interactivo, la de R es su “extensibilidad” gracias a que es de código abierto: la extensibilidad permite incorporar nuevos programas, inclusive escritos en otros lenguajes, que pueden ejecutarse en R. Los usuarios pueden escribir sus propias funciones y

compartirlas. Nuevas funciones pueden escribirse usando la sintaxis de S. Los resultados que la función devuelve, son objetos. R se diferencia de S además en el hecho de poseer **paquetes** (archivos que contienen programas). La función de los paquetes no es sólo conformar una unidad de información con programas, sino que también su rol es poder distribuirlos: los paquetes se pueden compartir. Esto vuelve al lenguaje sumamente extensible, habiendo al momento 7084 paquetes disponibles⁶ en Comprehensive R Archive Network (CRAN) (<http://cran.r-project.org/>).

5.3.1. Propiedades de R como lenguaje orientado a objetos

R es un lenguaje y un entorno de programación. Como lenguaje, posee las siguientes características:

- Es un **lenguaje interpretado** (como Java) y **no compilado** (como C, C++, Fortran, Pascal, etc.), esto significa que los comandos escritos en el teclado son ejecutados directamente sin necesidad de construir ejecutables.
- Es un **lenguaje orientado a objetos**: un objeto es una entidad que posee **identidad** (un objeto es diferenciable de otros objetos: posee un nombre), **estado** (datos asociados al objeto) y **métodos** (operaciones que se pueden realizar con el mismo). Las variables, datos, funciones, resultados, etc., se guardan en la memoria activa de la computadora en forma de objetos con un nombre específico. El usuario puede modificar o manipular estos objetos con operadores (aritméticos, lógicos, y comparativos) y funciones (que a su vez son objetos). Todas las acciones en R se realizan con objetos que son guardados en la memoria activa sin usar archivos temporales. Los resultados también son objetos. La identidad de un objeto es independiente de su estado y no cambia. Un objeto pertenece a una **clase**. La clase de un objeto determina su comportamiento y propiedades⁷. Una clase o un objeto posee **métodos**, operaciones que el objeto puede hacer⁸. Todos los objetos de una clase comparten los mismos métodos. A su vez, un objeto conoce su clase y cómo implementar un método⁹. Toda clase u

⁶ Al mes de agosto de 2015.

⁷ Una analogía sería pensar en *P. alba* como una “clase” con propiedades, en este caso la “clase” es “especie”, y un individuo concreto como un “objeto”.

⁸ Por ejemplo, un objeto de la clase “vector” se puede sumar, restar, etc.

⁹ Esto permite por ejemplo, definir el método “suma” para distintas clases: un método suma para un numero, y un método suma para dos matrices de igual tamaño, que sume cada celda y devuelva una matriz.

objeto tiene **atributos**¹⁰. Finalmente, las clases poseen **herencia**, es decir, una clase puede heredar los atributos y operaciones de una clase superior¹¹.

- En R hay tres formas de definir una clase: **clases S3**, **clases S4** y **clases de referencia**. Cada programador puede elegir un estilo de definir clases según su gusto. El estilo S3 es el más antiguo, e informal, pero el más simple para implementar. El estilo S4 tiene más formalidad y rigurosidad que S3, permitiendo generar programas estructurados. El estilo de clases de referencia, es similar al usado por Java y C ++ , y es particularmente útil cuando se requieren programas dinámicos.

Un objeto puede ser creado con el operador “asignar” el cual se denota como una flecha construida con el signo menos y el símbolo “>” o “<” dependiendo de la dirección en que asigna el objeto¹². Por ejemplo, se puede asignar el valor 5 a la variable (objeto) *n*:

```
> n <- 5
> n
[1] 5
> 5 -> n
> n
[1] 5
```

Si el objeto ya existe, su valor anterior es borrado después de la asignación (la modificación afecta solo objetos en memoria, no a los datos en el disco). El valor asignado puede ser el resultado de una operación y/o de una función:

```
> n <- rnorm(1)
> [1] 2.208807
```

La función *rnorm*(1) genera un dato al azar muestreado de una distribución normal con media 0 y varianza 1.

¹⁰ Por ejemplo, tipo de objeto, longitud, etc.

¹¹ Por ejemplo, la especie *P. alba* posee los atributos del género *Prosopis*, y atributos específicos que diferencian a *P. alba* de *P. flexuosa*.

¹² En gris se muestran comandos de R.

Hay clases básicas de objetos que se pueden usar para hacer operaciones y construir otros objetos: **factor**¹³, **función**¹⁴, **lista**¹⁵, **marco de datos**¹⁶ (**data frame**), **matriz** y **vector**.

Más detalles acerca de las funciones básicas de R pueden encontrarse en el manual “R para principiantes” (http://cran.r-project.org/doc/contrib/rdebuts_es.pdf).

13 Variable categórica con niveles.

14 Una función se define de modo intuitivo, con la palabra “function”, una o más variables entre paréntesis, y la función correspondiente a continuación, por ejemplo: `mi_funcion <- function(x){x+2}`

15 Una lista es un vector genérico, donde cada elemento puede ser un objeto distinto; los vectores en cambio poseen elementos sencillos de tipo numérico o carácter.

16 Un marco de datos (*data frame*) está compuesto por un conjunto de vectores similares a los que componen las columnas de una matriz, pero aquí cada vector puede ser de distinta clase, a diferencia de las matrices y los vectores, donde los elementos tienen que ser de una clase determinada, por ejemplo, todos de clase numérica o carácter.

Capítulo II

Objetivos e hipótesis

II. OBJETIVOS E HIPÓTESIS

1. OBJETIVO GENERAL

La sección Algarobia del género *Prosopis* comprende especies arbóreas, la mayoría de las cuales constituyen importantes recursos naturales en las regiones áridas y semiáridas de América. En Argentina, su explotación ha sido fundamentalmente extractiva, lo que condujo a una alarmante reducción de los bosques nativos a lo largo de los últimos dos siglos. Los estudios poblacionales en esta sección han sido enfocados principalmente a escalas gruesas; para comprender mejor y modelar los procesos evolutivos en Algarobia, es necesario realizar investigaciones a escala fina, utilizando las herramientas informáticas disponibles en la actualidad.

El objetivo general de esta tesis es analizar en *Prosopis alba*, una especie de la sección Algarobia, la estructura poblacional a escala fina por medio de marcadores neutros hipervariables y caracteres cuantitativos, y la asociación que ambos tipos de datos presentan con variables geográficas y ecológicas, incorporando información proveniente de un huerto experimental y desarrollando herramientas informáticas adecuadas para los análisis a esta escala.

2. OBJETIVOS ESPECÍFICOS

El estudio se llevó a cabo con muestras provenientes de la provincia de Santiago del Estero. Se analizaron *a*) individuos seleccionados por tolerancia y sensibilidad a la salinidad, de los cuales los tolerantes se clonaron y plantaron en un huerto¹⁷ y *b*) individuos provenientes de una población natural¹⁸.

La distribución de esta especie en el área de estudio se encuentra fragmentada; la extensión original de los bosques naturales se halla reducida a un conjunto de parches y corredores. Entre los parches que pueden encontrarse actualmente, algunos representan restos de poblaciones antiguas y

17 Los individuos sensibles se desarrollaron en vivero con el objetivo de tener un contraste para los datos moleculares de los tolerantes; los mismos no fueron plantados a campo al momento del presente estudio.

18 La población natural se encuentra ubicada en el área de riego del río Dulce, en el centro de la provincia. Los individuos sensibles provienen de árboles madre de la misma zona; los tolerantes, de árboles madre ubicados al sur de la provincia.

II. Objetivos e hipótesis

otros recolonizaciones recientes y renovales. En este contexto resulta de interés analizar las consecuencias de la fragmentación del hábitat sobre la distribución de la variación genética a escala fina.

Los objetivos específicos de esta tesis son:

1. Estudiar con herramientas morfométricas y moleculares individuos seleccionados por tolerancia a la salinidad, utilizando como contraste para los datos moleculares un conjunto de individuos sensibles a la salinidad. Este punto puede dividirse en las siguientes tareas:

- 1.1. Analizar la variación morfológica de caracteres cuantitativos en individuos tolerantes a la salinidad.
- 1.2. Determinar si estos caracteres *a)* se encuentran relacionados con la capacidad de crecimiento en medios salinos y *b)* poseen una base genética.
- 1.3. Calcular la heredabilidad en sentido amplio de estos caracteres.
- 1.4. Comparar los patrones ISSR (Roser, 2010) y SSR, obtenidos con individuos contrastantes (tolerantes y sensibles a la salinidad).
- 1.5. Evaluar el uso de marcadores moleculares como herramientas predictivas en la detección de individuos tolerantes a la salinidad.

Estos objetivos se desarrollan en el capítulo III.

2. Estudiar a escala fina individuos de una población natural. Este punto puede dividirse en las siguientes tareas:

- 2.1. Desarrollar un paquete estadístico para el análisis de datos fenotípicos, genotípicos y ambientales en biología de poblaciones, con especial énfasis en el estudio genético-poblacional a escala fina.
- 2.2. Analizar si la estructura genética espacial presente en la población corresponde a la esperada bajo un modelo de aislamiento por distancia, estudiando la variación de una estima de coancestría con la distancia geográfica entre pares de individuos.
- 2.3. Evaluar en los patrones moleculares de la población la presencia de señales que permitan inferir cambios demográficos recientes.

II. Objetivos e hipótesis

- 2.4. Explorar los patrones espaciales a escala fina en caracteres cuantitativos (rasgos foliares y de crecimiento), de forma global y local.
- 2.5. Evaluar si algunos de los caracteres analizados se encuentran espacialmente autocorrelacionados a la escala estudiada.
- 2.6. Estudiar posibles relaciones entre los patrones espaciales de las variables cuantitativas, genéticas y ambientales.

Estos objetivos se desarrollan en el capítulo V.

3. HIPÓTESIS

1. Hipótesis para el experimento con individuos de *P. alba* seleccionados por tolerancia a la salinidad (capítulo III)

1. La variación observada en los caracteres cuantitativos analizados en los individuos tolerantes del huerto posee una base genética.
2. Los individuos estudiados muestran diferenciación para marcadores neutros en base a su tolerancia a la salinidad.
3. Esta variación podría ser explicada al menos en parte por el proceso de selección, y no meramente por un efecto de estructuración poblacional.

2. Hipótesis para la población natural de *P. alba* (capítulo V)

1. La estructura genética espacial presente en la población natural corresponde a la esperada bajo un modelo de aislamiento por distancia.
2. Los patrones moleculares de la población permiten inferir la presencia de señales de cambios demográficos recientes.
3. Los caracteres cuantitativos analizados se encuentran espacialmente autocorrelacionados a escala fina.
4. Los patrones espaciales de las variables cuantitativas muestran relaciones con variables genéticas y ambientales.

Capítulo III

Estudios genéticos y morfométricos en un huerto experimental con clones de *Prosopis alba* seleccionados por tolerancia a la salinidad



III. ESTUDIOS GENÉTICOS Y MORFOMÉTRICOS EN UN HUERTO EXPERIMENTAL CON CLONES DE *PROSOPIS ALBA*, SELECCIONADOS POR TOLERANCIA A LA SALINIDAD

1. RESUMEN

La especie *Prosopis alba* mostró en estudios previos una alta tolerancia a la salinidad, lo que permitiría utilizarla en programas de forestación de suelos salinos. En el presente capítulo se continuó la investigación iniciada en la tesis de licenciatura de Roser (2010), en la que se caracterizó, por medio de técnicas morfométricas y moleculares, una plantación de 21 clones tolerantes a la salinidad (ST) replicados en bloques al azar. Para el estudio molecular se utilizaron marcadores ISSR, analizando además mediante esta técnica un conjunto de 20 individuos sensibles a la salinidad (SS) crecidos en vivero, con el fin de comparar de modo contrastante los patrones moleculares obtenidos. Con los datos del análisis de ISSR se estudió si el proceso de selección utilizado (tolerancia a la salinidad) fue efectivo en generar una diferenciación entre los individuos tolerantes y sensibles, detectable mediante marcadores moleculares. Todos estos análisis se ampliaron en el presente capítulo para concluir la investigación. Se estudiaron para ello los patrones de heredabilidad de los rasgos de hoja, previamente medidos en los individuos ST, y se genotiparon los árboles ST y SS mediante la técnica SSR. Los resultados obtenidos con ISSR y SSR se compararon para evaluar la consistencia entre ambas caracterizaciones moleculares. La mayor parte de los rasgos morfológicos revelaron diferencias altamente significativas entre los clones ST. Los caracteres largo, ancho y área del foliólulo, y número de pares de foliólulos por pinna mostraron los valores de heredabilidad más altos. El análisis de ISSR permitió detectar 89 loci, 91% de ellos polimórficos al 1%. Los grupos ST y SS difirieron entre sí significativamente por la frecuencia de 22 de estos loci a nivel individual, siendo 12 de ellos significativos a nivel matriz. El análisis de 6 loci SSR mostró un alto polimorfismo y diferencias altamente significativas entre los grupos ST y SS. El análisis de coancestría de estos dos grupos indicó que la diferenciación molecular que presentan no puede ser explicada solamente en base al parentesco. El análisis de coinerencia mostró consistencia en los agrupamientos moleculares basados en ISSR y SSR.

2. INTRODUCCIÓN

El impacto negativo de la salinidad sobre el crecimiento de plantas, en zonas áridas irrigadas y no irrigadas del mundo, continúa siendo un problema importante. Mundialmente se estima que más de 800 millones de hectáreas se encuentran afectadas por salinidad (Rains, 1991) y las actividades agrícolas tienden a agravar este escenario (Flowers y Yeo, 1995). La necesidad de incrementar la productividad en suelos salinos ha estimulado la investigación en torno a la selección de genotipos tolerantes. Epstein y col. (1979) realizaron los primeros esfuerzos para identificar plantas tolerantes a una salinidad equivalente a la del agua de mar, para cultivos comerciales tales como trigo y cebada. Más recientemente, Qadir y col. (2008) reconocieron promisorias ciertas accesiones de los géneros *Acacia*, *Eucalyptus* y *Melaleuca* para actividades silvopastoriles en suelos salinos.

Argentina está considerado el tercer país luego de Rusia y Australia en términos de suelos afectados por la salinidad (Lavado, 2008; Taleisnik y López Launestein, 2011). La mayor parte del trabajo local orientado a la tolerancia a la salinidad en especies forestales ha sido enfocado sobre flora nativa y exótica, principalmente pertenecientes a los géneros *Prosopis* (Ahmad y col., 1994; Baker y col., 1995; Cony y Trione, 1998; Velarde y col., 2003; Meloni y col., 2008a; Meloni y Martinez, 2009), *Schinopsis* (Carnevale y col., 2004; Meloni y col., 2008b), *Vitis* (Cavagnaro y col., 2006) y *Atriplex* (Aiazzi y col., 2002; Taleisnik y López Launestein, 2011).

Un trabajo inicial que evaluó la tolerancia a la salinidad en *Prosopis* (Felker y col., 1981), encontró que accesiones de *P. tamarugo* y *P. pallida* de Hawai tenían la habilidad de crecer en una salinidad equivalente a la del agua de mar. Un trabajo posterior (Rhodes y Felker, 1988), que examinó una amplia base de germoplasma, encontró que *P. juliflora* tenía la más alta combinación de tolerancia a la salinidad y crecimiento entre las especies consideradas, y que *P. alba* podría tener la suficiente variabilidad para permitir seleccionar plantas individuales capaces de crecer en salinidades equivalentes a la del agua de mar.

Dado que las especies de *Prosopis* son polinizadas por insectos y mayormente de fecundación cruzada (Hunziker y col., 1986; Bessega y col., 2000; Felker y col. 2001; Bessega y col., 2011), es esperable que las semillas del mismo árbol madre tengan alta probabilidad de haber sido producto

III. Estudios genéticos y morfométricos en un huerto experimental

de la cruza con múltiples plantas masculinas, resultando en una alta variabilidad entre la progenie. Debido a esto, en *Prosopis* la propagación clonal de plántulas seleccionadas es una herramienta importante para el mejoramiento genético en base a características útiles, tales como la alta tolerancia a la salinidad. Trabajos previos reportaron clones de *P. juliflora* para forma erecta y crecimiento rápido en Haití (Wojtusik y col., 1993), para producción de biomasa y forma en India (Harsh y col., 1996), y clones multipropósito del complejo *P. pallida*/*P. juliflora* en Perú (Alban y col., 2002).

En tiempos recientes, ha tomado relevancia el esfuerzo por establecer plantaciones de *P. alba*, siguiendo distintos criterios de selección, entre ellos la tolerancia a la salinidad. Algunas de las primeras pruebas de progenie de *P. alba* (Felker y col. 1983a,b) se llevaron a cabo en el sur de California usando familias de medios hermanos colectadas por Simpson y Solbrig en 1977. Felker y col. (2001) seleccionaron 12 clones individuales de *P. alba* en base a la producción y al sabor de las vainas, y al crecimiento en altura. Posteriormente, Velarde y col. (2003) obtuvieron 21 clones de *P. alba* tolerantes a la salinidad, provenientes de semillas colectadas en 9 procedencias de un área salina de Santiago del Estero. Ewens y col. (2012) demostraron diferencias altamente significativas entre estos clones para biomasa, diámetro y altura, luego de 5 años de crecimiento a campo.

El proceso de identificación clonal en varias especies forestales se ha basado tradicionalmente en el estudio de características morfológicas y fenológicas. La morfología de hoja, espina y fruto son el criterio principal para la identificación de especies de la sección Algarobia del género *Prosopis*, por lo que es esperable que estos caracteres cuantitativos también puedan ser útiles para la caracterización clonal. Los marcadores moleculares son otra de las herramientas útiles para este propósito. En los últimos años se utilizaron exitosamente diferentes tipos de marcadores moleculares en el estudio de especies forestales. Entre las técnicas disponibles, los marcadores ISSR (Inter-Simple Sequence Repeats) han mostrado ser una herramienta capaz de proveer un medio poderoso, rápido, simple, reproducible y poco costoso en la caracterización y análisis de la variabilidad genética de accesiones, y en la identificación de cultivares y variedades (Kumar y col., 2009; Monte Corvo y col., 2011) y clones de diferentes especies forestales (De Pasquale y col., 2006; Chen y col., 2011). Los SSR (Simple Sequence Repeats) o microsatélites son otro tipo de marcadores ampliamente utilizados para este tipo de estudios, debido a su naturaleza codominante, reproducibilidad y alto contenido informativo (UPOV, 1995; Marcucci Poltri y col., 2003).

III. Estudios genéticos y morfométricos en un huerto experimental

En este capítulo se analizó la variación de caracteres cuantitativos y moleculares en individuos tolerantes a la salinidad de *P. alba*. La investigación aquí presentada constituye un punto de partida para futuros estudios genómicos en este importante recurso natural de regiones áridas.

3. OBJETIVOS

- a. Analizar la variación morfológica de caracteres cuantitativos en individuos tolerantes a la salinidad.
- b. Determinar si estos caracteres *i)* se encuentran relacionados con la capacidad de crecimiento en medios salinos y *ii)* poseen una base genética.
- c. Calcular la heredabilidad en sentido amplio de estos caracteres.
- d. Comparar los patrones ISSR (Roser, 2010) y SSR, obtenidos con individuos contrastantes (tolerantes y sensibles a la salinidad).
- e. Evaluar el uso de marcadores moleculares como herramientas predictivas en la detección de individuos tolerantes a la salinidad.

4. MATERIALES Y MÉTODOS

4.1. Obtención de individuos tolerantes a la salinidad

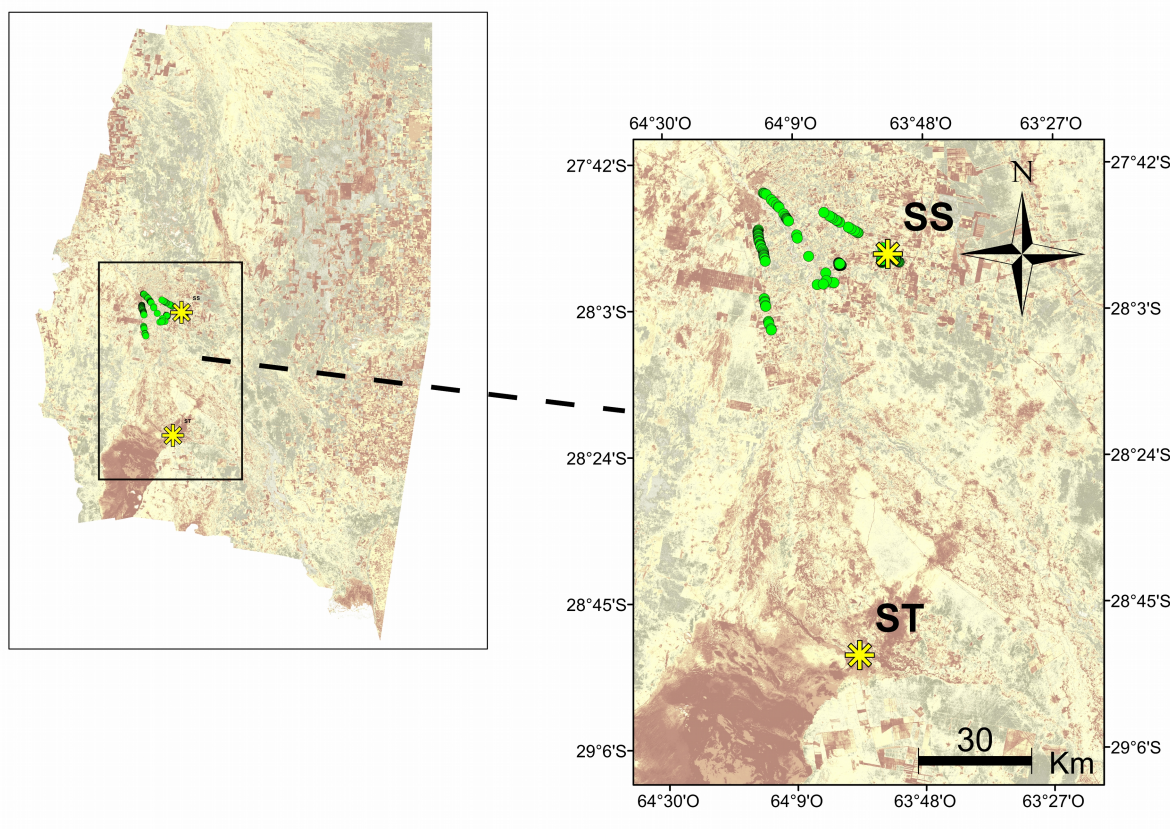
Se utilizaron 21 clones tolerantes a la salinidad (clones ST, Tabla 3.1), seleccionados por Velarde y col. (2003) y Felker y col. (2008). Los árboles madre utilizados para el ensayo se obtuvieron a partir de semillas pertenecientes a 7 familias de distinto origen (Tabla 3.1), colectadas en la llanura de inundación del río Saladillo (Santiago del Estero, 28°52'36.85"S - 63°58'46.36"O, Figura 3.1). A pesar de que estas familias derivaran de árboles identificados taxonómicamente como *P. alba*, cuatro de los clones (ST3, ST7, ST12 y ST15) exhibieron características morfológicas

III. Estudios genéticos y morfométricos en un huerto experimental

híbridas, lo que sugiere que el donante del polen pudo haber sido *P. ruscifolia* o *P. vinalillo* (considerado un híbrido natural entre estas dos especies).

TABLA 3.1 Códigos individuales y abreviaturas (Abr.) usadas en esta parte para los árboles tolerantes y sensibles. Los individuos tolerantes se codificaron con las abreviaturas dadas en Felker y col. (2008) (O: origen, B: bloque en sistema hidropónico, P: árbol en sistema hidropónico). Los individuos sensibles se codificaron siguiendo una nomenclatura similar (M: árbol madre, B: bloque en sistema hidropónico, P: árbol en sistema hidropónico).

Tolerantes		Sensibles	
Árbol	Abr.	Árbol	Abr.
O2B4P1	ST1	M1B12P6	SS1
O7B3P11	ST2	M3B1P3	SS2
O7B4P4	ST3	M1B2P2	SS3
O6B1P12	ST4	M4B9P2	SS4
O6B1P11	ST5	M4B2P6	SS5
O7B4P11	ST6	M8B15P2	SS6
O4B1P6	ST7	M1B9P7	SS7
O9B4P14	ST8	M8B13P3	SS8
O3B4P8	ST9	M4B5P5	SS9
O6B4P4	ST10	M8B13P8	SS10
O3B1P13	ST11	M9B15P4	SS11
O3B4P7	ST12	M5B4P3	SS12
O8B3P15	ST13	M9B9P7	SS13
O3B4P12	ST14	M5B13P7	SS14
O7B4P6	ST15	M7B5P6	SS15
O6B1P13	ST16	M7B10P5	SS16
O7B3P3	ST17	M5B1P4	SS17
O7B3P10	ST18	M3B12P3	SS18
O8B3P14	ST19	M9B14P6	SS19
O7B3P8	ST20	M9B8P3	SS20
O7B4P8	ST21	M9B7P7	SS21
-----	-----	M7B3P5	SS22



Tras los procesos de selección y propagación vegetativa, Felker y col. (2008) plantaron los clones en un huerto experimental (diseño completamente aleatorizado con 5 bloques) ubicado en la Estación Experimental Fernández (Santiago del Estero) en enero de 2004 (Figura 3.2). Cada individuo se encontró representado a través de un clon por bloque, para poder tener en cuenta efectos de bloque en la diferenciación fenotípica entre clones. El suelo del huerto experimental mostró en un análisis previo un gradiente de pH natural, con un rango promedio de 9.35 en el bloque 1 hasta 10.3 en el bloque 5 (Figura 3.3).



FIGURA 3.2 Huerto experimental con los clones tolerantes a la salinidad en la Estación Experimental Fernández, Santiago del Estero (imagen tomada de Roser, 2010).

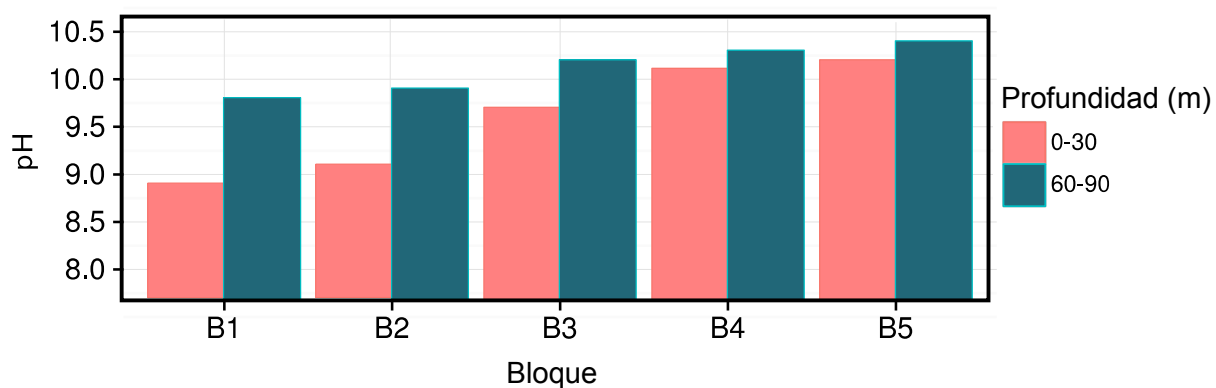


FIGURA 3.3 pH en los bloques del huerto experimental a 0-30 y 60-90 m de profundidad (rango de pH a profundidad promedio: 9.35-10.3). Las muestras de suelo fueron analizadas por INTA-Estación Experimental Santiago del Estero (información provista por el Ing. Mauricio Ewens).

4.2. Obtención de individuos sensibles a la salinidad

Se seleccionaron 22 plantines sensibles a la salinidad¹⁹ (individuos SS, Tabla 3.1) en octubre de 2007, provenientes de semillas colectadas en el Departamento de Robles (Santiago del Estero, 27°55'04"S- 63°53'43"O, Figura 3.1). Para la selección, se utilizaron cajas negras de plástico que se dividieron en 128 celdas, cada una con 30 cm³ de capacidad, distribuidas en arreglos de 8 x 16. Cada cavidad se rellenó con una mezcla de vermiculita-arena (1:1) y se pusieron dos semillas pre-germinadas por celda, para asegurar la presencia de plántula; a los 4 días de la emergencia en el caso que ambas sobrevivieran, se eliminó una. De esta manera se contó al inicio del experimento con 1280 plantines.

Las bandejas se colocaron dentro de una cuba de 163 x 114 x 15.5 cm (Figura 3.4), inundada diariamente con una solución nutritiva con la siguiente composición (mg/l): N— 190, P—35, K— 210, Mg—45, B—0.2, Cu—0.1, Fe—1, Mn—0.5, Mo—0.005, Ca—150 y Zn—0.15, S—70, B— 0.5, Mo—0.05. El exceso de solución era drenado en un tanque de 500 l, desde el cual la solución se bombeaba nuevamente para rellenar la cuba. La temperatura del aire se midió diariamente y la conductividad eléctrica de la solución se registró una vez por semana. El experimento se basó en la habilidad de los plantines para crecer bajo un medio con diferentes conductividades eléctricas.



FIGURA 3.4 Sistema utilizado para la obtención de los individuos sensibles a la salinidad (imagen tomada de Roser, 2010).

¹⁹ La selección fue realizada por el Ing. Mauricio Ewens en la Estación Experimental Fernández, Santiago del Estero.

III. Estudios genéticos y morfométricos en un huerto experimental

Se registraron a intervalos regulares cuatro mediciones de altura (h_1 a h_4), tal como se describe en la Figura 3.5. Para ello, se midió cada plantín del cuello al ápice con una regla graduada al milímetro más cercano. La habilidad de crecer durante la primera fase se estimó como $a = h_2 - h_1$, durante la segunda como $b = h_3 - h_2$, y durante la tercera como $c = h_4 - h_3$. Para una planta sensible a la salinidad, $b - a < 0$, ya que el crecimiento se detiene luego de incrementar la conductividad del medio. En contraste, la reducción de la conductividad puede determinar que el plantín crezca de nuevo durante la tercera fase y la diferencia $c - b$ puede ser positiva. Una diferencia $c - b < 0$ indica que el plantín no es capaz de recuperarse ante una condición de baja conductividad eléctrica. En resumen, cuando $(b - a)/(c - b) \leq 0$, la planta se consideró sensible a la salinidad.

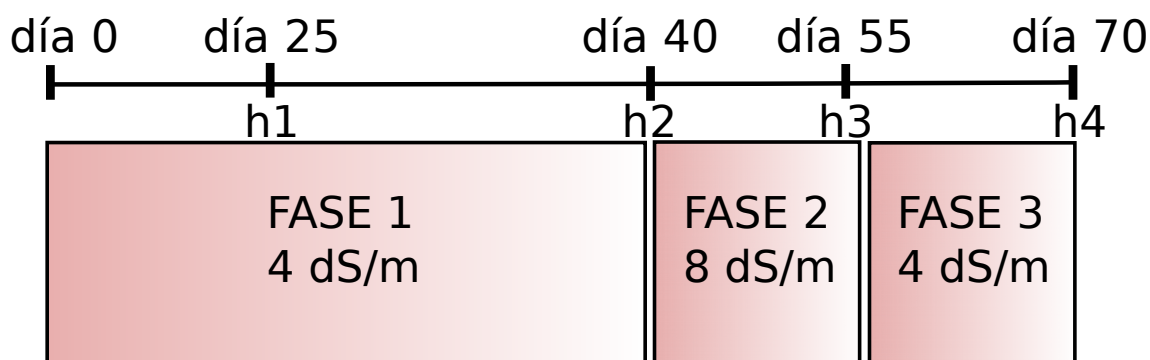


FIGURA 3.5 Representación esquemática del sistema hidropónico utilizado en la selección de los clones SS. Entre los días 0 y 40 (fase 1) la conductividad de la solución suministrada se mantuvo baja (4 dS/m); entre los días 41 y 55 (fase 2) la conductividad fue incrementada a 8 dS/m; y entre los días 56 y 70 (fase 3) la conductividad fue reducida nuevamente a 4 dS/m. Se tomaron mediciones de altura en los días 25 (h_1), 40 (h_2), 55 (h_3) y 70 (h_4). Las plantas fueron clasificadas como sensibles si no crecieron, crecieron más despacio o mostraron muerte del ápice durante la fase 2, y crecieron nuevamente durante la fase 3.

Al final del experimento, se transfirieron las plantas seleccionadas a potes plásticos de 5 litros, y se pusieron a crecer en invernadero. Las características morfológicas de los individuos SS no fueron registradas, ya que al momento del estudio no habían sido plantados a campo, y se encontraban en una etapa de crecimiento diferente a la de los individuos ST, ambos factores dificultando la comparación de los rasgos cuantitativos entre individuos contrastantes (ST-SS).

4.3. Métodos morfométricos

4.3.1. Colección de datos morfométricos

Para cada clon, se tomaron muestras de hojas adultas en dos estratos (superior e inferior) y de frutos, que se herborizaron en papel de diario. Con el material de hoja se realizaron mediciones para 11 rasgos de morfología foliar: longitud del peciolo (PEL), número de pares de foliólulos por pinna (NLP), longitud de las pinnas (PIL), número de pinnas (NPI), longitud (LEL), ancho (LEW), relación largo/ancho (LEL/LEW), falcado (LEF), área (LEA), ápice (LEX) y relación ápice/área total (LEX/LEA) de los foliólulos. El procedimiento utilizado para registrar la información se encuentra detallado en el Apéndice 1 (Protocolo 1, página 155).

4.3.2. Análisis estadístico de los datos morfométricos

Las diferencias morfológicas entre clones tolerantes a la salinidad se evaluaron mediante ANOVA y análisis multivariado de la varianza (MANOVA). Los componentes de varianza se estimaron usando máxima verosimilitud restringida (REML), que es un método flexible aplicable a diseños desbalanceados. Para ello se utilizó el siguiente modelo lineal generalizado (GLM):

$$y_{ijk} = \mu + b_i + s_j + (b:s)_{ij} + I_k + \varepsilon_{ijk} \quad \text{(Expresión 4.1)}$$

Donde y_{ijk} representa cada observación individual, μ es la media general, b_i es el efecto fijo bloque, s_j es el efecto fijo estrato, $(b:s)_{ij}$ es la interacción entre los efectos bloque y estrato, I_k es el efecto aleatorio clon, y ε_{ijk} es el efecto aleatorio residual (error).

La heredabilidad en sentido amplio (H^2) se obtuvo a partir de las estimas de varianza calculadas por REML, mediante la siguiente expresión:

$$\hat{H}^2 = \frac{\sigma_E^2}{\sigma_E^2 + \sigma_D^2} \quad \text{(Expresión 4.2)}$$

Donde σ_E^2 y σ_D^2 son los componentes de varianza entre y dentro (residual) de clones,

respectivamente.

Los análisis de ANOVA y MANOVA se llevaron a cabo con el paquete *stats* de R²⁰ (R core team, 2012). El análisis de los componentes de varianza mediante REML se realizó con el paquete *nlme* 3.1-109 (Pinheiro y col., 2013) de R.

4.4. Métodos de análisis molecular

4.4.1. Obtención de datos moleculares

Para los individuos ST y SS se colectó material de hoja joven para el análisis de ADN, que se secó inmediatamente en sílica gel y se almacenó posteriormente a temperatura ambiente hasta el momento de la extracción. Las muestras de ADN se analizaron mediante marcadores ISSR y SSR, utilizando respectivamente los Protocolos 2-A (página 157) y 2-B (página 158) del Apéndice 1. Para el análisis de ISSR, se seleccionaron los cebadores **ISSR1** (AG)₈Y, **ISSR2** (AC)₈G y **AE2** (CA)₈G, que mostraron un alto nivel de bandas por individuo y gran reproducibilidad entre experimentos. Para el análisis de SSR, se seleccionaron seis cebadores: **Mo05** (CT)₃T(CT)₂, **Mo07** (GC)₈, **Mo08** (AC)₉, **Mo09** (TG)₁₇, **Mo13** (GT)₁₀CT(GT)₂ y **Mo16** (CA)₁₂. Estos cebadores, desarrollados previamente para *P. chilensis*, mostraron amplificación cruzada con otras especies (Mottura y col., 2005), funcionando con éxito en *P. alba* (Bessega y col., 2009).

4.4.2. Análisis estadístico de los datos moleculares

A. Análisis de los datos de ISSR

La frecuencia de bandas ISSR entre los individuos ST y SS se comparó por medio de chi-cuadrado, estimando la significación con un test de permutación. Los valores *P* obtenidos se ajustaron por comparaciones múltiples con el método FDR (Benjamini y Hochberg, 1995). Este análisis se realizó con el paquete *stats* de R.

20 En este capítulo se utilizó R versión 2.15.0.

A.1. Análisis de escalado multidimensional (MDS)

Con el conjunto de datos originales se obtuvo una matriz de distancia euclidiana entre los individuos. A partir de esta matriz, se realizó un análisis de MDS con el paquete *smacof* (De Leeuw y Mair, 2009) de R. El método implementado en este paquete está basado en la minimización del estrés métrico, un estadístico de bondad de ajuste, basado en la suma de las diferencias cuadráticas entre el valor de distancia observado y el esperado para cada par de puntos. La estabilidad de la solución del análisis se examinó mediante jackknife sobre las configuraciones. El resultado se graficó para observar si los individuos tienden a agruparse de acuerdo a la tolerancia a la salinidad.

A.2. Análisis entre grupos (BGA)

El análisis de asociación entre los patrones ISSR y la habilidad de tolerar condiciones de alta salinidad es dificultoso, debido a la presencia de más loci que muestras (individuos analizados). La mayor parte de las clasificaciones supervisadas están limitadas al requerimiento de un mayor número de casos (individuos) que variables (marcadores). Por ejemplo, para utilizar correctamente el análisis discriminante tradicional (DA), se deben poseer más casos que variables, idealmente por un factor de 10 o más. Los métodos BGA (Dolédéc y Chessel, 1987) y DA están basados en el mismo principio: encontrar una combinación lineal de variables que definan la dirección en el espacio multidimensional, a lo largo de la cual se maximice la varianza entre los grupos (Truntzer y col., 2007). BGA sin embargo permite superar la limitación relacionada con el número de casos, ya que puede ser usado cuando el número de variables es mayor (Culhane y col., 2002). El análisis BGA se llevó a cabo mediante el paquete *ade4* (Chessel y col., 2004; Dray y Dufour, 2007) de R.

A.3. Análisis de árboles de clasificación

Se realizó un análisis de árboles de clasificación (Breiman y col., 1993), el cual permite identificar la estructura presente en conjuntos de datos de gran dimensión. Para comparar los clones ST y SS mediante este método, se consideró la tolerancia a la salinidad como la variable respuesta (y) y a los patrones ISSR como las variables predictivas (x_i). El análisis de árboles de clasificación consistió en obtener, a partir de todo el conjunto de predictores potenciales, la variable x_i (locus) con mayor capacidad de predecir y (tolerancia a la salinidad), mediante una partición recursiva de la

III. Estudios genéticos y morfométricos en un huerto experimental

muestra (ver Foulkes, 2009, páginas 157-179). El análisis se llevó a cabo con el paquete *rpart* (Therneau y Atkinson, 2011) de R.

B. Análisis de los datos de SSR

Para los marcadores SSR se calculó el número de alelos (A) para cada locus y promedio por locus, y los valores promedio por locus de heterocigosis observada (H_o) y esperada bajo equilibrio de Hardy-Weinberg (H_E) mediante un estimador insesgado (Nei y Chesser, 1983) (Apéndice 2, página 162) con el paquete *hierfstat* 0.04-10 (Goudet, 2005) en R.

B.1. Escalado multidimensional y análisis entre grupos de los marcadores SSR

La asociación de los patrones SSR con la tolerancia a la salinidad se analizó por MDS y BGA aplicando las mismas metodologías previamente descritas para el análisis de ISSR (secciones 4.4.2, A.1 y A.2).

B.2. Análisis de coancestría

Con el fin de determinar si la asociación entre los datos moleculares y la tolerancia a la salinidad puede atribuirse a relaciones entre los individuos dentro de los grupos ST y SS, se estimaron los valores de coancestría entre y dentro de grupos a partir de los patrones SSR. Esto se realizó aplicando el modelo *admixture F-model* de Metropolis-Hastings (Karhunen y Ovaskainen, 2012), con el paquete *RAFM* de R.

4.4.3. Consistencia entre los marcadores ISSR y SSR

La consistencia de la diferenciación molecular brindada por ambos tipos de marcadores (ISSR y SSR) se evaluó mediante dos métodos. Por un lado, se compararon las matrices de ISSR y SSR mediante el test de Mantel (1967), utilizando distancias euclidianas. Por otro lado, se aplicó un Análisis de Coinercia (CIA) (Dolédéc y Chessel, 1994; Culhane, 2003; Dray y col., 2003), para identificar tendencias y estudiar la estructura común entre las dos matrices. Esta aproximación permite comparar pares de fuentes de datos para el mismo conjunto de muestras. El método se basa

en la obtención de ejes que definen las direcciones de máxima covarianza entre ambos conjuntos de datos y en el estudio de la geometría global de la nube de puntos en el espacio definido por estos ejes. El análisis se realizó con el paquete *ade4* de R.

5. RESULTADOS

5.1. Análisis de datos morfométricos

Las estadísticas básicas de las características medidas en los 21 clones ST se muestran en la Tabla 3.2. El coeficiente de variación promedio (*cv*) dentro de cada clon ST toma el valor más alto para PEL (*cv* = 0.440) y el más bajo para LEF (*cv* = 0.028).

TABLA 3.2 Estadísticas básicas (media y coeficiente de variación) de los rasgos morfológicos medidos en los árboles de *P. alba* tolerantes a la salinidad (ST).

Individuo	LEW		LEX		LEX/LEA		LEA		LEF		LEL/LEW	
	Media	cv	Media	cv	Media	cv	Media	cv	Media	cv	Media	cv
ST1	0.188	0.151	0.928	0.078	0.206	0.093	0.136	0.299	0.951	0.026	4.398	0.127
ST2	0.182	0.117	0.942	0.092	0.203	0.095	0.139	0.252	0.954	0.026	4.835	0.119
ST3	0.271	0.185	0.872	0.070	0.194	0.116	0.301	0.422	0.948	0.025	4.711	0.135
ST4	0.185	0.144	0.924	0.082	0.203	0.105	0.144	0.215	0.951	0.025	4.997	0.133
ST5	0.194	0.123	0.901	0.068	0.200	0.084	0.175	0.255	0.950	0.021	5.471	0.096
ST6	0.172	0.110	0.912	0.079	0.200	0.092	0.131	0.174	0.960	0.022	5.106	0.112
ST7	0.357	0.224	0.826	0.071	0.186	0.182	0.465	0.391	0.939	0.032	4.623	0.138
ST8	0.183	0.138	0.908	0.084	0.195	0.131	0.166	0.210	0.955	0.023	6.055	0.157
ST9	0.146	0.155	0.994	0.099	0.212	0.115	0.101	0.254	0.953	0.027	5.571	0.178
ST10	0.200	0.123	0.951	0.217	0.199	0.101	0.181	0.247	0.959	0.029	5.233	0.132
ST11	0.151	0.128	0.951	0.105	0.207	0.130	0.110	0.236	0.953	0.025	5.787	0.146
ST12	0.357	0.244	0.833	0.081	0.187	0.159	0.506	0.319	0.942	0.057	5.032	0.187
ST13	0.205	0.107	0.902	0.073	0.197	0.085	0.167	0.176	0.952	0.026	4.722	0.160
ST14	0.163	0.151	0.970	0.085	0.207	0.113	0.112	0.320	0.960	0.026	4.798	0.130
ST15	0.414	0.214	0.781	0.072	0.177	0.214	0.785	0.375	0.938	0.029	6.010	0.192
ST16	0.186	0.128	0.945	0.090	0.205	0.124	0.147	0.228	0.954	0.024	5.038	0.131
ST17	0.162	0.145	0.965	0.095	0.210	0.140	0.117	0.239	0.956	0.027	5.206	0.113
ST18	0.199	0.201	0.916	0.090	0.201	0.124	0.181	0.297	0.953	0.023	5.366	0.151
ST19	0.206	0.122	0.898	0.076	0.199	0.121	0.165	0.229	0.952	0.045	4.561	0.115
ST20	0.180	0.152	0.933	0.078	0.208	0.122	0.138	0.287	0.948	0.035	4.991	0.115
ST21	0.193	0.090	0.907	0.080	0.195	0.089	0.177	0.223	0.952	0.022	5.626	0.215
Media	0.214	0.150	0.912	0.089	0.200	0.121	0.216	0.269	0.951	0.028	5.149	0.142

TABLA 3.2 (continuación).

Individuo	LEL		PIL		PEL		NLP		NPI	
	Media	cv	Media	cv	Media	cv	Media	cv	Media	cv
ST1	0.821	0.173	98.033	0.184	25.950	0.344	36.567	0.249	2.267	0.230
ST2	0.875	0.156	89.414	0.185	24.241	0.699	34.793	0.148	2.207	0.223
ST3	1.282	0.261	91.708	0.242	25.292	0.387	22.833	0.199	1.750	0.253
ST4	0.915	0.138	90.933	0.152	29.500	0.327	31.133	0.112	2.100	0.229
ST5	1.060	0.149	105.533	0.196	25.586	0.187	32.800	0.157	1.931	0.192
ST6	0.873	0.095	91.944	0.225	30.167	0.312	36.722	0.094	2.500	0.206
ST7	1.628	0.201	91.793	0.182	25.569	0.347	14.000	0.138	1.103	0.281
ST8	1.093	0.137	111.167	0.168	24.600	0.378	39.333	0.165	2.600	0.192
ST9	0.807	0.181	91.333	0.163	25.300	0.341	39.900	0.147	2.167	0.175
ST10	1.043	0.158	112.300	0.147	19.917	0.268	37.200	0.106	2.067	0.177
ST11	0.867	0.157	85.133	0.129	25.650	0.435	38.233	0.111	2.433	0.207
ST12	1.749	0.184	79.633	0.347	19.467	0.910	13.467	0.148	1.367	0.359
ST13	0.958	0.121	108.633	0.156	29.400	0.579	30.400	0.151	2.267	0.230
ST14	0.780	0.186	86.933	0.166	22.267	0.455	37.167	0.194	2.433	0.207
ST15	2.447	0.218	96.917	0.291	30.550	1.047	9.900	0.216	1.033	0.177
ST16	0.929	0.131	86.633	0.190	23.250	0.366	33.833	0.196	2.300	0.203
ST17	0.839	0.141	90.733	0.212	28.483	0.292	38.167	0.199	2.400	0.235
ST18	1.050	0.155	96.600	0.165	25.567	0.366	35.767	0.141	2.300	0.203
ST19	0.938	0.155	95.900	0.252	23.933	0.468	31.367	0.178	2.000	0.322
ST20	0.891	0.161	84.033	0.171	17.250	0.460	30.667	0.143	2.500	0.203
ST21	1.078	0.209	83.250	0.318	21.667	0.276	29.917	0.188	2.750	0.226
Media	1.092	0.165	93.741	0.202	24.934	0.440	31.151	0.161	2.118	0.225

Los ANOVA individuales, realizados para comparar cada característica entre clones ST, mostraron que las diferencias para todos los caracteres son significativas, con excepción de PEL (Tabla 3.3). El MANOVA considerando todas las características demostró diferencias altamente significativas entre clones ($Wilks = 4.50 \cdot 10^{-5}$, $P < 2.2 \cdot 10^{-16}$).

Los análisis de componentes de varianza y de H^2 (heredabilidad en sentido amplio) para cada carácter, indicaron que aquellos con los más altos valores fueron LEL, LEW, LEA y NLP (Tabla 3.3). En ningún caso el intervalo de confianza para H^2 incluyó al cero, indicando valores significativos para todos los caracteres excepto PEL, que mostró el segundo valor más bajo de H^2 (0.014).

III. Estudios genéticos y morfométricos en un huerto experimental

TABLA 3.3 Resultados de ANOVA, componentes de varianza y heredabilidad en sentido amplio (H^2) para los caracteres morfométricos analizados en los árboles tolerantes a la salinidad (ST).

Análisis	Carácter	LEA	LEF	LEL	LEL/ LEW	LEW	LEX	LEX/ LEA	NLP	NPI	PEL	PIL
ANOVA	F	42.800	2.604	44.555	7.799	48.408	7.205	8.429	43.070	17.094	1.328	5.799
	P	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.189	0.000
	Paj	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.189	0.000
Entre clones	estima	0.028	2.6 10 ⁻⁵	0.160	0.212	0.005	0.003	6.6 10 ⁻⁵	76.882	0.209	7.169	70.219
	inferior	0.015	1.1 10 ⁻⁵	0.086	0.112	0.003	0.001	3.3 10 ⁻⁵	41.052	0.109	2.392	33.826
	superior	0.053	5.9 10 ⁻⁵	0.297	0.401	0.010	0.005	1.3 10 ⁻⁴	143.984	0.398	21.489	145.763
Dentro de clones	estima	0.009	7.9 10 ⁻⁴	0.046	0.555	1.7 10 ⁻³	0.008	6.1 10 ⁻⁴	27.391	0.224	151.203	355.724
	inferior	0.008	7.5 10 ⁻⁴	0.043	0.521	1.6 10 ⁻³	0.007	5.8 10 ⁻⁴	24.381	0.199	134.575	316.646
	superior	0.010	8.5 10 ⁻⁴	0.048	0.591	1.8 10 ⁻³	0.008	6.5 10 ⁻⁴	30.772	0.252	169.886	399.626
H^2	estima	0.759	0.032	0.778	0.276	0.758	0.245	0.097	0.737	0.482	0.045	0.165
	inferior	0.613	0.013	0.639	0.159	0.612	0.138	0.049	0.572	0.303	0.014	0.078
	superior	0.862	0.073	0.874	0.435	0.861	0.395	0.184	0.855	0.666	0.138	0.320

NOTA: F estadístico F del ANOVA, P significación del ANOVA P_{aj} . (P ajustado) valores de P ajustados por comparaciones múltiples con el método FDR, H^2 heredabilidad en sentido amplio, *inferior* y *superior* valores inferior y superior, respectivamente, para el intervalo de confianza del 95% del valor estimado. Los valores significativos se muestran en **negrita**.

5.2. Análisis de ISSR

El análisis de ISSR permitió identificar 89 bandas altamente repetibles (Figura 3.6). Un 91% de los loci identificados (81) resultaron polimórficos al 1%.

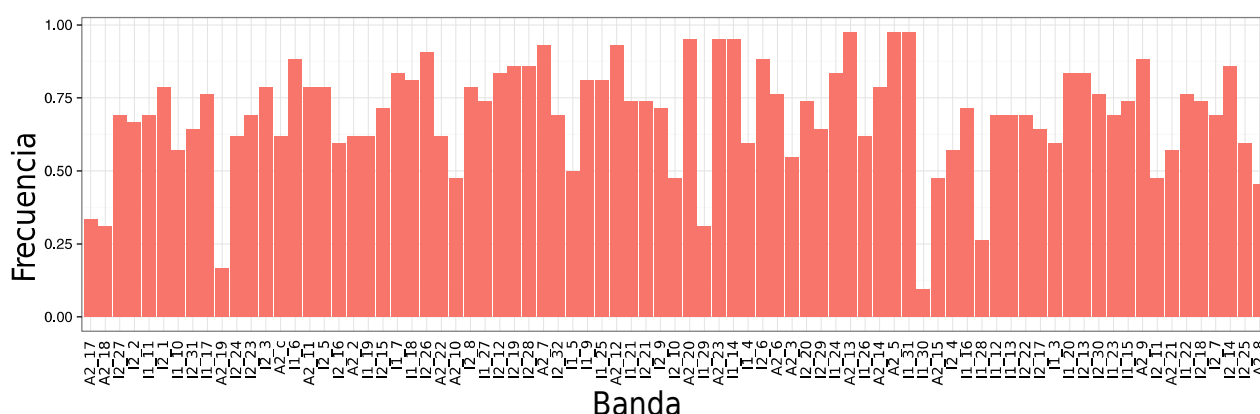


FIGURA 3.6 Frecuencias de bandas para los marcadores ISSR (individuos SS y ST). El eje de las abscisas indica el nombre (arbitrario) dado a cada banda observada.

III. Estudios genéticos y morfométricos en un huerto experimental

La comparación de frecuencias de bandas entre los grupos ST y SS por medio de chi-cuadrado, indicó que las diferencias fueron significativas o altamente significativas para 22 loci a nivel individual y para 12 a nivel tabla (valores P corregidos por comparaciones múltiples) (Tabla 3.4).

TABLA 3.4 Frecuencias de los fenotipos dominantes de los loci ISSR en los individuos de *P. alba* tolerantes (ST) y sensibles (SS) a la salinidad, y significación de la comparación entre grupos por medio de pruebas de chi-cuadrado. Sólo se muestran aquellos loci con diferencias significativas entre grupos a nivel individual. Los valores significativos se muestran en negrita.

Locus	ST	SS	χ^2	P	P. aj
A2_17	0.700	0.000	23.100	5.00 10⁻⁰⁴	0.008
A2_18	0.650	0.000	20.710	5.00 10⁻⁰⁴	0.008
I1_11	1.000	0.409	17.116	5.00 10⁻⁰⁴	0.008
I2_2	0.350	0.955	17.230	5.00 10⁻⁰⁴	0.008
I2_27	0.350	1.000	20.710	5.00 10⁻⁰⁴	0.008
I1_10	0.300	0.818	11.486	1.50 10⁻⁰³	0.015
I2_1	0.550	1.000	12.600	1.50 10⁻⁰³	0.015
I2_31	0.400	0.864	9.808	1.50 10⁻⁰³	0.015
I1_17	0.550	0.955	9.451	2.50 10⁻⁰³	0.022
I2_24	0.850	0.409	8.636	4.50 10⁻⁰³	0.036
A2_19	0.350	0.000	9.240	5.00 10⁻⁰³	0.037
I2_3	0.600	0.955	7.821	6.50 10⁻⁰³	0.044
I2_23	0.900	0.500	7.843	8.50 10⁻⁰³	0.053
A2_c	0.400	0.818	7.769	1.10 10⁻⁰²	0.064
I2_5	0.950	0.636	6.121	1.60 10⁻⁰²	0.086
A2_11	0.950	0.636	6.121	1.80 10⁻⁰²	0.091
I1_6	0.750	1.000	6.243	2.30 10⁻⁰²	0.110
I1_19	0.800	0.455	5.301	2.80 10⁻⁰²	0.121
I2_16	0.400	0.773	6.041	2.85 10⁻⁰²	0.121
A2_2	0.800	0.455	5.301	3.20 10⁻⁰²	0.130
I1_7	0.700	0.955	4.887	4.05 10⁻⁰²	0.153
I2_15	0.550	0.864	5.050	4.15 10⁻⁰²	0.153

NOTA: P valor de significación para la prueba de chi-cuadrado, P_{aj} (P ajustado) valor de significación para la prueba de chi-cuadrado, ajustado por múltiples comparaciones con el método FDR. Los valores significativos se muestran en negrita.

5.2.1. Análisis de escalado multidimensional

El análisis MDS para los datos de ISSR (Figura 3.7) mostró que el eje 1 tiende a separar a los individuos SS (a la izquierda) de los ST (a la derecha). Los individuos con características híbridas (posibles *P. vinallillo*) se disponen en el gráfico de modo disperso entre los individuos ST y no presentaron una tendencia particular. El estrés métrico (basado en 327 iteraciones) fue bajo (0.11) y la estabilidad evaluada por medio de jackknife fue alta (varianza entre/varianza total = 0.99).

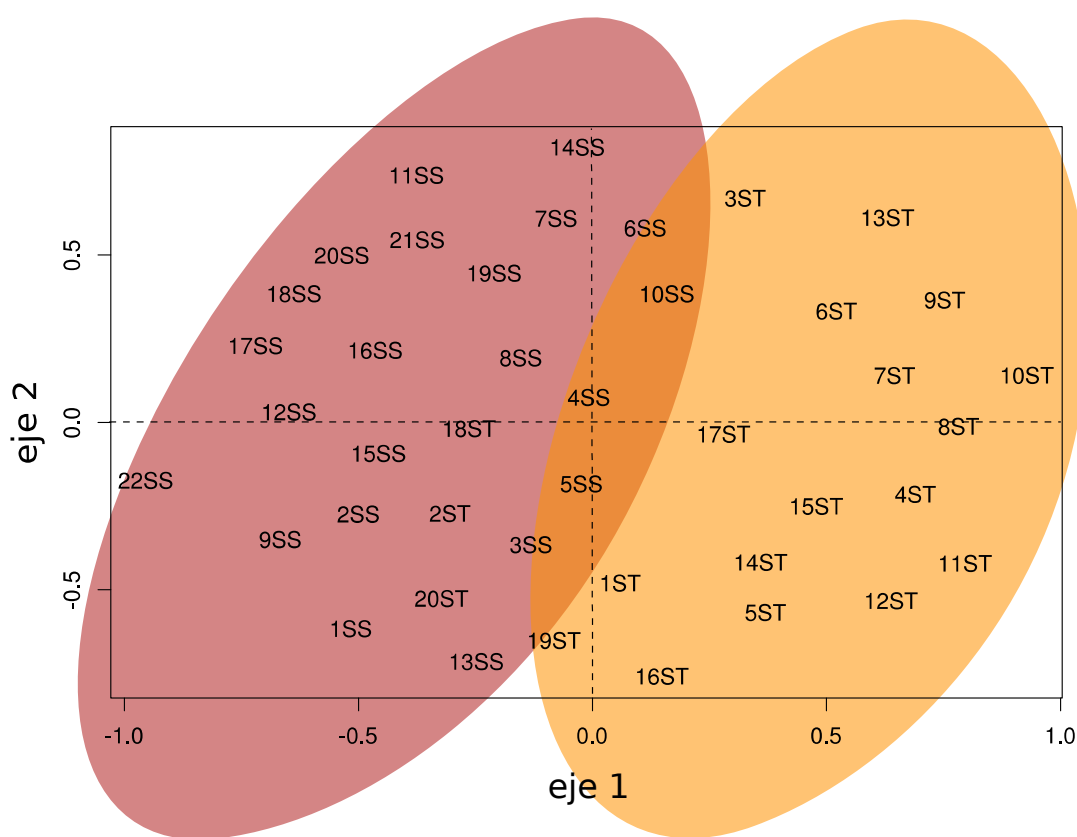


FIGURA 3.7 Gráfico de los individuos tolerantes (ST) y sensibles (SS) a la salinidad de acuerdo a los dos primeros ejes del análisis de escalado multidimensional (MDS), basado en la matriz de distancias euclidianas de los marcadores ISSR. Los colores indican las zonas del gráfico donde se encuentran de modo aproximado los individuos SS (morado) y ST (naranja).

5.2.2. Análisis entre grupos

El análisis BGA basado en los datos de ISSR (Figura 3.8) produjo un eje principal, mostrando diferencias altamente significativas entre los grupos ST y SS (prueba de Monte Carlo basada en 9999 permutaciones, $P = 10^{-4}$).

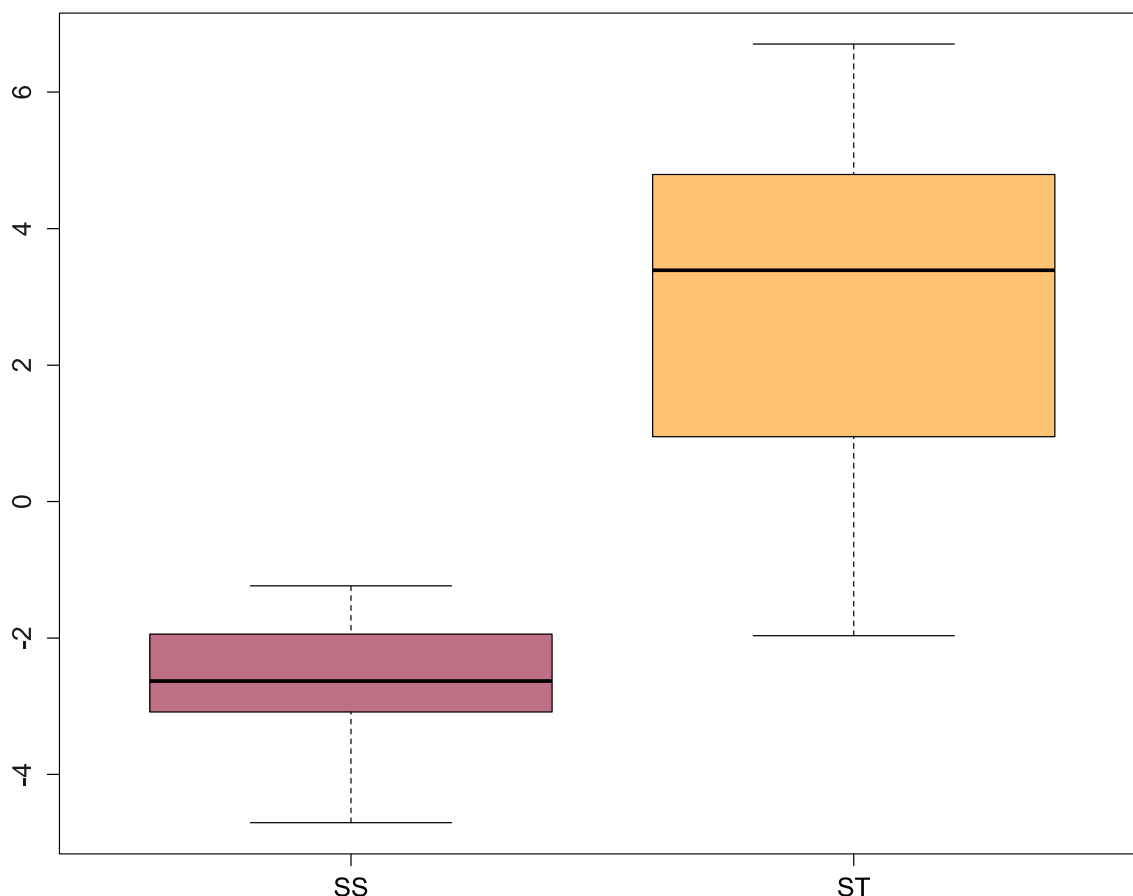


FIGURA 3.8 Gráfico de cajas, mostrando la distribución de los individuos tolerantes (ST) y sensibles (SS) a la salinidad de *P. alba* de acuerdo al primer eje del análisis entre grupos del conjunto de datos ISSR. La línea negra horizontal dentro de cada caja representa la mediana. La parte superior e inferior de la caja indican respectivamente el primer y el tercer cuartil. Las líneas verticales debajo y encima de la caja indican los valores mínimo y máximo, respectivamente.

Los loci con la más alta contribución a la diferenciación entre los individuos tolerantes y sensibles fueron A2_17, A2_18, y A2_19.

5.2.3. Análisis de árboles de clasificación

El algoritmo de árboles de clasificación mostró que la banda A2_17 diferenció con el mayor grado de significación a los dos grupos, encontrándose ausente en todos los individuos SS y presente en el 70% de los ST. Los tres análisis (prueba de chi-cuadrado, BGA y árboles de clasificación) fueron consistentes en identificar a A2_17 como el locus con la más alta contribución a la diferenciación entre los individuos SS y ST.

5.3. Análisis de datos SSR

Los 6 loci SSR estudiados fueron polimórficos al 1%, mostrando entre 2 y 7 alelos, con un promedio de 4.5 alelos por locus. Los valores de heterocigosis observada y esperada promedio por locus fueron $H_o = 0.397$ y $H_e = 0.516$, respectivamente. La Figura 3.9 muestra las frecuencias para cada uno de los marcadores.

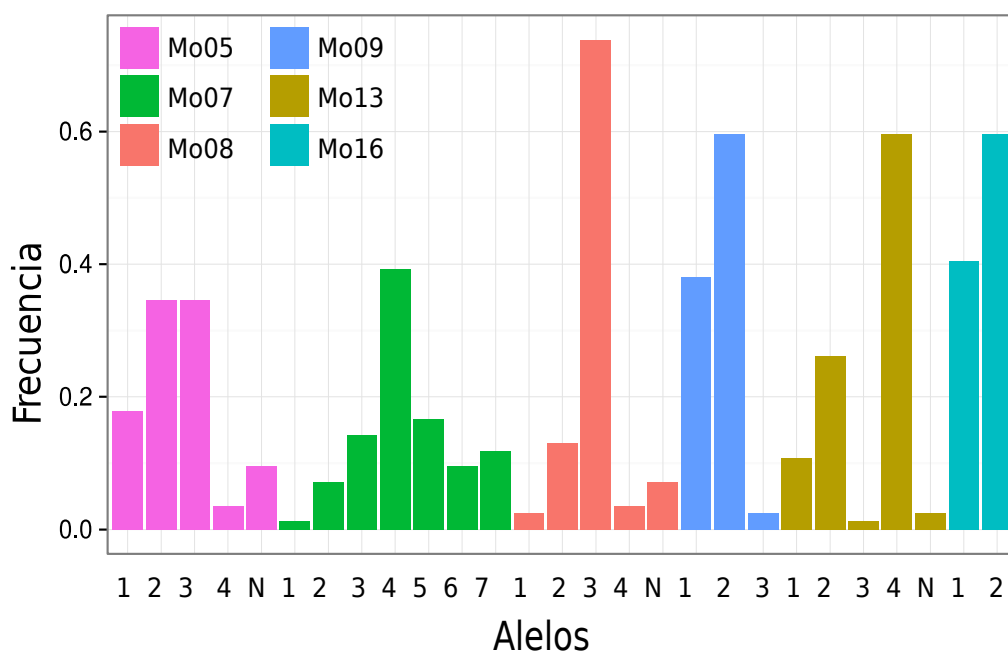


FIGURA 3.9 Frecuencias alélicas para los marcadores SSR (individuos SS y ST). La letra “N” indica un alelo nulo para el locus correspondiente. Los números en el eje de las abscisas indican el nombre (arbitrario) dado a cada alelo observado.

5.3.1. Análisis de escalado multidimensional

El análisis MDS para los datos de SSR (Figura 3.10) fue consistente con el basado en los datos de ISSR (Figura 3.7). Los individuos SS mostraron una tendencia a agruparse a la derecha y los ST a la izquierda del gráfico, y no se observaron tendencias particulares respecto a los individuos con caracteres híbridos, estando dispersos entre los individuos ST.

El estrés métrico (basado en 494 iteraciones) fue bajo (0.07) y la estabilidad evaluada mediante jackknife fue alta (varianza entre/varianza total = 0.87).

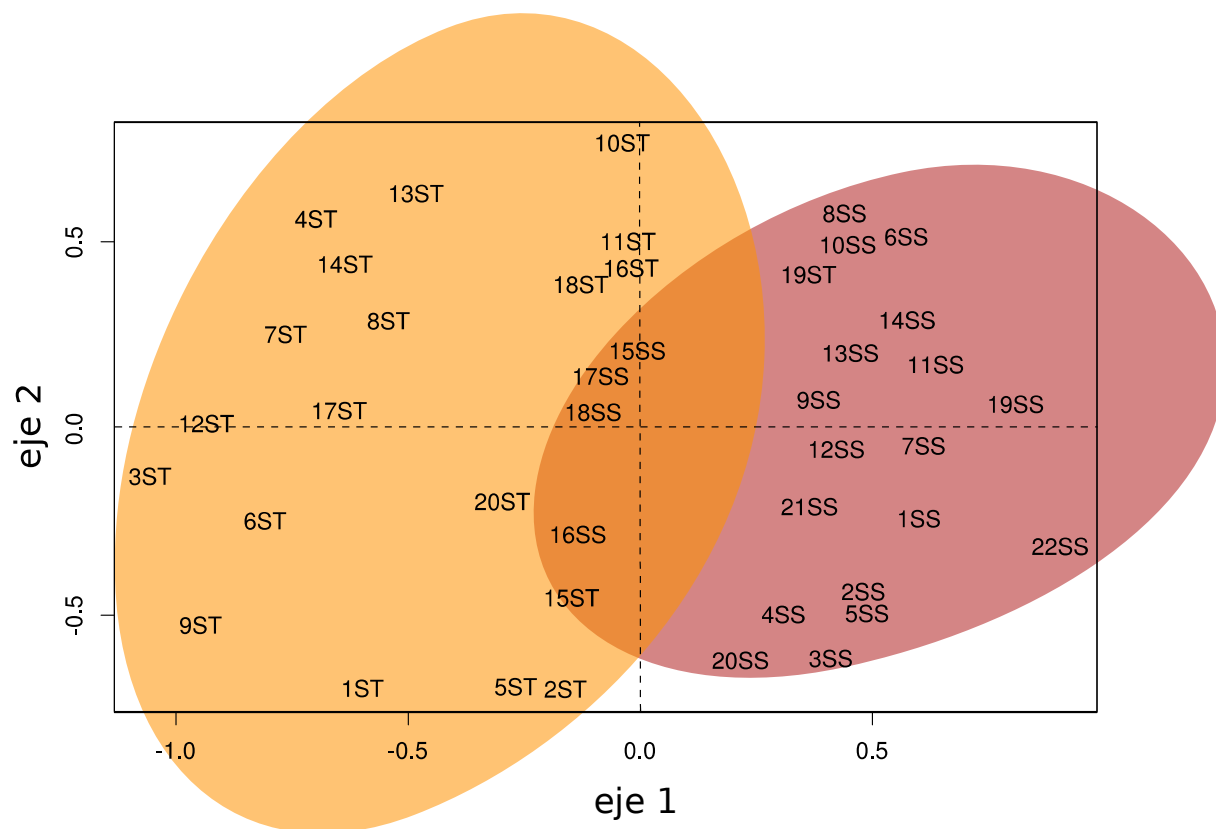


FIGURA 3.10 Gráfico de los individuos tolerantes (ST) y sensibles (SS) a la salinidad de acuerdo a los dos primeros ejes del análisis de escalado multidimensional (MDS) basado en la matriz de distancias euclidianas de los marcadores SSR. Los colores indican las zonas del gráfico donde se encuentran de modo aproximado los individuos SS (morado) y ST (naranja).

5.3.2. Análisis entre grupos

El análisis BGA basado en los datos de SSR (Figura 3.11) mostró diferencias altamente significativas entre los grupos ST y SS (prueba de Monte Carlo basada en 9999 permutaciones, $P=10^{-4}$). Los alelos con la más alta contribución a la diferenciación entre los individuos ST y SS fueron Mo09.1, Mo16.1, Mo16.2 y Mo09.2 (Figura 3.9).

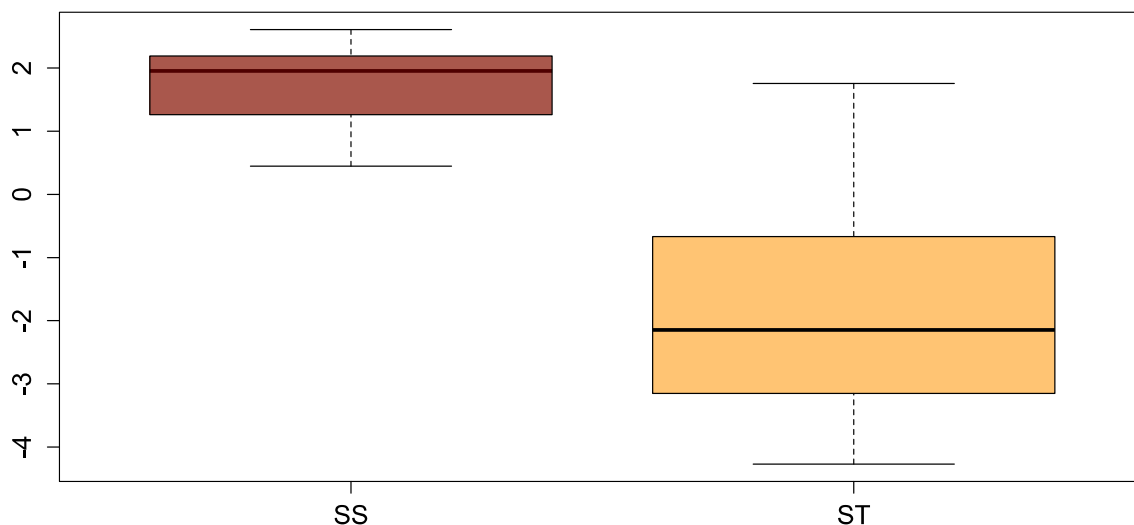


FIGURA 3.11 Gráfico de cajas, mostrando la distribución de los individuos tolerantes (ST) y sensibles (SS) a la salinidad de acuerdo al primer eje del análisis BGA para los datos SSR. La línea negra horizontal dentro de cada caja representa la mediana. La parte superior e inferior de la caja indican respectivamente el primer y el tercer cuartil. Las líneas verticales debajo y encima de la caja indican los valores mínimo y máximo, respectivamente.

5.3.3. Análisis de coancestría

El análisis de coancestría para las clases ST y SS (Tabla 3.5), dentro y entre grupos, sugiere que la variación en tolerancia a la salinidad a gran escala tiende a encontrarse más bien dentro de las familias. A pesar de que la estima más alta se observó dentro del grupo sensible, la coancestría dentro del grupo tolerante es similar a la observada entre los grupos. Además, los intervalos de confianza de las estimas de coancestría se solapan ampliamente, indicando que los individuos dentro del grupo tolerante no se encuentran significativamente más relacionados entre sí que los individuos dentro del otro grupo.

III. Estudios genéticos y morfométricos en un huerto experimental

TABLA 3.5 Estimaciones de coancestría entre individuos pertenecientes al mismo grupo (Dentro) y diferentes grupos (Entre) de tolerancia a la salinidad (sensibles-tolerantes).

Coancestría		Mediana	Intervalo de Confianza	
			inferior	superior
Dentro	tolerantes	0.196	0.123	0.440
	sensibles	0.386	0.241	0.600
Entre		0.145	0.084	0.257

5.4. Congruencia entre los conjuntos de datos ISSR y SSR

Las estructuras de las matrices de distancias entre individuos obtenidas a partir de los datos de ISSR y SSR, fueron consistentes de acuerdo al test de Mantel ($r = 0.315$, $P = 10^{-04}$). Por otra parte, el análisis de coherencia, realizado entre los PCA con los datos de ISSR y SSR, mostró que estos dos conjuntos de datos son similares ($RV = 0.437$, basado en 9999 permutaciones, $P = 2 \cdot 10^{-04}$). El gráfico de coherencia de la Figura 3.12, sugiere que los individuos más diferenciados en términos de tolerancia a la salinidad y patrones moleculares serían de un lado ST9 y ST12, y del otro SS22.

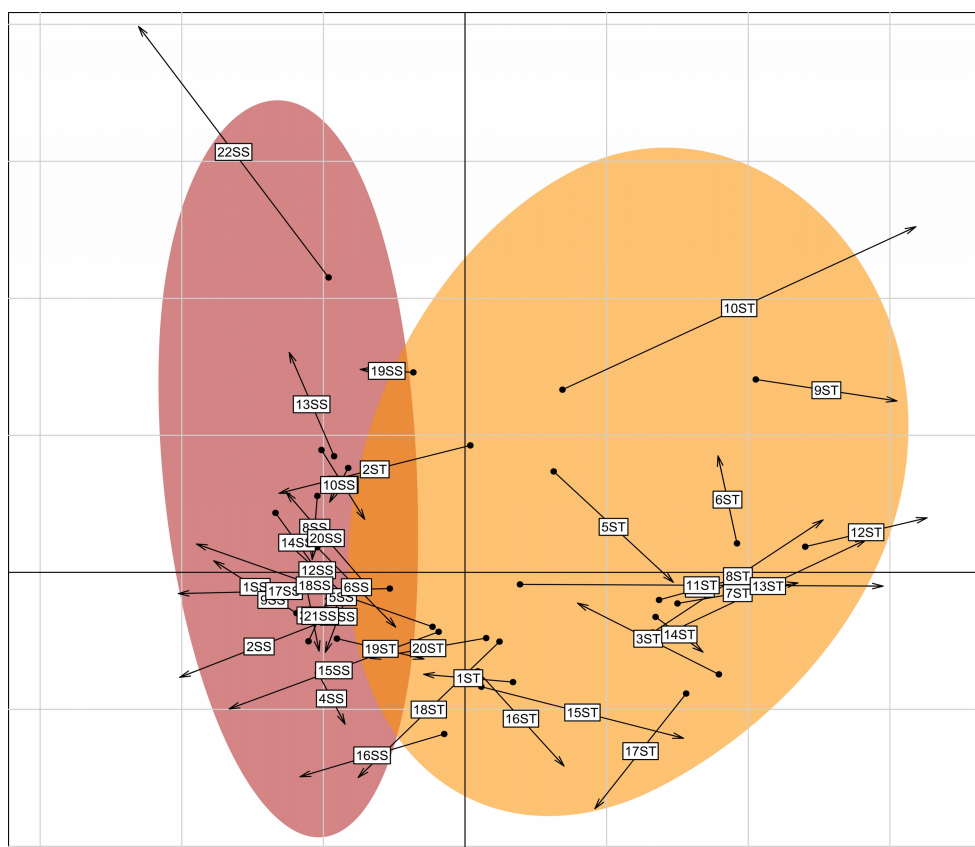


FIGURA 3.12 Análisis de coinercia para los patrones SSR e ISSR en los individuos analizados. Los puntos representan la proyección de los individuos en los dos primeros ejes del PCA basado en el conjunto de datos de SSR, mientras que las puntas de las flechas representan la proyección correspondiente basada en los datos de ISSR. Los ejes horizontal y vertical corresponden respectivamente al primer y segundo eje de los análisis de PCA. Cada punto y punta de flecha se encuentran unidos por una línea, cuya longitud es proporcional a la divergencia entre muestras en los dos conjuntos de datos. Los colores indican las zonas del gráfico donde se encuentran de modo aproximado los individuos SS (morado) y ST (naranja).

6. DISCUSIÓN

La forestación con especies tolerantes a condiciones salinas, principalmente aquellas fijadoras de nitrógeno, es una de las alternativas productivas más atractivas para restaurar suelos degradados por efecto de la salinidad (Qadir y col., 2008), adoptada en diferentes partes del mundo. Las especies fijadoras de nitrógeno del género *Prosopis* son promisorias para este propósito en regiones subtropicales (Dutton, 1989; Geesing y col., 2000), debido a su importancia ecológica y económica. Entre otros atributos, se pueden mencionar una alta tolerancia a condiciones salinas, presencia de variedad de adaptaciones a diferentes ambientes y un alto valor económico de diversos productos (Cony, 1995; Felker y col., 2008), teniendo potencial para conformar el eje biológico de nuevos sistemas productivos (Roig, 1993; Fagg y Stewart, 1994).

La sección Algarobia incluye las especies multipropósito del género más valiosas desde el punto de vista productivo, capaces de proveer madera para la fabricación de muebles, leña, carbón vegetal, forraje y materias primas para la elaboración de alimentos destinados al consumo humano, entre otras cosas (Pasiecznik y col., 2001). Cony (1995) propuso que el uso de especies multipropósito de árboles, tales como *P. chilensis* y *P. flexuosa*, debería ser considerado en programas de reforestación en la eco-región del Monte, mientras que el uso de *P. alba* sería promisorio para la eco-región Chaqueña (Felker y col., 2001, 2008). Las especies de esta sección poseen además un importante potencial para ser incluidas en programas de mejoramiento genético, tal como revelaron distintos estudios, que sugieren una alta heredabilidad de diferentes características fisiológicas y de crecimiento (Cony, 1996; Cony y Trione, 1998; Felker y col., 2001; Bessega y col., 2009, 2011).

Velarde y col. (2003) y Felker y col. (2008) seleccionaron y clonaron individuos de *P. alba* capaces de crecer bajo condiciones salinas y los plantaron en un huerto experimental con un diseño de bloques al azar. En este capítulo se estudiaron estos clones con el fin de conocer la proporción de variación morfológica explicada por causas genéticas y ambientales en el huerto, y evaluar la capacidad de los marcadores de identificar individuos tolerantes a la salinidad.

Los análisis univariados mostraron, para la mayor parte de las características morfológicas, diferencias altamente significativas entre clones que no pueden ser explicadas solamente por causas

III. Estudios genéticos y morfométricos en un huerto experimental

ambientales. La única excepción fue la longitud del peciolo (PEL), que mostró una alta variación ambiental, tal como pudo evidenciarse por su alto coeficiente de variación (cv) dentro de los clones. La comparación multivariada, considerando todas las características morfométricas, también mostró diferencias altamente significativas entre clones. El análisis de los componentes de la varianza y la heredabilidad en sentido amplio (H^2) para cada carácter indicaron que los rasgos con las más alta heredabilidad fueron NPI, NLP, LEW, LEA, y LEL (estimas de H^2 : 0.482, 0.737, 0.758, 0.759 y 0.778, respectivamente). Este resultado indica que estos caracteres pueden ser útiles para realizar selección sobre caracteres cuantitativos con importancia adaptativa. Ewens y col. (2012) también detectaron diferencias significativas entre estos clones para biomasa, diámetro y altura, con estimas de heredabilidad entre 0.45 y 0.59.

Bessegua y col. (2009) estudiaron los mismos caracteres del presente capítulo, en un ensayo de polinización abierta con individuos no seleccionados, establecido en San Carlos (Santiago del Estero), a unos 38 km al norte del huerto aquí analizado. En su trabajo obtuvieron estimas de heredabilidad en sentido estricto altas para todos los rasgos foliares estudiados. Se pudo observar sin embargo una sobreestimación general de los valores.

El cálculo de parámetros genéticos de variación de caracteres cuantitativos es dificultoso en especies forestales, debido a la complejidad de los diseños experimentales y al largo tiempo que insumen. Por esta razón, en los últimos años los marcadores moleculares se han empezado a utilizar de modo creciente para asistir estudios morfológicos clásicos sobre caracteres cuantitativos, y para caracterizar clones y cultivares de especies domesticadas.

Los estudios moleculares mostraron que los genotipos tolerantes tendieron a diferenciarse de los sensibles en los gráficos de MDS, indicando que los marcadores SSR e ISSR son eficientes para discriminar entre los genotipos ST y SS. Este resultado es consistente con el obtenido por Zeng y col. (2004) en arroz, quienes observaron diferencias significativas en caracteres fisiológicos, relacionados con la tolerancia a la salinidad entre diferentes grupos identificados por medio de marcadores microsatélites.

El análisis BGA mostró una clara diferenciación entre los individuos ST y SS. En particular, tres loci ISSR (A2_17, A2_18 y A2_19) y cuatro alelos SSR (MO09.1, M016.1, MO16.2 y

III. Estudios genéticos y morfométricos en un huerto experimental

MO09.2) dieron cuenta de la mayor parte de la diferenciación entre estos grupos. El análisis de árboles de clasificación para los marcadores ISSR fue consistente con el análisis BGA en identificar A2_17 como el locus con la más alta contribución a la diferenciación entre grupos.

El análisis de coancestría dentro y entre los grupos SS y ST mostró similitud en la coancestría promedio dentro del grupo ST y entre individuos de diferentes grupos. Este hecho y la variación en la tolerancia a la salinidad observada dentro de familias, indican que la diferenciación molecular entre grupos no puede ser explicada solamente en base a fundamentos históricos o familiares.

El análisis de coinercia indicó que los resultados obtenidos a partir de diferentes marcadores moleculares resultaron consistentes, soportando el uso de las técnicas ISSR y SSR como herramientas para asistir en la selección temprana de propiedades fisiológicas útiles. La habilidad para seleccionar plantas basada en el genotipo antes que en el fenotipo es extremadamente atractiva, debido a que se puede evitar muchos de los problemas asociados con la selección fenotípica (Ibitoye y Akin-Idowu, 2010). En particular, la selección de genotipos superiores de *P. alba* es complicada, debido a la necesidad de seleccionar para caracteres multipropósito que compiten por fotosintatos, esto es, crecimiento rápido del tronco, características deseable para la producción de madera y alta producción de vainas (Felker y col., 2001). Debido a que marcadores ligados a caracteres de interés pueden asistir en la selección de progenies portadoras de alelos deseables (Ibitoye y Akin-Idowu, 2010), el resultado aquí obtenido es promisorio y estimula la profundización de los estudios moleculares.

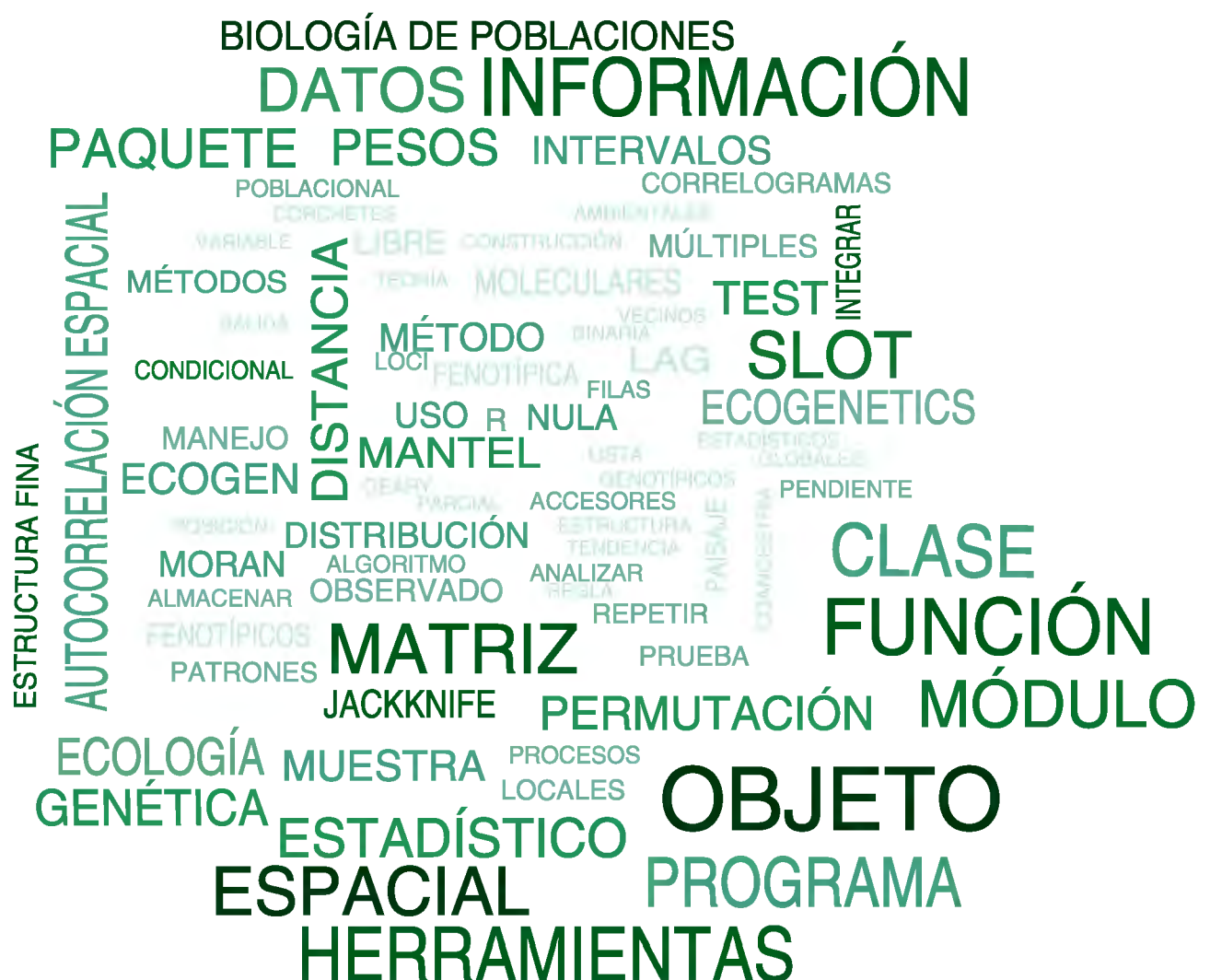
La información genómica de *P. alba* es aún insuficiente para llevar a cabo análisis de asociación amplios, ya que se encuentran disponibles pocos marcadores genéticos. En los últimos años, la cantidad de marcadores disponibles se incrementó a partir de los desarrollos de Bessega y col. (2013) y Pomponio y col. (2015), pero se requiere de un número aún mayor. Los clones aquí analizados son los únicos caracterizados molecularmente para una propiedad fisiológica particular. En la actualidad, se están estudiando características importantes de productividad, tales como propiedades de la madera en especies de *Prosopis* (Pometti y col., 2009, 2010) pero no hay datos a nivel poblacional que puedan relacionarse con marcadores moleculares.

III. Estudios genéticos y morfométricos en un huerto experimental

En resumen, en este capítulo se presentó la primera contribución de marcadores moleculares potencialmente útiles para la identificación de regiones genómicas asociadas con diferentes propiedades fisiológicas en *P. alba*. El desarrollo de un alto número de SSR y de otros marcadores es un paso inicial para estudios genómicos de identificación clonal, pruebas de asociación y mapeo genético.

Capítulo IV

Desarrollo de un paquete estadístico de R para el análisis de datos fenotípicos, genotípicos y ambientales en biología de poblaciones



IV. DESARROLLO DE UN PAQUETE ESTADÍSTICO DE R PARA EL ANÁLISIS DE DATOS FENOTÍPICOS, GENOTÍPICOS Y AMBIENTALES EN BIOLOGÍA DE POBLACIONES

1. RESUMEN

La integración entre la ecología y la genética es un hecho que se ha dado de modo cada vez mayor en las últimas décadas, a la par de las nuevas tecnologías que fueron surgiendo. Un manejo integral de la información poblacional proveniente de distintas fuentes, puede facilitar la realización de estudios espaciales integrales, que involucren datos fenotípicos, genotípicos y ambientales. Herramientas geoestadísticas, que permitan el estudio de variables autocorrelacionadas, pueden ser útiles para el análisis de este tipo de información. A continuación se presenta un paquete estadístico de R elaborado en esta tesis, denominado “EcoGenetics”, útil para el manejo y análisis de información espacial en poblaciones, con especial énfasis en el estudio a escala fina.

2. INTRODUCCIÓN

Los procesos microevolutivos han sido clásicamente estudiados por la genética de poblaciones, luego de la aceptación de la teoría mendeliana de la herencia y el posterior desarrollo de la Teoría Sintética de la Evolución en la década del 40 (Dobzhansky, 1937; Huxley y col., 1942; Mayr, 1942; Simpson, 1944). Desde los 60, la biología de poblaciones emergió como la fusión de la ecología y la genética de poblaciones (Westoby, 2006). El advenimiento de los datos genéticos moleculares para poblaciones en los 60 y las técnicas de PCR y secuenciación en los 80, revolucionaron esta última ciencia. La teoría neutral desarrollada a fines de los 60 (Kimura, 1983) introdujo un importante marco teórico para el análisis de los datos moleculares, dentro de un paradigma evolutivo. La gran cantidad de información genética disponible con la secuenciación masiva y el análisis de SNP, el crecimiento en estadística computacional y en modelos de simulación con el uso de computadoras de escritorio, el uso de internet y el software libre, y la utilización de datos georreferenciados en Sistemas de Información Geográfica (SIG), están produciendo importantes cambios en nuestra capacidad para entender la dinámica de las poblaciones en esta era.

La genética de poblaciones ha examinado tradicionalmente la relación entre los procesos evolutivos y los patrones espacio-temporales de variación genética modelados por ellos. Junto a otros métodos novedosos, el análisis de autocorrelación espacial (AE) ha resultado en una herramienta poderosa para integrar datos genéticos, con información fenotípica y ambiental, de un modo espacialmente explícito. La AE es la propiedad de pares de objetos en sitios distintos de ser más o menos similares (AE positiva y negativa, respectivamente) que lo esperable por azar; esto, trasladado al campo de la genética de poblaciones, implica la asociación entre el parentesco de individuos tomados de a pares con la distancia geográfica. Los trabajos de Sokal y Oden (1978a, 1978b) y de Sokal (1979) fueron pioneros en incorporar esta herramienta en un contexto poblacional y evolutivo, con el análisis de los modelos de aislamiento por distancia (Wright, 1943; Malécot, 1948). Desarrollos más recientes dieron lugar a indicadores de AE locales (Getis y Ord, 1992; Anselin, 1995; Ord y Getis, 1995; Sokal y col., 1998; Double y col., 2005).

En este siglo, subdisciplinas y tópicos han emergido como producto de los nuevos patrones tecnológicos. La genética del paisaje (Manel y col., 2003) es un nuevo área de estudio que integra la

genética de poblaciones, la ecología del paisaje y la geoestadística (Storfer y col., 2007), cuyo objetivo es analizar la interacción entre características del paisaje y procesos microevolutivos (Manel y col., 2003). Nuevas direcciones de investigación se han generado también con otros avances recientes, por ejemplo a partir de aquellos mostrando un espectro de fenómenos moleculares epigenéticos heredables a través de generaciones, capaces de alterar los patrones temporales, espaciales y de abundancia de la expresión genética, con consecuencias sobre el fenotipo (Rapp y Wendel, 2005).

Diferentes programas disponibles a la fecha son capaces de analizar datos espaciales, ecológicos o genéticos, pero la necesidad de integrar estas disciplinas requiere de programas que permitan interconectar la información desde un aspecto práctico. En este capítulo se presenta un paquete para biología de poblaciones llamado “EcoGenetics”²¹, diseñado en el lenguaje y ambiente estadístico R (R Core Team, 2015). El objetivo del paquete es proveer métodos para integrar, manejar y analizar, en particular a una escala fina, información poblacional espacialmente explícita, proveniente de distintas fuentes.

3. OBJETIVOS

Desarrollar un paquete estadístico para el análisis de datos fenotípicos, genotípicos y ambientales en biología de poblaciones, con especial énfasis en el estudio genético-poblacional a escala fina.

4. MÉTODOS

4.1. Estructura básica del paquete

La estructura básica del paquete se muestra en la Figura 4.1. El mismo posee cinco módulos. El **módulo base** está compuesto por funciones generales (modelos lineales múltiples, utilidades de conversión entre formatos, etc.). El **módulo espacial general** realiza análisis espaciales globales y locales (test de Mantel, test de Moran, etc.). Estos análisis utilizan una matriz de pesos provista por

²¹ Disponible en: <http://cran.r-project.org/web/packages/EcoGenetics>

el **módulo de pesos espaciales**²². Los pesos contenidos en estas matrices pueden definirse en base a distancias (individuos incluidos en un rango de distancia determinado) o en base a los vecinos más cercanos (los j vecinos más cercanos del individuo i). El **módulo de lag espacial** es capaz de generar correlogramas de distinto tipo y variogramas empíricos. Este módulo utiliza una secuencia de matrices de pesos provistas por el módulo de pesos espaciales. El **módulo de prueba** realiza análisis de permutación, bootstrap y jackknife, como parte de los procesos internos de las funciones incluidas en el paquete.

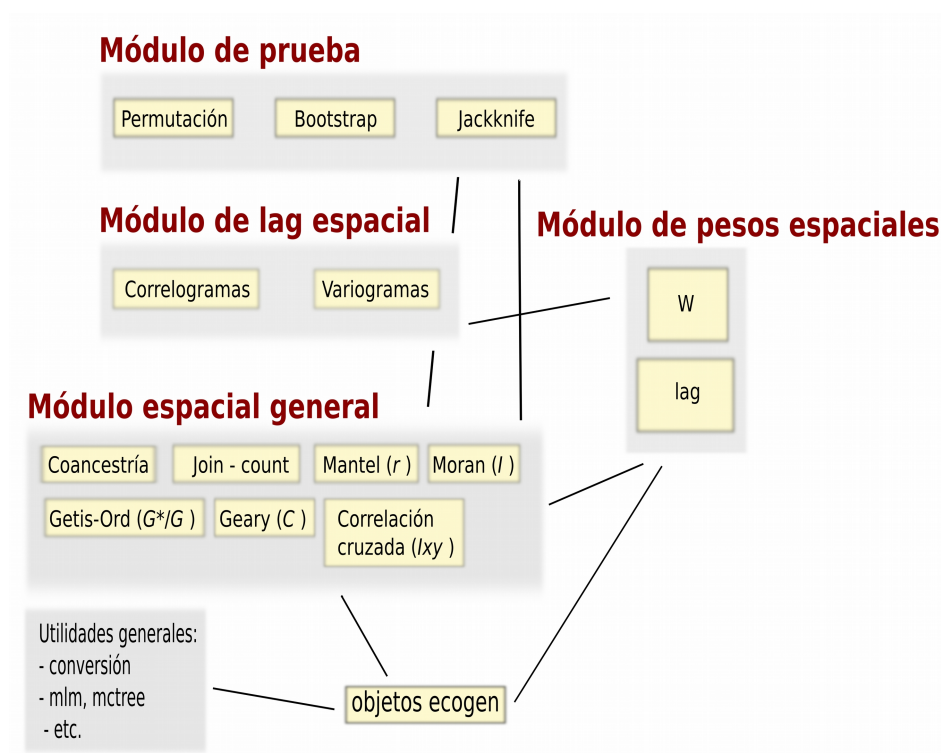


FIGURA 4.1 Estructura básica del paquete EcoGenetics.

²² Los pesos son, como el nombre indica, valores por los que se pondera una variable, como cuando por ejemplo se calcula un promedio ponderado; en el caso que aquí compete, el peso multiplica a una variable espacial.

4.2. Algoritmos de análisis espacial

A. Módulo espacial general

Este módulo contiene tres programas principales: **eco.gsa**, **eco.mantel** y **eco.lsa**.

El programa **eco.gsa**, calcula estadísticos espaciales globales (I de Moran, C de Geary, *Join-count* e I_{xy} bivariado de Moran). Requiere para ello una matriz de pesos espaciales. El programa **eco.mantel** calcula los test de Mantel ordinario (Mantel, 1967) y parcial (Smouse y col., 1986). Tanto **eco.gsa** como **eco.mantel** poseen el siguiente algoritmo:

1. calcular el valor observado del estadístico correspondiente
2. replicar k veces la función, para obtener una distribución nula, permutando aleatoriamente la posición de los individuos en cada ocasión
3. comparar el valor observado del estadístico con la distribución nula usando el módulo de prueba
4. repetir 1 a 4 en el caso de **eco.gsa**, cuando el análisis involucra múltiples variables

Para el paso 2 del test de Mantel, las filas y las correspondientes columnas de una de las matrices son permutadas k veces. Esto es equivalente a reordenar de modo aleatorio las filas en la matriz original (Legendre y Legendre, 2012). Para el test de Mantel parcial, el paso 2 es calculado con el método propuesto por Smouse y col. (1986), donde para obtener el estadístico entre las matrices M_1 y M_2 , condicionado por M_3 , se permuta M_1 un número k de veces, manteniendo las otras matrices sin cambios.

El programa **eco.lsa** calcula estadísticos espaciales locales (I_i de Moran, C_i de Geary, G_i^* y G_i de Getis-Ord). El algoritmo empleado en el programa es el siguiente:

1. calcular el valor observado del estadístico para los individuos i_1 a i_n
2. para un individuo i_i , recalculer k veces el estadístico para obtener una distribución nula, permutando aleatoriamente (de modo libre o condicional) la posición de los individuos en cada ocasión

3. comparar el valor observado de del estadístico para i_t con la distribución nula de i_t , utilizando el módulo de prueba
4. mientras $t < n$, tomar al individuo i_{t+1} y repetir los pasos 2 a 3
5. aplicar una corrección por pruebas múltiples si corresponde
6. repetir 1 a 5 cuando el análisis involucra múltiples variables

En el caso de los estadísticos de Moran y Geary locales, Sokal y col. (1998) sugirieron el uso de una permutación condicional (Anselin, 1995). Esto implica la permutación de $N-1$ individuos, usando al de prueba como pivote. Para G_i y G_i^* , Sokal y col. (1998) sugirieron el uso de permutación condicional y permutación libre para el modelo nulo, respectivamente. El programa aplica por defecto el tipo de permutación recomendada, sin embargo, puede haber razones para elegir una sobre otra, basados en la naturaleza de los datos a analizar (Sokal y col., 1998). Los valores de P pueden ser corregidos por comparaciones múltiples.

B. Módulo de lag espacial

Los métodos incluidos en el módulo de lag espacial están basados en la descomposición de la variabilidad espacial de los datos entre clases de distancias. Para construir estas clases, el módulo de pesos puede usar intervalos de distancia calculados bajo diferentes reglas: *a*) intervalos fijos (por ejemplo, cada 50 m), *b*) número fijo de individuos en cada clase, *c*) número predefinido de clases, *d*) intervalos de distancia determinados por un vector, *e*) intervalos de distancia para un dado rango dado por la regla de Sturges (Sturges, 1926) o el método de Freedman-Diaconis (Freedman y Diaconis, 1981), *f*) distintas combinaciones de estos casos²³. Cada uno de estos métodos puede ser usados en la construcción de correlogramas y variogramas. Para una regla determinada y para cada clase, se calcula una matriz de pesos espaciales.

Un correlograma es un gráfico de valores de correlación espacial (por ejemplo, de I de Moran) en función de la distancia entre pares de individuos, en forma de clases discretas. Diferentes tipos de correlogramas están soportados por el paquete:

- Para autocorrelación global, I , C e I_{xy} (función **eco.correlog**)

²³ Ver la función *eco.lagweight* en <http://cran.r-project.org/web/packages/EcoGenetics/EcoGenetics.pdf>

- Correlogramas de Mantel y Mantel parcial (función **eco.cormantel**)
- Para datos genéticos en particular, correlogramas basados en coancestría, y en particular con el coeficiente f_{ij} (Loiselle y col., 1995; Kalisz y col., 2001) (función **eco.malecot**)

En el caso de los correlogramas I , C e Ixy , para cada clase de distancia d , el estadístico se calcula con N' = número de conexiones incluidas en d , y una matriz de pesos binaria $w = w_\delta$, siendo w_δ una matriz con deltas de Kronecker, donde para cada celda w_{ij} se cumple que:

$$w_{ij} = \begin{cases} 1, & \text{si } distancia(i, j) \in d \\ 0, & \text{caso contrario} \end{cases} \quad (\text{Expresión 7.1})$$

Por defecto el algoritmo utiliza la matriz binaria, aunque la misma puede estandarizarse a suma 1 por fila.

El algoritmo básico de estos tres programas es el siguiente:

1. calcular los d_m intervalos de distancia bajo una determinada regla de construcción de intervalos
2. calcular la matriz de pesos espaciales w para los intervalos d_1 a d_m
3. calcular los valores observados del estadístico para los intervalos d_1 a d_m
4. para d_t , recalculan k veces el estadístico para obtener una distribución nula, permutando aleatoriamente la posición de los individuos en cada ocasión
5. comparar el valor observado del estadístico para d_t , con la distribución nula de d_t , usando el módulo de prueba y calculando el intervalo de confianza (IC) de la hipótesis nula
6. mientras $t < m$, tomar el intervalo i_{t+1} y repetir los pasos 4 y 5
7. aplicar una corrección por comparaciones múltiples si corresponde
- 7'. para **eco.correlog** y **eco.cormantel**, repetir 1 a 7 cuando el análisis involucra múltiples variables

En el programa **eco.cormantel**, el test se calcula para cada intervalo entre una matriz de distancia M (en el caso del test de Mantel, son dos matrices, M_1 y M_2) y la matriz binaria con deltas de Kronecker (w) de la Expresión 7.1 (Oden y Sokal, 1986; Sokal y col., 1986).

En el programa **eco.malecot**, se realizan los pasos 1 a 7, con 5 pasos adicionales:

8. para cada clase, calcular el *IC* (intervalo de confianza) del estadístico mediante jackknife sobre los loci
9. calcular la pendiente de la regresión de los valores de coancestría entre pares de individuos en función del logaritmo de la distancia entre ellos
10. significar la pendiente, realizando un test de Mantel entre las matriz de coancestría y la matriz del logaritmo de la distancia entre pares de individuos
11. calcular el *IC* de la pendiente mediante jackknife sobre los loci
12. calcular el estadístico *Sp* y su *IC* usando el *IC* de la pendiente

4.3. Estructura de los objetos “ecogen”

El paquete EcoGenetics fue diseñado con la ventaja de la programación orientada a objetos que ofrece R. La mayor parte de los análisis que se pueden realizar usando el paquete tienen clases asignadas con métodos gráficos, de presentación en la pantalla, etc. Para el manejo de las matrices básicas de datos, se creó la clase “ecogen”. Este objeto permite un manejo más eficiente de la información en las distintas etapas del análisis, facilitando el pre-procesamiento y almacenamiento. En general, el uso de las herramientas de análisis de los módulos es independiente de los objetos ecogen, para darle flexibilidad al usuario, aunque para el manejo óptimo del paquete es recomendable su utilización. Un objeto genérico de esta clase está compuesto por los siguientes slots²⁴: un slot **XY**, para almacenar las coordenadas de los individuos; un slot **P**, para almacenar la información fenotípica; un slot **G**, para la información genotípica; un slot **E** para la información ambiental; un slot **S**, conteniendo clases asignadas a los individuos; un slot **A**, que contiene en forma de frecuencias alélicas la matriz genética; un slot **C**, para guardar una matriz personalizada, y un slot **OUT**, conteniendo una lista para almacenar resultados. Adicionalmente, un objeto genind²⁵ modificado, que actúa de intermediario entre los datos del usuario y los objetos ecogen, se encuentra almacenado en un slot no visible por el usuario final. Como los objetos ecogen pueden contener las matrices básicas de datos y los resultados, toda la información puede ser organizada en un solo objeto.

24 Slots (“ranuras”) son elementos (objetos), que conforman un objeto S4. Se puede acceder de modo genérico a cada slot de un objeto con el símbolo “@” seguido de su nombre.

25 Tipo de objeto del paquete *adegenet* (Jombart, 2008), para almacenar y procesar información genética.

IV. Desarrollo de un paquete de R

La clase “ecogen” posee un conjunto de métodos asociados para el manejo de la información. La construcción de un nuevo objeto de esta clase está dada por la función homónima²⁶:

```
require(EcoGenetics)
data(eco.test)
eco <- ecogen(XY = coordinates, P = phenotype, G = genotype, E =
environment, S = as.data.frame(structure))
```

Durante la asignación de una matriz al slot G, la información genética es pre-procesada (eliminación de loci no informativos, eliminación opcional de individuos no informativos y ordenamiento opcional de alelos en caso de marcadores codominantes - útil para análisis que requieran un único código por genotipo). Los métodos asociados a los objetos ecogen se desarrollarán en el ejemplo a continuación.

5. APLICACIONES

EcoGenetics incluye datos simulados para ejemplificar el uso del paquete:

```
require("EcoGenetics")
data(eco.test)

ls()
[1] "coordinates" "eco" "environment" "genotype" "phenotype" "structure"
```

²⁶ Se muestran en gris los comandos ingresados en R, y en amarillo las salidas de texto.

IV. Desarrollo de un paquete de R

Con los datos cargados en el espacio de trabajo, un nuevo objeto `ecogen` es creado de la siguiente manera:

```
eco <- ecogen(XY = coordinates, P = phenotype, G = genotype, E =
environment, S = as.data.frame(structure))
eco
```

|| ECOGEN CLASS OBJECT ||

Access to slots: <ecoslot.> + <name of the slot> + <(name of the object)>
See help("EcoGenetics accessors")

slot XY:	--> 225 x 2	coordinates
slot P:	--> 225 x 8	phenotypic variables
slot G:	--> 225 x 10	loci >> ploidy: 2 codominant
slot A:	--> 225 x 40	alleles
slot E:	--> 225 x 6	environmental variables
slot S:	--> 225 x 1	structure >> 1 structure found
slot C:	--> 0 x 0	variables
slot OUT:	--> 0	results

Todos los objetos S4 del paquete `EcoGenetics` tienen un conjunto de "accesores" asignados. Un accesor es una función que permite, como la palabra sugiere, "acceder" al contenido de los slots, actuando como interfaz entre los objetos y los usuarios finales²⁷. Los métodos utilizables con este tipo de objetos son los siguientes:

1. Acceso al contenido de los slots usando accesorios genéricos o equivalentemente corchetes dobles ("[")

```
ecoslot.XY(eco); eco[["XY"]]
ecoslot.P(eco); eco[["P"]]
ecoslot.G(eco); eco[["G"]]
ecoslot.A(eco); eco[["A"]]
ecoslot.E(eco); eco[["E"]]
ecoslot.S(eco); eco[["S"]]
ecoslot.C(eco); eco[["C"]]
ecoslot.OUT(eco); eco[["OUT"]]
```

²⁷ Ver `help("EcoGenetics accessors")` en el paquete `EcoGenetics`.

2. Método de asignación a los slots usando accesorios genéricos o equivalentemente corchetes dobles ("[[")

```
eco.temp <- ecogen(XY = coordinates, P = phenotype)
eco.temp
ecoslot.G(eco.temp, order.G = TRUE) <- genotype

## esto tiene su equivalencia usando dobles corchetes
eco[["G", order.G = TRUE]] <- genotype
ecoslot.E(eco.temp) <- environment
## idéntico a eco[["E"]] <- environment

ecoslot.S(eco.temp) <- structure
## idéntico a eco[["S"]] <- structure

eco.temp
```

3. Método de recorte usando el corchete simple ("[")

```
eco.sub <- eco[1:50]
eco.sub
```

4. Intersección entre dos objetos

```
eco1 <- eco
merged <- eco.merge(eco, eco1)
merged
```

5. Recorte del objeto, reteniendo sólo los individuos de un grupo determinado de la slot S (en este caso, "1")

```
eco.subS <- eco.subset(eco, "pop", 1)
eco.subS
```

6. Método de unión por fila (para dos o más objetos)

```
eco1 <- eco
rownames(eco1[["P"]]) <- 226:450 # nombres duplicados no son permitidos
eco.r <- eco.rbind(eco, eco1) # algunas matrices con nombres duplicados
```

7. Método de unión por columna (para dos o más objetos)

```
eco1 <- eco
eco.c <- eco.cbind(eco, eco1)
eco.c
```

8. Método para ordenar un objeto por filas, usando como referencia el orden en los datos del slot XY

```
eco1 <- eco
ecoslot.P(eco1) <- eco[["P"]][sample(1:173), ] # filas desordenadas en el
                                                    # slot P
ordered <- eco.order(eco1)
ordered[["P"]]                                # slot P ordenado
```

Test de Mantel para las matrices fenotípica y ambiental:

```
set.seed(10)
mantel.PE <- eco.mantel(dist(eco[["P"]]), dist(eco[["E"]]), nsim = 99)
```

```
#####
Mantel test
#####
> Correlation coefficient used --> Pearson
> Number of simulations --> 99
> Alternative --> greater
> P-value --> 0.01
> Observed value --> 0.187
> Expected value --> -0.0047
```

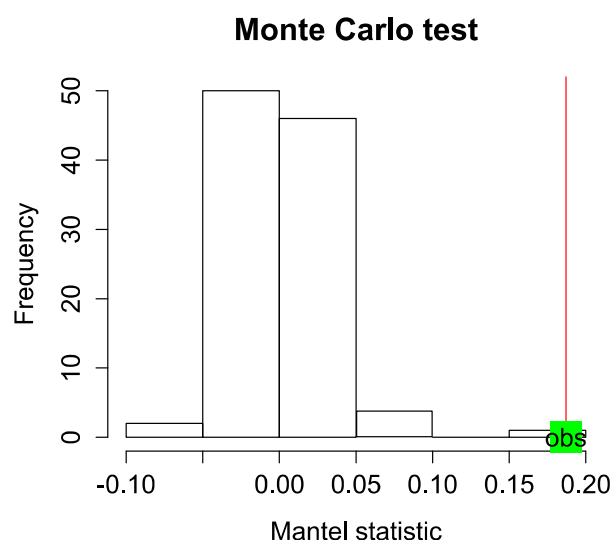


FIGURA 4.2 Salida de la función *eco.mantel* tal como se muestra en R: distribución nula y valor observado para el test de Mantel entre las matrices P y E del objeto “eco”.

IV. Desarrollo de un paquete de R

El objeto “eco2” incluido en el paquete contiene patrones de puntos en una grilla de 30 x 30 píxeles, simulando datos fenotípicos:

```
data(eco2)
par(mfrow=c(1,3))
image(matrix(eco2[["P"]][,1], 30, 30))
image(matrix(eco2[["P"]][,2], 30, 30))
image(matrix(eco2[["P"]][,3], 30, 30))
```

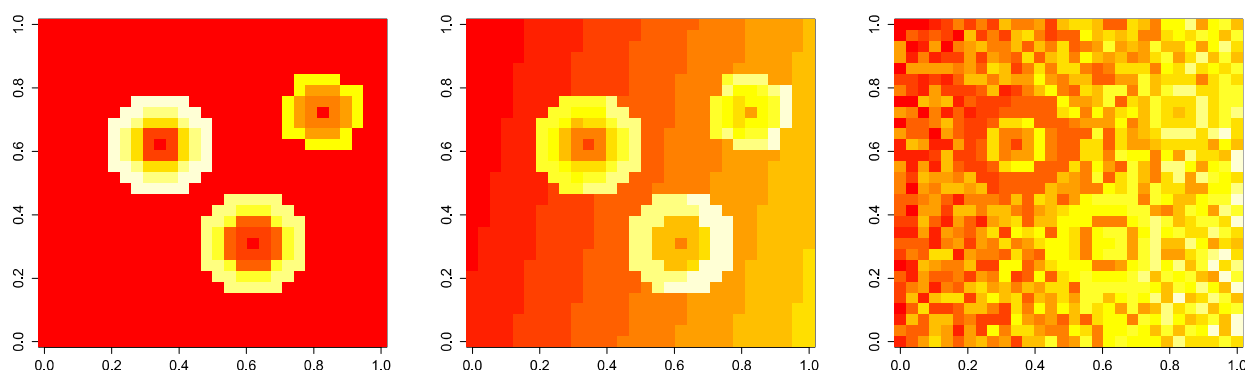


FIGURA 4.3 Imágenes de las primeras tres variables del slot P en el objeto eco2, tal como se muestra en R. La primera variable posee un patrón local, mientras que la segunda y la tercera incluyen gradientes lineales globales. Estos últimos pueden removerse con un análisis de superficies de tendencias²⁸ (Borcard y col., 2011) usando la función **eco.detrend**, realizando una regresión de cada rasgo (z) sobre el polinomio de grado n de las coordenadas espaciales (x, y). En este ejemplo, se usa el polinomio de grado 1,

Para remover las tendencias globales sobre la segunda de las imágenes de la Figura 4.3, que posee una tendencia en sentido sudeste-noroeste, se utiliza el polinomio de primer grado de las coordenadas espaciales:

$$z_i = a_1 + a_2x + a_3y$$

²⁸ Esto se describe con más detalle en el capítulo V, página 135

IV. Desarrollo de un paquete de R

A continuación se implementa este método:

```
det <- eco.dettrend(eco2[["P"]][,2], eco2[["XY"]][,1:2], degree = 1)
sim.det <- data.frame(matrix(0, 900, 3))
sim.det[,1:2] <- eco2[["XY"]]
sim.det[, 3] <- ecoslot.RES(det)
par(mfrow = c(1,1))
image(matrix(sim.det[,3], 30, 30))
```

La figura 4.4 muestra la imagen resultante al remover la tendencia global:

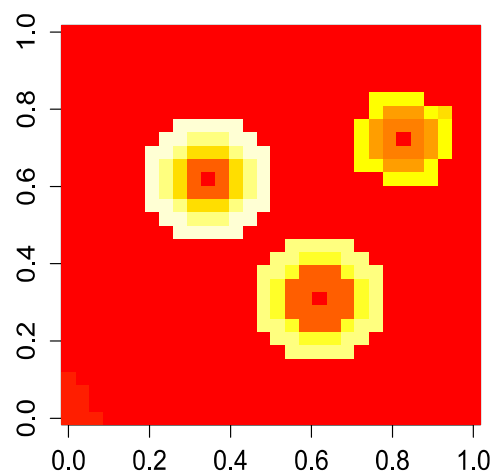


FIGURA 4.4 Imagen de la segunda variable de la slot P en el objeto eco2, tras realizar una remoción de la tendencia global, tal como se muestra en R.

La tendencia fue removida y los datos ahora se encuentran listos para ser analizados, por ejemplo, calculando un correlograma de Moran:

```
cordata <- eco.correlog(Z = sim.det[,3], XY = sim.det[,1:2], int = 2,
smax = 30)
plot(cordata)
```

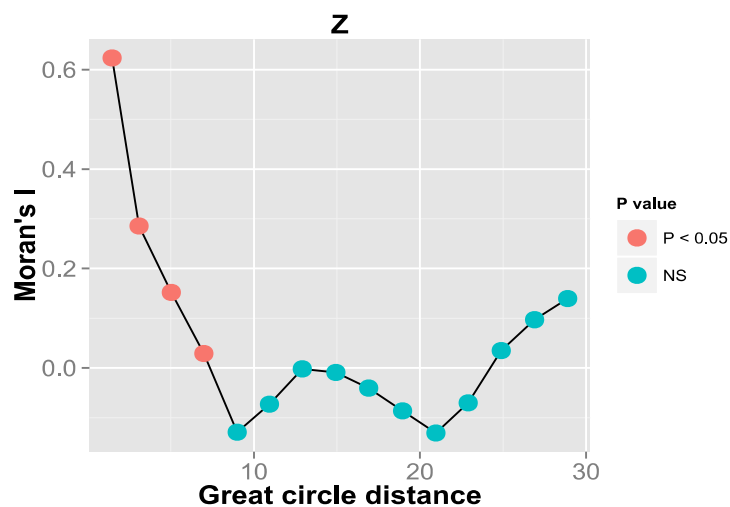


FIGURA 4.5 Salida de la función *eco.correlog*, tal como se muestra en R.

La variable muestra una tendencia decreciente desde valores positivos a negativos, y luego se estabiliza. Los patrones locales pueden analizarse con el estadístico G_i^* :

```
## matriz de pesos con 20 vecinos más cercanos
weight.con <- eco.weight(sim.det[,1:2], method = "knearest", k = 20,
self = TRUE)
weight.con
```

```
#####
spatial weights
#####
Method --> knearest - min
Parameters --> (k = 20)
Row-standardization --> FALSE
Non-zero (non-self) links --> 11.8 %
Number of individuals --> 900
Individuals with non-zero (non-self) links --> 100 %
Average (non-self) links per individual --> 105.8
Average distance between connected individuals --> 4.203
```

```
get <- eco.lsa(var = sim.det[,3], con = weight.con, method = "G*",
nsim = 10)
plot(get, significant = FALSE)
```

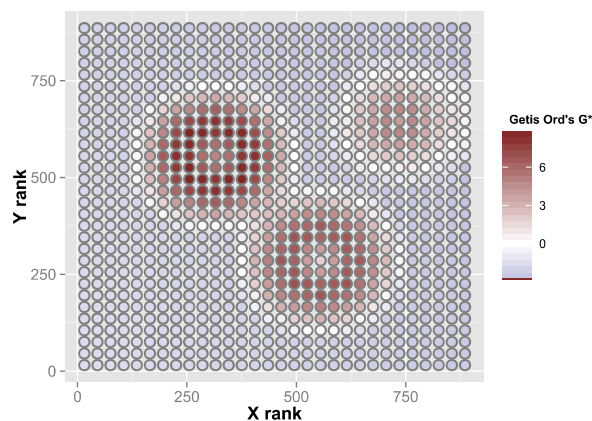


FIGURA 4.6 Salida de la función *eco.lsa*, tal como se muestra en R.

Los resultados pueden ser almacenados en el objeto “ecogen” creado:

9. Almacenamiento de resultados mediante accesorios genéricos o equivalentemente corchetes dobles (["["])

```
ecoslot.OUT(eco2) <- list(get, cordata, weight.con)

## el uso del accesor OUT es equivalente al doble corchete
eco2[["OUT"]] <- list(cordata, weight.con)
```

Para almacenar varios resultados, los mismos deben asignarse como lista, tal como se muestra en el panel superior ($\text{list}(x_1, x_2, \dots, x_n)$, para n distintos objetos).

Se pueden ver los resultados ordenados alfabéticamente:

```
ecoslot.OUT(eco2)
```

	OBJECTS	CLASSES
1	cordata	eco.correlog
2	cordata.1	eco.correlog
3	get	eco.lsa
4	weight.con	eco.weight
5	weight.con.1	eco.weight

IV. Desarrollo de un paquete de R

Tal como puede observarse, el panel muestra los objetos almacenados alfabéticamente y sus clases. El slot OUT puede manejarse como una lista convencional.

Especificando un objeto, se puede acceder al mismo:

```
ecoslot.OUT(eco2, "cordata")
```

Se pueden remover resultados del objeto:

10. Método de remoción de resultados

```
eco2 <- eco.remove(eco2, cordata)  
ecoslot.OUT(eco2)
```

Finalmente, el espacio de trabajo puede limpiarse, almacenando sólo el objeto creado, manteniendo sólo una estructura de información con los datos de entrada y los resultados:

```
eco.clear(eco2)
```

Capítulo V

Estudios con una población natural
fragmentada de *Prosopis alba*

V. ESTUDIOS CON UNA POBLACIÓN NATURAL FRAGMENTADA DE *PROSOPIS ALBA*

1. DESCRIPCIÓN DEL ÁREA DE ESTUDIO

El paisaje analizado en este capítulo comprende sitios ubicados en el cono aluvial del río Dulce o en sus cercanías. En la Figura 5.1 se muestra una imagen del área, confeccionada con una imagen del satélite Landsat 8 (bandas 7-5-3) y los datos de altitud del satélite ASTER (ASTER GDEM V2). En esta región, coexisten áreas urbanas, forestales y agrícolas, debido a que el agua es abundante y se encuentra distribuida de modo regular, factor que permitió históricamente el poblamiento humano y el desarrollo de actividades productivas. El río Dulce²⁹ es el recurso principal de recarga del cono aluvial. Un sistema de irrigación compuesto por una red de canales distribuye el agua a través del cono y el área circundante. La profundidad de la napa freática en las zonas irrigadas va de 1 a 6 m bajo la superficie, y el acuífero más superficial alcanza una profundidad de 15 m. La recarga del mismo también se produce por precipitación estacional y escurrimiento (Lindbäck y Sjölin, 2006).

29 Este río tiene una baja salinidad promedio (aproximadamente 0.36 dS/m, según Lieveld [2005]).



FIGURA 5.1 Imagen de Landsat 8 (2013) (bandas 7-5-3) con un modelo de elevación digital de exageración vertical 40x.

De acuerdo al sistema de clasificación Köppen-Geiger (Kottek y col., 2006), la región estudiada posee un clima semiárido cálido (Bsh). Su microtopografía es heterogénea, presentando parches variables en fertilidad, salinidad y textura (Aristide, 2010).

2. RECOLECCIÓN DE MUESTRAS

Para la presente investigación, se define a la "población" de estudio desde un punto de vista instrumental, como la unidad evolutiva, compuesta por individuos que se aparean entre sí en un área geográfica determinada. Se supone que la muestra de individuos es representativa de una población mayor, y que es adecuada para estudiar los fenómenos evolutivos que son emergentes a la escala de análisis. La hipótesis nula de modelo poblacional es H_0 : los individuos estudiados pertenecen a una

única población. Sería equivalente a una muestra tomada de una población de extensión desconocida.

Se analizó una muestra total de 186 individuos adultos, ubicados en los márgenes de las rutas nacionales 1, 9 y 34 (Figura 5.2), que conjuntamente consideradas se denominaron “sitio R” y corresponden a un remanente de la población ancestral en el área, y en dos parches con cobertura arbórea secundaria de colonización reciente, previamente usados con fines agrícolas y actualmente sin uso productivo, que fueron denominados sitios “P” y “A” (Figura 5.2). A partir de la examinación visual de estos sitios, se determinó una cobertura arbórea dispersa en los corredores a lo largo de los márgenes de las rutas, mientras que los sitios P y A presentaron cobertura cerrada y abierta, respectivamente.

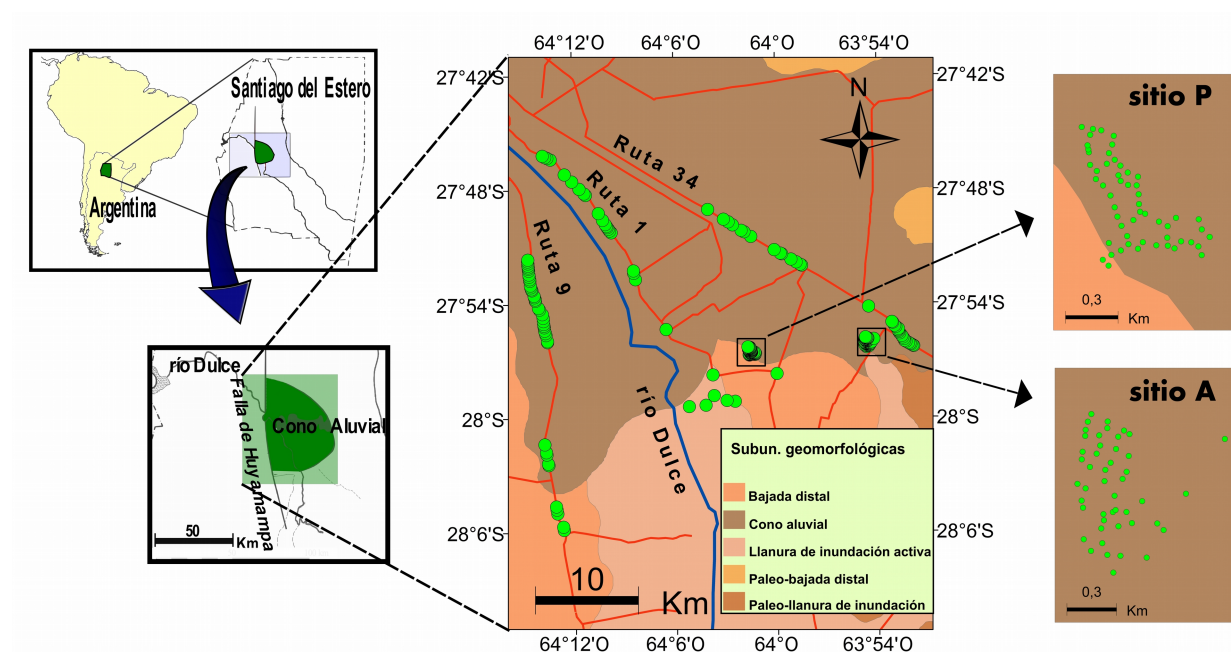


FIGURA 5.2 Esquema del área de muestreo. La leyenda indica los colores que corresponden a las subunidades geomorfológicas descritas en Angueira, 2007. Los puntos verdes representan los árboles muestreados en el área completa (imagen central) y en los sitios P (imagen superior derecha) y A (imagen inferior derecha).

Se muestrearon árboles a una distancia mínima de 30 m (Figura 5.3), excepto en tres casos (17.0, 20.5, y 23.5 m). De cada árbol individual se tomó un ejemplar de hojas adultas y frutos, que se herborizó en papel de diario, y se colectó material de hoja joven para el análisis de ADN, que se

secó inmediatamente en sílica gel y se almacenó posteriormente a temperatura ambiente hasta el momento de la extracción.

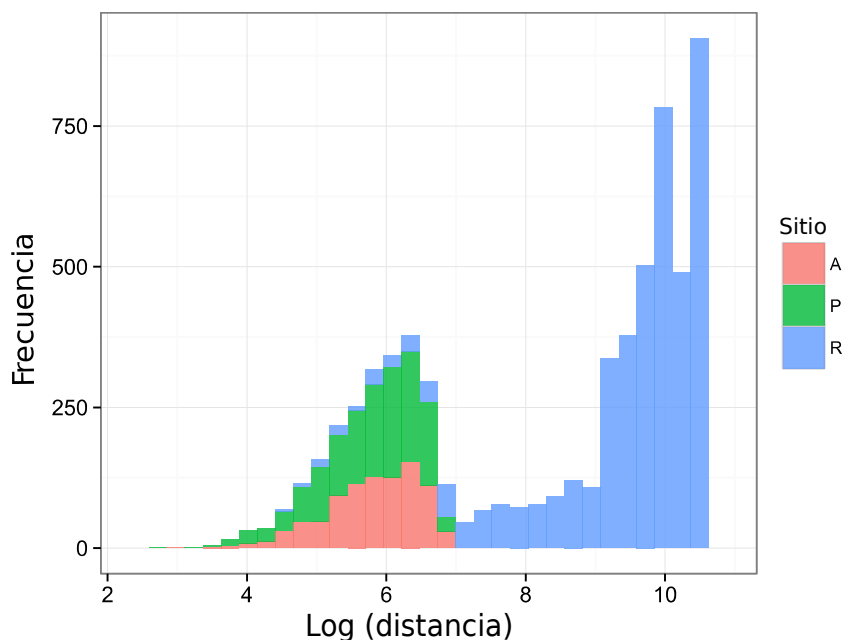


FIGURA 5.3 Histograma (frecuencia absoluta) para el logaritmo de la distancia en metros, entre todos los pares posibles de individuos estudiados, según los tres sitios de muestreo considerados (A, P y R).

A cada individuo se le midió el diámetro a la altura del pecho (DAP) con cinta diamétrica, graduada al milímetro más cercano. En este estudio sólo se incluyeron individuos con un valor de DAP mayor a 11 cm. Estos valores de DAP se utilizaron como un estimador de la edad de los árboles.

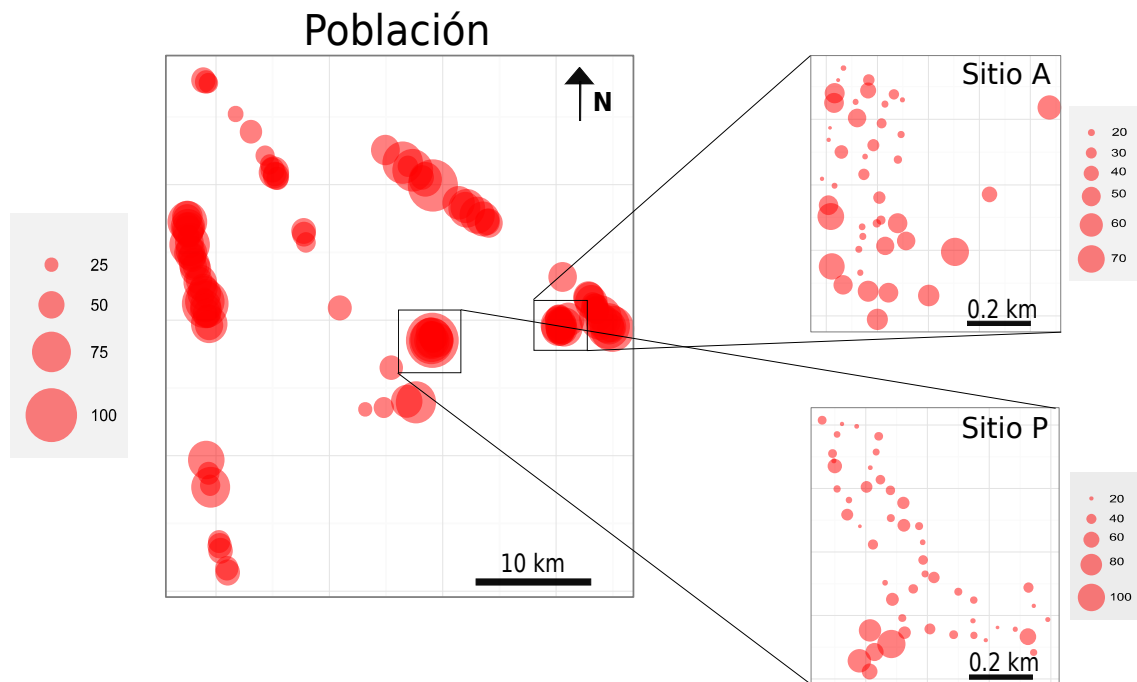


FIGURA 5.4 Diámetro a la altura del pecho (DAP) de los individuos muestreados (círculos rojos) en la población entera (imagen de la izquierda) y en los sitios A y P (imágenes de la derecha). El tamaño de los círculos es proporcional a los valores de DAP de los individuos, tal como se detalla en las leyendas.

Capítulo V Primera parte

Estudio de la estructura genética espacial a escala fina



PRIMERA PARTE: ESTUDIO DE LA ESTRUCTURA GENÉTICA ESPACIAL A ESCALA FINA

1. RESUMEN

La estructuración genética espacial a escala fina en plantas es un fenómeno que puede ocurrir cuando los individuos que componen una población presentan una distancia de dispersión baja en comparación con la extensión del área que en conjunto cubren. En árboles, la estructura a escala fina estaría determinada principalmente por la dispersión a través de las semillas y el polen. Cuando la dispersión se encuentra limitada dentro de poblaciones continuas, es esperable que el bajo flujo génico resultante se manifieste en un proceso de diferenciación genética local (Cavers y col., 2005; Hardy y col., 2006), generándose de este modo una estructura poblacional acorde a un modelo de aislamiento por distancia (Wright, 1943; Malécot, 1948). En esta parte se analizó la estructura genética fina de la población estudiada utilizando 7 marcadores microsatélites. Los objetivos fueron evaluar el ajuste a un modelo de aislamiento por distancia, mediante un estimador del coeficiente de coancestría, y analizar la presencia de señales de cambios demográficos recientes cuantificando el desvío de la heterocigosis esperada bajo equilibrio de Hardy-Weinberg (H_E) del valor esperado bajo equilibrio mutación-deriva (H_{EQ}). Los análisis mostraron que 1) no hay evidencia de discontinuidades en la población, 2) la diferenciación entre los sitios considerados es baja, 3) los valores de coancestría tienden a decrecer con la distancia geográfica entre pares de individuos y 4) hay indicios de una posible expansión o más bien de un proceso de inmigración reciente en uno de los sitios de estudio. Estos resultados sugieren que, a la escala espacial estudiada, el modelo de aislamiento por distancia sería el que mejor representa la estructura genética de la población y que algunos procesos demográficos podrían ser detectados a partir de marcadores moleculares con diseños de estudio adecuados.

2. INTRODUCCIÓN

La estructura genética espacial (EGE) puede definirse como la distribución no aleatoria de genotipos en un espacio determinado (Vekemans y Hardy, 2004), y es una consecuencia de combinaciones complejas de procesos actuando a diferentes escalas. La baja distancia de dispersión que los árboles suelen presentar, en comparación con la extensión espacial de las poblaciones, puede promover una distribución no aleatoria de la variación intra-poblacional, generando patrones espaciales genéticos a escala fina (Vekemans y Hardy, 2004; Lopez-Gallego y O'Neil, 2010). En el caso de especies perennes con ciclos de vida largo, la EGE de juveniles representa primariamente los patrones de dispersión de las semillas y del polen (Heywood, 1991; Vekemans y Hardy, 2004; Van Rossum y Triest, 2006; Yamagishi y col., 2007). La dispersión de semillas influencia mucho más la estructura poblacional que la del polen (Chung y col., 2003). Cuando es localizada, puede resultar en individuos relacionados germinando uno cerca del otro, determinando de este modo alta EGE (Pardini y col., 2008). Este fenómeno de dispersión localizada puede darse hacia sitios específicos, como cuando frutos con múltiples semillas son dispersados de modo simultáneo, o también producirse en un entorno de árboles madre (Fuchs y Hamrick, 2010).

El modelo de aislamiento por distancia (Wright, 1943; Malécot, 1948) puede ser utilizado para predecir los patrones esperados de EGE en una población, cuando los individuos que la componen se aparean al azar dentro de un vecindario definido, pero poseen una posibilidad restringida de aparearse con otros miembros más distantes, debido a limitaciones en la dispersión (Sokal y Wartenberg, 1983). Bajo aislamiento por distancia se supone que sólo la dispersión a corta distancia y la deriva genética afectan una población (Barbujani, 1987). La predicción del modelo es una disminución gradual de la coancestría con el incremento de la distancia geográfica entre individuos (Malécot, 1959; Morton y col., 1968; Loiselle y col., 1995). Este patrón fue cuantificado como una relación lineal entre la coancestría y el logaritmo natural de la distancia geográfica en poblaciones bidimensionales con dispersión isotrópica (Maruyama, 1977; Rousset, 1997; Hardy y Vekemans, 1999; Rousset, 2000). De este modo, mientras mayor sea el solapamiento entre individuos adyacentes, más alta será la probabilidad de que sus descendientes sean hermanos.

Las técnicas de análisis de autocorrelación espacial (capítulo I, página 35) son particularmente útiles para el estudio de la variación genética a escala fina. Dado que el modelo teórico de aislamiento por distancia predice un patrón de disminución exponencial en la coancestría con la distancia entre pares de individuos (función $\phi(d)$), este fenómeno puede estudiarse de modo empírico mediante correlogramas³⁰, estimando la coancestría por medio de un coeficiente de correlación.

Los árboles son generalmente muy diversos y caracterizados por niveles de variación más altos dentro de las poblaciones que entre ellas (Jump y Penuelas, 2007). El flujo génico tiende a reducir la diferenciación entre poblaciones, pero a su vez distintos factores pueden influenciar la distribución de la diversidad genética dentro de las poblaciones, tales como el sistema de apareamiento (Vekemans y Hardy, 2004), la densidad poblacional de adultos (Hamrick et al., 1993), la selección de microhábitat (Levin y Kerster, 1974; Epperson, 1993), la estructura demográfica (Chung y col., 2003) y eventos históricos, que incluyen los disturbios y procesos de colonización (Parker y col., 2001; Premoli y Kitzberger, 2005; Pardini y Hamrick, 2008). En el caso particular de eventos fundadores, su efecto sobre la diversidad genética depende de número y origen de los individuos migrantes (Slatkin, 1977; Hamrick y Trapnell, 2011). Poblaciones fundadas por pocos individuos de un número limitado de poblaciones incrementan la diferenciación entre poblaciones, mientras que poblaciones fundadas por muchos individuos de poblaciones genéticamente diversas, pueden reducir esta diferenciación (Slatkin, 1977; Hamrick y Trapnell, 2011) (Figura 5.5).

30 Un correlograma es un gráfico de los valores de correlación para una serie de clases de distancia entre pares de individuos. Para cada clase se obtiene un valor promedio de correlación incluyendo a todos los individuos pertenecientes a esa categoría.

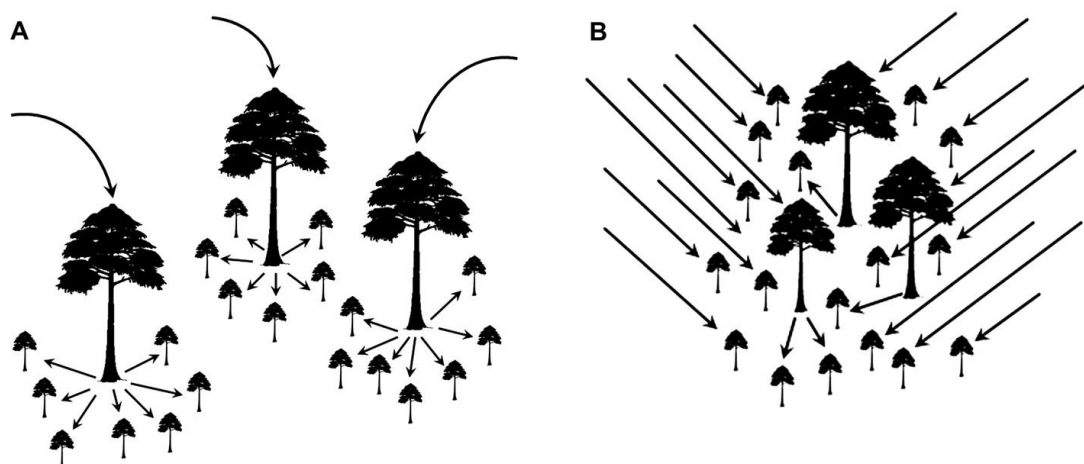


FIGURA 5.5 Dos modelos de expansión poblacional luego de un efecto fundador. Las plantas de mayor tamaño representan los fundadores originales mientras que las de menor tamaño subsiguientes reclutas. En **(A)** la mayor parte de los reclutas son progenie de los fundadores originales, resultando en una alta estructura genética espacial. En **(B)** la mayor parte de los reclutas son producto de inmigración hacia la población desde múltiples fuentes, resultando en una baja estructura genética espacial. Extraído de Hamrick y Trapnell (2011).

Los métodos moleculares son útiles para inferir la historia demográfica de una población, permitiendo estudiar procesos tales como expansiones recientes, reducciones en tamaños censales o en el tamaño poblacional efectivo (Busch y col., 2007). El número de alelos declina más rápidamente durante un evento de cuello de botella que la heterocigosis, mientras que lo opuesto (un exceso de alelos) puede observarse durante un proceso de expansión poblacional o de influjo de alelos raros por inmigración (Maruyama y Fuerst, 1985; Luikart y Cornuet, 1998). Esta propiedad puede utilizarse para poner a prueba la hipótesis de cambios recientes del tamaño poblacional o de ocurrencia de eventos de inmigración.

En esta parte se analizó la estructura genética de la población de *P. alba* a escala fina mediante marcadores microsatélites, poniendo a prueba un modelo de aislamiento por distancia y estudiando la presencia de señales de cambios demográficos recientes.

3. OBJETIVOS

- a. Analizar si la estructura genética espacial presente en la población corresponde a la esperada bajo un modelo de aislamiento por distancia, estudiando la variación de una estima de coancestría con la distancia geográfica entre pares de individuos.
- b. Evaluar en los patrones moleculares de la población la presencia de señales que permitan inferir cambios demográficos recientes.

4. MATERIALES Y MÉTODOS

4.1. Aislamiento de ADN y genotipado - Análisis de SSR

Para la extracción y el análisis de los marcadores SSR, se siguió el Protocolo 2-B del Apéndice 1 (página 158). Se seleccionaron siete marcadores microsátélites:

- Cuatro desarrollados para *P. chilensis* (Mottura y col., 2005) y amplificados exitosamente en *P. alba* (Bessegga y col., 2009): **Mo05** (CT)₃T(CT)₂, **Mo07** (GC)₈, **Mo09** (TG)₁₇ y **Mo13** (GT)₁₀CT(GT)₂.
- Tres desarrollados más recientemente para *P. alba* (Bessegga y col., 2013): **GL12** (CT)₅(ATCT)₃(CT)₅, **GL24** (AC)₁₁ y **GL 26** (GT)₈.

Los individuos que no mostraron bandas tras repetidas verificaciones en geles de poliacrilamida³¹, se codificaron como homocigotas para un alelo nulo.

³¹ El proceso de verificación consistió en los siguientes pasos: primero, reamplificación contra testigos usando distintas condiciones de PCR (concentración, temperatura de annealing), y segundo, tras no observar producto de PCR, nueva extracción de ADN y nueva amplificación contra testigos, verificando la ausencia de producto en los individuos con alelos nulos y la presencia en los testigos positivos.

4.2. Análisis de datos

4.2.1. Estimación de la diversidad genética

Para el análisis de la diversidad genética, se estimaron la heterocigosis observada (H_o) y la esperada bajo equilibrio de Hardy-Weinberg (H_E) por locus y promedio por locus, mediante un estimador insesgado (Nei y Chesser, 1983) (Apéndice 2, página 162) con el paquete hierfstat 0.04-14 (Goudet, 2005) en R 3.1.2³². La diversidad genética equivale aquí a la frecuencia de heterocigotas.

Para evaluar la distorsión en la segregación independiente de los loci (hipótesis nula: los loci segregan de modo independiente), el desequilibrio de ligamiento se analizó para todos los pares posibles usando la prueba G (Sokal y Rohlf, 1981). El exceso de homocigotas (hipótesis nula: no hay exceso de homocigotas) para cada loci se evaluó mediante la prueba U (Rousset y Raymond, 1995). Ambos análisis se llevaron a cabo en Genepop 4.0.10, estimando la significación mediante MCMC (cadenas de Markov de Monte Carlo) con los valores por defecto. Las estimas obtenidas se corrigieron por comparaciones múltiples con el método FDR (False Discovery Rate) de Benjamini y Hochberg (1995) en R.

Además de la heterocigosis, se calculó la riqueza alélica (R_A), que es un conteo del número de alelos observados en una muestra. Para obtener el valor para cada sitio individual, el cálculo se efectuó bajo rarefacción³³, usando un tamaño de muestra g , fijado al menor tamaño muestral. Este análisis se realizó con el paquete ARES (Van Loon y col., 2007), en una versión actualizada a R 3.1.2 por Leandro Roser. El número promedio de alelos por locus (A) se calculó como R_A/L , donde L es el número de loci analizados.

³² Todos los análisis de este capítulo utilizan la versión 3.1.2 de R.

³³ El número de alelos (riqueza alélica) en una población es una medida dependiente del tamaño muestral. El método de rarefacción (Hurlbert, 1971) permite comparar estadísticos provenientes de distintos tamaños muestrales. La riqueza alélica bajo rarefacción es el número esperado de alelos en una muestra con g individuos (tamaño muestral estandarizado) en vez de N ($g \leq N$) (Kalinowski, 2004). El paquete ARES permite obtener los valores de riqueza alélica bajo rarefacción con intervalos de confianza, utilizando una estima basada en momentos para los casos de extrapolación (Colwell y col., 2004; Mao y col., 2004; Van Loon y col., 2007).

4.2.2. Análisis de estructura genética

La estructura genética de la población se analizó mediante los estadísticos F (Wright, 1951), tal como fueron definidos por Weir y Cockerham³⁴ (1984). Para el análisis jerárquico, se tomaron los sitios A, P y R (Figura 5.2, página 102) como tres muestras distintas, y se estimó el valor de F_{IS} para cada sitio y de modo global. Se calculó también el coeficiente F_{ST} , global y entre pares de sitios. Estos análisis se llevaron a cabo con el paquete *diveRsity* (Keenan y col., 2013) en R 3.1.2. Los intervalos de confianza (IC) se obtuvieron mediante bootstrap, con 999 réplicas a través de individuos y loci. Las estimas se consideraron significativas cuando los IC no incluyeron el valor 0.

El análisis de la variación jerárquica (entre grupos, entre muestras dentro de grupos, y dentro de muestras), se realizó también mediante Análisis Molecular de la Varianza³⁵ (AMOVA) (Michalakis y Excoffier, 1996), con el paquete *poppr* (Kamvar y col., 2014) en R. Este paquete utiliza como medida de distancia el número de diferencias entre dos muestras a través de los loci, tal como fue definido en Excoffier y col. (1992). El nivel de significación de los estadísticos ϕ se obtuvo a través de pruebas de Monte Carlo con 999 permutaciones.

La existencia de estructura espacial críptica se analizó con el método bayesiano implementado en STRUCTURE 2.3.4 (Pritchard y col., 2000), asumiendo un modelo “admixture” con frecuencias alélicas correlacionadas (Falush y col., 2003). K se varió de 1 a 10, corriendo 10 réplicas de cada valor, con 750000 MCMC, descartando las 250000 primeras. La salida del programa se analizó en STRUCTURE HARVESTER (Earl y col., 2012).

4.2.3. Análisis de coancestría

La probabilidad de que los loci sean idénticos por descendencia en individuos pertenecientes a una muestra puede ser estimada mediante alelos idénticos en estado, con los coeficientes de coancestría. Se obtuvo una matriz de coancestría entre individuos con el paquete EcoGenetics, mediante el coeficiente de Nason (f_{ij}) (Loiselle y col., 1995; Kalisz y col., 2001). Este es un estimador de la coancestría³⁶, que mide la correlación en frecuencias de los alelos homólogos (a) a

34 Ver capítulo I, página 30.

35 Ver capítulo I, página 31.

36 Puede verse que es básicamente un coeficiente de correlación (primer término), con una corrección por el error muestral (segundo término) (Fenster y col., 2003).

través de los loci analizados (l), entre los pares de individuos i y j . Para el alelo a del locus l , el valor del coeficiente entre los individuos i -ésimo y j -ésimo se calcula de la siguiente manera (Loiselle y col., 1995; Kalisz y col., 2001; Hardy y col., 2003):

$$f_{ij}(al) = \frac{\sum_i \sum_j (p_{ial} - \bar{p}_{al})(p_{jal} - \bar{p}_{al})}{\bar{p}_{al} * (1 - \bar{p}_{al})} + \frac{1}{2(n_l - 1)} \quad (\text{Expresión 5.3})$$

Donde $\bar{p}_{al}$ es la frecuencia media del a -ésimo alelo del l -ésimo locus, p_{ial} y p_{jal} son las frecuencias del a -ésimo alelo del l -ésimo locus en los i -ésimo y j -ésimo individuos, respectivamente (valores de la matriz básica), y n_l es el número de copias de genes definidas en la muestra para el l -ésimo locus (el número de individuos sin datos faltantes para el locus l , multiplicado por la ploidía). El coeficiente multilocus se obtuvo ponderando los valores de f_{ij} para cada alelo a por su valor de $\text{PIC} = \bar{p}_{al} * (1 - \bar{p}_{al})$ (Heuertz y col., 2003).

Se construyó un correlograma considerando 9 clases de distancias entre individuos, con 250 pares de individuos incluidos en cada una, en el rango 0-900 m. Para cada una de estas clases, se promediaron los valores de f_{ij} y se calculó un intervalo de confianza del 95% para el valor de f_{ij} promedio obtenido bajo aleatorización (hipótesis nula de ausencia de estructura genética para cada clase), permutando la ubicación geográfica de los individuos 999 veces (Loiselle y col., 1995; Kalisz y col., 2001; Vekemans y Hardy, 2004).

Bajo aislamiento por distancia en un espacio bidimensional, la coancestría tiende a decrecer de modo aproximadamente lineal con el logaritmo de la distancia espacial (Maruyama, 1977; Rousset, 1997; Hardy y Vekemans, 1999). Para estudiar la hipótesis de ausencia de EGE, Vekemans y Hardy (2004) sugirieron calcular la pendiente (b) de la regresión entre f_{ij} y $\ln(d_{ij})$ (el logaritmo de la distancia que separa los individuos i y j). De este modo se obtiene un solo estadístico utilizando toda la información en una prueba única. El valor de b se calculó en el rango 0-900 m, significando la correlación entre f_{ij} y $\ln(d_{ij})$ mediante una prueba de Mantel (Rousset, 1997; Vekemans y Hardy, 2004). Esta consistió en la permutación de la ubicación geográfica de los individuos 999 veces, usando todos los pares con $d_{ij} < 900$ m para el cálculo del estadístico estandarizado de Mantel. Los errores estándar de b (SE_b), y de $f_{(1)}$ (la coancestría de la primera clase), se obtuvieron mediante jackknife sobre los loci (Vekemans y Hardy, 2004; Hardy col., 2006).

4.2.4. Análisis demográfico

Para estudiar la presencia de señales moleculares de cambios demográficos recientes en cada uno de los sitios (A, P y R), se utilizó un método cuantitativo con el programa BOTTLENECK 1.2.02 (Piry y col., 1999). La prueba incluida en BOTTLENECK asume que las reducciones en el tamaño poblacional hacen declinar más rápidamente el número de alelos que la heterocigosis (Maruyama y Fuerst, 1985); por lo tanto, H_E (la heterocigosis esperada bajo equilibrio de Hardy-Weinberg) se vuelve mayor que la heterocigosis esperada bajo equilibrio mutación-deriva (H_{EQ}), calculada en base al número de alelos ($H_E > H_{EQ}$), observándose el patrón opuesto en el caso de un proceso de expansión poblacional o de influjo de alelos raros por inmigración ($H_E < H_{EQ}$) (Cornuet y Luikart, 1996; Luikart y Cornuet, 1998). El programa compara los valores de H_E y H_{EQ} para un tamaño poblacional y de mutación dados (simulaciones dadas por coalescencia bajo un modelo de mutación particular). Para definir este último, se compararon 3 modelos³⁷: el modelo de alelos infinitos (IAM) (Kimura y Crow, 1964), el modelo de mutación por pasos (SMM) (Ohta y Kimura, 1973) y el modelo de mutación de dos fases (TPM) (Di Rienzo y col., 1994), con 95% de SMM y una varianza de 12 entre múltiples pasos, tal como fue recomendado por Piry y col. (1999). Se supone que el modelo TPM es el que mejor ajusta a la evolución de los microsatélites (Di Rienzo y col., 1994). La significación de cualquier desvío del equilibrio mutación-deriva se evaluó mediante una prueba de rangos con signo de Wilcoxon de dos colas (Cornuet y Luikart, 1996), haciendo un ajuste FDR por comparaciones múltiples. También se incluyó una aproximación cualitativa. En esta prueba se evaluó la distribución de frecuencias alélicas; si algún tipo de desviación es observada de la forma normal en “L” que se espera en las poblaciones (corrimiento de la moda hacia valores intermedios), debe sospecharse un cuello de botella reciente (Luikart y col., 1998).

Para obtener más información acerca de los patrones sugeridos por BOTTLENECK, se dividieron las muestras de A, P y R en clases diamétricas, de aproximadamente 20 individuos cada una. Para los sitios A y P se consideraron dos clases: individuos con valores de DAP (diámetro a la altura del pecho) (1) menor y (2) mayor a 37 cm. Para el sitio R se consideraron 4 clases: individuos

³⁷ Los modelos mutacionales pueden usarse para derivar el número esperado de alelos en una población con la heterocigosis observada. En el **modelo IAM**, cada mutación crea aleatoriamente un alelo (esto es, se altera el número de repeticiones en microsatélites). El número de repeticiones no indica parecido filogenético. En el **modelo SMM**, nuevos alelos surgen a partir de otros por el agregado/eliminación de una repetición. De este modo, alelos que difieren en longitud por una repetición, están más relacionados entre sí que los que difieren por varias, habiendo una relación entre la cercanía a un ancestro reciente y la cantidad de repeticiones. El **modelo TPM** es una extensión del SMM; además de eventos de pérdida o ganancia de una repetición, en este modelo también se dan alteraciones en grandes números de repeticiones (para mayor detalle, ver Oliveira y col., 2006).

con DAP (1) menor a 44, (2) entre 44 y 53, (3) entre 53 y 61, (4) y mayores a 61 cm. Para cada clase se calculó la riqueza alélica (R_A) bajo rarefacción y su intervalo de confianza con el paquete ARES (página 110), interpolando a un tamaño de muestra de 18 individuos.

5. RESULTADOS

5.1. Diversidad genética

Los 7 loci estudiados fueron polimórficos, con 5-10 alelos cada uno y una media de 6.3 alelos por locus (Figura 5.6).

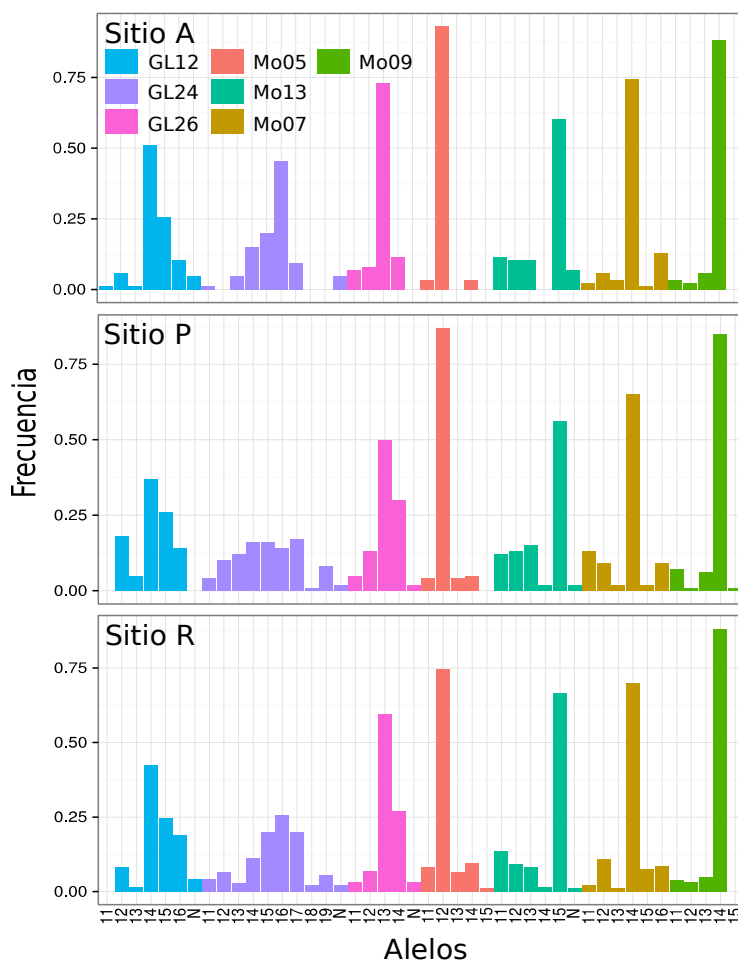


FIGURA 5.6 Frecuencias alélicas para los tres sitios de estudio. La letra “N” indica un alelo nulo. Los números en el eje de las abscisas indican el nombre (arbitrario) dado a cada alelo observado. Barras ubicadas en la misma posición del eje de las abscisas indican los mismos alelos en los tres gráficos. La leyenda en la imagen superior indica el color que le corresponde a cada marcador.

Los valores de diversidad genética se muestran en la Tabla 5.1. Para el análisis de rarefacción, el tamaño muestral en los sitios A, P y R se fijó en $N = 43$, el tamaño mínimo de individuos a través de poblaciones y loci. Tal como puede observarse en la Tabla 5.1, el análisis con ARES indicó que los intervalos de riqueza alélica no se solapan entre los sitios A y R.

TABLA 5.1 Valores de diversidad genética para la población y para los sitios de muestreo. Las letras para la columna R_A indican si los intervalos de confianza se solapan.

Unidad	H_o	H_E	R_A	A
Sitio A	0.409	0.460	36.000 ^a	5.143
Sitio P	0.500	0.568	40.101 ^{a, b}	5.729
Sitio R	0.468	0.540	39.815 ^b	5.688
Población	0.463	0.570	44	6.285

NOTA: *Unidad* unidad de análisis, H_o heterocigosis observada, H_E heterocigosis esperada, R_A riqueza alélica para un tamaño de 43 individuos en los sitios A, P y R y para la totalidad de los individuos (Población), A número promedio de alelos por locus bajo rarefacción, *a-b-c* coeficientes indicando diferencias significativas entre pares de sitios para el análisis de riqueza alélica (las diferencias son significativas cuando los intervalos de confianza no se solapan), A número promedio de alelos por locus.

La prueba U fue significativa y mostró un exceso de homocigotas ($P < 0.05$), tanto para la población en conjunto como para cada sitio (Tabla 5.2). En ninguno de los dos casos se observó evidencia de desequilibrio de ligamiento.

TABLA 5.2 Valores de significación para la prueba U en la población y en los sitios de muestreo. Los valores significativos se muestran en negrita.

Marcador	Población	Sitios		
		A	P	R
Mo05	0.389	0.069	1.000	0.713
Mo07	0.000	0.070	0.046	0.013
Mo09	0.580	1.000	0.157	0.485
Mo13	0.000	0.002	0.002	0.000
GL12	0.000	0.070	0.037	0.001
GL24	0.000	0.000	0.000	0.002
GL26	0.000	0.800	0.041	0.000
Global	0.000	0.000	0.000	0.000

5.2. Análisis de estructura genética

El coeficiente de endogamia (F_{IS}) mostró un exceso de homocigotas para la población y para los sitios (Tabla 5.3).

TABLA 5.3 coeficientes de endogamia (F_{IS}) para la población y cada sitio (A, P, R) individual. Los valores significativos se muestran en negrita.

Marcador	Población	Sitios		
		A	P	R
Mo05	0.004	0.297	-0.095	-0.046
Mo07	0.127	0.014	0.154	0.138
Mo09	0.000	-0.087	0.033	0.013
Mo13	0.145	0.100	0.114	0.176
GL12	0.052	-0.134	0.055	0.112
GL24	0.200	0.483	0.170	0.068
GL26	0.234	-0.061	0.157	0.353
Global	0.130	0.100	0.110	0.127

El valor de F_{ST} global fue bajo pero significativo ($F_{ST} = 0.015$), sugiriendo diferenciación entre sitios. Los valores de F_{ST} entre pares de sitios, mostraron valores significativos para las comparaciones A-P ($F_{ST} = 0.069$) y A-R ($F_{ST} = 0.043$) Los resultados de AMOVA mostraron la mayor proporción de la variación dentro de los individuos (Tabla 5.4).

TABLA 5.4 Resultados del análisis de AMOVA. Todos los valores de φ presentaron $P = 0.001$.

Fuente de variación	g.l.	% var	φ
Entre sitios (φ_{ST})	2	1.428	0.014
Entre individuos dentro de sitios (φ_{IS})	183	12.291	0.125
Dentro de individuos	186	86.281	-
Total	371	100	-

NOTA: g.l. grados de libertad, % var componentes de la varianza (porcentual), φ coeficiente phi del AMOVA para cada nivel jerárquico.

El análisis con STRUCTURE no mostró evidencia para rechazar la hipótesis de una población genéticamente continua. STRUCTURE HARVESTER detectó $K = 1$ como valor más probable. El método de ΔK no puede evaluar si $K = 1$ es la mejor solución, por esta razón se seleccionó el modelo en función de la verosimilitud. De las 10 iteraciones para cada K desde $K = 1$ a $K = 10$, las 10 correspondientes a $K = 1$ tomaron los valores de verosimilitud más altos.

5.3. Análisis de coancestría

El análisis de coancestría mostró una tendencia decreciente en los valores de f_{ij} en función de las clases de distancia en el rango 0-900 m (Figura 5.7). El valor de $f_{(1)}$ obtenido fue de 0.0347 (95% CI: 0.0140 - 0.0561). La regresión de f_{ij} en función de $\ln(d_{ij})$ mostró una pendiente $b = -0.0072$ (95% CI: (-0.0146) – (-0.0003), $P = 0.044$).

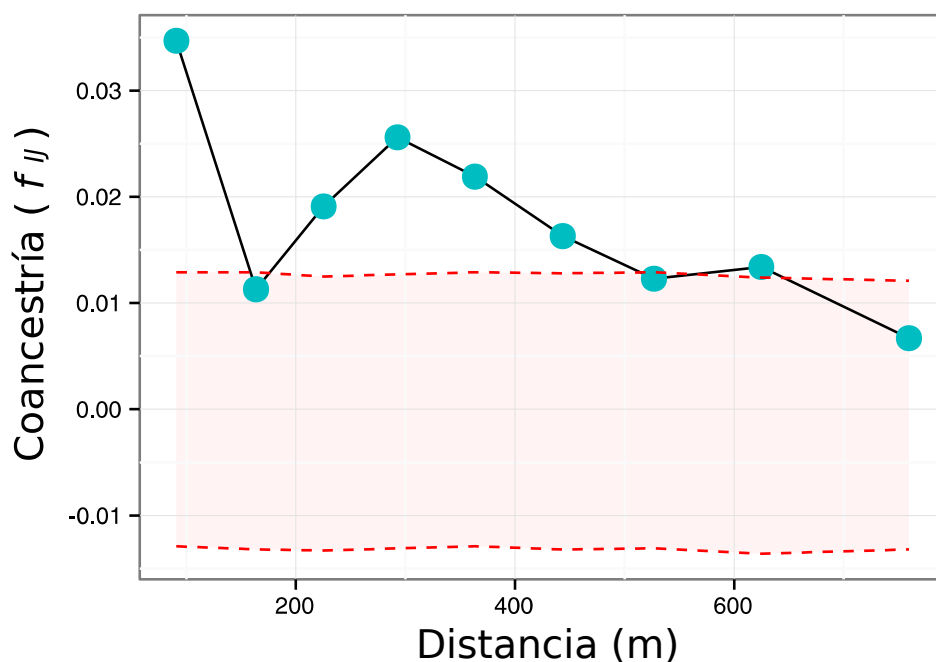


FIGURA 5.7 Coancestría en función de la distancia promedio por clase entre pares de individuos. El área delimitada por las líneas punteadas denota el intervalo de confianza del 95% para los valores obtenidos bajo aleatorización (hipótesis nula: los valores de parentesco promedio por clase no son diferentes de lo esperado por azar).

5.4. Análisis demográfico

El análisis con el programa BOTTLENECK mostró una deficiencia significativa de heterocigotas ($H_E < H_{EQ}$) bajo los modelos TPM y SMM en el sitio A ($P < 0.01$) y en R ($P \leq 0.02$), y sólo bajo SMM en el sitio P ($P = 0.04$). Los tres sitios mostraron ajuste a una distribución asimétrica con forma en “L” en el gráfico de frecuencias alélicas (Figura 5.8).

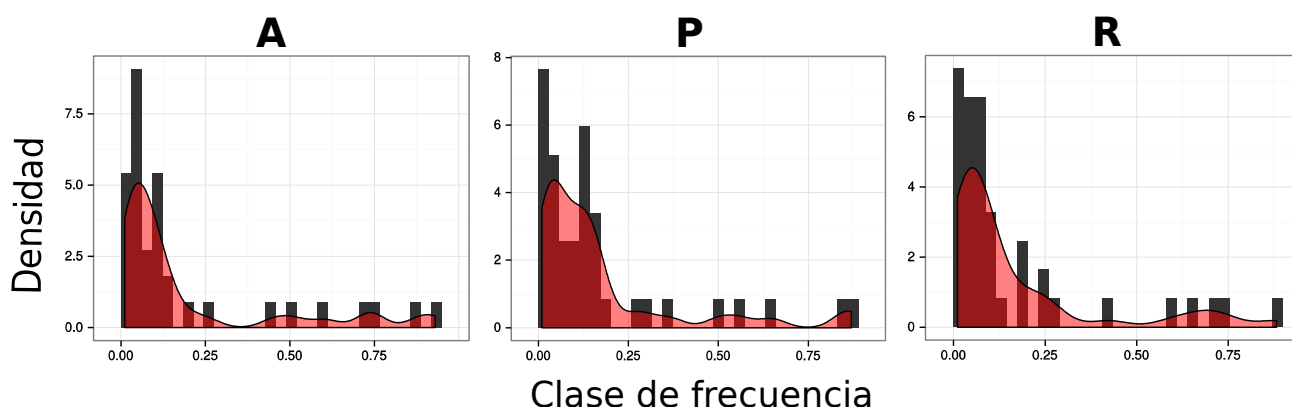


FIGURA 5.8 Histogramas para las clases de frecuencias alélicas, considerando todos los loci, en cada uno de los sitios estudiados. Los 3 casos muestran histogramas con forma en "L".

La Figura 5.9, que muestra los histogramas de los valores de DAP para cada sitio, sugiere la presencia de una distribución bimodal en A. Esta distribución separa la población en dos grupos etarios (individuos con DAP menor y mayor a 37 cm, aproximadamente).

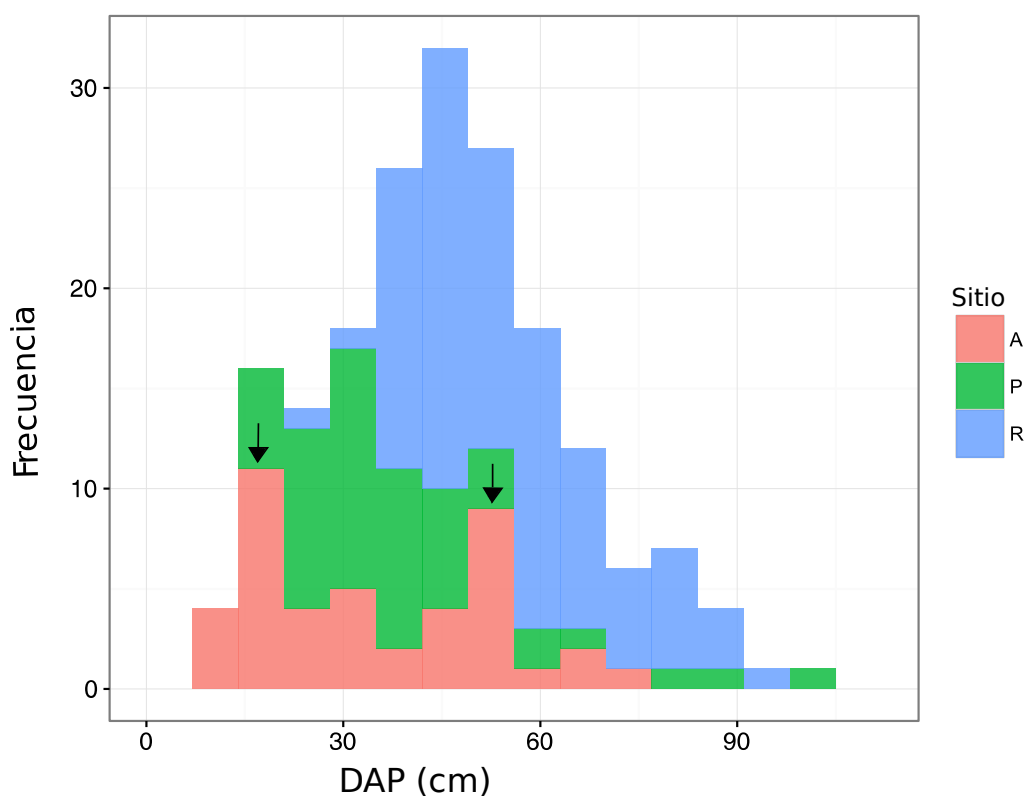


FIGURA 5.9 Histograma (frecuencia absoluta) para los valores de DAP (diámetro a la altura del pecho, en cm) en los individuos analizados según los tres sitios de muestreo considerados (A, P y R). Las flechas marcan los valores más frecuentes de la distribución bimodal en A.

La comparación mediante rarefacción de la riqueza alélica (Tabla 5.5 y Figura 5.10) mostró que los individuos con DAP menor a 37 cm en el sitio A (A1) tienen menor número de alelos que aquellos con DAP menor de 37 cm en P (P1) y de 44 cm en R (R1), mientras que P1 y R1 no se diferencian entre sí en el número de alelos.

TABLA 5.5 Valores de riqueza alélica para cada una de las clases diamétricas consideradas

Muestra	N	Rango de DAP (cm)	R_A	IC_{INF}	IC_{SUP}
A1	25	DAP < 37	27.863	24.694	31.032
A2	18	DAP ≥ 37	32.000	27.581	36.419
P1	29	DAP < 37	35.479	31.425	39.534
P2	21	DAP ≥ 37	34.253	29.889	38.618
R1	22	DAP ≤ 44	36.335	32.737	39.932
R2	22	44 < DAP ≤ 53	32.719	29.763	35.675
R3	21	53 < DAP ≤ 61	32.668	26.872	38.464
R4	28	DAP > 61	33.942	30.800	37.083

NOTA: *Muestra* grupos de DAP, *N* número de individuos en la muestra, *Rango de DAP* rango de DAP en la muestra, R_A riqueza alélica calculada bajo rarefacción, IC_{INF} valor inferior del intervalo de confianza de la riqueza alélica, IC_{SUP} valor superior del intervalo de confianza de la riqueza alélica. Los intervalos de los pares A1-P1 y A1-R1 no se solapan.

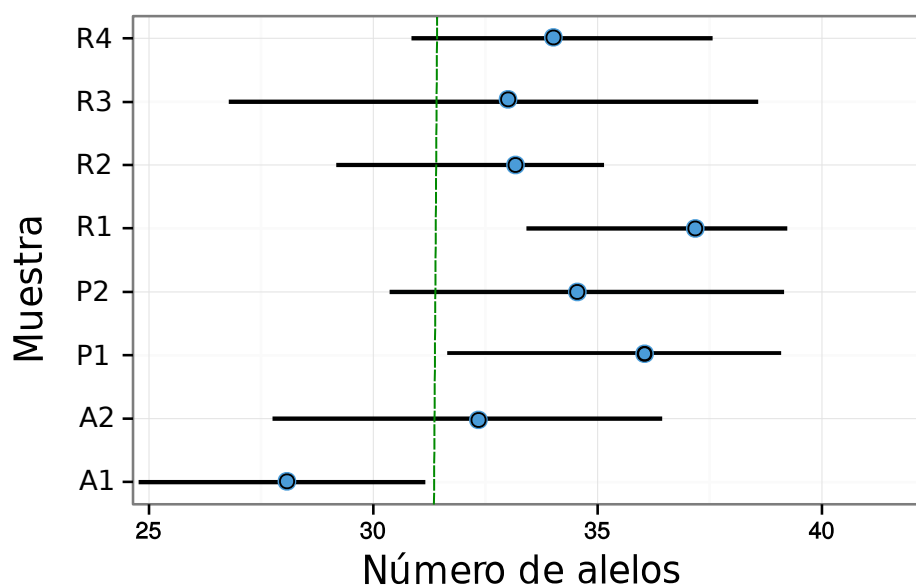


FIGURA 5.10 Valores observados de riqueza alélica (círculos celestes) e intervalos de confianza para las clases diamétricas consideradas para los distintos sitios en el análisis demográfico. Pares de intervalos que no se solapan, indican diferencias significativas al 95% (pares A1-P1 y A1-R1, el no solapamiento de estos pares de intervalos es marcado por la línea vertical verde que los separa).

6. DISCUSIÓN

Los niveles de diversidad genética estimados como heterocigosis en este capítulo fueron altos ($H_o = 0.463$ y $H_e = 0.570$). Estudios previos en la sección Algarobia indican que las poblaciones son muy variables, pero que la diferenciación entre poblaciones y especies es baja (Saidman, 1985; Saidman y Vilardi, 1987, 1993; Saidman y col., 1998; Bessega y col., 2000; Ferreyra y col., 2004, 2007, 2010, 2013). En este capítulo se encontró una diferenciación genética baja pero significativa entre sitios de muestreo. La mayor parte de la variación se halló en el nivel jerárquico más bajo (individuos), hecho esperable en una especie de fecundación cruzada. La diferenciación entre zonas medida a través del valor de F_{ST} global fue baja. Los valores del F_{IS} globales mostraron un exceso general de homocigotas. Trabajos previos mostraron un exceso significativo de homocigotas en poblaciones de esta sección usando diversas técnicas (Saidman, 1985, 1986; Saidman y Vilardi 1987; Saidman 1988, 1990; Saidman y Vilardi, 1993; Keys y Smith, 1994; Verga, 1995; Saidman y col. 1997; Bessega y col, 2011, 2015b). En el presente caso, cuando todos los individuos fueron agrupados en una sola población, se pudo observar un desvío significativo del equilibrio de Hardy Weinberg (exceso de homocigotas) para más de la mitad de los loci, con un F_{IS} global significativo y positivo. Esto podría sugerir efecto Wahlund (Castric y col., 2002), endogamia (Wright, 1921) o un artefacto causado por alelos nulos (DeWoody y col., 2006); la primera y la segunda causa son plausibles, ya que la población mostró subestructuración ($F_{ST} > 0$ significativo), y que un trabajo previo en la zona aquí analizada (Bessega y col., 2012) encontró una endogamia biparental baja de aproximadamente un 5%, que se debería al sistema de dispersión de semillas, lo que es congruente con el valor relativamente bajo de coancestría para individuos cercanos (Van Rossum y Triest, 2006) hallado en este capítulo ($f_{(1)} = 0.0347$). Los valores de F_{IS} globales para cada sitio también mostraron exceso de homocigotas; sin embargo, a nivel loci se observaron pocos valores significativos, para exceso y defecto de homocigotas, sin una tendencia clara hacia alguno de estos casos.

Bajo escenarios de largas distancias de dispersión por semillas con flujo génico limitado, no debería presentarse estructura genética en la población, ya que los individuos emparentados estarían bastante dispersos. En cambio, bajo dispersión restringida siempre se espera una estructura genética a escala fina, incluso si el movimiento de polen es aleatorio y de larga distancia, debido a que los hermanos tendrían un patrón espacial agregado. En *Prosopis*, tanto la dispersión por polen como

por semillas se encuentra limitada. Estos dos factores son claves en el establecimiento de estructura genética espacial (Wright, 1943; Schnabel y Hamrick, 1995; Streiff y col., 1998; Vekemans y Hardy, 2004). Dado que la especie de *Prosopis* son entomófilas (Genise y col., 1990), el polen no es usualmente capaz de migrar largas distancias, lo que favorece el cruzamiento recíproco entre plantas cercanas (Bessega y col., 2000). En el área de estudio, Bessega y col. (2012) detectaron una distancia de polinización promedio de 5 a 30 m, con un vecindario de polinización efectivo de 22 m de radio alrededor de cada árbol madre. El sistema de dispersión de semillas endozoico en especies de *Prosopis*, mediado por pequeños herbívoros nativos y ganado (Hafez, 1962; Mares y col. 1977), determina que las semillas de la misma planta madre sean comidas, transportadas y finalmente depositadas todas juntas en las heces de los animales. Este mecanismo causa la germinación de hermanos o medios hermanos en áreas estrechas (Keys, 1993; Bessega y col., 2000, Bessega y col., 2012), y en consecuencia un elevado parentesco entre árboles cercanos, que tenderían a compartir alelos idénticos por descendencia con mayor probabilidad que los árboles distantes.

En este estudio se encontró una disminución significativa en los valores de coancestría estimados con el aumento de la distancia entre individuos. Este patrón observado suele ser interpretado como limitación en la dispersión de semillas (Dick y col., 2008). El análisis con STRUCTURE por otra parte no mostró evidencia para rechazar la hipótesis nula de un solo grupo genético habitando el área estudiada. Estos dos resultados en conjunto sugieren un proceso de aislamiento por distancia, generado por un efecto local. El valor de F_{ST} global, significativo pero bajo, también da soporte en esta dirección. Los resultados estarían así indicando que la diferenciación de la población a través del espacio sería más bien gradual que discreta. En el caso de la coancestría para individuos cercanos, la estima obtenida en este trabajo ($f_{(1)} = 0.0347$) se mostró congruente con los valores observados en sistemas de apareamiento exogámicos, y está cerca del valor esperado de 0.0312 para el coeficiente de coancestría de primos de segundo grado (Crow y Kimura, 1970), tal como muestran trabajos previos en *Manilkara huberi* (Azevedo y col., 2007) y *Dalbergia nigra* (Buzatti y col., 2012).

A pesar de que los efectos sobre la variación medida como heterocigosis requieren largos períodos de tiempo para ser observados, los cambios demográficos tales como reducciones y expansiones poblacionales recientes, o eventos de inmigración con influjo de alelos raros, pueden ser detectados en períodos de tiempo corto, a través del desvío de la heterocigosis respecto al valor

esperado bajo equilibrio mutación-deriva (Luikart y Cornuet, 1998). La prueba con signo de Wilcoxon no detectó señales de exceso de heterocigotas en este estudio, tal como se esperaría en caso de presentarse cuellos de botella recientes. En cambio, sí se detectó un defecto de heterocigotas ($H_E < H_{EQ}$) en cada uno de los tres sitios considerados (bajo TPM y SMM en los sitios A y R; sólo bajo SMM en P), sugiriendo posibles eventos de expansión poblacional o de influjo de alelos raros a través de inmigración. En concordancia con estos resultados, no se observó una alteración significativa en la distribución de frecuencias alélicas en ninguno de los tres sitios respecto a la forma en “L” esperada en situación de equilibrio. Generalmente se espera observar una distorsión en la forma esperada a causa de la pérdida de alelos raros (Luikart y col., 1998). Esto concuerda con la ausencia de señales de cuellos de botella detectada por BOTTLENECK.

En el caso particular del sitio A, el histograma del diámetro a la altura del pecho (DAP) de los árboles mostró una distribución bimodal, con individuos de menor y de mayor DAP (siendo este un estimador de la edad). El análisis de la riqueza alélica mostró un valor de riqueza alélica bajo rarefacción menor en los árboles más jóvenes del sitio A que en los individuos jóvenes de los otros dos sitios (P y R). Estos datos, en conjuntos con los resultados de BOTTLENECK y la distribución de frecuencias alélicas, conforman evidencia a favor de la hipótesis que pudo haber un proceso de inmigración reciente hacia este parche. Un posible modelo explicativo para el patrón genético observado es que el sitio A, previamente usado con fines de producción agrícola, podría haber sido colonizado mediante un proceso de dispersión reciente, originado en un muestreo no representativo, donde un bajo número de árboles contribuiría a la población de semillas (Davies y col., 2010). Debido a la pérdida de alelos raros durante eventos de dispersión hacia sitios específicos, es esperable como consecuencia una riqueza alélica más baja en las poblaciones de semillas y en los individuos reclutados. La distribución bimodal en los valores de DAP, que sugiere la existencia de dos cohortes habitando este sitio (una de menor y otra de mayor edad) y la menor riqueza alélica en los individuos más jóvenes del sitio A, da soporte a esta conjetura. El ganado presente en la zona, capaz de transportar frutos conteniendo hermanos y medios hermanos, podría haber sido un factor relevante en la generación de la estructura genética observada en este parche. Bajo escenario similares al observado, los resultados de BOTTLENECK podrían indicar eventos inmigratorios recientes.

La diversidad genética puede aparecer estructurada espacialmente a diferentes escalas, tal como a nivel paisaje, población o entre individuos cercanos, debido a diferentes procesos ecológicos y genéticos operando en las poblaciones naturales (Zeng y col., 2012). En cada nivel de estudio, se requiere un diseño adecuado para apreciar los procesos que corresponden a esa escala y poder utilizar adecuadamente las herramientas de análisis disponibles. En la presente investigación, el posible evento de inmigración hacia el sitio A pudo inferirse a partir del muestreo local intensivo que se realizó allí. El estudio del modelo de aislamiento por distancia a partir de la coancestría también pudo llevarse a cabo debido a que se dispuso de gran cantidad de individuos a pequeña escala.

En conclusión, en esta parte se analizaron los patrones de EGE a escala fina en *P. alba*. La población estudiada posee una alta diversidad genética. Los patrones espaciales sugieren, que a la escala estudiada, la estructuración puede deberse a un efecto local por flujo génico restringido, tal como predice el modelo de aislamiento por distancia. Los análisis realizados con el programa STRUCTURE, los índices F de Wright y el estudio de la coancestría mostraron resultados complementarios. Por último, la posible ocurrencia de un posible evento de inmigración reciente en el sitio A, indica que estos procesos se pueden detectar mediante diseños de muestreo adecuados.

Los resultados de este capítulo evaluaron de modo sistemático y por primera vez los patrones espaciales a escala fina en una especie de *Prosopis*, mediante un modelo de aislamiento por distancia, técnicas de autocorrelación espacial y estudio de cambios demográficos recientes, resaltando la complejidad de la dinámica mostrada por poblaciones de distribución continua en paisajes modelados por el hombre.

Estudio de la estructura espacial de caracteres cuantitativos a escala fina



SEGUNDA PARTE: ESTUDIO DE LA ESTRUCTURA ESPACIAL DE CARACTERES CUANTITATIVOS A ESCALA FINA

1. RESUMEN

La presencia de autocorrelación espacial (AE) es un fenómeno frecuente en paisajes heterogéneos, donde tanto la similitud genética entre individuos causada por coancestría, como la variación ambiental a nivel de microhábitat, son capaces de generar patrones espaciales a escala fina en caracteres cuantitativos. En este capítulo se examinó la presencia de esta clase de patrones en caracteres cuantitativos de la población estudiada. El trabajo se realizó combinando información fenotípica con datos de sensores remotos. Se midieron 11 rasgos de hoja y el diámetro a la altura del pecho (DAP) en cada individuo. Se extrajo además, para cada punto (árbol) georreferenciado, los valores de NDVI, tipo de suelo, precipitación anual y evapotranspiración potencial anual. Con la información obtenida, se estudió *i*) la relación entre caracteres y condiciones ambientales, y *ii*) la presencia de patrones de AE. Los patrones encontrados, sugieren una distribución no aleatoria de variación fenotípica en el área de estudio, lo que podría asociarse, al menos en parte, con flujo génico restringido. Los resultados de este capítulo muestran cómo la presencia de aislamiento por distancia podría jugar un rol importante en la formación de patrones espaciales a escala fina en rasgos foliares y de crecimiento de *P. alba*.

2. INTRODUCCIÓN

2.1. Elementos conceptuales para el estudio a escala fina

El análisis espacial es un campo de estudio en crecimiento dentro de la ecología y otras disciplinas de las ciencias biológicas. Las dimensiones temporal y espacial, inherentes a todo fenómeno natural, son fundamentales en los procesos ecológicos, siendo su estudio sistemático bastante reciente. Un concepto geoestadístico clave es el de **autocorrelación espacial** (AE), es decir, la propiedad de pares de objetos espaciales de ser más similares (AE positiva) o menos similares (AE negativa) entre sí que lo esperable por azar. La autocorrelación es una propiedad general de variables ecológicas (Legendre y Legendre, 2012), que puede generar problemas a los análisis estadísticos clásicos, ya que las observaciones autocorrelacionadas no son independientes entre sí³⁸.

Otra pieza conceptual fundamental es la **escala**, tal como mencionara Levin en su importante trabajo de 1992, “The problem and pattern of scale in ecology”. Desde el punto de vista del muestreo, la escala de un estudio espacial comprende tres elementos (Figura 5.11): el **tamaño de grano** (el tamaño promedio de la unidad de muestreo), la **extensión** (la dimensión espacial máxima donde las observaciones son hechas y que las incluye) y el **intervalo de muestreo** (intervalo promedio entre las unidades de estudio, comúnmente denominado *lag*) (Dungan y col., 2002; Borcard y col., 2011; Legendre y Legendre, 2012).

38 En un modelo de regresión, por ejemplo, la varianza es estimada de modo sesgado en presencia de AE (algo que puede ser común con datos espaciales y/o temporales), manteniéndose los coeficientes de regresión sin sesgo. Esto altera la validez de los test de significación y las medidas de ajuste (Anselin, 1995; Anselin y Griffith, 1988).

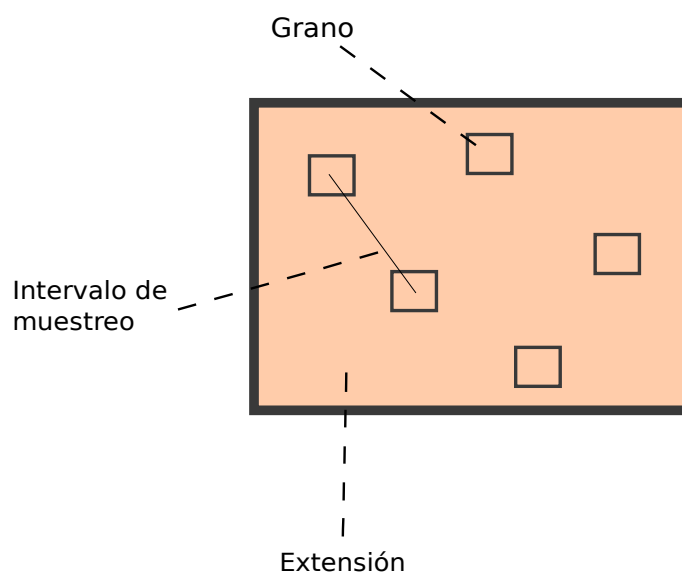


FIGURA 5.11 Componentes de la escala. Cada rectángulo indica el *grano*. La *extensión* es el área en color. El *intervalo de muestreo* es la distancia promedio entre los centroides de cada rectángulo. Adaptado de Legendre y Legendre (2012).

Estos factores afectan la complejidad o la variabilidad (es decir, la **heterogeneidad**) de propiedades ecológicas espaciales en cualquier sistema (Li y Reynolds, 1995; Dungan y col., 2002). Los patrones observables en un estudio son **relativos a la escala utilizada**. De este modo, manteniendo un tamaño de grano constante, un incremento en la extensión incorpora más heterogeneidad al sistema (Figura 5.12); manteniendo la extensión constante, un incremento del grano puede conducir a una disminución de la varianza entre granos y a un aumento de la varianza contenida dentro de los mismos (Figura 5.12) (Wiens, 1989). Por estas razones, para un mejor entendimiento de los procesos ecológicos y genéticos, es necesario generar modelos a diferentes escalas de análisis de un modo integrativo.

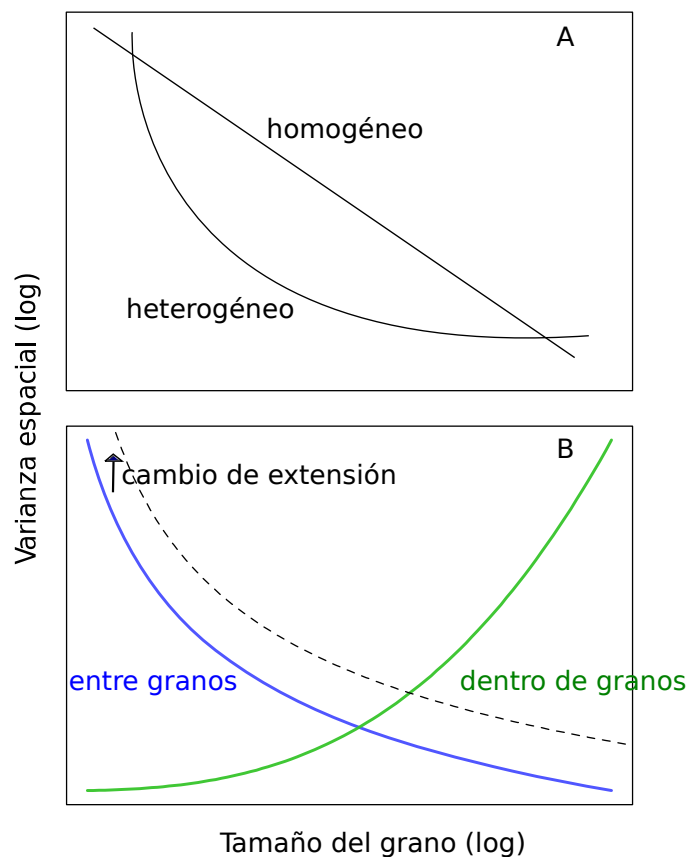


FIGURA 5.12 (A) Con el aumento del grano, la varianza total del sistema decrece, de modo diferente sin embargo para áreas homogéneas y heterogéneas. Esto se encuentra relacionado con los componentes de variación *dentro* y *entre granos* (muestras). **(B)** Con el incremento del tamaño del grano, menos varianza es debida a diferencias *entre granos* y más varianza es incluida *dentro de granos*. Al incrementarse la extensión de la investigación, también se incrementaría el componente *entre granos* de la varianza, agregando nuevos tipos de parches al paisaje estudiado, pero la varianza *dentro de granos* no sería afectada. Imagen y explicación extraídos de Wiens, 1989.

2.2. Elementos para el estudio a escala fina en *Prosopis alba*

Los mecanismos de dispersión de polen y semillas, son factores importantes capaces de generar patrones de AE a través de la restricción en el flujo génico (Sokal y col., 1987). Caracteres cuantitativos con base genética y alta heredabilidad podrían mostrar patrones que reflejen la estructura genética de la población subyacente, descrita con marcadores moleculares, tal como se vio en varios trabajos (Bocquet-Appel y Sokal, 1989; Argyres y Smith, 1991; Hardy y col., 2000).

V. Estudios con una población natural - estructura espacial de caracteres cuantitativos a escala fina

Otra causa de estructura a escala fina es la heterogeneidad ambiental, que puede promover procesos de adaptación local, plasticidad fenotípica y demográficos (Baythavong, 2011; Paccard y col., 2013), generándose de este modo patrones espaciales autocorrelacionados en variables genéticas y cuantitativas. Entre los distintos tipos de heterogeneidad ambiental, la variación a nivel de micrositio puede tener una gran influencia en la estructura espacial en árboles al producir patrones de AE positiva entre individuos vecinos (Zhang y col., 2004).

Reportes previos en el género *Prosopis* mostraron asociación entre marcadores moleculares y variables ambientales (Ferreira y col., 2010) y gran variación geográfica no neutral en caracteres foliares y de forma, tales como la altura de los árboles o el ancho de la copa (Darquier y col., 2013). Diversos autores llevaron también a cabo investigaciones acerca de la forma en que covarían la morfología foliar y variables ambientales en *Prosopis* (Bleckmann y col., 1980; Nilsen y col., 1986; Vilela, 1993; Vilela y Palacios, 1997; Villagra, 1998; Villagra y Cavagnaro, 2006). Estudios de la morfología en *P. alba* mostraron que en esta especie la forma de la hoja es modelada por factores genéticos y ambientales (Bessega y col., 2009, 2012; Ewens y col., 2012; Ferreira y col., 2013; Bessega y col., 2015b).

3. OBJETIVOS

- a. Explorar los patrones espaciales a escala fina en caracteres cuantitativos (rasgos foliares y de crecimiento), de forma global y local.
- b. Evaluar si algunos de los caracteres analizados se encuentran espacialmente autocorrelacionados a la escala estudiada.
- c. Estudiar posibles relaciones entre los patrones espaciales de las variables cuantitativas, genéticas y ambientales.

4. MATERIALES Y MÉTODOS

4.1. Colección de datos de morfométricos y de crecimiento

Se midieron en total 11 caracteres de morfología foliar en cada árbol individual: longitud del peciolo (PEL), número de pares de foliólulos por pinna (NLP), longitud de las pinnas (PIL), número de pinnas (NPI); longitud (LEL), ancho (LEW), relación largo/ancho (LEL/LEW), falcado (LEF), área (LEA), ápice (LEX) y relación ápice/área total (LEX/LEA) de los foliólulos. La metodología utilizada para la obtención de los datos se detalla en el Apéndice 1 (Protocolo 1, página 155). Se emplearon también los datos de diámetro basal a la altura del pecho (DAP) de cada árbol. La distribución estadística de cada variable a través de los sitios de estudio se analizó con el test de Kruskal-Wallis. Las comparaciones entre pares de sitios se realizó mediante el test de Mann-Whitney. La homogeneidad de varianzas para cada rasgo entre los sitios de estudio se evaluó mediante el test de Fligner-Killeen. Para corregir los valores P por comparaciones múltiples, se utilizó el método False Discovery Rate (FDR) (Benjamini y Hochberg, 1995). Estos análisis se llevaron a cabo con el paquete *stats* en R 3.1.3³⁹ (R Core Team, 2014). Para los datos de DAP se construyó además un mapa de valores predichos a través de todo el área de estudio. Se utilizó para este propósito el método de interpolación “Kriging Bayesiano Empírico⁴⁰” (Empirical Bayesian Kriging), en ArcGIS 10.1 (ESRI, 2012). Este método tiene, respecto a los modelos de kriging clásicos, las ventajas de manejar un conjunto de variogramas, en vez de uno solo, pudiendo ser utilizado para interpolar datos no estacionarios (Krivoruchko, 2012).

³⁹ En este capítulo se utiliza la versión 3.1.3 de R.

⁴⁰ **Kriging** se denomina a un conjunto de técnicas de interpolación espacial, donde a los valores conocidos f , se les asigna un peso w para predecir los valores $F(x)$ en los sitios x : $F(x) = \sum w_x f_x$. En los métodos clásicos, estos pesos se eligen de modo de obtener un estimador insesgado y que minimice el error cuadrático medio. Para obtener los pesos se utilizan variogramas en los cálculos. El **variograma empírico** de un conjunto de observaciones está dado por la función $\gamma(d) = (\sum (y_i - y_j)^2)/2n$. El cálculo se realiza para una serie de intervalos de distancia d , usando para cada d los n pares de individuos (i, j) pertenecientes al intervalo. Tras obtener el variograma empírico, se realiza un ajuste a un **variograma teórico**, y finalmente se obtienen los pesos para la interpolación. El método bayesiano aquí empleado (*Empirical Bayesian Kriging*) es un método de kriging mucho más sofisticado que los tradicionales, que permite capturar la variación a nivel local. Para ello el algoritmo divide al área de estudio en zonas, resultando el mapa predicho en un “mosaico” de interpolaciones locales. A cada una de las zonas se les aplica un método iterativo que implica 1) el ajuste inicial de los valores observados a un variograma teórico, 2) simulación de datos bajo este variograma, 3) ajuste de las simulaciones a un variograma teórico usando el teorema de Bayes 3) iteración a través de los pasos (2) y (3). De este modo, se obtiene un espectro de variogramas mediante una aproximación bayesiana.

4.2. Análisis de NDVI

El NDVI⁴¹ (índice de vegetación normalizado) es la combinación aritmética del rojo (λ : 0.58 - 0.68 μm) y del infrarrojo cercano (IRc) (λ : 0.72 - 1.10 μm) del espectro electromagnético. Este índice se utiliza aquí como un estimador de la densidad del dosel. El NDVI puede distinguir la biomasa verde de otras superficies, ya que la clorofila de estos tejidos absorbe las ondas de la región del rojo para la fotosíntesis y refleja las del infrarrojo cercano (IRc), debido a la dispersión generada por la estructura interna de la hoja (Tucker, 1979). Los valores de NDVI correlacionan positivamente con la biomasa, la densidad del follaje y el área foliar (Wilson y col., 2002); por lo tanto, se espera una diferencia en el NDVI entre sitios con diferentes grados de cobertura verde (Washington-Allen y col., 2008).

Se obtuvieron 26 imágenes libres de nubes y corregidas a nivel terreno (L1T⁴²) del satélite Landsat 5, (path 230, row 79) abarcando un período temporal entre los años 2002 y 2009, tal como se describe en el Apéndice 1 (Protocolo 3, página 159). La precisión geodésica típica de las imágenes L1T se encuentra dentro de los 30 m (Gitelson y col., 2012). Las imágenes se convirtieron a la proyección UTM (Universal Transverse Mercator) y se recortaron posteriormente al área de estudio. Dado que el satélite tomó las imágenes en diferentes momentos del año y diferentes horas del día, las mismas poseen condiciones atmosféricas y ángulos solares variables en estado nativo. Estos factores alteran las longitudes de onda roja e IRc, influenciando las estimas de NDVI (Huete y Tucker, 1991). Para minimizar los efectos atmosféricos causados por la absorción y la dispersión de las ondas electromagnéticas (Du y col., 2002), se realizó una corrección siguiendo el método COST (COSine Theta) (Chavez, 1996) descrito en el Apéndice 1 (Protocolo 3, página 159). Con la serie de 26 imágenes corregidas, se calculó el NDVI promedio para cada píxel de 30 m x 30 m en el área de estudio. Se espera de este análisis que los valores de NDVI promediados a través del tiempo distingan entre diferentes tipos de cobertura.

Los cálculos de NDVI se llevaron a cabo con el paquete EcoGenetics (Roser y col., 2015) en R. Los valores obtenidos se compararon globalmente entre sitios mediante el test de Kruskal-Wallis y entre pares de sitios con el test de Mann-Whitney. Los niveles de significación se ajustaron por

41 Para más detalles acerca del índice ver página 45 del Capítulo 1.

42 Dentro de los niveles de pre-procesamiento incluido en las imágenes, las L1T (Level-1 Terrain corrected) poseen precisión radiométrica, geométrica y topográfica. Ver descripción en: http://landsat.usgs.gov/descriptions_for_the_levels_of_processing.php

V. Estudios con una población natural - estructura espacial de caracteres cuantitativos a escala fina

comparaciones múltiples por medio del método FDR. Estos análisis se realizaron con el paquete *stats* de R. Para obtener un resumen gráfico de los valores de NDVI, se calculó un mapa de predicción con Kriging Bayesiano Empírico en ArcGIS 10.1.

4.3. Asociación entre variables cuantitativas y ambientales

4.3.1. Obtención de datos

Los valores de tipo de suelo, precipitación y evapotranspiración para cada individuo se extrajeron de mapas de curvas de nivel, obtenidos del programa SIGSE 2.0 (Angueira y col. 2007). Para la extracción de los valores correspondientes a la posición de cada árbol se utilizó el método de interpolación “topo to raster” en ArcGIS 10.1. La Figura 5.13 muestra una sección de mapa edafológico perteneciente a la provincia de Santiago del Estero, a una escala 1:500.000. Los individuos se muestrearon en áreas de las unidades cartográficas⁴³ AEtc-34 (complejo de suelos) y MNai-7 (asociación de suelos) (Tabla 5.6).

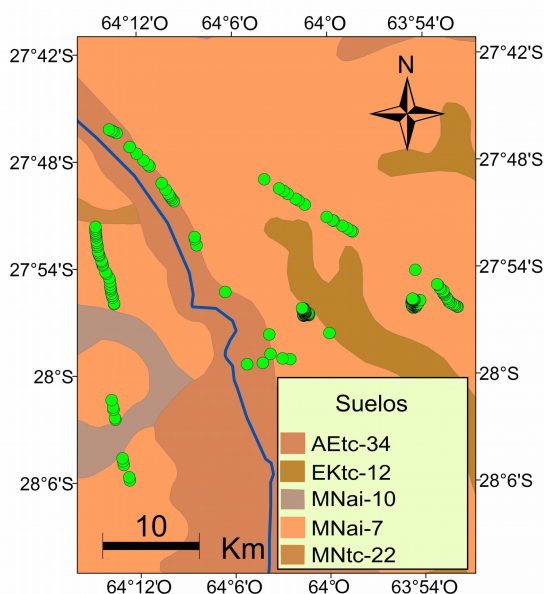


FIGURA 5.13 Sección de mapa edafológico de Santiago del Estero (1:500.000). Los puntos verdes indican los individuos muestreados. Los mismos pertenecen a las unidades cartográficas AEtc-34 y MNai-7.

⁴³ La **unidad cartográfica** corresponde a una porción del espacio ocupada por el suelo (lo que es distinto a una unidad taxonómica de suelo). Dentro de una unidad cartográfica, pueden distinguirse unidades taxonómicas “puras”. Este tipo de unidad cartográfica se denomina **consociación** y corresponde a una sola unidad taxonómica. En los casos que varios tipos de suelos conformen una unidad, la misma puede ser una **asociación** o un **complejo**. En las asociaciones, cada unidad taxonómica integrante ocupa un lugar en el paisaje y en un mapa una escala menor las mismas podrían separarse. En los complejos en cambio, las unidades taxonómicas están íntimamente ligadas y no es posible separarlas tal como en las asociaciones.

V. Estudios con una población natural - estructura espacial de caracteres cuantitativos a escala fina

TABLA 5.6 Propiedades de las unidades cartográficas de suelo MNai-7 y AEtc-34.

Unidad	Tipo de suelo	Factores limitantes generales	Subgrupo de suelo	Comp. (%)	Posición del subgrupo en el paisaje	Factores limitantes del subgrupo
MNai-7	asociación	1.climáticos	Haplustoles arídicos (molisol)	60	explanada general	1. clima 2. algo excesivamente drenado
		2. baja capacidad de retención de humedad	Torripsamientos típicos (entisol)	30	relieves planos convexos de pequeñas montañas	1. baja capacidad de retención de humedad
			Haplustol salídico (molisol)	10	micro-relieves cóncavos	1. salinidad 2. drenaje imperfecto
AEtc-34	complejo	1.drenaje	Natrualfes típicos (alfisol)	60	llanuras aluviales	1. sodicidad 2. salinidad 3. drenaje pobre
		2.sodicidad	Ustifluventes típicos (entisol)	30	llanuras aluviales	1. drenaje imperfecto
			Ustifluventes ácuicos (entisol)	10	zonas de escurrimiento	1. drenaje pobre

NOTA: *Subgrupo de suelo* subgrupos de los suelos incluidos en cada unidad de acuerdo a USDA (Soil Survey Staff, 1999), *Comp. (%)* fracción porcentual de cada subgrupo de suelo en la unidad correspondiente.

4.3.2. Análisis de árboles de inferencia condicional

Con el propósito de estudiar la relación entre los caracteres cuantitativos y las variables ambientales, se realizó un análisis de árboles de inferencia condicional⁴⁴ (Hothorn y col., 2006) con el paquete *party* (Hothorn y col., 2006) de R. Este análisis no requiere supuestos *a priori* acerca de la distribución de los datos. Como variables independientes se utilizaron datos del tipo de suelo (*SUE*), el NDVI temporal promedio calculado (*NDVI*), precipitación (*PP*) y evapotranspiración potencial anual (*ETP*). La resolución de las imágenes provenientes del satélite Landsat 5, determina un grano de 30 m para el NDVI, permitiendo utilizar a esta variable como un indicador ambiental local. El grano es intermedio para la clasificación de suelos y el mayor para las variables climáticas. El modelo de regresión utilizado fue el siguiente:

⁴⁴ El método de árboles de inferencia condicional estima una relación de regresión mediante una partición recursiva de la muestra. La regresión se realiza de modo jerárquico, dividiendo sucesivamente y de modo binario el conjunto de datos a partir de los predictores. El algoritmo comprende los siguientes pasos: 1) probar la hipótesis nula global de independencia entre las variables predictoras y la variable respuesta. Parar si la hipótesis no puede ser rechazada, en caso contrario seleccionar la variable predictora que presente la más fuerte asociación con la respuesta, 2) realizar una partición binaria de la muestra en función de la variable predictora, 3) repetir recursivamente los pasos (1) y (2) (ver la descripción incluida en el paquete *party* con el comando: *help("ctree")*).

$$y_i = \mu + SUE_i + NDVI_i + PP_i + ETP_i + \varepsilon_i \quad (\text{Expresión 5.4})$$

Donde y_i representa cada observación individual (un escalar representando el fenotipo univariado del individuo i , μ es el promedio general, SUE_i es el efecto fijo de suelo, $NDVI_i$, PP_i , y ETP_i son los efectos fijos del NDVI promedio, la precipitación y la evapotranspiración potencial, respectivamente, y ε_i es el error residual aleatorio.

4.3.3. Evaluación de la robustez del análisis con árboles de inferencia condicional ante la presencia de autocorrelación espacial en las variables cuantitativas

A pesar de que los árboles de inferencia condicional son robustos a la mayor parte de los problemas que pueden presentar los estudios estadísticos, son aún sensibles a la AE, que infla el error de tipo I (Czúcz y col., 2011). La presencia de AE en los residuos puede ser utilizada como una herramienta diagnóstica para evaluar si un modelo no se parametrizó adecuadamente (Dale y Fortin, 2005). La AE en los residuos de cada modelo de inferencia condicional se analizó mediante correlogramas para el estadístico I de Moran (Moran, 1950; Sokal y Oden, 1978a; Dormann y col., 2007) con el paquete EcoGenetics. Los valores de I normalmente se encuentran dentro del rango ± 1 , con $E(I) = -1 / (N - 1)$ bajo ausencia de AE (Cliff y Ord, 1981). El análisis se realizó usando 9 clases de distancias íter-individuales, con 250 pares de individuos en cada una, en el intervalo 0-900 m. La significación para cada intervalo se obtuvo mediante un test de permutación de individuos a través de sitios, utilizando 9999 repeticiones. Para corregir los valores P por comparaciones múltiples, se utilizó el método Holm- Bonferroni secuencial (HSB) (Legendre y Legendre, 2012). Los pesos espaciales⁴⁵ utilizados (pesos binarios) se obtuvieron calculando una secuencia de matrices binarias simétricas, en las cuales para cada celda representando la conexión entre los pares de individuos (i, j) , se asignó el valor de 1 si la distancia entre ellos se encontraba dentro de un rango específico de distancias, y 0 en caso contrario.

⁴⁵ Sea un conjunto de objetos $x_1, x_2, \dots, x_n$; una matriz de pesos espaciales es una matriz cuadrada de dimensión $n \times n$, donde cada fila $x_1, \dots, x_n$ y cada columna $x_1, \dots, x_n$ representa un objeto y cada celda representa la conexión entre los pares de objetos (intersección de filas y columnas). Las celdas contienen un valor (peso) asignado a la conexión. En una matriz binaria, las celdas poseen el valor de 1 si los pares de objetos (x_i, x_j) están conectados y 0 en caso contrario. Esto por ejemplo anula aquellos valores correspondientes a pares no conectados (multiplicación por $w_{ij} = 0$) en un índice como el I de Moran. Otros esquemas de pesos diferentes del binario son comunes, como por ejemplo el que asigna a cada conexión un valor inverso a la distancia entre los pares de objetos relacionados (las conexiones entre objetos distantes poseen menor peso que aquellas entre objetos cercanos).

4.4. Análisis de superficies de tendencia

4.4.1. Modelado de los datos

Los patrones espaciales son con frecuencia el resultado de efectos de primer y de segundo orden. Los **efectos de primer orden** corresponden a la variación en la media de un proceso sobre un área y pueden ser identificados con la intensidad de tendencias globales o de escala gruesa, mientras que los **efectos de segundo orden** son el resultado de la presencia de AE e implican que las desviaciones en la media de un proceso se dan de modo más frecuente entre sitios cercanos y es equivalente a efectos locales o de escala fina (Sabel y col., 2007; Dale y Fortin, 2014). La variabilidad de un conjunto de datos espaciales puede modelarse de este modo a partir de la suma de tres componentes de varianza (Plant, 2012):

1. Un componente "estructural" o determinístico, que puede reflejar una tendencia, una variación a gran escala, o ambas
2. Un componente aleatorio espacialmente autocorrelacionado
3. Un componente aleatorio, representando variación no correlacionada y el error estadístico

Explícitamente, el modelo puede escribirse de la siguiente manera:

$$Y(x, y) = T(x, y) + A(x, y) + e(x, y) \quad (\text{Expresión 5.5})$$

Donde x e y representan las coordenadas de cada punto, $T(x, y)$ la variación debida a la tendencia determinística, $A(x, y)$ la variación que refleja la autocorrelación espacial, y $e(x, y)$ el error aleatorio, no correlacionado. Por ejemplo, T podría estar midiendo una variable como precipitación, pendiente o aspecto, A podría estar modelando variables a pequeña escala, autocorrelacionadas, y el error e mediría el resto de la variación no correlacionada.

Cuando los datos espaciales muestran una tendencia en su distribución, análisis geoestadísticos tales como los correlogramas para el estadístico I de Moran podrían no ser adecuados, ya que es posible que reflejen esos patrones a escala gruesa, en vez de los presentes a escala fina, conduciendo a identificaciones erróneas de procesos ecológicos (Freeman y Moisen,

2006). Para separar los componentes de variación en un modelo y estudiar la variación a escala fina, primero se tiene que estimar la tendencia $T(x, y)$ y sustraerla de los datos (Plant, 2012). La cantidad $Y(x, y) - T(x, y) = A(x, y) + e(x, y)$ se denomina **datos libres de tendencia** (*detrended data*). Esta cantidad se espera que tenga un casi nulo componente global conocido.

Con el propósito de remover las estructuras globales detectadas en los datos, se realizó un análisis de superficies de tendencias (*trend surface analysis*) (Lichstein, 2002; Legendre y Legendre, 2012). Este método consiste en efectuar una estimación de la tendencia de los datos, por medio de una regresión sobre una expansión polinómica de las coordenadas geográficas. Para mostrar el funcionamiento de la técnica, el paquete EcoGenetics contiene un ejemplo con variables simuladas. La función `eco.detrend`⁴⁶ incluida en el paquete puede realizar un análisis de superficies de tendencias.

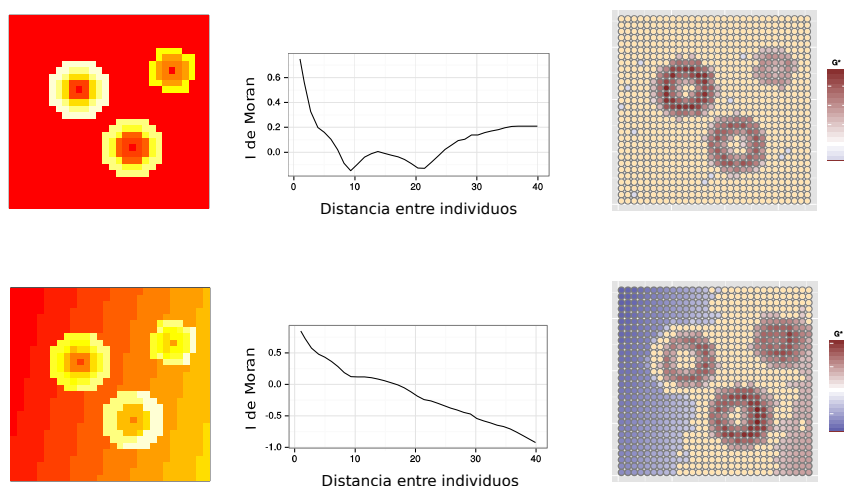


FIGURA 5.14 En las imágenes de la izquierda (inferior y superior) se muestra un patrón simulado en una grilla rectangular de 15 x 15 (225 puntos). La matriz (imagen) superior contiene los valores de una variable con valores discretos (rojo: 1, amarillo: 2, etc). Con estos patrones, se representan tres parches (círculos) y un entorno. Esto corresponde a un patrón con tendencias locales. La matriz (imagen) inferior representa el mismo conjunto de datos, a los que se les sumó una segunda variable en forma de gradiente lineal en sentido sudeste-noroeste. El análisis con `eco.detrend` transforma la variable inferior en la superior al realizar una interpolación con polinomios de primer grado, estimando $Y(x, y) - T(x, y)$. Las figuras del medio representan los correlogramas correspondientes a cada uno de los dos patrones. En el primer caso (arriba) se ve un correlograma que se estabiliza en el valor 0, a una distancia de 10 unidades aproximadamente, y luego vuelve a subir levemente con el aumento de la distancia. El segundo caso (abajo) es característico de variación clinal (correlograma monótonico). El valor en el que los gráficos cruzan al eje x corresponde al tamaño promedio del parche (ver otros ejemplos similares en Legendre y Fortin, 1989; Dale y Fortin, 2014, página 128 y ss.). Las imágenes de la derecha muestran los patrones de autocorrelación local calculados con el estadístico G_i^* (Ord y Getis, 1995).

⁴⁶ Para mas detalles, ver la ayuda para la función en EcoGenetics con el comando `help("eco.detrend")`.

V. Estudios con una población natural - estructura espacial de caracteres cuantitativos a escala fina

Con los datos de las variables cuantitativas se utilizó el mismo principio explicado en la Figura 5.14. Dada la estructura del terreno, con una dirección de variación espacial importante en sentido sudeste-noroeste, implicando variación en altitud (Figura 1.5, página 18), precipitación (Figura 1.4-B, página 17) y geomorfológica (Figura 1.7, página 20), con el río Dulce corriendo de modo aproximadamente perpendicular a esta dirección (Figura 1.6, página 19), se modela aquí la variación ambiental global como un gradiente espacial. Para ello se efectuó una regresión para cada variable morfométrica (z) sobre el polinomio de primer grado de las coordenadas geográficas proyectadas (x, y):

$$z = a_1 + a_2x + a_3y \text{ (Expresión 5.6)}$$

Previamente al análisis, las coordenadas x e y se centraron (media = 0) pero no se estandarizaron a varianza 1 (Legendre y Legendre, 2012). El modelo más parsimonioso para cada rasgo se seleccionó mediante el criterio de Akaike (Akaike, 1974), reteniendo los residuos como las nuevas variables de estudio.

4.4.2. Evaluación de la capacidad del análisis de superficies de tendencias para remover las tendencias globales conocidas de las variables cuantitativas

Para evaluar si la técnica fue capaz de remover las tendencias globales conocidas de las variables cuantitativas, se efectuó un proceso de verificación consistente en analizar con árboles de inferencia condicional los residuos del análisis de superficies de tendencias. Para ello, se utilizó el modelo de la Expresión 5.4, con las variables SUE , $NDVI$, ETP y PP como predictoras. Se espera aquí que el análisis de superficies de tendencia haya sido capaz de remover de las variables cuantitativas la relación observada con variables ambientales globales conocidas.

4.5. Análisis espacial a escala fina

4.5.1. Análisis espacial global

En la disciplina del análisis espacial, se utiliza el término *global* para hacer referencia al estudio de la estructura de los datos a través del sitio de análisis entero, mientras que el término

V. Estudios con una población natural - estructura espacial de caracteres cuantitativos a escala fina

local hace referencia al estudio de un subconjunto de los datos disponibles, con el fin de investigar la presencia de focos calientes (“hot spots”) de autocorrelación espacial, que pueden existir a pesar de no presentarse una estructura global significativa (Anselin, 1995; Sokal y col., 1998).

Para el análisis global, los patrones de AE de las variables cuantitativas sin tendencia (residuos del análisis de superficies de tendencia) se analizaron mediante correlogramas. En esta parte se asume que las variables pueden modelarse como procesos estacionarios de segundo orden⁴⁷. Se utilizaron dos coeficientes diferentes (fórmulas disponibles en Apéndice 2, página 163), capaces de proveer respectivamente información uni, o bivariada:

1. El coeficiente I de Moran, que permite evaluar la AE en variables individuales. El mismo se utilizó para estudiar la AE en cada variable cuantitativa residual.
2. El coeficiente I_{xy} bivariado de Moran (Wartenberg, 1985; Reich y col., 1994), que permite evaluar la correlación espacial entre pares de variables. El mismo se utilizó para evaluar las relaciones entre cada variable cuantitativa residual con AE significativa y el NDVI.

Para los análisis se utilizaron 9 clases de distancias ínter-individuales, con 250 pares de individuos en cada una, en el intervalo 0-900 m. La significación para cada intervalo se realizó con un test de permutación de individuos a través de sitios, utilizando 999 repeticiones. Para corregir los valores P por comparaciones múltiples, se utilizó el método Holm- Bonferroni secuencial⁴⁸ (HSB) (Legendre y Legendre, 2012). Los pesos espaciales espaciales empleados fueron binarios. En el caso de los correlogramas bivariados (I_{xy}), además de las 9 clases mencionadas, se incluyó la distancia $d = 0$, en la que I_{xy} es equivalente al estadístico de correlación de Pearson (r) (Saracco y col., 2004).

47 Un proceso se dice estacionario de segundo orden si *a*) la esperanza y la varianza son uniformes en el área de estudio, y *b*) la covarianza depende sólo de la longitud y la orientación del vector entre dos puntos, no de su posición en el área (Legendre y Legendre, 2012; Dale y Fortin, 2014). Se supone que las mediciones de morfología foliar, en una misma especie y en el área estrecha estudiada, pertenecen a una única distribución. Vale aclarar que el hecho de que un proceso se requiera estacionario, es una propiedad del modelo, y no de los datos en si mismos (Dale y Fortin, 2014); por ello es incorrecto homologar homogeneidad a estacionariedad.

48 Se hace una corrección “Holm-Bonferroni” aplicando el siguiente algoritmo secuencial (descrito en Legendre y Legendre, 2012): test sobre la primera clase: sin corrección, test sobre las primeras dos clases: corregir para 2 pruebas simultáneas, test para las n primeras clases: corregir para n pruebas simultáneas.

4.5.2. Análisis espacial local

Con el propósito de detectar paquetes locales de dependencia espacial (Getis y Ord, 1992), los patrones de AE en las variables cuantitativas sin tendencia, se analizaron mediante el estadístico G_i^{*49} (Ord y Getis, 1995) (Apéndice 2, pag 166).

Los análisis a nivel local requieren definir una matriz de pesos espaciales; cuando la misma está basada en vecinos más cercanos, esto implica fijar una cantidad k de vecinos para cada individuo. Para una observación individual, el valor absoluto de G_i^* crece en presencia de autocorrelación a medida que la distancia entre observaciones se incrementa (es decir, al incrementar k), llegando a un punto en el que el estadístico se maximiza, asumiendo entonces que se alcanzó el límite de la autocorrelación espacial una vez que $|G_i^*|$ comienza a disminuir. De este modo se toma a este tamaño para el que $|G_i^*|$ es máximo como umbral de corte (Getis y Griffith 2002). Basándose en este principio, se escribió un script en R que se describe en la sección 1 del Apéndice 3 (página 165); se eligió a partir del mismo el valor $k = 18$, tal como se muestra en la sección 2 del mismo apéndice (página 167). Para el análisis con $k = 18$, se calcularon los valores de G_i^* y su significación con 9999 permutaciones y ajuste FDR⁵⁰. Los análisis se realizaron con el paquete EcoGenetics y el script mencionado (disponible para uso público mediante solicitud al autor a: learoser@gmail.com).

49 Este coeficiente permite evaluar la presencia de patrones de autocorrelación positiva, distinguiendo zonas donde tienden a agruparse valores similares (grupos con autocorrelación espacial local). Estas zonas pueden ser de dos tipos: focos fríos o calientes, que representan, respectivamente, áreas de contagio donde hay bajos o altos valores de una variable de estudio. Para ello mide la asociación entre los valores de un atributo en una unidad espacial y sus unidades vecinas (esto es, el nivel de agrupamiento de un atributo determinado en torno a una de dichas unidades). Las unidades de estudio pueden ser píxeles en un mapa; para un análisis poblacional, pueden representar individuos o localidades.

50 La aplicación de este tipo de corrección permite obtener una situación intermedia entre la esperada tras una corrección conservativa y aquella que resultaría sin corrección (Caldas de Castro y Singer, 2006).

5. RESULTADOS

5.1. Estadísticas básicas

Las estadísticas básicas de las hojas y DAP de los árboles de los sitios A, P y R (Figura 5.2, página 102) se describen en la Tabla 5.7.

TABLA 5.7 Media y desvío estándar para los caracteres cuantitativos y el NDVI, en la población completa, y en los sitios A, P y R.

Caracter	Población	Sitio		
		A	P	R
PEL	3.355 ± 0.741	3.275 ± 0.746	3.209 ± 0.700	3.471 ± 0.749
PIL	9.551 ± 1.434	9.042 ± 1.363 ^a	9.201 ± 1.432 ^a	9.976 ± 1.352 ^b
NLP	30.243 ± 3.709	30.048 ± 3.830	29.505 ± 3.621	30.728 ± 3.666
NPI	2.031 ± 0.248	2.059 ± 0.217	1.988 ± 0.220	2.042 ± 0.275
LEW	0.180 ± 0.027	0.174 ± 0.026 ^a	0.165 ± 0.028 ^a	0.191 ± 0.022 ^b
LEX	0.862 ± 0.021	0.862 ± 0.020	0.866 ± 0.021	0.860 ± 0.021
LEX/LEA	0.197 ± 0.006	0.198 ± 0.005	0.197 ± 0.005	0.196 ± 0.006
LEA	0.164 ± 0.048	0.149 ± 0.039 ^a	0.139 ± 0.044 ^a	0.184 ± 0.047 ^b
LEF	0.930 ± 0.009	0.924 ± 0.010 ^a	0.933 ± 0.008 ^b	0.931 ± 0.009 ^b
LEL	1.072 ± 0.193	1.021 ± 0.144 ^a	0.978 ± 0.173 ^a	1.147 ± 0.194 ^b
LEL/LEW	6.003 ± 0.769	5.942 ± 0.658	5.963 ± 0.655	6.054 ± 0.871
DAP	45.917 ± 18.519	35.173 ± 17.833 ^a	38.257 ± 17.871 ^a	55.002 ± 14.322 ^b
NDVI	0.446 ± 0.129	0.441 ± 0.104 ^a	0.601 ± 0.079 ^b	0.365 ± 0.075 ^c

NOTA: Diferentes símbolos (a, b, c) indican diferencias significativas de a pares ($P < 0.001$) entre medias, de acuerdo al test de Mann-Whitney.

Se encontraron diferencias altamente significativas para PIL, LEW, LEA, LEF, LEL y DAP entre sitios ($\chi^2 = 15.80, 29.68, 32.49, 19.88, 26.53$ y 49.56 , respectivamente, $g.l. = 2$, $P < 0.001$ en todos los casos, test de Kruskal-Wallis). Para PIL, LEW, LEA, LEL, y DAP, se detectaron diferencias significativas entre los pares de sitios R-P y R-A, y para LEF entre A-R y A-P ($P < 0.02$ en todos los casos, test de Mann-Whitney). NLP y NPI mostraron valores constantes a través de sitios, y PIL, LEW, LEA y LEL, valores menores en los sitios A y P, lo que sugiere la presencia de hojas de menor tamaño promedio en estos dos sitios, respecto al tamaño promedio observado en el sitio R. El test de Fligner-Killeen mostró que la varianza es homogénea entre los tres grupos considerados para los caracteres analizados, con la excepción de DAP ($\chi^2 = 6.822$, $P = 0.033$).

V. Estudios con una población natural - estructura espacial de caracteres cuantitativos a escala fina

Los valores de DAP de los árboles muestreados, representados en el mapa de predicción de la Figura 5.15A, variaron entre 11.78 y 102.81 cm. En general, árboles más jóvenes se encontraron en los sitios con cobertura vegetal secundaria (A y P). Esta variable presentó diferencias altamente significativas entre sitios ($\chi^2 = 49.56$, *g.l.* = 2, $P < 0.001$, test de Kruskal-Wallis) y en comparaciones de a pares entre R-A y R-P ($P < 0.001$, test de Mann-Whitney) (Tabla 5.7).

5.2. Análisis de NDVI

El promedio temporal de NDVI difirió significativamente entre los sitios muestreados identificados en la Figura 5.15B ($\chi^2 = 96.67$, *g.l.* = 2, $P < 0.001$, test de Kruskal-Wallis) (Tabla 5.7). Las comparaciones múltiples entre todos los pares de sitios mostraron diferencias significativas ($P < 0.001$, test de Mann-Whitney) (Tabla 5.7). Los márgenes de las rutas (R) presentaron los valores más bajos de la variable, el sitio A valores intermedios y P los más altos. Esta tendencia puede observarse en el mapa de predicción de la Figura 5.15B.

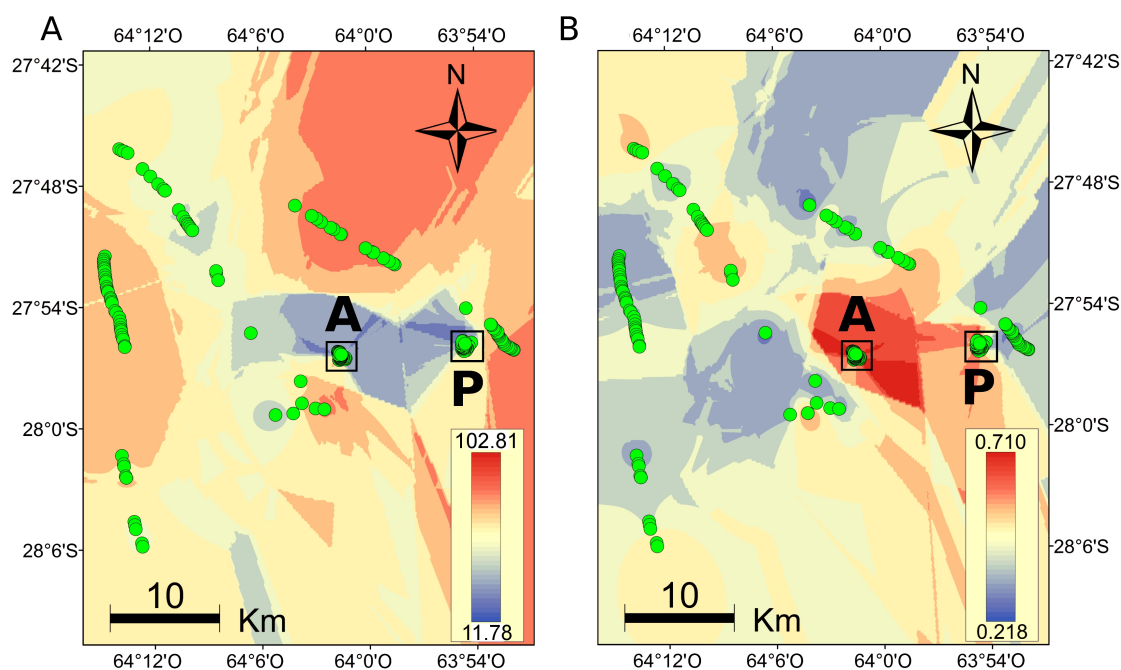


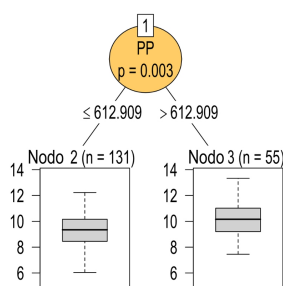
FIGURA 5.15 Superficies de predicción construidas mediante kriging bayesiano empírico. Los círculos verdes indican árboles individuales. Cada figura contiene, en la esquina inferior derecha, el gradiente de color asociado con el valor mínimo (número inferior) y máximo (número superior) de la correspondiente variable. (A) Superficie de predicción para los valores de DAP; (B) Superficie de predicción para NDVI.

5.3. Asociación entre los caracteres cuantitativos y las variables ambientales

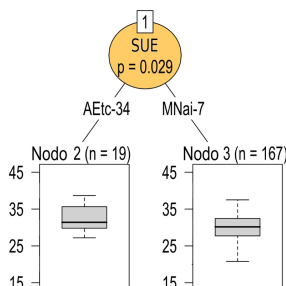
5.3.1. Resultados del modelo

Los árboles de inferencia condicional (Figura 5.16) para las variables originales mostraron que, con la excepción de PEL, NPI, y LEX, todos los rasgos estaban asociados de modo significativo o altamente significativo con al menos una variable ambiental. Un umbral de NDVI de aproximadamente 0.47- 0.50 se mostró vinculado con la primera partición para LEW, LEA y DAP (Figura 5.16), y la proporción de individuos en cada nodo definido por el umbral de NDVI difirió entre los sitios A, P y R (Tabla 5.8). Para LEW y LEA, más del 95% de los individuos muestreados en los márgenes de las rutas (R) fueron asignados a la partición con valores de NDVI por debajo del umbral mencionado, y para DAP esta asignación fue de un 89% (Tabla 5.8). Para los tres caracteres, el 90 % de los individuos del sitio P fueron asignados a la partición con NDVI más elevado (Tabla 5.8). Para LEW y LEA, el 65 % de los individuos del sitio A fueron asignados a la partición con NDVI más bajo, mientras que para DAP la asignación fue de un 58 % (Tabla 5.8).

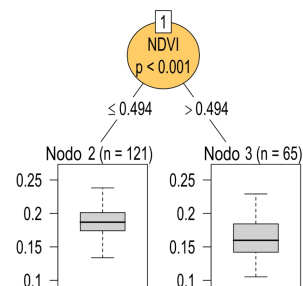
V. Estudios con una población natural - estructura espacial de caracteres cuantitativos a escala fina



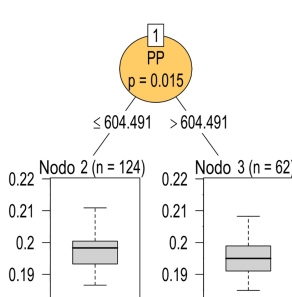
PIL



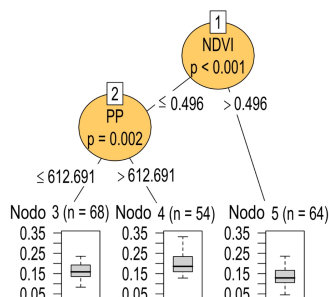
NLP



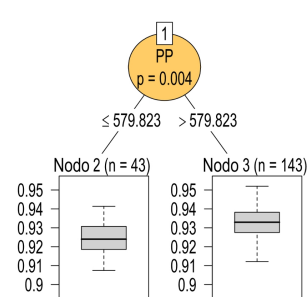
LEW



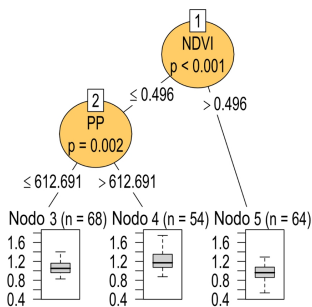
LEX/LEA



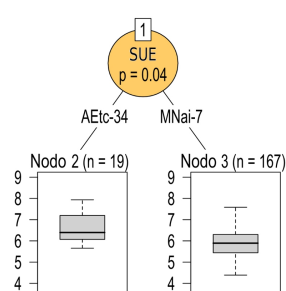
LEA



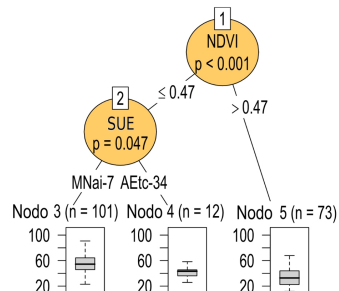
LEF



LEL



LEL/LEW



DAP

V. Estudios con una población natural - estructura espacial de caracteres cuantitativos a escala fina

TABLA 5.8 Número de individuos asignados a un nodo particular en el análisis de árboles de clasificación (ver Figura 5.16). Los rasgos incluidos en la tabla son aquellos con particiones significativas en el análisis.

Sitio	PIL		NLP		LEW		LEX/ LEA		LEA			LEF		LEL			LEL/ LEW		DAP		
	n2	n3	n2	n3	n2	n3	n2	n3	n3	n4	n5	n2	n3	n3	n4	n5	n2	n3	n2	n3	n4
A	43	0	0	43	28	15	43	0	28	0	15	26	15	28	0	15	0	43	25	0	18
P	50	0	0	50	5	45	50	0	5	0	45	0	50	5	0	45	0	50	5	0	45
R	38	55	19	74	88	5	31	62	35	54	4	+17	76	35	54	4	19	74	71	12	10

NOTA: A, P, R sitios de muestreo; n1, n2, n3, n4, n5 nodos (1 a 5).

Ninguno de los rasgos mostró relación significativa con ETP. Para los caracteres relacionados con el tamaño del foliólulo (LEW, LEL, LEA), el NDVI fue un buen descriptor.

5.3.2. Evaluación de la robustez del análisis con árboles de inferencia condicional ante la presencia de autocorrelación espacial en las variables cuantitativas

Los residuos de los modelos de regresión condicional para PEL y DAP mostraron valores de autocorrelación significativos ($P < 0.05$) para las primeras clases de los correlogramas de Moran (no mostrados), lo que estaría indicando una falta de poder en los modelos para describir estas dos variables. Para LEW, LEA, LEF y LEL, los residuos de los modelos de regresión condicional no mostraron AE, indicando que en estos casos el análisis fue robusto.

5.4. Evaluación de la capacidad el análisis de superficies de tendencias para remover las tendencias globales conocidas de las variables cuantitativas

Los residuos del análisis de superficie de tendencia (“variables sin tendencia”) mostraron ausencia de relación significativa con la precipitación o el tipo de suelo. La única excepción a esto fue DAP, que mostró en sus residuos efectos significativos con el tipo de suelo ($P = 0.001$). El efecto significativo del NDVI (descriptor local) no fue removido en los residuos de ninguna de las variables que lo incluían como predictor en el punto 5.3.1 (LEW, LEA, LEL, y DAP; $P < 0.001$ para la asociación en las variables residuales).

5.5. Análisis espacial a escala fina

5.5.1. Análisis espacial global

Los correlogramas I de Moran, para los rasgos sin tendencias globales (residuos) con AE significativa (PEL, LEW, LEA, LEF, LEL y DAP) se muestran en la Figura 5.17. Los valores más altos de AE se encontraron debajo de los 400 m en todos los análisis, con la excepción de LEA.

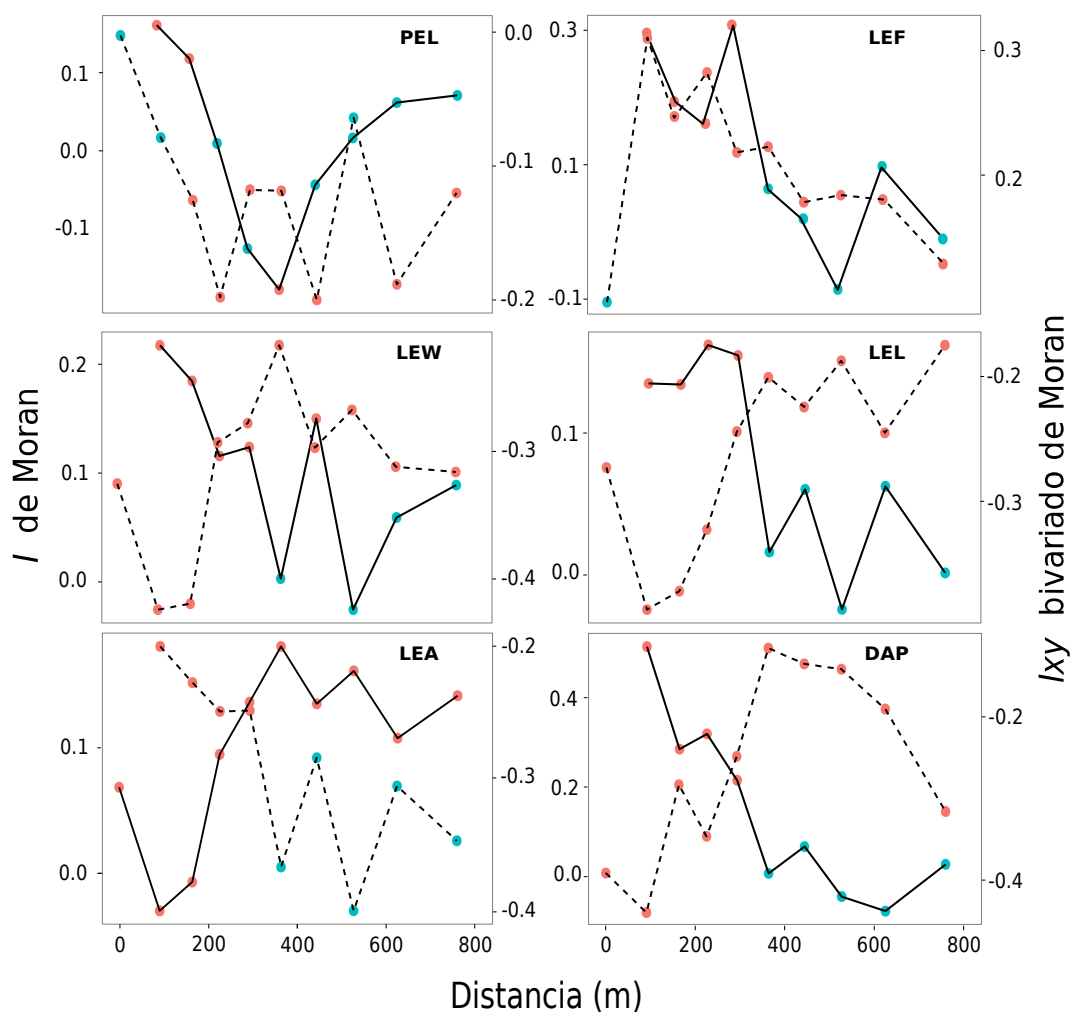


FIGURA 5.17 Correlogramas I de Moran (eje y a la izquierda de cada gráfico, líneas continuas) para los rasgos sin tendencia, con autocorrelación espacial significativa, y correlogramas cruzados I_{xy} de Moran (eje de las ordenadas a la derecha de cada gráfico, líneas punteadas) entre los rasgos sin tendencia y el NDVI. Los colores de los círculos indican la significación de los coeficientes respectivos luego de la corrección por comparaciones múltiples Holm-Bonferroni secuencial (rojo: $P < 0.05$, azul: no significativo).

V. Estudios con una población natural - estructura espacial de caracteres cuantitativos a escala fina

Los correlogramas bivariados I_{xy} (Figura 5.17) mostraron una tendencia decreciente para PEL y LEF, y creciente para LEW, LEA, LEL y DAP. La tendencia de la curva en los gráficos de I_{xy} para LEW, LEA, LEL, y DAP fue inversa a la evidenciada por los gráficos de I para esos mismos rasgos (Figura 5.17). Esto indica que en áreas más densas, con altos valores de NDVI, el tamaño de la hoja tendió a ser menor.

5.5.2. Análisis espacial local

LEW, LEA, LEF, LEL y DAP mostraron valores significativos de autocorrelación a nivel local bajo corrección FDR. El coeficiente G_i^* detectó la presencia de patrones no aleatorios en la distribución de los rasgos foliares y para DAP, con individuos rodeados por otros con morfología foliar y diámetro similar, respectivamente, formado focos fríos y calientes (es decir, agregados de individuos con valores altos y bajos de las variables) (Figura 5.18). Los patrones locales de focos fríos y calientes corresponden con las diferencias generales encontradas entre los sitios, mostrada en la Tabla 5.7. Para LEW, LEA, LEL y DAP, las rutas presentaron focos calientes, mientras que los sitios A y P presentaron focos fríos. LEF mostró focos con valores bajos en A y altos en P.

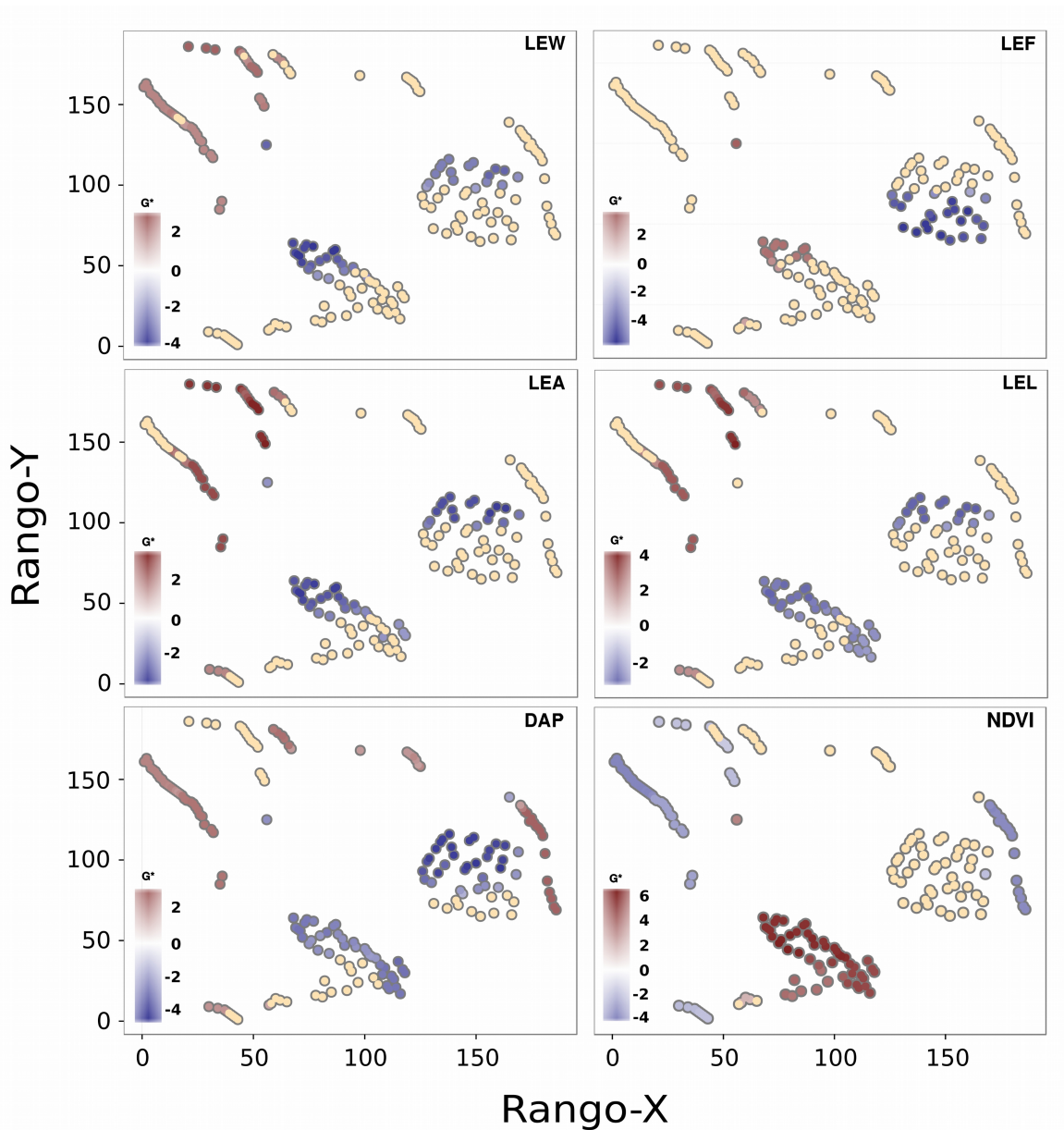


FIGURA 5.18 Resultado del análisis con el estadístico G_i^* para los rasgos sin tendencia global y para NDVI. Los valores de longitud y latitud se ordenaron de menor a mayor y se tomó el número de orden de cada una de estas dos variables como coordenadas en una grilla de $N \times N$, con N siendo el número de individuos. El objetivo de este procedimiento fue graficar a los individuos como círculos distribuidos regularmente, de modo tal que pueda observarse a todos en un mismo gráfico. Los círculos amarillos indican individuos con valores no significativos de G_i^* luego de corrección FDR por comparaciones múltiples. Los gradientes de color en las leyendas indican los valores asignados a los individuos (círculos) con valores significativos de G_i^* luego de la corrección FDR (azul: bajos valores de G_i^* , rojo: altos valores de G_i^*).

6. DISCUSIÓN

La autocorrelación espacial es un fenómeno que ocurre cuando los valores que toma una variable no son independientes entre sí (Dormann y col., 2007), presentándose la estructura de correlación como una función de la proximidad espacial de los sitios/objetos donde la variable es medida. La heterogeneidad espacial, por otra parte, puede ser definida como la diversidad y la configuración de los componentes ecológicos en un área determinada (Fortin y col., 2012). Estos conceptos son dependientes de la escala de observación, tal como la mayor parte de los fenómenos ecológicos. Por esta razón, es importante analizar patrones a diferentes escalas espaciales, identificando los procesos principales subyacentes a cada nivel de análisis para un mejor entendimiento de la interdependencia entre la dinámica poblacional y del paisaje. Este conocimiento puede ser útil para la optimización del manejo racional y de los programas de conservación de especies con valor agronómico, tal como *P. alba*.

Los resultados de esta parte sugieren que la heterogeneidad presente en el paisaje estudiado se encuentra reflejada en los caracteres cuantitativos analizados. Antes de la remoción de los efectos causados por la ubicación geográfica (relacionados con tendencias o efectos de primer orden), los datos de morfología foliar mostraron una asociación significativa con variables globales. Este hecho podría indicar una variación conjunta ambiente-fenotipo foliar a escala gruesa. Luego de remover el efecto de la ubicación geográfica sobre las variables, la estructura espacial persistió en 5 de estos caracteres (PEL, LEW, LEA, LEF y LEL) y en DAP, tal como fue revelado por el análisis global con correlogramas univariados de Moran, indicando la presencia de autocorrelación espacial o efectos de segundo orden. El análisis con el estadístico G_i^* mostró la existencia de autocorrelación espacial positiva a nivel local (grupos de individuos con valores similares) para LEW, LEA, LEF, LEL y DAP.

La dependencia espacial presentada por los caracteres con AE significativa, refleja el hecho de que los individuos cercanos resultaron más similares entre si en los valores de estos rasgos que los distantes. Al menos dos factores importantes pueden dar cuenta de estos patrones de AE, que no son mutuamente excluyentes: la presencia de aislamiento por distancia y la heterogeneidad ambiental.

V. Estudios con una población natural - estructura espacial de caracteres cuantitativos a escala fina

En especies de *Prosopis*, tanto la dispersión por polen como por semillas está limitada (Bessegga y col., 2012), y esto puede promover una disminución en el parentesco de árboles con el incremento de la distancia espacial que los separa. El polen no es hábil usualmente para migrar largas distancias debido a la condición entomófila del género (Genisse y col., 1990), lo que favorece los cruzamientos entre individuos cercanos. El transporte mediado por mamíferos, que ingieren los frutos, pero transportan hermanos o medios hermanos que luego son depositados con las heces, (Keys, 1993; Bessegga y col., 2000, 2012) determina que individuos altamente emparentados germinen en áreas cercanas. Un trabajo previo con *P. alba* en la región del presente análisis (Bessegga y col., 2012), encontró que la mayor parte de la variación genética se encuentra entre familias dentro de zonas, con una distancia de polinización promedio en el rango de 5 a 30 m, y con una exogamia casi completa, con un grado bajo (aproximadamente un 5%), pero significativo de endogamia biparental. El análisis con marcadores microsatélites del capítulo V de esta tesis mostró que: 1) hay una relación inversa entre la distancia de los individuos y la coancestría en el rango 0-900 m, que no puede ser explicada por azar, 2) no hay evidencia para rechazar la hipótesis nula de continuidad en la población y 3) la mayor parte de la variación se encuentra en el menor nivel jerárquico analizado mediante AMOVA (individuos). La similitud observada entre los patrones de AE de los caracteres morfológicos y moleculares en la población de estudio, sugiere que la AE presente en los datos genéticos sería un factor causal explicando parte de los patrones de AE significativos de los caracteres morfológicos analizados.

El capítulo III de esta tesis en el que se analizó un huerto con clones de *P. alba* originados a partir de árboles madre del sur de Santiago del Estero, mostró que bajo las condiciones experimentales utilizadas, los caracteres con más alta heredabilidad en sentido amplio significativa fueron LEL ($H^2 = 0.778$), LEW ($H^2 = 0.758$) y LEA ($H^2 = 0.759$) (tres caracteres altamente correlacionados y ligados al tamaño y forma de los foliólulos), mientras que LEF mostró el valor más bajo ($H^2 = 0.032$). También mostró que estos valores no pueden explicarse sólo en base al parentesco. Estos valores de heredabilidad indican que para LEL, LEW y LEA, el componente genético explica una parte de la varianza mostrada en el huerto, y que para LEF, el ambiente es una causa explicativa más importante de esta variación. Estos árboles mostraron además en un trabajo previo (Ewens, 2012) valores altos de heredabilidad para biomasa ($H^2 = 0.45$), diámetro ($H^2 = 0.53$) y altura ($H^2 = 0.59$). En conjunto, estos resultados de heredabilidad dan soporte a la posibilidad de que, en la población natural, un componente de varianza genética explique la autocorrelación en

V. Estudios con una población natural - estructura espacial de caracteres cuantitativos a escala fina

LEL, LEW, LEA y DAP, y un componente ambiental la autocorrelación en LEF. Las conclusiones sobre heredabilidad deben tomarse con recaudo sin embargo, ya que la muestra del huerto es pequeña y representa individuos que sobrevivieron a un tratamiento con agua salina.

El presente capítulo mostró a través de correlogramas bivariados que la densidad de la cobertura verde (medida como NDVI) se encuentra inversamente relacionada con el tamaño foliar. Los árboles de inferencia condicional confirmaron esto, mostrando que los individuos del sitio P (con hojas más pequeñas) tienden a ser clasificados en una categoría con NDVI más alto. Esto podría deberse tanto a la similitud genética entre los individuos debido al proceso de dispersión, como a respuestas ante el ambiente (plasticidad). La heterogeneidad en la disponibilidad de recursos (tipo de suelo, niveles de humedad, contenido de nutrientes) a través de distancias relativamente cortas, dividen a un hábitat en un mosaico de micrositios (Tripathi y Kahn, 1990), y esto puede ser una razón que genere patrones de AE (Fajardo y McIntire, 2007), afectando la fisiología y el desarrollo de los árboles y promoviendo la diferenciación fenotípica entre sitios. Ansley y col. (1998) observaron por ejemplo que en *P. glandulosa* el crecimiento y el tamaño de las hojas eran significativamente mayores en bajas que en altas densidades forestales luego de la poda. Para poder evaluar posibles efectos ambientales sobre los rasgos, se requieren futuros estudios ecofisiológicos analizando la relación del ambiente con el fenotipo a un nivel micro-espacial.

En conclusión, los resultados de este capítulo sugieren que la forma de la hoja y el tamaño de *P. alba* se encuentran espacialmente estructurados a escala fina en el paisaje fragmentado estudiado. El análisis permitió revelar la existencia de patrones fenotípicos locales (AE positiva), que podrían encontrarse asociados con flujo génico limitado. La gran variabilidad genética descrita en esta tesis y en investigaciones previas de estructura poblacional de *P. alba* y otras especies de *Prosopis* (Ferreira y col., 2007, 2010; Bessega y col., 2012), y los patrones de coancestría dentro de los sitios aquí muestreados, podrían explicar parte de los patrones de AE en base a un modelo de aislamiento por distancia mediado por flujo génico restringido. Este modelo operaría a través de la limitación en la distancia del movimiento del polen y las semillas. La heterogeneidad espacial y el parentesco entre individuos jugarían roles importantes en la expresión del fenotipo promedio en *P. alba* a escala espacial fina.

Capítulo VI

Discusión general y conclusiones

VI. DISCUSIÓN GENERAL Y CONCLUSIONES

El género *Prosopis* comprende un conjunto de especies leñosas, en su mayoría formas arbóreas y arbustivas, distribuidas en ecosistemas áridos y semiáridos del planeta. En Argentina, el género se distribuye a lo largo y ancho del territorio nacional, exceptuando al sur de la Patagonia y la provincia de Misiones (Burkart, 1976; Hunziker y col., 1986).

La sección Algarobia del género incluye especies con amplia distribución geográfica, dispersión endozoica de las semillas, y mayormente de fecundación cruzada (Solbrig y Cantino, 1975; Simpson, 1977; Bessega y col., 2000, 2012). La variabilidad genética está contenida en gran parte dentro de las poblaciones, habiendo una divergencia baja entre poblaciones de una misma especie y entre especies. El concepto de especie es confuso en la sección Algarobia, ya que los individuos clasificados dentro de diferentes grupos pueden cruzarse en zonas de simpatria, dejando descendencia fértil (Hunziker y col., 1986).

Prosopis alba (algarrobo blanco) es una especie representante de la sección Algarobia, con árboles que poseen gran adaptación a las regiones semiáridas del NE argentino. Actualmente su distribución abarca bosques relativamente prístinos, reducidos a algunas zonas de Chaco y Formosa; parches remanentes de la extensa distribución original, especialmente en Santiago del Estero, usualmente rodeados de áreas con uso productivo (agroganadero o forestal); y zonas de colonización relativamente reciente, como consecuencia de modificaciones antrópicas asociadas con cambios en los cursos de agua por la construcción de diques y aparición de hábitats abiertos luego del desmonte. En Santiago del Estero las poblaciones naturales de algarrobos blancos están prácticamente reducidas a pequeñas áreas protegidas entre ríos, a los costados de los caminos o mantenidas en pequeños parches en fincas donde proveen alimento y sombra al ganado.

Estudios con poblaciones naturales en esta especie indicaron que la mayor parte de la variabilidad genética se encuentra dentro de las mismas, con una limitada diferenciación regional (Ferreira y col., 2007, 2010; Bessega y col., 2012, 2014). La selección natural podría ser un factor importante en la generación de patrones morfológicos a escalas gruesas (Bessega y col., 2014). En ensayos se pudo estimar asimismo una alta heredabilidad para rasgos relacionados con la forma y el tamaño de hojas y espinas, y con rasgo de crecimiento (altura y diámetro del tronco) (Bessega y col. 2009, 2011; Ewens y col., 2012). Estudios sobre el sistema de fecundación (Bessega y col., 2000) y

VI. Discusión general y conclusiones

el análisis de la estructura fina a nivel familiar (Bessega y col., 2012) mostraron que esta especie es principalmente de fecundación cruzada y que la dispersión del polen sería limitada (entre 5 y 30 m).

En la presente tesis se estudiaron individuos de *P. alba*, originarios de Santiago del Estero, mediante un doble enfoque. En primer lugar, se analizó un huerto experimental con individuos tolerantes a la salinidad. Para contrastar los datos moleculares del huerto se seleccionaron además individuos sensibles a la salinidad, que se desarrollaron en vivero. En segundo lugar, se estudió una población natural a escala fina. Para ello se tomaron muestras en dos parches, y en los márgenes de tres rutas. Dos elementos de la escala, la extensión, y la distancia entre muestras (en particular en los parches, con muestreo intensivo), fueron determinantes a la hora de elegir las metodologías y los marcos teóricos de análisis.

La diversidad genética puede aparecer estructurada espacialmente a diferentes escalas, tal como a nivel paisaje, población o entre individuos cercanos, debido a diferentes procesos ecológicos y genéticos operando en las poblaciones naturales (Zeng y col., 2012). Gran parte de los estudios genéticos de estructura poblacional con árboles han sido desarrollados hasta tiempos recientes principalmente a escalas gruesas. Un enfoque a escala fina puede enriquecer las teorías y modelos disponibles, aportando puntos de vista novedosos para el estudio poblacional. La incorporación explícita de diferentes escalas de análisis en los modelos biológicos es importante, al considerar que las inferencias hechas son escala-dependientes.

A partir de lo visto a lo largo de los capítulos de esta tesis, se puede concluir lo siguiente:

- Los estudios realizados en un huerto experimental permitieron observar una diferenciación genética con marcadores ISSR y SSR, en individuos seleccionados por tolerancia a la salinidad. Los marcadores mostraron resultados similares y pudieron predecir el fenotipo de los individuos.
- Se obtuvieron altos valores de heredabilidad en sentido amplio para rasgos foliares en el huerto, que no pueden explicarse sólo por causas ambientales.

VI. Discusión general y conclusiones

- Las herramientas informáticas generadas permitieron un adecuado manejo de la información durante el desarrollo del estudio a escala fina. Estas herramientas, que involucraron principalmente programas diseñados para el estudio integral de la autocorrelación, tanto global como local, son valiosas para estudios poblacionales. La generalización de métodos de autocorrelación local es una novedad promisorio para futuros estudios genéticos y ecológicos.
- El diseño de muestreo utilizado con la población natural permitió observar fenómenos a escala fina, tales como la caída de la coancestría con la distancia entre individuos, la ocurrencia de cambios demográficos recientes y la presencia de patrones locales en los rasgos morfológicos.
- La estructura genética a escala fina mostró ajuste a un modelo de aislamiento por distancia. Esta situación (aislamiento por distancia) se presenta cuando los individuos se aparean al azar dentro de un vecindario, pero están limitados de aparearse con otros miembros más distantes (Wright, 1943). En árboles, los mecanismos de dispersión del polen y las semillas pueden generar patrones de este tipo (Sokal y col., 1987).
- Los rasgos de morfología foliar estudiados mostraron patrones de autocorrelación espacial similares al detectado para el conjunto de marcadores SSR. Esto sugiere que el fenómeno de aislamiento por distancia observado en los datos genéticos tendría una contraparte en las variables morfológicas medidas. No pueden descartarse sin embargo otros procesos adicionales, tales como plasticidad fenotípica en respuesta a efectos de micrositio, o adaptación local a la configuración del paisaje.
- Teniendo el recaudo de considerar el tamaño muestral estudiado en el huerto experimental y que los individuos representan sobrevivientes a un tratamiento con agua salina, los altos valores de heredabilidad observados para los rasgos foliares, sugieren la posibilidad de una base genética en la estructura espacial que estos mostraron a escala fina.

La presente investigación generó un aporte para el estudio de *P. alba* a escala fina, desde un enfoque genético y ecológico. Este tipo de análisis podría ser aplicado tanto en otras especies de *Algarobia* como en otros modelos forestales. Es importante asimismo profundizar el desarrollo de

VI. Discusión general y conclusiones

metodologías, algoritmos y programas para el estudio de datos autocorrelacionados en modelos biológico-poblacionales, sacando provecho de las herramientas informáticas disponibles al momento, tales como un enfoque bayesiano del problema y el poder de cálculo de R. En particular, los modelos espaciales de autocorrelación a nivel local, poco explorados al momento en ecología y genética, pueden revelar interesantes fenómenos de estructuración poblacional a escala fina, y son promisorios para la elaboración de investigaciones futuras.

Apéndices

APÉNDICES

APÉNDICE 1: PROTOCOLOS UTILIZADOS PARA LA OBTENCIÓN DE DATOS MORFOMÉTRICOS, MOLECULARES Y AMBIENTALES

Protocolo 1: obtención de datos de morfología foliar

Se midieron en total 11 rasgos de morfología foliar en cada árbol individual: longitud del peciolo (PEL), número de pares de foliólulos por pinna (NLP), longitud de las pinnas (PIL), número de pinnas (NPI), longitud (LEL), ancho (LEW), relación largo/ancho (LEL/LEW), falcado (LEF), área (LEA), ápice (LEX) y relación ápice/área total (LEX/LEA) de los foliólulos. Las mediciones se efectuaron luego de secar el material recolectado a campo.

Algunos de los rasgos se muestran en la Figura 7.1A. LEF es la relación l/f , donde l representa la longitud de un segmento recto que va desde la base hasta la punta del foliólulo, y f representa la longitud del foliólulo siguiendo una línea curva que pasa a través de su parte media (Figura 7.1B). LEX es la relación t/s , donde t es el área del tercio superior del foliólulo y s es el área de un rectángulo con la misma dimensión (ancho y largo) que t (Figura 7.1C).

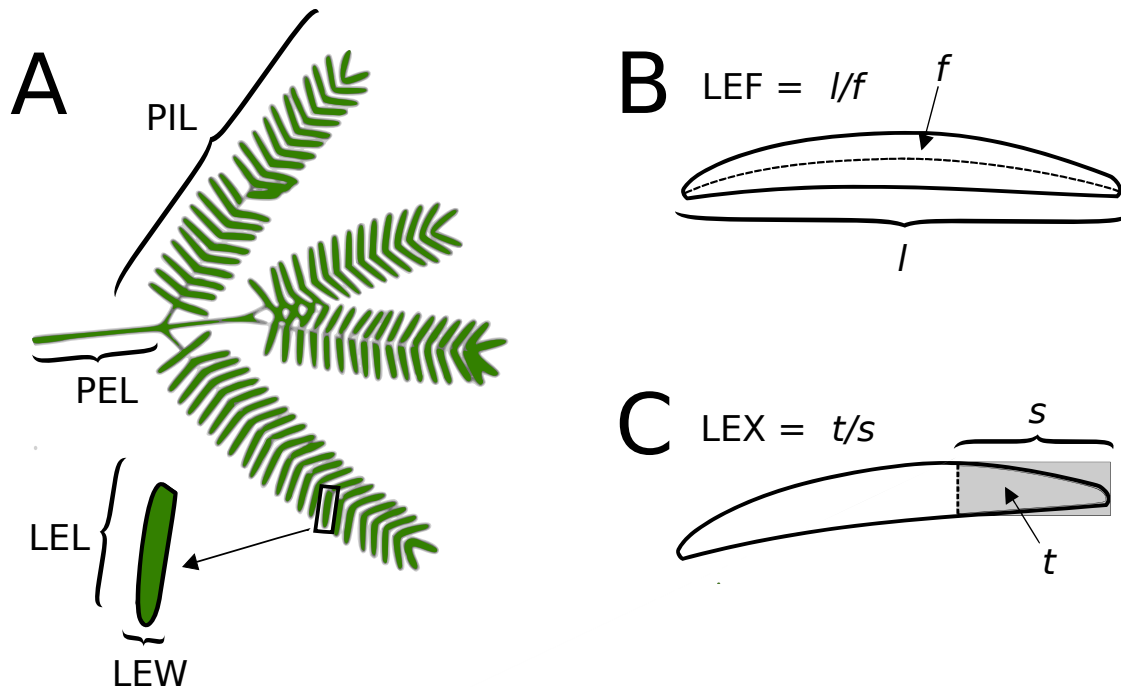


Figura 7.1 Algunos caracteres de morfología foliar medidos y registrados para cada planta individual. **(A)** longitud del peciolo (PEL), longitud de la pinna (PIL), longitud del foliólulo (LEL), ancho del foliólulo (LEW); **(B)** descripción de mediciones para estimar el falcado del foliólulo (LEF); **(C)** descripción de mediciones para estimar el ápice del foliólulo (LEX). l = distancia desde la base a la punta del foliólulo; f = longitud desde la base a la punta del foliólulo siguiendo una línea curva a lo largo del medio del foliólulo; t = área del tercio superior del foliólulo; s = área de un rectángulo con las mismas dimensiones que t .

PEL y PIL se midieron con una regla graduada al milímetro más cercano. NPI y NLP se contaron de modo individual. LEL, LEL/LEW, LEF, LEX y LEX/LEA se midieron a partir de imágenes escaneadas y analizadas con el programa HOJA1.1 (disponible del autor: A. Verga, INTA-IFFIVE, verga@gmail.com). Para las mediciones de LEL, LEL/LEW, LEF, LEX y LEX/LEA, se eligieron 20 foliólulos de modo aleatorio de diferentes hojas y áreas del follaje. Para los otros rasgos, se tomó una medición por hoja.

De cada árbol muestreado se guardaron ejemplares en el Laboratorio de Genética de Especies Leñosas (GEEL), Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires.

Protocolo 2: obtención de datos moleculares

Para los análisis de ADN, se recolectaron hojas jóvenes de los individuos adultos a estudiar. Este material se secó inmediatamente en sílica gel y almacenado a temperatura ambiente hasta el momento de la extracción. El ADN genómico se extrajo usando un kit de extracción *DNeasy Plant Mini Kit* (QIAGEN Inc., Valencia, CA, USA) siguiendo las instrucciones del fabricante. Las muestras se guardaron a -20°C hasta posterior análisis.

A. Análisis de ISSR

El análisis de ISSR se realizó siguiendo el protocolo de Zietkiewicz y col. (1994). Se probaron inicialmente distintos cebadores, para identificar bandas polimórficas y bien amplificadas. Las reacciones de PCR se llevaron a cabo en un volumen final de 50 µl, con la siguiente composición: 10-30 ng de ADN, una concentración 0.8 µM de cebador (Invitrogen, Carlsbad, CA, USA), Taq ADN polimerasa 0.3 U (Invitrogen), Buffer (Tris HCl 10 mM, pH 8.3 y KCl 50 mM), 0.20 mM de cada dNTP (Biodynamics) y MgCl₂ 1.2 mM. La amplificación se hizo en un termociclador *MyCycler Thermal Cycler System* (Bio-Rad), con el siguiente programa: una etapa inicial a 94°C por 90 seg., 30 ciclos de desnaturalización a 94°C por 30 seg., alineamiento a 48°C por 45 seg., extensión a 72°C por 90 seg., y una extensión final de 5 min. a 72°C. Se tomaron las precauciones usuales necesarias para prevenir la contaminación de los experimentos de PCR con fragmentos previamente amplificados. La confiabilidad de los experimentos de PCR se testeó a través de distintos controles usados rutinariamente, un control sin cebador, uno sin Taq ADN polimerasa y uno sin ADN genómico. No se observaron productos de amplificación en ninguno de estos controles negativos. Para los individuos que no presentaron producto de PCR para alguno de los marcadores, se repitió la amplificación contra testigos positivos con el ADN original.

Los productos de ISSR se separaron por medio de electroforesis a potencia constante de 60 W por 3.5 horas en un aparato para electroforesis *Model S2* (Gibco BRL Sequencing System, Life Technologies) en geles de poliacrilamida al 4% (m/v) conteniendo 5 M de urea en un buffer 1X de TBE (Tris 89 mM, ácido bórico 89 mM y EDTA 2 mM, pH 8). En dos calles de cada corrida se incluyó un marcador de peso molecular de 100 pb (Invitrogen). Los geles se tiñeron con nitrato de plata (Bassam y col. 1991). Cada muestra de ADN se corrió en tres geles diferentes, observándose

el mismo patrón. En cada gel, se analizaron individuos de grupos morfológicos previamente predefinidos con el fin de evitar errores en la comparación entre grupos atribuibles a errores experimentales.

La matriz básica de datos de ISSR se construyó codificando, para cada individuo (filas), la presencia/ausencia de cada banda (columnas) de los cebadores como 1 o 0, respectivamente.

B. Análisis de SSR

Las condiciones de amplificación por PCR se llevaron a cabo en un volumen final de 50 μ l, conteniendo 10-30 ng de ADN, una concentración 0.6 μ M de cada cebador, Taq ADN polimerasa 0.3 U (Invitrogen), Buffer (Tris HCl 10 mM, pH 8.3 y KCl 50 mM), 0.20 mM de cada dNTP, y $MgCl_2$ 1.5 mM. Para las amplificaciones se utilizó un termociclador *Progene* (Techne Cambridge Ltd., Duxford, Cambridge, UK.), donde el perfil de ciclado constó de una desnaturalización inicial a 94 °C por 5 min., seguida de 30 ciclos de desnaturalización a 94°C por 45 seg., un ciclo de annealing con temperatura cebador-específica (58°C-64°C) por 45 seg., y un ciclo de elongación a 72°C por 45 seg.; y un ciclo final de elongación a 72°C por 10 min. La confiabilidad de los experimentos de PCR se testeó usando controles negativos de modo rutinario: un control sin cebador, uno sin Taq ADN polimerasa y uno sin ADN genómico. No se observaron productos de amplificación en ninguno de estos tres casos. Para los individuos que no presentaron producto de PCR para alguno de los marcadores, se repitió la amplificación contra testigos positivos con el ADN original.

Se separaron 7 μ l de producto de PCR mediante electroforesis en un aparato *Model S2* (Gibco BRL Sequencing System, Life Technologies) en un gel de poliacrilamida 6% (m/v) conteniendo urea 5 M en un buffer 1X de TBE (Tris 89 mM, ácido bórico 89 mM y EDTA 2 mM, pH 8). Se incluyó un marcador de peso molecular de 10 pb DNA Ladder (Invitrogen) por duplicado en cada corrida electroforética. Los geles se tiñeron con una solución de nitrato de plata (Bassam y col. 1991) y se repitieron por triplicado para cada marcador. Para los individuos que no presentaron producto de PCR para alguno de los marcadores, se repitió la amplificación contra testigos positivos con el ADN original y, en caso de no observar un patrón de bandas, se extrajo el ADN y se repitió el proceso de amplificación para cerciorar la presencia de alelos nulos.

Se construyó una matriz básica de datos de individuos por alelos, con cada celda representando la proporción (frecuencia) de los alelos de cada clase (A_i) para cada locus del individuo I , y los posibles valores asignados fueron: 1 (si el genotipo de I era A_iA_i), 0.5 (si el genotipo era A_iA_j con $j \neq i$) o 0 (si el genotipo era A_jA_k , con $j \neq i$ y $k \neq i$). Siendo frecuencias, la suma de los valores para los alelos de cada locus en el individuo I es 1.

Protocolo 3: análisis de NDVI

Para el análisis de NDVI se seleccionaron las imágenes de Landsat 5 de las siguientes fechas: 11-01-2002, 17-04-2002, 30-01-2003, 20-04-2003, 06-11-2003, 19-07-2004, 05-09-2004-09, 08-11-04, 27-01-2005, 10-10-2005, 15-02-2006, 26-08-2006, 13-10-2006, 01-01-2007, 25-05-2007, 26-06-2007, 12-07-2007, 13-08-2007, 05-02-2008, 16-09-2008, 06-01-2009, 22-01-2009, 11-03-2009, 12-04-2009, 02-08-2009 y 06-11-2009.

Para procesar la información de las bandas 3 (TM3) y 4 (TM4), los valores de número digital (DN) se convirtieron primero a radiancia al sensor usando los metadatos de las imágenes, usando la siguiente expresión (Chavez 1989, Chander et al. 2009):

$$L_j = G_i \cdot DN_j + B_i \quad (\text{Expresión 7.1})$$

Donde:

- L_j = radiancia al sensor ($\text{W m}^{-2} \text{sr}^{-1}$) para el píxel j
- G_i = ganancia ($\text{W m}^{-2} \text{sr}^{-1} \text{mm}^{-1}$) para la banda i . Obtenido de Chander y col. (2009)
- B_i = error ($\text{W m}^{-2} \text{sr}^{-1} \text{mm}^{-1}$) para la banda i . Obtenido de Chander y col. (2009)
- DN_j = número digital para el píxel j

Los valores de radiancia al sensor, se convirtieron en reflectancia a nivel superficie, usando el principio de substracción de objeto oscuro (asumiendo una reflectancia superficial del 1% para objetos oscuros) con el método COST (Chavez, 1996) mediante la siguiente expresión (Chavez, 1989; Goslee 2011):

$$ref = \frac{\pi \cdot (Lsat_i - Lhaze_i) \cdot d^2}{TAUv \cdot (ESUN_i \cdot \cos(\Theta) \cdot TAUz + Edown)} \quad (\text{Expresión 7.2})$$

Donde:

- ref = reflectancia espectral de la superficie
 - $Lsat_i$ = radiancia total medida por el satélite en el tope de la atmósfera ($Wm^{-2} sr^{-1} \mu m^{-1}$) para cada banda i
 - $Lhaze_i$ = radiancia resultante de la interacción de la radiación electromagnética con los componentes de la atmósfera ($Wm^{-2} sr^{-1} \mu m^{-1}$) para cada banda i
 - d = distancia Tierra - Sol normalizada (A.U)
 - $TAUv$ = transmitancia atmosférica para el blanco a través del sensor (adimensional). $\cos(tv)$ (donde tv = ángulo de visión), es usado como una aproximación de $TAUv$
- Para Landsat, $TAUv = 1$, porque el sensor posee un ángulo de visión a nadir (Chavez, 1996)
- $ESUN_i$ = constante solar exoatmosférica ($Wm^{-2} sr^{-1} \mu m^{-1}$) (corregida por la distancia solar) para cada banda i . Valores obtenidos de Chander y col. (2009)
 - Θ = ángulo cenital solar (radianes)
 - $TAUz$ = transmitancia atmosférica en la dirección de la iluminación (adimensional). $\cos(\Theta)$ es usado como una aproximación (Chavez, 1996)
 - $Edown$ = radiancia atmosférica descendente ($m^{-2} \mu m^{-1}$). Chavez (1996) asume $Edown = 0$ (es decir, ignora este factor)

Es poco probable que las imágenes contengan píxeles que sean verdaderamente negros, por lo que se aplicó una corrección asumiendo un 1% de reflectancia real en estas áreas oscuras (Song et al. 2001; Goslee 2011):

$$Lhaze = Lmin - \frac{0.01 \cdot ESUN_i \cdot \cos(\Theta) \cdot TAUz}{\pi \cdot d^2} \quad (\text{Expresión 7.3})$$

Donde L_{min} es el valor umbral obtenido para cada banda (“starting haze value”), considerado el mínimo de radiancia (Chander et al., 2009). Como en la Expresión 7.2, $\cos(\theta)$ es usado como una aproximación de TAU_z .

Finalmente, el NDVI (Tucker 1979) se obtuvo para cada fecha de la siguiente manera:

$$NDVI = \frac{TM\ 4 - TM\ 3}{TM\ 4 + TM\ 3} \quad (\text{Expresión 7.4})$$

Luego de calcular las imágenes de NDVI corregida atmosféricamente para cada fecha, una Imagen de NDVI promedio se obtuvo con la serie de 26 imágenes usadas en el análisis. El valor asignado a cada píxel en la imagen resultante se estimó a partir de la imagen promedio calculada sobre un área de 3 por 3 píxeles.

APÉNDICE 2: FÓRMULAS DE RELEVANCIA PARA EL DESARROLLO DE LA TESIS

1. Cálculo de la heterocigosis observada (H_o) y esperada (H_E)

La heterocigosis observada (H_o) se calculó de la siguiente manera:

$$H_o = 1 - \sum_k \left(\sum_i (P_{kii}) / np \right) \quad \text{(Expresión 7.5)}$$

Donde P_{kii} representa la proporción del homocigota para el alelo i en la muestra k y np el número de muestras. El valor promedio por locus se obtuvo promediando los valores de H_o para cada locus individual.

La heterocigosis esperada (H_E) bajo equilibrio de Hardy-Weinberg (también llamada diversidad genética dentro de las poblaciones) se calculó mediante la siguiente expresión:

$$H_E = \frac{\tilde{n}}{(\tilde{n} - 1)} * \left[1 - \sum_i \tilde{p}_i^2 - H_o / 2 \tilde{n} \right] \quad \text{(Expresión 7.6)}$$

Donde $\tilde{n} = np / \sum_k 1/n_k$ representa la media armónica del número de individuos a través de las muestras y $\tilde{p}_i^2 = \sum_k p_{ki}^2 / n$, con p_{ki} siendo la frecuencia del alelo i en la muestra k .

2. Indicadores de autocorrelación espacial

2.1. Indicadores globales

A. *I* de Moran (Moran, 1950)

$$I = \frac{N \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{X})(x_j - \bar{X})}{W \sum_{i=1}^n (x_i - \bar{X})^2} \quad \text{con } i \neq j \quad (\text{Expresión 7.7})$$

Donde N es el número de individuos. $\bar{X}$ es el valor promedio de la variable, x_i es el valor de un rasgo en un sitio particular. x_j es el valor de un rasgo en otro sitio, w_{ij} es el peso espacial entre los individuos i y j , y $W = (\sum_{i=1}^n \sum_{j=1}^n w_{ij})$. Los pesos espaciales reflejan la proximidad entre las áreas i y j .

B. *I*_{xy} de Moran (Reich y col., 1994)

$$I_{xy} = \frac{N \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{X})(y_j - \bar{Y})}{W \sqrt{\sum_{i=1}^n (x_i - \bar{X})^2 \sum_{i=1}^n (y_i - \bar{Y})^2}} \quad (\text{Expresión 7.8})$$

Donde N es el número de individuos, $\bar{X}$ es el valor promedio de un primer rasgo, $\bar{Y}$ es el valor promedio de un segundo rasgo, x_i es el valor del primer rasgo en un sitio particular, x_j es el valor del primer rasgo en otro sitio, y_i es el valor del segundo rasgo en un sitio particular, y_j es el valor del segundo rasgo en otro sitio, w_{ij} es el peso espacial entre los individuos i y j , y $W = (\sum_{i=1}^n \sum_{j=1}^n w_{ij})$. Los pesos espaciales reflejan la proximidad entre las áreas i y j .

2.2. Indicadores locales: G_i^* de Getis y Ord (Ord y Getis, 1995)

Para la unidad i , G_i^* se define como⁵¹:

$$G_i^* = \frac{\sum_{j=1}^n w_{ij} x_j - \bar{X} \sum_{j=1}^n w_{ij}}{S \sqrt{\frac{N \sum_{j=1}^n w_{ij}^2 - \left(\sum_{j=1}^n w_{ij} \right)^2}{N-1}}} \quad (\text{Expresión 7.9})$$

Aquí N es el número de individuos, $\bar{X}$ es el valor promedio de la variable en el área de estudio, $S = \sqrt{\left(\sum_{j=1}^n x_j^2 / N \right) - \bar{X}^2}$ es el desvío estándar de la variable en el área de estudio, x_i es el valor de una variable x un sitio particular, x_j es el valor de la variable en otro sitio, y w_{ij} es el peso espacial entre los individuos i y j . Los pesos espaciales reflejan la proximidad entre las áreas i y j .

El numerador es una suma de diferencias respecto a la media de una variable x . El denominador es un número estrictamente positivo que varía de acuerdo al número de unidades en la región analizada. G_i^* está distribuido de modo asintótico como una variable normal $N(0, 1)$. Valores positivos (negativos) del estadísticos indican la presencia de valores altos (bajos) de una variable x en torno a una unidad espacial i . Para una única unidad de análisis, el estadístico se denomina $G_i^*(0)$ y es simplemente el valor normalizado de la variable $G_i^*(0) = (x_i - \bar{X}) / S$. Un valor positivo (negativo) de $G_i^*(0)$ indica el atributo medido en la unidad i es mayor (menor) que la media, dada por todas las unidades.

⁵¹ Esta es la versión normalizada del índice, tal como fue reelaborada por Ord y Getis (1995).

APÉNDICE 3: UN MÉTODO PARA LA SELECCIÓN DE PESOS ESPACIALES PARA G_i^*

1. Descripción del método

Los análisis a nivel local requieren definir una matriz de pesos. Una posibilidad para fijar estos pesos es crear un variograma para cada variable y utilizar el valor donde el gráfico alcanza una meseta (Getis y Griffith, 2002). Otra posibilidad es calcular el estadístico G_i^* para una serie de tamaños crecientes de vecinos cercanos/distancias. A medida que la distancia entre observaciones se incrementa, el valor G_i^* crece en presencia de autocorrelación. De este modo se llega a un punto donde se maximiza la autocorrelación, y una vez que G_i^* comienza a disminuir se asume que se alcanzó el límite de la autocorrelación espacial y se elige este tamaño como el umbral de corte (Getis y Griffith, 2002). Para cada variable, un agregado de valores autocorrelacionados hace incrementar el valor de $|G_i^*|$. Luego los valores de autocorrelación empiezan a decrecer, al incorporar un número creciente de valores no autocorrelacionados. Teniendo en cuenta esta propiedad, Getis (1995) y Le Gallo y Ertur (2003) sugieren elegir el tamaño del vecindario (la distancia de corte o la cantidad de vecinos mas cercanos) que maximice la suma absoluta de G_i^* para todas las observaciones. Se genera aquí un método empírico partiendo de la posibilidad mencionada por Getis (1995), con dos refinaciones de la técnica. Primero, se hace un promedio con los valores de $|G_i^*|$ significativos, en vez de sumar directamente todos los valores de $|G_i^*|$. Esto puede concebirse como un promedio ponderado, utilizando un peso binario (1 si G_i^* es significativo para i , 0 en caso contrario), donde el denominador es la cantidad de elementos significativos (suma de pesos iguales a 1). El propósito de esto es filtrar para cada variable los valores no significativos a nivel tabla, tras una corrección por comparaciones múltiples.

El algoritmo consiste en los siguientes pasos:

1. Calcular una matriz de pesos binarios basada en los **k -vecinos más cercanos** desde $k = 1$ hasta $k = k_{max}$. Esto implica construir para cada k una matriz binaria W , donde a cada celda (x_i, x_j) se le asigna el valor de 1 si x_j es parte de los k vecinos más cercanos de x_i , y 0 en caso contrario. Para G_i^* asignar adicionalmente a la celda (x_i, x_i) el valor de 1 (se considera igual a 1 el peso de una variable con si misma).

2. Para cada valor de k , realizar un análisis de autocorrelación local exploratorio mediante el estadístico G_i^* , evaluando la significación de los valores observados por medio de **permutaciones libres**, tal como fue sugerido por Sokal y col. (1998). Corregir los valores de P individuales por comparaciones múltiples en caso de ser necesario.
3. Para cada valor de k , **obtener un promedio ponderado del valor absoluto** de todos los individuos i con autocorrelación significativa, **ponderado en función de los valores significación**: $G^*_P(k) = \sum_i \beta_i |G_i^*(k)| / \sum_i \beta_i$, donde β_i representa un vector de pesos binarios tal que $\beta_i = 1$ si $G_i^*(k)$ es significativo y 0 en caso contrario.
4. Dividir cada valor $G^*_P(k)$ por el máximo entre ellos, $G^*_{MAX} = \max(G^*_P(k))$, de modo de obtener una **estimador estandarizado** que varíe entre 0 y 1, y que tome el valor 1 para algún k .

Se define entonces al indicador G^*_{OPT} como:

$$G^*_{OPT}(k) = \frac{\sum_i \beta_i |G_i^*(k)| / \sum_i \beta_i}{\max(\sum_i \beta_i |G_i^*(k)| / \sum_i \beta_i)}$$

Que puede también escribirse como:

$$G^*_{OPT}(k) = \frac{G^*_P(k)}{G^*_{MAX}(k)}$$

Se analizan los puntos de G^*_{OPT} que describen valores máximos o puntos de ensilladura como informativos, y posibles candidatos que optimizan el valor de k . Suponiendo que se dispone de múltiples variables, se pueden agregar los siguientes pasos adicionales:

5. Calcular G^*_{OPT} para cada una de las n variables. Descartar las variables para las cuales en el paso 3 se obtenga $\sum_i \beta_{ki} = 0$ para todo k (ningún resultado significativo).

6. Suponiendo que la autocorrelación en las variables es debida a un mismo fenómeno, promediar para cada valor de k los valores estandarizados de las variables retenidas, obteniendo una función $\bar{G}^*_{OPT}$. Evaluar aquellos valores para los cuales $\bar{G}^*_{OPT}$ es máxima o tiene un punto de ensilladura.

2. Aplicación del método al modelo de estudio

Se calculó, para cada rasgo morfológico y para los valores de NDVI, el índice G^*_{OPT} utilizando tamaños crecientes de k , desde $k = 1$ hasta $k = 40$. Los valores P se obtuvieron mediante 9999 permutaciones y una corrección FDR, lo que permite obtener una situación intermedia entre la esperada tras una corrección conservativa y aquella que resultaría sin corrección (Caldas de Castro y Singer, 2006). Tal como puede observarse en la Figura 7.2, de las variables que muestran valores significativos bajo corrección FDR, y exceptuando a PEL y NLP, el resto tiende a un valor estable aproximadamente para $k = 18$, que es el valor elegido por lo tanto para el análisis local.

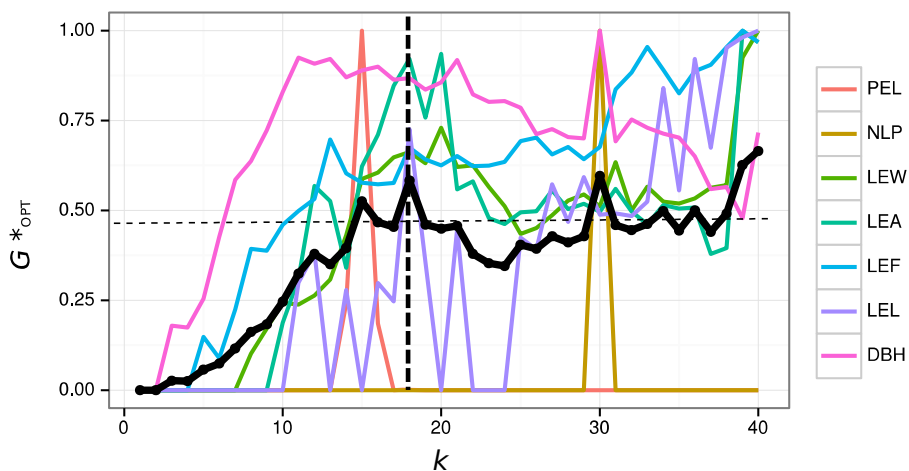


Figura 7.2 Estadístico G^*_{OPT} en función de la cantidad de vecinos próximos (k). La línea vertical punteada indica el valor elegido para el análisis local ($k = 18$). La línea negra muestra los valores del estimador $\bar{G}^*_{OPT}$. La línea horizontal punteada indica el valor para el cual $\bar{G}^*_{OPT}$ se estabiliza.

Bibliografía

BIBLIOGRAFÍA

- Ahmad, R., Ismail, S., Moinuddin, M. y Shaheen, T.** (1994). Screening of mesquite (*Prosopis* spp) for biomass production at barren sandy areas using highly saline water for irrigation. *Pakistan Journal of Botany* 26, 265–282.
- Aiazzi, M., Carpane, P., Di Rienzo, J. y Argüello, J.** (2002). Effects of salinity and temperature on the germination and early seedling growth of *Atriplex cordobensis* Gandoger et Stuckert (Chenopodiaceae). *Seed Science and Technology* 30, 329–338.
- Akaike, H.** (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on* 19, 716–723.
- Alban, L., Matorel, M., Romero, J., Grados, N., Cruz, G. y Felker, P.** (2002). Cloning of elite, multipurpose trees of the *Prosopis juliflora/pallida* complex in Piura, Peru. *Agroforestry Systems* 54, 173–182.
- Angueira, C., Prieto, D., López, J. y Barraza, G.** (2007). Sistema de Información Geográfica de Santiago del Estero (SigSE 2.0). Instituto Nacional de Tecnología Agropecuaria, Argentina.
- Anselin, L.** (1995). Local indicators of spatial association-LISA. *Geographical Analysis* 27, 93–115.
- Anselin, L. y Griffith, D.** (1988). Do spatial effects really matter in regression analysis? *Papers in Regional Science* 65, 11–34.
- Ansley, R., Trevino, B. y Jacoby, P.** (1998). Intraspecific competition in honey mesquite: leaf and whole plant responses. *Journal of Range Management* 51, 345–352.
- ESRI** (2011). ArcGIS Desktop. Release 10.1 Environmental Systems Research Institute (ESRI). Redlands. California.
- Argyres, A. y Schmitt, J.** (1991). Microgeographic genetic structure of morphological and life history traits in a natural population of *Impatiens capensis*. *Evolution* 45, 178–189.
- Arístide, P.** (2010). Procesos históricos de cambio en la apropiación del territorio en Figueroa (Santiago del Estero, Argentina, Chaco semiárido). Tesis de Maestría. Universidad Internacional de Andalucía, Andalucía.
- Asrar, G., Fuchs, M., Kanemasu, E. y Hatfield, J.** (1984). Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat. *Agronomy Journal* 76, 300–306.
- AT&T staff** (2013). From S to R: 35 Years of AT&T Leadership in Statistical Computing. Disponible en: http://www.research.att.com/articles/featured_stories/2013_09/201309_SandR.htmlfbid=jdyjLsQqFCt [accedido: febrero de 2015].
- Azevedo, V., Kanashiro, M., Ciampi, A. y Grattapaglia, D.** (2007). Genetic structure and mating system of *Manilkara huberi* (Ducke) A. Chev., a heavily logged Amazonian timber species. *Journal of Heredity* 98, 646–654.
- Baker, A., Sprent, J. y Wilson, J.** (1995). Effects of sodium chloride and mycorrhizal infection on the growth and nitrogen fixation of *Prosopis juliflora*. *Symbiosis (Rehovot)* 19, 39–51.
- Barbujani, G.** (1987). Autocorrelation of gene frequencies under isolation by distance. *Genetics* 117, 777–782.
- Bassam, B., Caetano-Anollés, G. y Gresshoff, P.** (1991). Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry* 196, 80–83.

- Baythavong, B.** (2011). Linking the spatial scale of environmental variation and the evolution of phenotypic plasticity: selection favors adaptive plasticity in fine-grained environments. *The American Naturalist* 178, 75–87.
- Benjamini, Y. y Hochberg, Y.** (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57, 289–300.
- Bessegga, C., Ferreyra, L., Julio, N., Montoya, S., Saidman, B. y Vilardi, J.** (2000). Mating system parameters in species of genus *Prosopis* (Leguminosae). *Hereditas* 132, 19–27.
- Bessegga, C., Saidman, B. y Vilardi, J.** (2005). Genetic relationships among American species of *Prosopis* (Leguminosae) based on enzyme markers. *Genetics and Molecular Biology* 28, 277–286.
- Bessegga, C., Vilardi, J. y Saidman, B.** (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.
- Bessegga, C., Saidman, B., Darquier, M., Ewens, M., Sánchez, L., Rozenberg, P. y Vilardi, J.** (2009). Consistency between marker-and genealogy-based heritability estimates in an experimental stand of *Prosopis alba* (Leguminosae). *American Journal of Botany* 96, 458–465.
- Bessegga, C., Saidman, B., Darquier, M., Ewens, M., Felker, P. y Vilardi, J.** (2011). Accuracy of dominant markers for estimation of relatedness and heritability in an experimental stand of *Prosopis alba* (Leguminosae). *Tree Genetics & Genomes* 7, 103–115.
- Bessegga, C., Pometti, C., Ewens, M., Saidman, B. y Vilardi, J.** (2012). Strategies for conservation for disturbed *Prosopis alba* (Leguminosae, Mimosoideae) forests based on mating system and pollen dispersal parameters. *Tree Genetics & Genomes* 8, 277–288.
- Bessegga, C., Pometti, C., Miller, J., Watts, R., Saidman, B. y Vilardi, J.** (2013). New microsatellite loci for *Prosopis alba* and *P. chilensis* (Fabaceae). *Applications in Plant Sciences* 1.
- Bessegga, C., Pometti, C., Ewens, M., Saidman, B. y Vilardi, J.** (2015a). Evidences of local adaptation in quantitative traits in *Prosopis alba* (Leguminosae). *Genetica* 143, 31–44.
- Bessegga, C., Pometti, C., Ewens, M., Saidman, B. y Vilardi, J.** (2015b). Improving initial trials in tree breeding using kinship and breeding values estimated in the wild: the case of *Prosopis alba* in Argentina. *New Forests* 46, 427–448.
- Bleckmann, C., Hull, H. y Hoshaw, R.** (1980). Cuticular ultrastructure of *Prosopis velutina* and *Acacia greggii* leaflets. *Botanical Gazette* 141, 1–8.
- Bocquet-Appel, J.-P. y Sokal, R.** (1989). Spatial autocorrelation analysis of trend residuals in biological data. *Systematic Biology* 38, 333–341.
- Borcard, D., Gillet, F. y Legendre, P.** (2011). *Numerical ecology with R*. Springer-Verlag, Nueva York.
- Bornet, B. y Branchard, M.** (2001). Nonanchored inter simple sequence repeat (ISSR) markers: reproducible and specific tools for genome fingerprinting. *Plant Molecular Biology Reporter* 19, 209–215.
- Breiman, L., Friedman, J., Stone, C. y Olshen, R.** (1993). *Classification and regression trees*. Chapman and Hall/CRC, Boca Raton, Florida.
- Bukhari, Y.** (1997). Cytoevolution of taxa in *Acacia* and *Prosopis* (Mimosaceae). *Australian Journal of Botany* 45, 879–891.
- Burghardt, A.** (1996). La identidad de cuatro especies de *Prosopis* L. expresada a través de sus patrones electroforéticos. *Mendeliana* 12, 38–50.
- Burghardt, A. y Palacios, R.** (1997). Electrophoretic characterization of the American sections of *Prosopis* L. (Leguminosae: Mimosoideae). *Bulletin of the International Group for the Study of Mimosoideae* 20, 71–83.

- Burkart, A.** (1976). A monograph of the genus *Prosopis* (Leguminosae subfam. Mimosoideae). *Journal of the Arnold Arboretum* 57, 219–249.
- Busch, J., Waser, P. y De Woody J.** (2007). Recent demographic bottlenecks are not accompanied by a genetic signature in banner-tailed kangaroo rats (*Dipodomys spectabilis*). *Molecular Ecology* 16, 2450–2462.
- Buzatti, R. de O., Ribeiro, R., Lemos Filho, J. y Lovato, M.** (2012). Fine-scale spatial genetic structure of *Dalbergia nigra* (Fabaceae), a threatened and endemic tree of the Brazilian Atlantic Forest. *Genetics and Molecular Biology* 35, 838–846.
- Cabrera, A.** (1971). Fitogeografía de la República Argentina. *Boletín de La Sociedad Argentina de Botánica* 14, 1–42.
- Cabrera, A.** (1976). Regiones fitogeográficas argentinas. Editorial Acme, Buenos Aires.
- Cabrera, A. y Willink, A.** (1973). Biogeografía de América latina. Monografía 13, Serie de Biología. OEA, Washington D.C., Estados Unidos.
- Caldas de Castro, M. y Singer, B.** (2006). Controlling the false discovery rate: a new application to account for multiple and dependent tests in local statistics of spatial association. *Geographical Analysis*, 38, 180–208.
- Campbell, J.** (2002). Introduction to remote sensing. Guilford Press, Nueva York, Estados Unidos.
- Carnevale, N., Alzugaray, C. y Lopez, D.** (2004). Effect of salinity on seedlings of two dominant tree species in the “Quebrachal” of *Schinopsis balansae* Engl. *Revista de Investigaciones de la Facultad de Ciencias Agrarias* 6, 43–49.
- Castric, V., Bernatchez, L., Belkhir, K. y Bonhomme, F.** (2002). Heterozygote deficiencies in small lacustrine populations of brook charr *Salvelinus fontinalis* Mitchell (Pisces, Salmonidae): a test of alternative hypotheses. *Heredity* 89, 27–35.
- Catalano, S., Vilardi, J., Tosto, D. y Saidman, B.** (2008). Molecular phylogeny and diversification history of *Prosopis* (Fabaceae: Mimosoideae). *Biological Journal of the Linnean Society* 93, 621–640.
- Cavagnaro, J., Ponce, M., Guzmán, J. y Cirrincione, M.** (2006). Argentinean cultivars of *Vitis vinifera* grow better than European ones when cultured in vitro under salinity. *Biocell* 30, 1–7.
- Cavers, S., Degen, B., Caron, H., Lemes, M., Margis, R., Salgueiro, F. y Lowe, A.** (2005). Optimal sampling strategy for estimation of spatial genetic structure in tree populations. *Heredity* 95, 281–289.
- Chander, G., Markham, B. y Helder, D.** (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment* 113, 893–903.
- Chavez, P.S.** (1989). Radiometric calibration of Landsat Thematic Mapper multispectral images. *Photogrammetric Engineering and Remote Sensing* 55, 1285–1294.
- Chavez, P.** (1996). Image-based atmospheric corrections-revisited and improved. *Photogrammetric Engineering and Remote Sensing* 62, 1025–1035.
- Chen, K., Ren, P., Ying, C., Jiang, Q. y Jia, X.** (2011). Genetic relationships among *Jatropha curcas* L. clones from Panzhuhua, China as revealed by RAPD and ISSR. *African Journal of Agricultural Research* 6, 2582–2585.
- Chessel, D., Dufour, A. y Thioulouse, J.** (2004). The ade4 package-I-One-table methods. *R News* 4, 5–10.
- Chung, MY., Nason, J., Epperson, B. y Chung, MG.** (2003). Temporal aspects of the fine-scale genetic structure in a population of *Cinnamomum insularimontanum* (Lauraceae). *Heredity* 90, 98–106.
- Cliff, A. y Ord, J.** (1973). Spatial autocorrelation. Pion. Londres.

- Cliff, A. y Ord, J..** (1981). Spatial processes: models and applications, Pion, Londres.
- Cockerham, C.** (1969). Variance of gene frequencies. *Evolution* 23, 72–84.
- Cockerham, C.** (1971). Higher order probability functions of identity of alleles by descent. *Genetics* 69, 235.
- Cockerham, C.** (1973). Analyses of gene frequencies. *Genetics* 74, 679–700.
- Cockerham, C. y Weir, B.** (1968). Sib mating with two linked loci. *Genetics* 60, 629.
- Cockerham, C. y Weir, B.** (1987). Correlations, descent measures: drift with migration and mutation. *Proceedings of the National Academy of Sciences* 84, 8512–8514.
- Colwell, R., Mao, C. y Chang, J.** (2004). Interpolating, extrapolating, and comparing incidence-based species accumulation curves. *Ecology*, 85, 2717–2727.
- Cony, M.** (1995). Reforestación racional de zonas áridas y semiáridas con árboles de múltiples propósitos. *Interciencia* 20, 249–253.
- Cony, M.** (1996). Genetic variability in *Prosopis flexuosa* DC, a native tree of the Monte phytogeographic province, Argentina. *Forest Ecology and Management* 87, 41–49.
- Cony, M. y Trione, S.** (1998). Inter-and intraspecific variability in *Prosopis flexuosa* and *P. chilensis*: seed germination under salt and moisture stress. *Journal of Arid Environments*, 40, 307–317.
- Cornuet, J. y Luikart, G.** (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144, 2001–2014.
- Coronel de Renolfi, M. y Gómez, A.** (2014). Estimación del costo del primer raleo de algarrobo blanco (*Prosopis alba*): metodología aplicada al caso de una plantación en Herrera, Santiago del Estero, Argentina. *Ingeniería Solidaria* 10, 19–27.
- Crow, J. y Kimura, M.** (1970). An introduction to population genetics theory. Evanston and London: Harper & Row, Publishers, Nueva York.
- Culhane, A., Perriere, G., Considine, E., Cotter, T. y Higgins, D.** (2002). Between-group analysis of microarray data. *Bioinformatics* 18, 1600–1608.
- Culhane, A., Perrière, G. y Higgins, D.** (2003). Cross-platform comparison and visualisation of gene expression data using co-inertia analysis. *BMC Bioinformatics* 4, 59.
- Czúcz, B., Gálhidy, L. y Mátyás, C.** (2011). Present and forecasted xeric climatic limits of beech and sessile oak distribution at low altitudes in Central Europe. *Annals of Forest Science* 68, 99–108.
- D'Surney, S., Shugart, L. y Theodorakis, C.** (2001). Genetic markers and genotyping methodologies: an overview. *Ecotoxicology* 10, 201–204.
- Dale, M. y Fortin, M.** (2014). Spatial analysis: a guide for ecologists. Segunda edición. Cambridge University Press, Cambridge.
- Darquier, M., Bessega, C., Cony, M., Vilardi, J. y Saidman, B.** (2013). Evidence of heterogeneous selection on quantitative traits of *Prosopis flexuosa* (Leguminosae) from multivariate Q ST–F ST test. *Tree Genetics & Genomes* 9, 307–320.
- Davies, S., Cavers, S., Finegan, B., Navarro, C. y Lowe, A.** (2010). Genetic consequences of multigenerational and landscape colonisation bottlenecks for a neotropical forest pioneer tree, *Vochysia ferruginea*. *Tropical Plant Biology* 3, 14–27.
- DeWoody, J., Nason, J. y Hipkins, V.** (2006). Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes*, 6, 951–957.

- Di Rienzo, A., Peterson, A., Garza, J., Valdes, A., Slatkin, M., y Freimer, N.** (1994). Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences* 91, 3166-3170.
- Dick, C., Etchelecu, G. y Austerlitz, F.** (2003). Pollen dispersal of tropical trees (*Dinizia excelsa*: Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rainforest. *Molecular Ecology* 12, 753-764.
- Dick, C., Hardy, O., Jones, F. y Petit, R.** (2008). Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Tropical Plant Biology* 1, 20-33.
- Dobzhansky, T.** (1937). Further data on the variation of the Y chromosome in *Drosophila Pseudoobscura*. *Genetics* 22, 340-346.
- Dolédéc, S. y Chessel, D.** (1987). Rythmes saisonniers et composantes stationnelles en milieu aquatique. I: Description d'un plan d'observation complet par projection de variables. *Acta Oecologica. Oecologia Generalis* 8, 403-426.
- Dolédéc, S. y Chessel, D.** (1994). Co-inertia analysis: an alternative method for studying species-environment relationships. *Freshwater Biology* 31, 277-294.
- Dormann, C., McPherson, J., Araújo M. y Bivand, R.** (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* 30, 609-628.
- Double, M., Peakall, R., Beck, N., Cockburn, A. y Sorenson, M.** (2005). Dispersal, philopatry, and infidelity: dissecting local genetic structure in superb fairy-wrens (*Malurus cyaneus*). *Evolution* 59, 625-635.
- Dray, S., Chessel, D. y Thioulouse, J.** (2003). Co-inertia analysis and the linking of ecological data tables. *Ecology* 84, 3078-3089.
- Dray, S. y Dufour, A.-B.** (2007). The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* 22, 1-20.
- Du, Y., Teillet, P. y Cihlar, J.** (2002). Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection. *Remote Sensing of Environment* 82, 123-134.
- Dungan, J., Perry, J., Dale, M., Legendre, P., Citron-Pousty, S., Fortin, M., Jakomulska, A., Miriti, M. y Rosenberg, M.** (2002). A balanced view of scale in spatial statistical analysis. *Ecography* 25, 626-640.
- Dutton, R.** (1989). *Prosopis* sp: a justification for their future research and development. En: Dutton, R. (Ed.). *A Research and development strategy for Prosopis*. Center for Overseas research and Development, Durham, Inglaterra. pp. 33-40.
- Earl, D.** (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4, 359-361.
- Epperson, B.** (1993). Recent advances in correlation studies of spatial patterns of genetic variation. *Evolutionary Biology*. 27, 95-155.
- Epperson, B.** (1999). Gustave Malécot, 1911-1998: Population Genetics Founding Father. *Genetics* 152, 477-484.
- Epperson, B. y Alvarez-Buylla, E.** (1997). Limited seed dispersal and genetic structure in life stages of *Cecropia obtusifolia*. *Evolution* 51, 275-282.
- Epperson, B. y Li, T.-Q.** (1997). Gene dispersal and spatial genetic structure. *Evolution* 672-681.
- Epstein, E., Kingsbury, R., Norlyn, J. y Rush, D.** (1979). Production of food crops and other biomass by seawater culture. En: *The Biosaline Concept: An Approach to the Utilization of Underexploited Resources*. Hollaender, A., Aller, J., Epstein, E., San Pietro, A., y Zaborsky, O. (Ed.). Springer, pp. 77-99.

- Ewens, M., Gezan, S. y Felker, P.** (2012). Five year field evaluation of *Prosopis alba* clones on pH 9–10 soils in Argentina selected for growth in the greenhouse at seawater salinities (45 dS m⁻¹). *Forests* 3, 95–113.
- Excoffier, L., Smouse, P. y Quattro, J.** (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491.
- Fagg, C. y Stewart, J.** (1994). The value of *Acacia* and *Prosopis* in arid and semi-arid environments. *Journal of Arid Environments* 27, 3–25.
- Fajardo, A. y McIntire, E.** (2007). Distinguishing microsite and competition processes in tree growth dynamics: an a priori spatial modeling approach. *The American Naturalist* 169, 647–661.
- Falush, D., Stephens, M. y Pritchard, J.** (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164, 1567–1587.
- Felker, P., Clark, P.R., Laag, A. y Pratt, P.** (1981). Salinity tolerance of the tree legumes: Mesquite (*Prosopis glandulosa* var. *torreyana*, *P. velutina* and *P. articulata*) Algarrobo (*P. chilensis*), Kiawe (*P. pallida*) and Tamarugo (*P. tamarugo*) grown in sand culture on nitrogen-free media. *Plant and Soil* 61, 311–317.
- Felker, P., Cannell, G., Clark, P., Osborn, J. y Nash, P.** (1983a). Biomass production of *Prosopis* species (mesquite), *Leucaena*, and other leguminous trees grown under heat/drought stress. *Forest Science* 29, 592–606.
- Felker, P., Cannell, G., Osborn, J., Clark, P. y Nash, P.** (1983b). Effects of irrigation on biomass production of 32 *Prosopis* (mesquite) accessions. *Experimental Agriculture* 19, 187–198.
- Felker, P., Lopez, C., Soulier, C., Ochoa, J., Abdala, R. y Ewens, M.** (2001). Genetic evaluation of *Prosopis alba* (algarrobo) in Argentina for cloning elite trees. *Agroforestry Systems* 53, 65–76.
- Felker, P., Ewens, M., Velarde, M. y Medina, D.** (2008). Initial evaluation of *Prosopis alba* Griseb clones selected for growth at seawater salinities. *Arid Land Research and Management* 22, 334–345.
- Fenster, C., Vekemans, X. y Hardy, O.** (2003). Quantifying gene flow from spatial genetic structure data in a metapopulation of *Chamaecrista fasciculata* (Leguminosae). *Evolution*, 57, 995–1007.
- Ferreira, L., Bessega, C., Vilardi, J. y Saidman, B.** (2004). First report on RAPDs patterns able to differentiate some Argentinean species of section Algarobia (*Prosopis*, Leguminosae). *Genetica* 121, 33–42.
- Ferreira, L., Bessega, C., Vilardi, J. y Saidman, B.** (2007). Consistency of population genetics parameters estimated from isozyme and RAPDs dataset in species of genus *Prosopis* (Leguminosae, Mimosoideae). *Genetica* 131, 217–230.
- Ferreira, L., Vilardi, J., Tosto, D., Julio, N. y Saidman, B.** (2010). Adaptive genetic diversity and population structure of the “algarrobo” [*Prosopis chilensis* (Molina) Stuntz] analysed by RAPD and isozyme markers. *European Journal of Forest Research* 129, 1011–1025.
- Ferreira, L., Vilardi, J., Verga, A., López, V. y Saidman, B.** (2013). Genetic and morphometric markers are able to differentiate three morphotypes belonging to Section Algarobia of genus *Prosopis* (Leguminosae, Mimosoideae). *Plant Systematics and Evolution* 299, 1157–1173.
- Flowers, T. y Yeo, A.** (1995). Breeding for salinity resistance in crop plants: where next? *Functional Plant Biology* 22, 875–884.
- Fortin, M., James, P.M., MacKenzie, A., Melles, S.J. y Rayfield, B.** (2012). Spatial statistics, spatial regression, and graph theory in ecology. *Spatial Statistics* 1, 100–109.
- Foulkes, A.** (2009). Applied statistical genetics with R: for population-based association studies. Springer Science & Business Media, New York.

- Freedman, D. y Diaconis, P.** (1981). On the histogram as a density estimator: L 2 theory. *Probability Theory and Related Fields* 57, 453–476.
- Freeman, E. y Moisen, G.** (2007). Evaluating kriging as a tool to improve moderate resolution maps of forest biomass. *Environmental Monitoring and Assessment* 128, 395–410.
- Fuchs, E. y Hamrick, J.** (2010). Spatial genetic structure within size classes of the endangered tropical tree *Guaiacum sanctum* (Zygophyllaceae). *American Journal of Botany* 97, 1200–1207.
- Galera, F.** (2000). Las especies del género *Prosopis* (algarrobos) de América Latina con especial énfasis en aquellas de interés económico. FAO. Disponible en: <http://www.fao.org/docrep/006/ad314s/ad314s00.HTM> [accedido: febrero de 2015].
- Galera, F., Trevisson, M. y Bruno, S.** (1992). *Prosopis* in Argentina: initial results on cultivation in greenhouses and orchards, and pod quality for food or feed of five native *Prosopis* species of Córdoba Province. FAO. Disponible en <http://www.fao.org/docrep/006/ad314s/AD314S02.htm> [accedido: febrero de 2015].
- Geary, R.** (1954). The contiguity ratio and statistical mapping. *The Incorporated Statistician* 5, 115–146.
- Geesing, D., Felker, P. y Bingham, R.** (2000). Influence of mesquite (*Prosopis glandulosa*) on soil nitrogen and carbon development: Implications for global carbon sequestration. *Journal of Arid Environments* 46, 157–180.
- Genise, J., Palacios, R., Hoc, P., Carrizo, R., Moffat, L., Mom, M., Agullo, M., Agullo, L., Picca, P. y Torregrosa, S.** (1990). Observaciones sobre la biología floral de *Prosopis* (Leguminosae, Mimosoideae). II. Fases florales y visitantes en el Distrito Chaqueño Serrano. *Darwiniana* 30, 71–85.
- Getis, A.** (1995). Spatial filtering in a regression framework: examples using data on urban crime, regional inequality, and government expenditures. En: Anselin, L y Rey, S. (Ed.). *Perspectives on spatial data analysis*. Springer Berlin Heidelberg, pp. 191–202.
- Getis, A.** (2008). A history of the concept of spatial autocorrelation: A geographer's perspective. *Geographical Analysis* 40, 297–309.
- Getis, A. y Griffith, D.** (2002). Comparative spatial filtering in regression analysis. *Geographical analysis*, 34, 130–140.
- Getis, A. y Ord, J.** (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis* 24, 189–206.
- Giménez, A., Ríos, N., Moglia, G., Hernández, P. y Bravo, S.** (2001). Estudio de magnitudes dendrométricas en función de la edad en *Prosopis alba* Griseb., Algarrobo blanco, Mimosaceae. *Revista Forestal Venezolana* 45, 175–183.
- Gitelson, A., Peng, Y., Masek, J., Rundquist, D., Verma, S., Suyker, A., Baker, J., Hatfield, J. y Meyers, T.** (2012). Remote estimation of crop gross primary production with Landsat data. *Remote Sensing of Environment* 121, 404–414.
- Goslee, S.** (2011). Analyzing remote sensing data in R: the landsat package. *Journal of Statistical Software*, 43, 1–25.
- Goudet, J.** (2005). Hierfstat, a package for R to compute and test hierarchical F statistics. *Molecular Ecology Notes*, 5, 184–186.
- Hafez E.** (1962) The behavior of domestic animals. Williams and Wilkins, Baltimore, Maryland, Estados Unidos.
- Hamrick, J., Murawski, D. y Nason, J.** (1993). The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. En: Fleming, E. y Estrada, A. (Ed.). *Frugivory and Seed Dispersal: Ecological and Evolutionary Aspects*, Springer Science & Business Media, pp. 281–297.
- Hamrick, J. y Trapnell, D.** (2011). Using population genetic analyses to understand seed dispersal patterns. *Acta Oecologica*, 37, 641–649.

- Hardy, O.** (2003). Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Molecular Ecology* 12, 1577–1588.
- Hardy, O. y Vekemans, X.** (1999). Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity* 83, 145–154.
- Hardy, O., Vanderhoeven, S., Meerts, P. y Vekemans, X.** (2000). Spatial autocorrelation of allozyme and quantitative markers within a natural population of *Centaurea jacea* (Asteraceae). *Journal of Evolutionary Biology*, 13, 656–667.
- Hardy, O., Maggia, L., Bandou, E., Breyné, P. y Caron, H.** (2006). Fine-scale genetic structure and gene dispersal inferences in 10 Neotropical tree species. *Molecular Ecology* 15, 559–571.
- Harris, P., Pasiecznik, N., Smith, S., Billington, J. y Ramí, L.** (2003). Differentiation of *Prosopis juliflora* (Sw.) DC. and *P. pallida* (H. & B. ex. Willd.) HBK using foliar characters and ploidy. *Forest Ecology and Management* 180, 153–164.
- Harsh, L., Tewari, J., Sharma, N. y Felker, P.** (1996). Performance of *Prosopis* Species in Arid Regions of India. Proceedings of the Symposium “*Prosopis*: semi-arid fuelwood and forage tree. Building Consensus for the Disenfranchised”. US National Academy of Sciences, pp. 11.
- Heuertz, M., Vekemans, X., Hausman, J.-F., Palada, M. y Hardy, O.** (2003). Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. *Molecular Ecology* 12, 2483–2495.
- Heywood, J.** (1991). Spatial analysis of genetic variation in plant populations. *Annual Review of Ecology and Systematics* 22, 335–355.
- Holsinger, K. y Weir, B.** (2009). Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nature Reviews Genetics* 10, 639–650.
- Hothorn, T., Hornik, K. y Zeileis, A.** (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15, 651–674.
- Huete, A. y Tucker, C.** (1991). Investigation of soil influences in AVHRR red and near-infrared vegetation index imagery. *International Journal of Remote Sensing* 12, 1223–1242.
- Hunziker, J., Poggio, L., Naranjo, C., Palacios, R. y Andrada, A.** (1975). Cytogenetics of some species and natural hybrids in *Prosopis* (Leguminosae). *Canadian Journal of Genetics and Cytology* 17, 253–262.
- Hunziker, J., Naranjo, C., Palacios, R. y Poggio, L.** (1977). Chromosomal cytology and hybridization. En: Simpson B. (Ed.), *Mesquite. Its Biology in Two Desert Scrub Ecosystems*. Hutchinson & Ross Inc., Dowden, Pensilvania, pp. 56–60.
- Hunziker, J., Saidman, B., Naranjo, C., Palacios, R., Poggio, L. y Burghardt, A.** (1986). Hybridization and genetic variation of Argentine species of *Prosopis*. *Forest Ecology and Management* 16, 301–315.
- Hurlbert, S.** (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecology*, 52, 577–586.
- Huxley, J.** (1942). *Evolution. The Modern Synthesis*. Evolution. George Allen & Unwin Ltd., Londres.
- Ibitoye, D. y Akin-Idowu, P.** (2010). Marker-assisted-selection (MAS): A fast track to increase genetic gain in horticultural crop breeding. *African Journal of Biotechnology* 9, 8889–8895.
- Ishida, Y.** (2009). Sewall Wright and Gustave Malécot on isolation by distance. *Philosophy of Science* 76, 784–796.
- Jacquard, A.** (1970). *Structures génétiques des populations*. Masson et Cie, Paris.
- Jiang, Z. y Huete, A.** (2010). Linearization of NDVI based on its relationship with vegetation fraction. *Photogrammetric Engineering & Remote Sensing* 76, 965–975.

- Jombart, T.** (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405.
- Jones, C., Edwards, K., Castaglione, S., Winfield, M., Sala, F., Van de Wiel, C., Bredemeijer, G., Vosman, B., Matthes, M., Daly, A., Brettschneider, R., Bettini, P., Buiatti, M., Maestri, E., Malcevschi, A., Marmiroli, N., Aert, R., Volckaert, G., Rueda, J., Linacero, R., Vazquez, A. y Karp, A.** (1997). Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding* 3, 381–390.
- Juarez-Munoz, J., Carrillo-Castaneda, G., Arreguin, R. y Rubluo, A.** (2002). Inter-and intra-genetic variation of four wild populations of *Prosopis* using RAPD-PCR fingerprints. *Biodiversity & 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, Córdoba, Argentina.
- Jump, A. y Penuelas, J.** (2007). Extensive spatial genetic structure revealed by AFLP but not SSR molecular markers in the wind pollinated tree, *Fagus sylvatica*. *Molecular ecology* 16, 925–936.
- Kabthimer, G.** (2012). Assessment of spatio-temporal patterns of NDVI in response to precipitation using NOAA-AVHRR rainfall estimate and NDVI data from 1996–2008, Ethiopia. Tesis de Maestría, Stockholm University, Estocolmo, Suecia.
- Kalinowski, S.** (2004). Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conservation Genetics* 5, 539–543.
- Kalisz, S., Nason, J., Hanzawa, F. y Tonsor, S.** (2001). Spatial population genetic structure in *Trillium grandiflorum*: the roles of dispersal, mating, history, and selection. *Evolution* 55, 1560–1568.
- Kamvar, Z., Tabima, J. y Grünwald, N.** (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2, e281.
- Karhunen, M. y Ovaskainen, O.** (2012). Estimating population-level coancestry coefficients by an admixture F model. *Genetics* 192, 609–617.
- Keenan, K., McGinnity, P., Cross, T., Crozier, W. y Prodöhl, P.** (2013). diveRsity: An R package for the estimation of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* 4, 782–788.
- Keys, R.** (1993). Mating systems and pollination biology of velvet mesquite (*Prosopis velutina* Wooten). University of Arizona, Arizona, Estados Unidos.
- Keys, R. y Smith, S.** (1994). Mating system parameters and population genetic structure in pioneer populations of *Prosopis velutina* (Leguminosae). *American Journal of Botany* 81, 1013–1020.
- Kesawat, M. y Kumar, B.** (2009). Molecular markers: It's application in crop improvement. *Journal of Crop Science and Biotechnology*, 12, 169–181.
- Kimura, M.** (1983). The neutral theory of molecular evolution. Cambridge University Press. Nueva York.
- Kimura, M. y Crow, J.** (1964). The number of alleles that can be maintained in a finite population. *Genetics* 49, 725.
- Kimura, M. y Weiss, G.** (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* 49, 561.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B. y Rubel, F.** (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift* 15, 259–263.
- Krause, A. y Olson, M.** (2005). The basics of S-Plus. Springer-Verlag, Nueva York, pp. 444.

- Krivoruchko, K.** (2012). Empirical bayesian kriging implemented in ArcGIS geostatistical analyst. ArcUser, The Magazine for Esri Software Users. ESRI 6–10.
- Kumar, R., Parthiban, K. y Rao, M.** (2009). Molecular characterization of *Jatropha* genetic resources through inter-simple sequence repeat (ISSR) markers. Molecular Biology Reports 36, 1951–1956.
- Lauer, D., Morain, S. y Salomonson, V.** (1997). The Landsat program: Its origins, evolution, and impacts. Photogrammetric Engineering and Remote Sensing 63, 831–838.
- Lavado, R., Taleisnik, E., Grunberg, K. y Santa María, G.** (2008). Visión sintética de la distribución y magnitud de los suelos afectados por salinidad en la Argentina. La Salinización de Suelos En La Argentina: Su Impacto En La Producción Agropecuaria. Editorial Universidad Católica de Córdoba, Córdoba.
- Lay, R.** (2005). Use of Landsat TM and ETM+ to describe intra-season change in vegetation, with consideration for wildlife management. Tesis doctoral, University of Northern British Columbia. Prince George, Canada.
- De Leeuw, J. y Mair, P.** (2009). Multidimensional scaling using majorization: the R package smacof. Journal of Statistical Software 31, 1–30.
- Le Gallo, J. y Ertur, C.** (2003). Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980–1995. Papers in regional science, 82, 175–201.
- Legendre, P. y Fortin, M.** (1989). Spatial pattern and ecological analysis. Vegetatio 80, 107–138.
- Legendre, P. y Legendre, L.** (2012). Numerical ecology. Tercera edición en inglés. Elsevier, Amsterdam.
- Lenney, M., Woodcock, C., Collins, J. y Hamdi, H.** (1996). The status of agricultural lands in Egypt: the use of multitemporal NDVI features derived from Landsat TM. Remote Sensing of Environment 56, 8–20.
- Levin, S.** (1992). The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. Ecology 73, 1943–1967.
- Levin, D. y Kerster, H.** (1974). Gene flow in seed plants. Evolutionary Biology 7, 139–220.
- Li, H. y Reynolds, J.** (1995). On definition and quantification of heterogeneity. Oikos 73, 280–284.
- Lichstein, J., Simons, T., Shriner, S. y Franzreb, K.** (2002). Spatial autocorrelation and autoregressive models in ecology. Ecological Monographs 72, 445–463.
- Lieveld, C.** (2005). The irrigation system as an Open System; Proyecto Rio Dulce and the Rio Dulce Basin. Tesis de Maestría. Delf University of Technology, Delft, Países Bajos.
- Lindbäck, K. y Sjölin, A.** (2006). Arsenic in groundwater in the south western part of the Rio Dulce alluvial cone, Santiago del Estero Province, Argentina. Tesis de Maestría, Department of Land and Water Resources Engineering KTH, Estocolmo, Suecia.
- Loiselle, B., Sork, V., Nason, J. y Graham, C.** (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). American Journal of Botany 82, 1420–1425.
- Lopez-Gallego, C. y O’Neil, P.** (2010). Life-history variation following habitat degradation associated with differing fine-scale spatial genetic structure in a rainforest cycad. Population Ecology 52, 191–201.
- Luikart, G. y Cornuet, J.** (1998). Empirical Evaluation of a Test for Identifying Recently Bottlenecked Populations from Allele Frequency Data. Conservation Biology 12, 228–237.

- Luikart, G., Allendorf, F., Cornuet, J. y Sherwin, W.** (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* 89, 238–247.
- Lynch, M.** (1988). Estimation of relatedness by DNA fingerprinting. *Molecular Biology and Evolution* 5, 584–599.
- Malécot, G.** (1948). *Mathématiques de l'hérédité*. Masson et Cie, Paris.
- Malécot, G.** (1955). Génétique de population: principes et applications. *Population* 10, 239–262.
- Malécot, G.** (1959). Les modèles stochastiques en génétique de population. Institut de Statistique de L'Université de Paris, Paris.
- Malécot, G.** (1973). Isolation by distance. En: Morton, E. (Ed.). *Genetic Structure of Populations*. University of Hawaii Press, Honolulu, Hawaii, pp. 72–75.
- Malhado, A., Malhi, Y., Whittaker, R., Ladle, R., Steege, H., Phillips, O., Butt, N., Aragão, L., Quesada, C., y Araujo-Murakami, A., Arroyo, L., Peacock J., Lopez-Gonzalez G., Baker T., Anderson L., Almeida S., Higuchi N., Killeen T., Monteagudo A., Neill D., Pitman N., Prieto A., Salom Vasquez-Martinez, R. y Laurance, W.** (2009). Spatial trends in leaf size of Amazonian rainforest trees. *Biogeosciences* 6, 1563–1576.
- Manel, S., Schwartz, M., Luikart, G. y Taberlet, P.** (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* 18, 189–197.
- Mantel, N.** (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research* 27, 209–220.
- Marcucci Poltri, S., Zelener, N., Traverso, J., Gelid, P. y Hopp, H.** (2003). Selection of a seed orchard of *Eucalyptus dunnii* based on genetic diversity criteria calculated using molecular markers. *Tree Physiology* 23, 625–632.
- Mares, M., Enders, F., Kingsolver, J., Neff, J. y Simpson, B.** (1977). *Prosopis* as a niche component. En: Simpson B. (Ed.), *Mesquite: Its Biology in Two Desert Scrub Ecosystems*. Hutchinson & Ross Inc., Dowden, Pensilvania, pp. 123–149.
- Maruyama, T.** (1971). Analysis of population structure: II. Two-dimensional stepping stone models of finite length and other geographically structured populations. *Annals of Human Genetics* 35, 179–196.
- Maruyama, T.** (1977). *Stochastic problems in population genetics*. Springer Verlag, Berlin.
- Maruyama, T. y Fuerst, P.** (1985). Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics* 111, 675–689.
- Mather, P. y Koch, M.** (2011). *Computer processing of remotely-sensed images: an introduction*. Cuarta edición Wiley-Blackwell, Chichester, Reino Unido, pp. 504.
- Matteucci, S. y Camino, M.** (2012). Protected Areas Isolation in the Chaco Region, Argentina. *Journal of Geography and Geology*, 4, 15.
- Mayr, E.** (1942). *Systematics and the origin of species*. Dover Press, Nueva York.
- Medina, C. y Cardemil, L.** (1993). *Prosopis chilensis* is a plant highly tolerant to heat shock. *Plant, Cell & Environment* 16, 305–310.
- Meloni, D. y Martínez, C.** (2009). Glycinebetaine improves salt tolerance in vinal (*Prosopis ruscifolia* Griesbach) seedlings. *Brazilian Journal of Plant Physiology* 21, 233–241.
- Meloni, D., Gulotta, M. y Martínez, C.** (2008a). *Prosopis ruscifolia* Griseb.(vinal) tolera concentraciones salinas equivalentes al agua de mar y excluye iones tóxicos de la parte aérea. *Quebracho* 16, 32–40.

- Meloni, D., Gulotta, M. y Martinez, C.** (2008b). Salinity tolerance in *Schinopsis* quebracho colorado: Seed germination, growth, ion relations and metabolic responses. *Journal of Arid Environments* 72, 1785–1792.
- Miah, G., Rafii, M., Ismail, M., Puteh, A., Rahim, H., Islam, K. y Latif, M.** (2013). A review of microsatellite markers and their applications in rice breeding programs to improve blast disease resistance. *International Journal of Molecular Sciences* 14, 22499–22528.
- Michalakis, Y. y Excoffier, L.** (1996). A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* 142, 1061–1064.
- Mitra, A., Yadav, B., Ganai, N. y Balakrishnan, C.** (1999). Molecular markers and their applications in livestock improvement. *Current Science Bangalore* 77, 1045–1053.
- Monte-Corvo, L., Goulão, L. y Oliveira, C.** (2001). ISSR analysis of cultivars of pear and suitability of molecular markers for clone discrimination. *Journal of the American Society for Horticultural Science* 126, 517–522.
- Mooney, H., Simpson, B. y Solbrig, O.** (1977). Phenology, morphology, physiology. En: Simpson B. (Ed.), *Mesquite: Its Biology in Two Desert Scrub Ecosystems*. Hutchinson & Ross Inc., Dowden, Pensilvania, pp. 26–45.
- Moran, P.** (1950). Notes on continuous stochastic phenomena. *Biometrika* 37, 17–23.
- Morton, N.** (1973a). Isolation by distance. En: Morton, N. (Ed.), *Genetic Structure of Populations*. University of Hawaii Press, Honolulu, Hawaii, pp. 76–77.
- Morton, N.** (1973b). Kinship and population structure. En: Morton, N. (Ed.), *Genetic Structure of Populations*. University of Hawaii Press, Honolulu, Hawaii, pp. 66–69.
- Morton, N., Miki, C. y Yee, S.** (1968). Bioassay of population structure under isolation by distance. *American Journal of Human Genetics* 20, 411.
- Moscattelli, G.** (1990). Atlas de Suelos de la República Argentina. Escala 1:500000 y 1:1000000. Tomo I y II. Secretaría de Agricultura, Ganadería y Pesca. INTA. Centro de Investigaciones de Recursos Naturales, Buenos Aires.
- Moscuzza, C., Pérez Carrera, A. y Fernández Cirelli, A.** (2003). Relación entre las actividades agropecuarias y la escasez de agua en la provincia de Santiago del Estero (Argentina). *El Agua En Iberoamérica. Aspectos de la Problemática de las Tierras Secas* 35–54.
- Mottura, M., Finkeldey, R., Verga, A. y Gailing, O.** (2005). Development and characterization of microsatellite markers for *Prosopis chilensis* and *Prosopis flexuosa* and cross-species amplification. *Molecular Ecology Notes* 5, 487–489.
- NASA** (2015). Landsat 9 is coming. Disponible en: <https://www.nasa.gov/press/2015/april/nasa-usgs-begin-work-on-landsat-9-to-continue-land-imaging-legacy/> [accedido: febrero de 2015].
- Nagylaki, T.** (1976). The decay of genetic variability in geographically structured populations. II. Theoretical Population Biology 10, 70–82.
- Naranjo, C., Poggio, L. y Zeiger, S.** (1984). Phenol chromatography, morphology and cytogenetics in three species and natural hybrids of *Prosopis* (Leguminosae-Mimosoideae). *Plant Systematics and Evolution* 144, 257–276.
- Nei, M. y Chesser, R.** (1983). Estimation of fixation indices and gene diversities. *Annals of Human Genetics* 47, 253–259.
- Nilsen, E., Virginia, R. y Jarrell, W.** (1986). Water relations and growth characteristics of *Prosopis glandulosa* var. Torreyana in a simulated phreatophytic environment. *American Journal of Botany* 73, 427–433.
- Obenschain, A., Williams, D. y Andary, J.** (1998). Landsat 7: extending the Landsat tradition into the next century. *Space Technology*, 18, 3–10.

- Oden, N. y Sokal, R. (1986). Directional autocorrelation: an extension of spatial correlograms to two dimensions. *Systematic Biology* 35, 608–617.
- Ohta, T. y Kimura, M. (1973). A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetical Research* 22, 201–204.
- Oliveira, E., Pádua, J., Zucchi, M., Vencovsky, R. y Vieira, M. (2006). Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology*, 29, 294–307.
- Ord, J. y Getis, A. (1995). Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis* 27, 286–306.
- Paccard, A., Vance, M. y Willi, Y. (2013). Weak impact of fine-scale landscape heterogeneity on evolutionary potential in *Arabidopsis lyrata*. *Journal of Evolutionary Biology* 26, 2331–2340.
- Palacios, R. y Bravo, L. (1981). Hibridación natural en *Prosopis* (Leguminosae) en la región chaqueña argentina. Evidencias morfológicas y cromatográficas. *Darwiniana* 23, 3–35.
- Palacios, R. y Bravo, L. (1984). On some Argentine species of *Prosopis* L. *Bulletin of the International Group for the Study of Mimosoideae* 12, 103–112.
- Pardini, E. y Hamrick, J. (2008). Inferring recruitment history from spatial genetic structure within populations of the colonizing tree *Albizia julibrissin* (Fabaceae). *Molecular Ecology* 17, 2865–2879.
- Parker, K., Hamrick, J., Parker, A. y Nason, J. (2001). Fine-scale genetic structure in *Pinus clausa* (Pinaceae) populations: effects of disturbance history. *Heredity* 87, 99–113.
- Pasiecznik N., Felker, P., Harris, P., Harsh, L., Cruz, G., Tewari, J., Cadoret, K. y Maldonado L. (2001). The *Prosopis Juliflora-Prosopis Pallida* Complex: A Monograph. Henry Doubleday Research Association, Coventry, Reino Unido.
- De Pasquale, F., Siragusa, M., Abbate, L., Tusa, N., De Pasquale, C. y Alonzo, G. (2006). Characterization of five sour orange clones through molecular markers and leaf essential oils analysis. *Scientia Horticulturae* 109, 54–59.
- Pérez-Carrera, A., Moscuza, C. y Fernández-Cirelli, A. (2008). Efectos socioeconómicos y ambientales de la expansión agropecuaria. Estudio de caso: Santiago del Estero, Argentina. *Revista Ecosistemas* 17, 5-15.
- Pinheiro, J., Bates, D., DebRoy, S. y Sarkar, D. (2007). Linear and nonlinear mixed effects models. R package version 3.1-57.
- Piry, S., Luikart, G. y Cornuet, J-P. (1999). BOTTLENECK: a program for detecting recent effective population size reductions from allele data frequencies. INRA, Montpellier, France.
- Pizzo, B., Pometti, C., Charpentier, J-P., Boizot, N. y Saidman, B. (2011). Relationships involving several types of extractives of five native Argentine wood species of genera *Prosopis* and *Acacia*. *Industrial Crops and Products* 34, 851–859.
- Plant, R. (2012). Spatial Data Analysis in Ecology and Agriculture Using R. CRC Press, Boca Raton, Florida.
- Pometti, C., Pizzo, B., Brunetti, M., Macchioni, N., Ewens, M. y Saidman, B. (2009). Argentinean native wood species: physical and mechanical characterization of some *Prosopis* species and *Acacia aroma* (Leguminosae; Mimosoideae). *Bioresource Technology* 100, 1999–2004.
- Pometti, C., Palanti, S., Pizzo, B., Charpentier, J-P., Boizot, N., Resio, C. y Saidman, B. (2010). Durability of five native Argentine wood species of the genera *Prosopis* and *Acacia* decayed by rot fungi and its relationship with extractive content. *Biodegradation* 21, 753–760.
- Pomponio, M., Acuña, C., Petreath, V., Lauenstein, D., Poltri, S. y Torales, S. (2015). Characterization of functional SSR markers in *Prosopis alba* and their transferability across *Prosopis* species. *Forest Systems* 24, 04.

- Propastin, P. y Kappas, M.** (2008). Reducing uncertainty in modeling the NDVI-precipitation relationship: a comparative study using global and local regression techniques. *GIScience & Remote Sensing*, 45, 47–67.
- Premoli, A. y Kitzberger, T.** (2005). Regeneration mode affects spatial genetic structure of *Nothofagus dombeyi* forests. *Molecular Ecology* 14, 2319–2329.
- Pritchard, J., Stephens, M. y Donnelly, P.** (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Qadir, M., Tubeileh, A., Akhtar, J., Larbi, A., Minhas, P. y Khan, M.** (2008). Productivity enhancement of salt-affected environments through crop diversification. *Land Degradation and Development* 19, 429.
- Queller, D. y Goodnight, K.** (1989). Estimating relatedness using genetic markers. *Evolution* 43, 258–275.
- Rains, D. y Choukr-Allah, R.** (1991). Salinity and alkalinity as an issue in world agriculture. En: Choukr Allah R. (Ed), *Plant salinity research. New challenges*. Institut Agronomique et Veterinaire Hassan II, Agadir, Marruecos.
- Ramachandra, T., Aithal, B. y Sanna, D.** (2012). Insights to urban dynamics through landscape spatial pattern analysis. *International Journal of Applied Earth Observation and Geoinformation* 18, 329–343.
- Rapoport, E., Ladio, A., Ghermandi, E. y Sanz, E.** (1998). Malezas comestibles. *Ciencia Hoy* 9, 30–43.
- Rapp, R. y Wendel, J.** (2005). Epigenetics and plant evolution. *New Phytologist* 168, 81–91.
- R Core Team** (2012). *R: A Language and Environment for Statistical Computing* (Vienna, Austria: R Foundation for Statistical Computing).
- R Core Team** (2014). *R: A Language and Environment for Statistical Computing* (Vienna, Austria: R Foundation for Statistical Computing).
- Reddy, M., Sarla, N. y Siddiq, E.** (2002). Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *Euphytica* 128, 9–17.
- Reich, R., Czaplewski, R. y Bechtold, W.** (1994). Spatial cross-correlation of undisturbed, natural shortleaf pine stands in northern Georgia. *Environmental and Ecological Statistics* 1, 201–217.
- Reynolds, H.** (1954). Some interrelations of the Merriam kangaroo rat to velvet mesquite. *Journal of Range Management* 7, 176–180.
- Reynolds, J., Weir, B. y Cockerham, C.** (1983). Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* 105, 767–779.
- Rhodes, D. y Felker, P.** (1988). Mass screening of *Prosopis* (mesquite) seedlings for growth at seawater salinity concentrations. *Forest Ecology and Management* 24, 169–176.
- Di Rienzo, A., Donnelly, P., Toomajian, C., Sisk, B., Hill, A., Petzl-Erler, M., Haines, G. y Barch, D.** (1998). Heterogeneity of microsatellite mutations within and between loci, and implications for human demographic histories. *Genetics* 148, 1269–1284.
- Ritland, K.** (1996). Estimators for pairwise relatedness and individual inbreeding coefficients. *Genetical Research* 67, 175–185.
- Roig, F.** (1993). Aportes a la etnobotánica del género *Prosopis*. En: IADIZA (Ed.), *Contribuciones Mendocinas a La Quinta Reunión de Regional para América Latina y el Caribe de La Red de Forestación Del CIID. Conservación y Mejoramiento de Especies del Género Prosopis*, Mendoza, Argetina. pp. 99–119.
- Roser, L.** (2010). Caracterización genética y morfológica de genotipos de *Prosopis alba* con vainas dulces y tolerancia a la salinidad. Tesis de Licenciatura, Universidad de Buenos Aires, Buenos Aires, Argentina.

- Roser, L., Ferreyra, L., Ewens, M., Vilardi, J. y Saidman, B.** (2014). Genetic and morphometric characterization of clones of *Prosopis alba*, Algarobia, selected for salt tolerance. *Tree Genetics & Genomes* 10, 383–397.
- Roser, L., Vilardi, J., Saidman, B. y Ferreyra, L.** (2015). EcoGenetics: Spatial Analysis of Phenotypic, Genotypic and Environmental Data. R package version 1.2.0-3. <https://cran.r-project.org/web/packages/EcoGenetics>.
- Rousset, F.** (1996). Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics* 142, 1357–1362.
- Rousset, F.** (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145, 1219–1228.
- Rousset, F.** (2002). Inbreeding and relatedness coefficients: what do they measure? *Heredity* 88, 371–380.
- Rousset, F. y Raymond, M.** (1995). Testing heterozygote excess and deficiency. *Genetics*, 140, 1413–1419.
- Sabel, C., Wilson, J., Kingham, S., Tisch, C. y Epton, M.** (2007). Spatial implications of covariate adjustment on patterns of risk: respiratory hospital admissions in Christchurch, New Zealand. *Social Science & Medicine* 65, 43–59.
- SAyDS** (2004). Atlas de los Bosques Nativos Argentinos. Secretaría de Ambiente y Desarrollo Sustentable de la Nación., Buenos Aires, pp. 243.
- Saidman, B.** (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.** (1986). Isoenzymatic studies of alcohol dehydrogenase and glutamate oxalacetate transaminase in four South American species of *Prosopis* and their natural hybrids. *Silvae Genetica* 35, 3-10.
- Saidman, B.** (1988). La electroforesis de isoenzimas para la medición de la variabilidad genética en especies de *Prosopis*. En: FAO - PIRB, (Ed.). *Prosopis en Argentina*. Documento preliminar elaborado para el I Taller Internacional sobre Recurso Genético y Conservación de Germoplasma en *Prosopis*. Facultad de Cs. Agrarias. U. N. Córdoba, Argentina. pp. 107-118.
- Saidman, B.** (1990). Isozyme studies on hybrid swarms of *Prosopis caldenia* and sympatric species. *Silvae Genetica* 39, 5–8.
- Saidman, B. y Vilardi, J.** (1987). Analysis of the genetic similarities among seven species of *Prosopis* (Leguminosae: Mimosoideae). *Theoretical and Applied Genetics* 75, 109–116.
- Saidman, B. y Vilardi, J.** (1993). Genetic variability and germplasm conservation in the genus *Prosopis*. Winrock-Oxford & IBH Publishing Co. PVT. Ltd., New Delhi, Bombay, Calcuta.
- Saidman, B., Vilardi, J., Montoya, S. y Poggio, L.** (1997). Genetic variability and ploidy level in species of *Prosopis* (Leguminosae, Mimosoideae). *Boletín de la Sociedad Argentina de Botánica* 32, 217–225.
- Saidman, B., Vilardi, J., Montoya, S., Dieguez, M., Hopp, H. y Puri, S.** (1998). Molecular markers: a tool for understanding the relationship among species of *Prosopis* (Leguminosae, Mimosoideae). En: Puri, S. (Ed.). *Tree Improvement: Applied Research and Technology Transfer*. Science Publishers Inc., New Hampshire, Estados Unidos. pp. 311–324.
- Saracco, J., Collazo, J. y Groom, M.** (2004). How do frugivores track resources? Insights from spatial analyses of bird foraging in a tropical forest. *Oecologia* 139, 235–245.
- Schnabel, A. y Hamrick, J.** (1995). Understanding the population genetic structure of *Gleditsia triacanthos* L.: the scale and pattern of pollen gene flow. *Evolution* 49, 921–931.
- SMIN** (2015). Provincia de Córdoba- unidades cartográficas de los suelos. Secretaría de Minería de la Nación. <http://www.mineria.gob.ar/estudios/irn/cordoba/X-6suec.asp>. [Accedido: febrero de 2015].

- Simpson, G.** (1944). Tempo and mode in evolution. Columbia University Press, Nueva York.
- Simpson, B.** (1977). Mesquite: its biology in two desert ecosystems. Dowden, Hutchinson & Ross Inc., Dowden, Pensilvania, Estados Unidos. pp. 250.
- Slatkin, M.** (1977). Gene flow and genetic drift in a species subject to frequent local extinctions. *Theoretical population biology* 12, 253–262.
- Slatkin, M.** (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* 47, 264–279.
- Slatkin, M.** (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139, 457–462.
- Smouse, P. y Peakall, R.** (1999). Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* 82, 561–573.
- Smouse, P., Long, J. y Sokal, R.** (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology* 35, 627–632.
- Soil Survey Staff** (1999). Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys. Natural Resources Conservation Service. U.S. Department of Agriculture Handbook 436.
- Sokal, R.** (1979). Ecological parameters inferred from spatial correlograms. En: Patil, G., y Rosenzweig, M. (Ed.). *Contemporary quantitative ecology and related ecometrics*. International Co-operative Publishing House, Fairland, Maryland, pp. 167–196.
- Sokal, R. y Menozzi, P.** (1982). Spatial autocorrelations of HLA frequencies in Europe support demic diffusion of early farmers. *American Naturalist* 119, 1–17.
- Sokal, R. y Oden, N.** (1978a). Spatial autocorrelation in biology: 1. Methodology. *Biological Journal of the Linnean Society* 10, 199–228.
- Sokal, R. y Oden, N.** (1978b). Spatial autocorrelation in biology. 2. Some biological implications and four applications of evolutionary and ecological interest. *Biological Journal of the Linnean Society* 10, 229–249.
- Sokal, R. y Rholf F.** (1981). *Biometry*. WH Freeman and Company. San Francisco, California.
- Sokal, R. y Wartenberg, D.** (1983). A test of spatial autocorrelation analysis using an isolation-by-distance model. *Genetics* 105, 219–237.
- Sokal, R., Smouse, P. y Neel, J.** (1986). The genetic structure of a tribal population, the Yanomama Indians. XV. Patterns inferred by autocorrelation analysis. *Genetics* 114, 259–287.
- Sokal, R., Oden, N. y Thomson, B.** (1997). A simulation study of microevolutionary inferences by spatial autocorrelation analysis. *Biological Journal of the Linnean Society* 60, 73–93.
- Sokal, R., Oden, N. y Thomson, B.** (1998). Local spatial autocorrelation in a biological model. *Geographical Analysis* 30, 331–354. Streiff
- Solbrig, O. y Bawa, K.** (1975). Isozyme variation in species of *Prosopis* (Leguminosae). *Journal of the Arnold Arboretum* 56, 398–412.
- Solbrig, O. y Cantino, P.** (1975). Reproductive adaptations in *Prosopis* (Leguminosae, Mimosoideae). *Journal of the Arnold Arboretum* 56, 185–210.
- Solbrig O., Bawa K., Carmen N., Hunziker J., Naranjo C., Palacios R., Poggio L. y Simpson B.** (1977). Patterns of variation. En: Simpson B. (Ed.). *Mesquite. Its Biology in Two Desert Scrub Ecosystems*. Hutchinson & Ross Inc., Dowden, Pensilvania, Estados Unidos. pp. 44–60.
- Song, C., Woodcock, C., Seto, K., Lenney, M. y Macomber, S.** (2001). Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote Sensing of Environment* 75, 230–244.
- Steibel, P. y Troiani, H.** (1999). El género *Prosopis* (Leguminosae) en la provincia de La Pampa. *Revista de la Facultad de Agronomía Universidad Nacional de La Pampa* 10, 25–48

- Storfer, A., Murphy, M., Evans, J., Goldberg, C., Robinson, S., Spear, S., Dezzani, R., Delmelle, E., Vierling, L. y Waits, L.** (2007). Putting the landscape in landscape genetics. *Heredity* 98, 128–142.
- Streiff, R., Labbe, T., Bacilieri, R., Steinkellner, H., Gloessl, J. y Kremer, A.** (1998). Within-population genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. *Molecular Ecology* 7, 317–328.
- Sturges, H.** (1926). The choice of a class interval. *Journal of the American Statistical Association* 21, 65–66.
- Taleisnik, E. y López Launestein, D.** (2011). Leñosas perennes para ambientes afectados por salinidad: Una sinopsis de la contribución argentina a este tema. *Ecología Austral* 21, 3–14.
- Teich, I., Verga, A. y Balzarini, M.** (2014). Assessing spatial genetic structure from molecular marker data via principal component analyses: A case study in a *Prosopis* sp. forest. *Advances in Bioscience and Biotechnology* 5, 89–99.
- Therneau, T., Atkinson, B. y Ripley, B.** (2011). rpart: Recursive Partitioning and Regression Trees. R package version 3.1-50.
- Thompson, E.** (1975). The estimation of pairwise relationships. *Annals of Human Genetics* 39, 173–188.
- Tortorelli, L.** (2009). Maderas y bosques argentinos. Orientación Gráfica Editora, Buenos Aires, Argentina.
- Trenchard, L., Harris, P., Smith, S. y Pasiecznik, N.** (2008). A review of ploidy in the genus *Prosopis* (Leguminosae). *Botanical Journal of the Linnean Society* 156, 425–438.
- Tripathi, R. y Khan, M.** (1990). Effects of seed weight and microsite characteristics on germination and seedling fitness in two species of *Quercus* in a subtropical wet hill forest. *Oikos* 57, 289–296.
- Truntzer, C., Mercier, C., Estève, J., Gautier, C. y Roy, P.** (2007). Importance of data structure in comparing two dimension reduction methods for classification of microarray gene expression data. *BMC Bioinformatics* 8, 90.
- Tucker, C.** (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* 8, 127–150.
- Tucker, C., Townshend, J. y Goff, T.** (1985). African land-cover classification using satellite data. *Science* 227, 369–375.
- UESRRD** (2015). Unidad Ejecutora del Servicio de Riego del Río Dulce. Disponible en: <http://www.riegoriordulce.gov.ar/> [accedido: febrero de 2015].
- UPOV** (1995). Working group on biochemical and molecular techniques and DNA-profiling in particular. Tercera sesión. Identification methods based on molecular techniques. Wageningen, Países Bajos.
- Van Loon, E., Cleary, D. y Fauvelot, C.** (2007). ARES: software to compare allelic richness between uneven samples. *Molecular Ecology Notes* 7, 579–582.
- Van Rossum, F. y Triest, L.** (2006). Fine-scale genetic structure of the common *Primula elatior* (Primulaceae) at an early stage of population fragmentation. *American Journal of Botany* 93, 1281–1288.
- Varshney, R., Graner, A. y Sorrells, M.** (2005). Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology* 23, 48–55.
- Vekemans, X. y Hardy, O.** (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology* 13, 921–935.
- Velarde, M., Felker, P. y Degano, C.** (2003). Evaluation of Argentine and Peruvian *Prosopis* germplasm for growth at seawater salinities. *Journal of Arid Environments* 55, 515–531.

- Verga, A.** (1995). Genetische untersuchungen an *Prosopis chilensis* und *Prosopis flexuosa* (Mimosaceae) im trockenen Chaco Argentinien. Göttingen Forstgenetische Berichte 19.
- Vilela, A.** (1993). Leaf anatomy of *Prosopis* (Leguminosae-Mimosoideae): Adaptive strategies to different environments in *Prosopis nigra*. Darwiniana 32, 99-107.
- Vilela, A. y Palacios, R.** (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.
- Villagra, P.** (1998). Comparación del comportamiento fitosociológico y ecofisiológico de *Prosopis argentina* y *P. alpataco* (Fabaceae, Mimosoideae). Tesis doctoral, Universidad Nacional de Cuyo, Mendoza, Argentina.
- Villagra, P. y Cavagnaro, J.** (2006). Water stress effects on the seedling growth of *Prosopis argentina* and *Prosopis alpataco*. Journal of Arid Environments 64, 390-400.
- Villagra, P., Giordano, C., Alvarez, J., Bruno Cavagnaro, J., Guevara, A., Sartor, C., Passera, C. y Greco, S.** (2011). Ser planta en el desierto: estrategias de uso de agua y resistencia al estrés hídrico en el Monte Central de Argentina. Ecología Austral 21, 29-42.
- Wahlund, S.** (1928). The combination of populations and the appearance of correlation examined from the standpoint of the study of heredity. Hereditas 11, 65-106.
- Wang, J.** (1997a). Effective size and F-statistics of subdivided populations. II. Dioecious species. Genetics 146, 1465-1474.
- Wang, J.** (1997b). Effective size and F-statistics of subdivided populations. I. Monoecious species with partial selfing. Genetics 146, 1453-1463.
- Wartenberg, D.** (1985). Multivariate spatial correlation: a method for exploratory geographical analysis. Geographical Analysis 17, 263-283.
- Washington-Allen, R., Ramsey, R., West, N. y Norton, B.** (2008). Quantification of the ecological resilience of drylands using digital remote sensing. Ecology and Society 13, 33.
- Weir, B. y Cockerham, C.** (1984). Estimating F-statistics for the analysis of population structure. Evolution 38, 1358-1370.
- Weir, B. y Hill, W.** (2002). Estimating F-statistics. Annual Review of Genetics 36, 721-750.
- Weir, B., Anderson, A. y Hepler, A.** (2006). Genetic relatedness analysis: modern data and new challenges. Nature Reviews Genetics 7, 771-780.
- Westoby, M.** (2006). Phylogenetic ecology at world scale, a new fusion between ecology and evolution. Ecology 87, S163-S165.
- Wiens, J.** (1989). Spatial scaling in ecology. Functional Ecology 3, 385-397.
- Wijsman, E. y Cavalli-Sforza, L.** (1984). Migration and genetic population structure with special reference to humans. Annual Review of Ecology and Systematics 15, 279-301.
- Williams, D., Goward, S. y Arvidson, T.** (2006). Landsat: Yesterday, Today, and Tomorrow. Photogrammetric Engineering & Remote Sensing 72, 1171-1178.
- Wilson, E. y Sader, S.** (2002). Detection of forest harvest type using multiple dates of Landsat TM imagery. Remote Sensing of Environment 80, 385-396.
- Wojtusik, T., Felker, P., Russell, E. y Benge, M.** (1993). Cloning of erect, thornless, non-browsed nitrogen fixing trees of Haiti's principal fuelwood species (*Prosopis juliflora*). Agroforestry Systems 21, 293-300.
- Wolfe, A.** (2000). ISSR Resource Website. Disponible en: <http://www.biosci.ohiostate.edu/~awolfe/ISSR/ISSR.html> [accedido: febrero de 2015].
- Wright, S.** (1921). Correlation and causation. Journal of Agricultural Research 20, 557-585.

- Wright, S.** (1940). Breeding structure of populations in relation to speciation. *The American Naturalist* 74, 232–248.
- Wright, S.** (1943). Isolation by distance. *Genetics* 28, 114.
- Wright, S.** (1946). Isolation by distance under diverse systems of mating. *Genetics* 31, 39.
- Wright, S.** (1951). The genetical structure of populations. *Annals of Eugenics* 15, 323–354.
- Wright, S.** (1952). The theoretical variance within and among subdivisions of a population that is in a steady state. *Genetics* 37, 312.
- Wright, S.** (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19, 395–420.
- Yamagishi, H., Tomimatsu, H. y Ohara, M.** (2007). Fine-scale spatial genetic structure within continuous and fragmented populations of *Trillium camschatcense*. *Journal of Heredity* 98, 367–372.
- Yip, P., Chau, C., Mak, C. y Kwan, H.** (2007). DNA methods for identification of Chinese medicinal materials. *Chinese Medicine* 2, 9.
- Zallochi E. y Palacios R.** (1990). Interpopulation variation in *Prosopis flexuosa* DC. from northern Chile. *Bulletin of the International Group for the Study of Mimosoideae* 18, 135–149.
- Zeng, X., Michalski, S., Fischer, M. y Durka, W.** (2012). Species diversity and population density affect genetic structure and gene dispersal in a subtropical understory shrub. *Journal of Plant Ecology* 5, 270–278.
- Zerda, H. y Tiedemann, J.** (2010). Dinámica temporal del NDVI del bosque y pastizal natural en el Chaco de la Provincia de Santiago del Estero, Argentina The temporal dynamic of NDVI, of forest and grassland in the Chaco Seco of Santiago del Estero province, Argentine. *Ambiência* 6, 13–24.
- Zhang, L. y Shi, H.** (2004). Local modeling of tree growth by geographically weighted regression. *Forest Science* 50, 225–244.
- Zietkiewicz, E., Rafalski, A. y Labuda, D.** (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20, 176–183.