

ASSOCIATION BETWEEN GENETIC AND PHENOTYPIC VARIABILITY WITH ADJUSTMENT FOR SPATIAL AUTOCORRELATION IN *Prosopis*

ASOCIACIÓN ENTRE VARIABILIDAD GENÉTICA Y FENOTÍPICA CON AJUSTE POR AUTOCORRELACIÓN ESPACIAL EN *Prosopis*

Teich I.^{1*}, Mottura M.², Verga A.², Balzarini M.¹

¹ Estadística y Biometría, Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba. CONICET.

² Instituto de Fisiología y Recursos Genéticos Vegetales (IFRGV). Centro de Investigaciones Agropecuarias (CIAP). Instituto Nacional de Tecnología Agropecuaria (INTA).

* Author for correspondence
ingridteich@gmail.com

ABSTRACT

Association studies that relate phenotypic and genetic information allow the detection of relevant interactions. In these studies it is necessary to consider the effects of spatial structure in the data, which is very common in nature, particularly in plant species. *Prosopis chilensis* and *P. flexuosa* are two hardwood arboreal species with environmental and economic importance. Significant gene flow occurs between them with the production of natural and fertile hybrids. Given that morphological characters are used to classify individual trees, it is of interest to explore the association between the phenotypic variables used in the taxonomic classification and genetic variability. In this work, multivariate associations between genetic and morphological variability were explored in a *P. chilensis* and *P. flexuosa* hybrid swarm considering the possible spatial autocorrelation. For this purpose the molecular (six polymorphic microsatellite markers) and the morphological (13 traits) characterizations of 81 adult trees in a 4,700 m² plot in Córdoba, Argentina were used. The spatial structures of genetic and phenotypic variability were analyzed at a fine spatial scale. Both characterizations were associated through a Mantel approach based on the correlation of distance matrixes, as well as with new approaches that include the complementary use of multivariate analysis and mixed linear models. A significant genotype-phenotype association, relevant for the management of *Prosopis* genetic resources was detected. Even with a weak spatial structure, the inclusion of the spatial autocorrelation in the analyses improved the detection and modeling of the association between the observed genetic and phenotypic variability.

Key words: Spatial Statistics, Moran Index, Partial Mantel Test, Canonical Correlations, Generalized Procrustean Analysis

RESUMEN

Los estudios de asociación entre datos genéticos y fenotípicos permiten detectar interacciones relevantes. En estos análisis es necesario considerar los efectos de la estructura espacial de los datos, la cual es muy común en la naturaleza, particularmente en plantas. *Prosopis chilensis* y *P. flexuosa* son dos especies arbóreas de importancia económica y ambiental entre las cuales se produce flujo génico y se originan híbridos fértiles. Para la clasificación de estas especies se utilizan caracteres morfológicos, siendo de interés conocer y cuantificar la asociación entre la variabilidad fenotípica caracterizada mediante las variables usadas en la clasificación taxonómica, y la variabilidad genética. En este trabajo se exploran asociaciones multivariadas entre la variabilidad genética y morfológica de un enjambre híbrido entre *P. chilensis* y *P. flexuosa*, considerando la posible autocorrelación espacial en los datos. Para esto se utilizó la caracterización genética (seis marcadores microsatélites) y morfológica (13 caracteres) en 81 individuos adultos ubicados en una parcela de 4.700 m² en Córdoba, Argentina. Se analizó la estructura espacial de la variabilidad genética y morfológica a escala fina y se asociaron ambas caracterizaciones mediante una aproximación de Mantel basada en la correlación de matrices de distancias y nuevas aproximaciones que incluyen el uso complementario de análisis multivariados y modelos lineales mixtos. Se encontró significativa asociación genotipo-fenotipo, información relevante para el ordenamiento del recurso genético algarrobo. Aún con la débil estructura espacial en los datos, la incorporación de la autocorrelación espacial mejoró la capacidad de detección y modelación de la asociación entre la variabilidad genética y fenotípica observada en el enjambre híbrido.

Palabras clave: Estadística Espacial, Índice de Moran, Correlación Parcial de Mantel, Análisis de Correlaciones Canónicas, Análisis de Procrustes Generalizado.

Fecha de recepción: 16/07/2015

Fecha de aceptación de versión final: 22/09/2015

INTRODUCCIÓN

Los estudios de asociación entre datos obtenidos a nivel genético y provenientes de otros niveles de organización permiten detectar interacciones relevantes de los genotipos. El estudio de la asociación de información fenotípica y genotípica se realiza frecuentemente con procedimientos univariados (Brachi *et al.*, 2010; Cockram *et al.*, 2010), considerando cada característica de interés separadamente. Sin embargo, la reemergencia de aproximaciones holísticas y multivariadas en los últimos años ha generado un cambio de perspectiva en la biología (Carrari *et al.*, 2006), probablemente facilitado por importantes avances en el desarrollo de tecnologías e instrumental para la caracterización multidimensional de los distintos niveles de organización. Así, las ciencias “ómicas” como la fenómica y la genómica, tienen en común la búsqueda del estudio integral de los seres vivos, para lo cual las asociaciones de distintos tipos de datos es crucial (Houle *et al.*, 2010). Las técnicas estadísticas del análisis multivariado permiten abordar los estudios de asociación multidimensional entre diferentes tipos de datos (Jombart *et al.*, 2009; Balzarini *et al.*, 2011; Teich *et al.*, 2011). Sin embargo, aún cuando las técnicas multivariadas clásicas permiten operar sobre variables correlacionadas, suponen independencia (no correlación) entre las unidades de análisis. Una cuestión fundamental que debe considerarse en cualquier estudio de asociación es la existencia de estructuras de correlación en los datos. Cuando se lleva a cabo un análisis de asociación sin considerar los efectos de dependencia de los datos, se aumenta el riesgo de detectar asociaciones espurias tanto como el de no detectar asociaciones (Keitt *et al.*, 2002). En la naturaleza es frecuente que la variabilidad genética presente estructuras espaciales, particularmente en especies de plantas (Vekemans y Hardy, 2004). Los datos fenotípicos también pueden estar estructurados espacialmente.

Dos procedimientos multivariados diseñados para cuantificar asociaciones no restringidas espacialmente son el análisis de correlaciones canónicas (Hotelling, 1933) y el análisis denominado Procrustes Generalizado (Gower, 1975). A pesar de que no se trata de técnicas de desarrollo reciente, su aplicación en el contexto de datos biológicos masivos sí lo es (Bramardi *et al.*, 2005). No obstante, estos procedimientos de análisis de asociaciones multivariadas se basan en modelos lineales generales que suponen datos independientes, por lo que podrían producir resultados errados bajo autocorrelación espacial.

Entre las aproximaciones para estudiar asociaciones multivariadas con datos estructurados espacialmente, se encuentra la prueba de Mantel Parcial (Smouse *et al.*, 1986). El estadístico de la prueba Mantel Parcial correlaciona la información que existe en dos matrices controlando el sesgo que podrían introducir variaciones de otro conjunto de variables (tercera matriz). Por ejemplo, puede ser usado para correlacionar matrices de diferencias de datos genéticos y datos fenotípicos controlando por una matriz de distancias espaciales entre las unidades de análisis. Algunos autores sugieren que cuando se trabaja con datos correlacionados, las pruebas de Mantel, tanto simple como parcial, podrían producir un exceso de tasa de error tipo I cuya magnitud se incrementa con la intensidad de la autocorrelación espacial (Raufaste y Rousset, 2001; Teich, 2012; Guillot y Rousset, 2013). Otra aproximación al análisis de datos correlacionados espacialmente que se encuentra cada vez más difundida, se basa en el uso de modelos lineales mixtos (MLM) (Littell *et al.*, 2006). Los MLM han permitido integrar conceptos relativos a la estructura de correlaciones de cualquier tipo de datos normales en un marco analítico flexible y unificado. Si bien podrían aplicarse a variables sintéticas de distribución continua, como puede ser una componente principal o una variable canónica, su aplicación en el estudio de asociaciones de tipo multivariado es prácticamente nula.

Las especies del género *Prosopis* (Mimosoideae) constituyen un recurso natural con características de interés debido a sus adaptaciones para crecer en suelos pobres, donde pocas especies pueden sobrevivir (Pasicznik *et al.*, 2001). En particular, *P. chilensis* y *P. flexuosa* son dos especies arbóreas de importancia económica y ambiental en nuestro país. La ocurrencia de flujo génico entre ambas y la producción de híbridos fértiles en ambientes naturales, aunque disturbados por el hombre, genera fenotipos intermedios que dificultan la determinación taxonómica de las especies y los híbridos (Saidman *et al.*, 2000; Mottura, 2006). Con este fin, Verga (2000) confeccionó una clave basada en taxonomía numérica de caracteres morfológicos de hoja y fruto. Asimismo, se desarrolló un conjunto de marcadores moleculares microsatélites (SSR) específico para estas especies (Mottura *et al.*, 2005). Dado que los caracteres morfológicos presentan alta variabilidad con el ambiente, es de interés conocer y cuantificar la asociación entre la variabilidad fenotípica caracterizada mediante las variables usadas en la clasificación, y la variabilidad genética en una escala espacial fina. Cuando el análisis conjunto de

la información genética y morfológica sugiere la existencia de un consenso entre ambas caracterizaciones, y de la existencia de unidades biológicas diferenciadas dentro del complejo de híbridos, los procesos microevolutivos que dieron origen a estas unidades pueden ser estudiados más eficientemente, otorgando valor tanto adaptativo como diagnóstico a los caracteres estudiados. El objetivo de este trabajo es explorar asociaciones multivariadas entre la variabilidad genética y morfológica en un enjambre híbrido de *Prosopis* sp., evaluando la posible estructuración espacial de la variabilidad.

MATERIALES Y MÉTODOS

Datos

Se utilizó una base de datos correspondiente a la caracterización genética y morfológica de un enjambre híbrido de *P. flexuosa* y *P. chilensis* ubicado en una parcela de 4.700 m² en el Parque Provincial y Reserva Forestal Chancaní, en el Oeste de la provincia de Córdoba, Argentina (Mottura, 2006). En dicho trabajo se georreferenciaron bidimensionalmente y caracterizaron 100 individuos adultos mediante 6 SSR y 20 caracteres morfológicos cuantitativos y cualitativos medidos en 10 hojas y 5 frutos por individuo. Para los fines del presente trabajo se utilizaron sólo aquellos individuos en los cuales los datos faltantes no superaran el 50 % de los marcadores moleculares, resultando en un total de 81 individuos. Asimismo se utilizaron sólo caracteres morfológicos cuantitativos y no redundantes, resultando en 13 variables morfológicas. Los individuos fueron clasificados en *P. flexuosa*, *P. chilensis* e híbridos en base a la clasificación propuesta por Verga (2000). Del total de individuos, 8 fueron clasificados como híbridos, 63 como *P. flexuosa* y 10 como *P. chilensis*. A continuación se resumen algunas características relevantes de la biología y ecología de las especies de estudio así como la descripción de la caracterización genética y morfológica.

Especies de estudio

Prosopis chilensis (Molina) Stuntz y *P. flexuosa* DC. pertenecen a la sección Algarrobia del género *Prosopis*. Ambas son diploides con un número cromosómico 2n=28. Morfológicamente ambas especies se diferencian principalmente por el tamaño de las pinnas y los foliólulos (más largos en *P. chilensis*), las espinas y los frutos. La polinización ocurre por medio de insectos, principalmente

abejas, y la dispersión de semillas es endozoócora. Ambas especies son nativas de América del Sur y en Argentina el rango de distribución natural de *P. chilensis* es más restringido que el de *P. flexuosa*, incluyendo las provincias fitogeográficas del Monte, el Espinal y el Chaco. Particularmente en la región del Chaco Árido ambas especies ocupan nichos ecológicos diferentes, debido a los mayores requerimientos de agua de *P. chilensis*. Sin embargo existen zonas de contacto entre ambas especies donde se registra la aparición de híbridos, particularmente en lugares disturbados por el hombre. La hibridación es común en la sección Algarrobia del género *Prosopis* (Hunziker *et al.*, 1986; Saidman, 1990; Vega y Hernández, 2005), en el cual existe una gran similitud genética entre las especies (Bessegga *et al.*, 2000; Saidman *et al.*, 2000). Estudios del sistema de apareamiento en el enjambre híbrido indican que el flujo génico entre los grupos es asimétrico y que la hibridación está acompañada por un proceso de introgresión (Mottura, 2006).

Caracteres morfológicos

Se utilizaron 13 caracteres morfológicos cuantitativos de hoja y fruto (Tabla 1), los cuales son utilizados para obtener grupos de individuos con características comunes en un grado de detalle mayor que el alcanzado mediante la sistemática clásica (Verga, 1995; 2000; Verga *et al.*, 2009). Si bien se seleccionaron aquellos caracteres menos redundantes existen altas correlaciones entre los mismos. Por ejemplo y como es de esperar el largo del foliólulo (LF) se encuentra altamente correlacionado con el área del foliólulo (AR), el área total de la hoja (AFOL), el largo de la pinna (LPI) y la relación largo-ancho del foliólulo. El único carácter que no se encuentra correlacionado con ningún otro rasgo es el falcado (FAL).

Caracteres moleculares

Los datos moleculares utilizados corresponden a la caracterización genética realizada mediante seis marcadores SSR (Mo05, Mo07, Mo08, Mo09, Mo13 y Mo16) desarrollados específicamente para *P. chilensis* (Mottura *et al.*, 2005). Dicha caracterización fue realizada a partir de ADN extraído de material foliar de cada árbol adulto (Mottura, 2006). Debido a que en muchos casos no hubo amplificación se seleccionaron aquellos individuos que fueron caracterizados con al menos cuatro de los seis marcadores. El número total de alelos sobre todos los individuos para los seis marcadores es 68 (Tabla 2).

El marcador más polimórfico es Mo16, con 15 alelos. El marcador Mo05 es el único que es más polimórfico en *P. chilensis* y en los híbridos que en *P. flexuosa*. El resto de los marcadores son muy polimórficos en *P. flexuosa* y poco polimórficos para *P. chilensis*, lo cual probablemente se deba a la diferencia en el número de individuos caracterizados en cada grupo. Sin embargo, existen alelos exclusivos tanto para *P. flexuosa* como para *P. chilensis*. Como es de esperar, en los híbridos no se detecta ningún alelo exclusivo. El marcador Mo07 tiene un alelo exclusivo de *P. chilensis* (209). Respecto a Mo08, todos los alelos se encuentran presentes en *P. flexuosa*, mientras que *P. chilensis* presenta 5 de los 10 alelos. El marcador Mo09 presenta cinco alelos muy poco frecuentes (203, 223, 227, 235 y 245) presentes sólo en *P. flexuosa* y con frecuencias menores a 0,01. En *P. chilensis* este marcador es muy poco polimórfico (dos alelos), siendo el alelo 209 muy frecuente en éste grupo. En la Tabla 2 puede observarse que la heterocigosidad observada para este marcador es la más alta ($H_o = 0,6$). Dos alelos del marcador Mo013 (244 y 246) y tres alelos de Mo16 (149, 17 y 173) no se registraron en *P. flexuosa*.

Análisis estadísticos

Para detectar la estructura espacial de los datos genéticos y morfológicos por separado se utilizaron dos aproximaciones multivariadas: la Prueba de Mantel (Mantel, 1967) y el cálculo del índice de autocorrelación espacial de Moran (Moran, 1950) de las primeras variables sintéticas obtenidas mediante Análisis de Componentes Principales. La prueba de Mantel es una de las herramientas estadísticas más utilizadas para evaluar la significancia estadística de la dependencia entre dos matrices de distancias. Para detectar estructura genética espacial (EGE) se evaluó la correlación entre una matriz de distancias genéticas y una matriz de distancias geográficas, mientras que para detectar estructura fenotípica espacial (EFE) se utilizaron las matrices de distancias morfológicas y de distancias geográficas. En todos los casos se utilizaron las distancias Euclídeas, calculadas sobre la matriz de frecuencias alélicas (distancias genéticas) sobre los datos morfológicos estandarizados (distancias morfológicas) y sobre las coordenadas geográficas (distancias geográficas). La evaluación de la significancia del estadístico de Mantel se realizó mediante valores-p obtenidos por permutación (9999 permutaciones). Los análisis se realizaron con las librerías “bioDist” (Ding *et al.*, 2008) y “vegan” (Oksanen *et al.*, 2011) en R (R Development Core Team, 2011). La

segunda aproximación metodológica (Novembre *et al.*, 2008; Teich *et al.*, 2014) consistió en realizar Análisis de Componentes Principales sobre las frecuencias alélicas y sobre las variables morfológicas y luego calcular la autocorrelación espacial de las primeras componentes obtenidas mediante el Índice de Moran utilizando la red de conexión de Gabriel (Gabriel y Sokal, 1969). La primera componente principal de los datos genéticos fue denotada por $CP1_g$ mientras que la primera componente principal de los datos morfológicos se denotó como $CP1_m$. Estos análisis se realizaron mediante las librerías de R “adeigenet” (Jombart, 2008) y “spdep” (Bivand *et al.*, 2011). A los fines de visualizar la estructura espacial, con las componentes principales cuya autocorrelación espacial resultó estadísticamente significativa, se realizó un mapa sintético de variabilidad mediante interpolación lineal de la variable sintética utilizando el inverso de la distancia de grado 2.

El análisis de asociación entre variabilidad genética y fenotípica se realizó con tres métodos multivariados: (1) Prueba de Mantel entre las matrices de distancias Euclídeas genéticas y morfológicas; (2) Análisis de Correlaciones Canónicas (ACC); y (3) Análisis de Procrustes Generalizado (APG). En las dos primeras aproximaciones la significancia fue evaluada tanto sin considerar la autocorrelación espacial como contemplando correlaciones espaciales. En el caso de la prueba de Mantel, la autocorrelación se contempló incorporando una matriz de distancias espaciales (Prueba de Mantel Parcial) (Smouse *et al.*, 1986). Para evaluar la significancia de la correlación entre un determinado par de variables canónicas se usaron dos MLM alternativos, y según el criterio de Akaike (AIC) se seleccionó el mejor modelo. Los MLM ajustados consideraron a la variable canónica construida con datos morfológicos como variable respuesta y, como variable regresora (dentro de los efectos fijos) a la variable canónica del mismo par, conformada a partir de datos genéticos. La diferencia entre ambos modelos se produjo al momento de especificar la estructura de la matriz de varianza y covarianza de los términos de error aleatorio. En el Modelo 1 la matriz error se supuso no correlacionada. En el Modelo 2 se especificó una matriz de tipo espacial, con decaimiento exponencial de las correlaciones entre datos de la variable canónica según se incrementa la distancia entre árboles. La raíz cuadrada del R^2 de este modelo se usó como el coeficiente de correlación canónica entre el par de variables canónicas analizadas y la significancia

de la asociación fue evaluada mediante el valor-p de la prueba de hipótesis de la pendiente del MLM ajustado. Los MLM se implementaron mediante la librería “lme4” (Bates *et al.*, 2015) en R a través de la interfaz que ofrece el programa estadístico Info-Gen (Balzarini y Di Rienzo, 2011). Además se estimó la “estructura canónica total”, *i.e.* las correlaciones simples entre las variables respuestas y las

variables canónicas. En el caso del APG, se cuantificó la asociación no necesariamente lineal entre datos genéticos y morfológicos del conjunto total de datos y de cada individuo en particular, de manera de luego comparar el porcentaje de consenso entre *P. chilensis*, *P. flexuosa* y los híbridos mediante un test de ANOVA y comparaciones a posteriori según el test de LSD.

Tabla 1. Caracteres morfológicos seleccionados y su abreviatura

NPI	Número de pares de pinnas
LPI	Largo de la pinna (mm)
NFO	Número de pares de foliólulos
DIFOL	Distancia entre foliólulos en la pinna (mm)
LF	Largo del foliólulo (mm)
L-AF	Relación Largo/Ancho del foliólulo
AR	Área del foliólulo (cm ²)
AFOL	Área total de la hoja (cm ²)
LPE	Largo del pecíolo
API	Relación entre el área del tercio superior del foliólulo respecto del rectángulo en que se inscribe la misma
G-AFR	Relación ancho/grosor de fruto
APT	Relación entre la superficie del tercio superior del foliólulo respecto del área total del foliólulo
FAL	Falcado (Relación entre la distancia desde el ápice hasta la base del foliólulo y la longitud de una línea que pasa por el centro del foliólulo)

Tabla 2. Polimorfismo y heterocigosidad de los seis marcadores moleculares en los 81 individuos del enjambre híbrido entre *P. flexuosa* y *P. chilensis*.

	Mo05	Mo07	Mo08	Mo09	Mo13	Mo16
Número de alelos (Na)	3	12	10	15	13	15
Número de alelos efectivos (Ne)	1,753	3,224	7,674	5,069	3,699	8,209
Heterocigosidad observada (Ho)	0,128	0,379	0,593	0,493	0,543	0,446
Heterocigosidad esperada (He)	0,430	0,690	0,870	0,803	0,730	0,878

RESULTADOS

Análisis de estructura espacial de la variabilidad genética y morfológica

No se detectó EGE en los individuos del enjambre híbrido de *Prosopis* sp. La Prueba de Mantel entre distancias genéticas y geográficas no resultó significativa ($r = -0,014$, $p = 0,6$) y ninguna de las dos primeras variables sintéticas ($CP1_g$ y $CP2_g$), que explican el 17 % y el 14 % de la variabilidad, presentó autocorrelación espacial estadísticamente significativa. Sin embargo, la primera componente principal de los datos morfológicos ($CP1_m$), que explica un 54 % de la variabilidad morfológica observada, presentó autocorrelación espacial estadísticamente significativa según el Índice de Moran ($I = 0,15$, $p = 0,02$). El coeficiente de correlación de la Prueba de Mantel fue de 0,062, siendo significativo con un nivel de significación del 10 % ($p = 0,1$). En la Tabla 3 se presentan los índices de autocorrelación de Moran y su significancia para cada una de las variables morfológicas originales y para la primera variable sintética morfológica. Entre las variables morfológicas con mayor autocorrelación espacial se encuentran aquellas que caracterizan el tamaño del foliólulo (LF, AR, API). La distribución espacial de la variabilidad morfológica se puede observar en el mapa sintético realizado con la $CP1_m$ (Figura 1) donde también se grafican los individuos según su posición geográfica.

Tabla 3. Autocorrelación espacial de las variables morfológicas y de las primeras componentes principales obtenidas mediante un ACP.

Variable morfológica	Índice de Moran	Valor p
NPI	0,06	0,20
LPI	0,12	0,06
NFO	0,03	0,31
DIFOL	0,15	0,03
LF	0,16	0,02
L-AF	0,05	0,22
AR	0,16	0,01
AFOL	0,07	0,15
LPE	-0,05	0,67
API	0,20	0,01
G-AFR	-0,02	0,55
APT	0,15	0,03
FAL	0,05	0,24
$CP1_m$	0,15	0,02

Asociación entre datos genéticos y fenotípicos

Cuando se evaluó la asociación entre la variabilidad genética y morfológica a través de la correlación entre matrices de distancias, la incorporación de la información espacial modificó la conclusión respecto a la significancia de la prueba de Mantel. En el caso en que no se incluyó la matriz de distancias geográficas (Prueba de Mantel Simple) la prueba de asociación arrojó un valor-p de 0,056. Sin embargo, al aplicar la Prueba de Mantel Parcial, que incluyó como matriz control a las distancias geográficas, la asociación entre las distancias morfológicas y las distancias genéticas resultó significativa ($p = 0,045$). El coeficiente de correlación fue de 0,103.

El Análisis de Correlaciones Canónicas también permitió concluir que la variabilidad genética se encuentra asociada a la variabilidad morfológica de los individuos del enjambre híbrido. Todas las correlaciones canónicas resultaron significativas tanto si su evaluación se realizó considerando la autocorrelación espacial en las variables canónicas morfológicas o no. En algunos casos la incorporación de la estructura espacial resultó en una mejora del ajuste logrado según el criterio de AIC. En la Tabla 4 se presentan los valores de los criterios de AIC usados para seleccionar el modelo más apropiado para evaluar la significancia de la correlación canónica con un MLM, en el

Tabla 4. Criterios Akaike (AIC) y prueba del cociente de verosimilitud (LRT) para modelos de regresión entre pares de variables canónicas (morfológicas y genéticas) con errores independientes y correlacionados espacialmente.

Par de variables canónicas	AIC	
	Errores Independientes	Errores correlacionados
1	-89,6	-89,6
2	-46,1	-44,7
3	-27,4	-28,0
4	5,2	7,1
5	63,6	65,6
6	79,4	80,8
7	116,9	118,9
8	129,5	131,5
9	138,0	138,0
10	171,8	173,8
11	193,7	191,7
12	194,6	196,6
13	218,3	220,3

contexto de posible autocorrelación de la variable respuesta. Todas las correlaciones canónicas resultaron significativas ($p < 0,0001$) y con coeficientes de correlación altos (Tabla 5). Estos resultados confirman una asociación estadísticamente significativa entre la diversidad genética y la morfológica en el enjambre híbrido. Las variables morfológicas que se correlacionan más fuertemente con el primer eje canónico (Eje 1_m), son el falcado (FAL), APT y API, las cuales caracterizan la forma del foliólulo, mientras que los alelos 240, 244 y 234 de Mo13 y el alelo 151 de Mo16 son los más correlacionados con el primer eje canónico (Tabla 6). Estas son las variables de mayor contribución en la correlación de la variabilidad genética y fenotípica. En el segundo par de ejes

canónicos las variables morfológicas más correlacionadas con la variable sintética son el área y el largo del foliólulo (AR y LF), así como API, mientras que el alelo 151 de Mo16 y el 226 de Mo13 son los más correlacionados con el segundo eje.

El consenso entre la caracterización molecular y fenotípica resultó del 67 %. El primer eje del APG resume un 58 % de la variabilidad del consenso y el segundo eje explica un 6 % (Figura 2). *P. flexuosa* fue el grupo que mostró la menor proporción de consenso entre la caracterización genética y morfológica (0,63) mientras que los híbridos y *P. chilensis* no se diferenciaron entre sí respecto a la proporción de consenso, con porcentajes de consenso entre ambos tipos de datos mayores al 70 % (Figura 3).

Tabla 5. Coeficientes de correlación canónica y significancia (valor p).

Correlación Canónica	Coefficiente de correlación	Valor p'
1	0,992	0,004
2	0,986	0,004
3	0,985	0,004
4	0,974	0,004
5	0,945	0,004
6	0,932	0,004
7	0,889	0,004
8	0,867	0,004
9	0,849	0,004
10	0,762	0,004
11	0,678	0,004
12	0,663	0,004
13	0,490	0,004

Tabla 6. Coeficientes de Pearson para la correlación entre el primer par de ejes canónicos y los caracteres morfológicos (A) y los 10 alelos más correlacionados (B).

(A)	Eje 1 _m
FAL	60,31
APT	49,0
AR	6,79
L_AF	0,95
NFO	0,26
DIFOL	0,23
LPE	0,06
AFOL	0,05
LPI	-0,07
G_AFR	-0,63
NPI	-1,22
LF	-5,4
API	-25,31
(B)	Eje 1 _g
Mo13.240	16,73
Mo16.151	14,07
Mo13.244	12,48
Mo13.234	10,56
Mo09.223	9,89
Mo07.197	9,22
Mo09.213	8,28
Mo13.206	7,72
Mo13.218	7,43
Mo07.209	-7,8

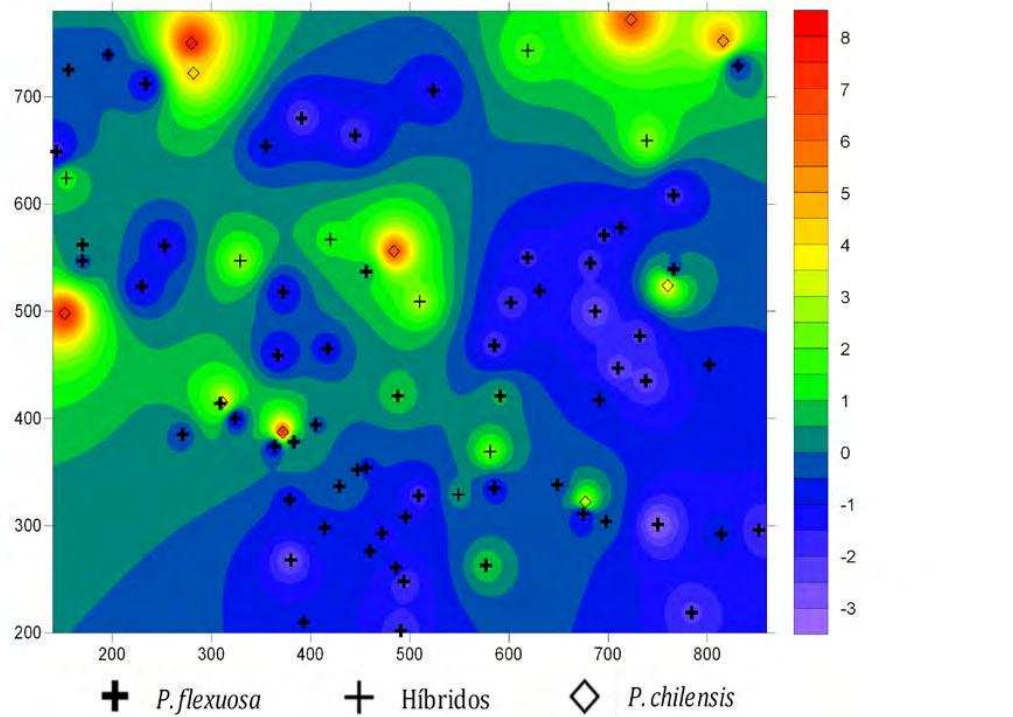


Figura 1. Mapa sintético de la variabilidad morfológica de un enjambre híbrido de *Prosopis* sp. obtenido mediante la interpolación de la CP_{1m}

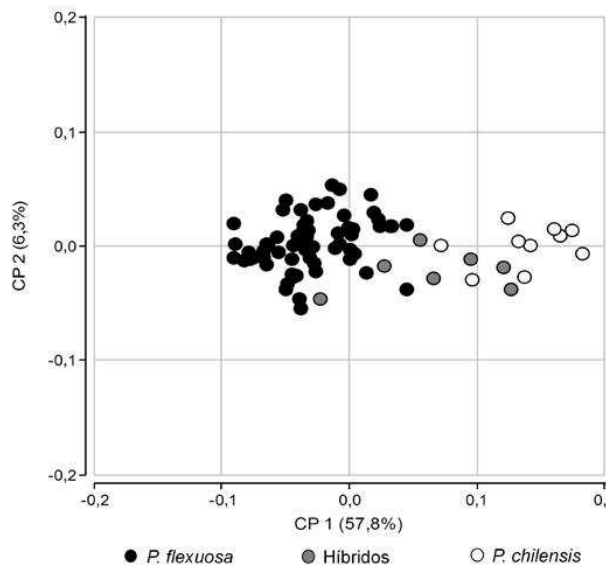


Figura 2. Ordenación de consenso entre la variabilidad genética y morfológica de 81 árboles pertenecientes a un enjambre híbrido de *Prosopis chilensis* y *P. flexuosa*. El gráfico sugiere una clara diferenciación entre los individuos de ambas especies y una mayor variabilidad entre los individuos clasificados como híbridos.

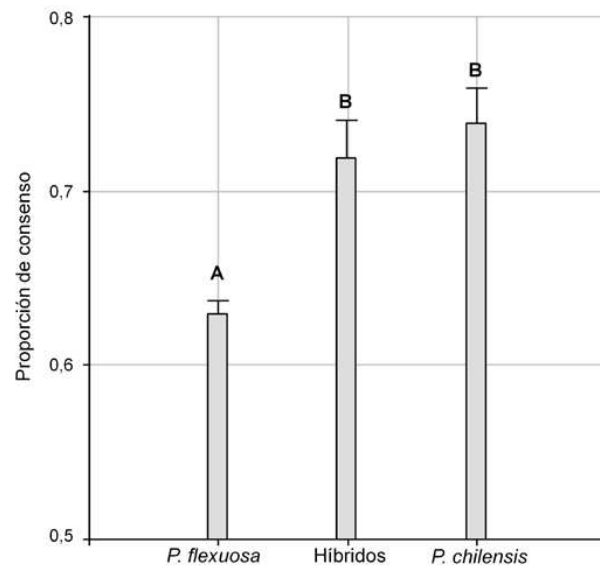


Figura 3. Proporción de consenso promedio (\pm EE) entre los datos genéticos y los datos morfológicos para los individuos de cada grupo. Letras distintas indican diferencias significativas entre las medias de consenso de los grupos morfológicos definidos *a priori*.

DISCUSIÓN

A pesar de que en especies de plantas la variabilidad genética suele estar estructurada espacialmente (Vekemans y Hardy, 2004), no se detectó EGE en el enjambre híbrido de *P. flexuosa* y *P. chilensis* analizado. La alta densidad de individuos (alrededor de 200 árboles/ha) podría ser motivo de la ausencia de estructura espacial de la variabilidad genética. En áreas con altas densidades de árboles, el solapamiento de las áreas de dispersión de semillas puede producir la mezcla de progenies de diferentes individuos. En esos casos se espera que la EGE sea débil (Hamrick *et al.*, 1993; Degen *et al.*, 2001). El estudio espacial de la variabilidad morfológica, en cambio, detectó estructura espacial. La asociación entre la variabilidad genética y morfológica fue detectada por los tres métodos empleados contribuyendo al proceso de diferenciación de unidades de uso y conservación de Algarrobo. A pesar de la débil estructura espacial en los datos, la incorporación de la autocorrelación espacial mejoró la capacidad de detección y modelación de la asociación entre la variabilidad genética y fenotípica observada en el enjambre híbrido. En el caso de la Prueba de Mantel modificó la significancia de la prueba y en el caso del ACC mejoró el ajuste del modelo que relaciona la variabilidad morfológica con la genética, expresadas ambas por variables canónicas.

Existe una larga discusión en la literatura sobre la pertinencia de la aplicación de la Prueba de Mantel Parcial, y en particular cuando ambos conjuntos de datos presentan autocorrelación espacial (Oden y Sokal, 1992; Dutilleul *et al.*, 2000; Raufaste y Rousset, 2001; Castellano y Balletto, 2002; Rousset, 2002; Legendre y Fortin, 2010; Guillot y Rousset, 2013). Claramente esta controversia no está resuelta aún y la Prueba de Mantel Parcial sigue siendo una de las más utilizadas para evaluar asociaciones bajo autocorrelación espacial. En nuestro trabajo la incorporación de la estructura espacial, aunque ésta era débil, permitió detectar asociaciones multivariadas entre datos genéticos y morfológicos que luego fueron confirmadas por el ACC. Otra forma de incluir la información espacial en estudios de asociación con ACC es la utilizada por Sork *et al.* (2010), donde estudian la asociación entre datos genéticos (microsatélites) y climáticos en bosques del hemisferio norte. Los autores determinaron la

contribución de variables geográficas a la construcción de una variable canónica climática particionando la suma de cuadrados total (no debida al error) de la variable canónica en tres componentes: una debida sólo a las coordenadas geográficas, otra sólo a variables climáticas y otra a la colinealidad entre ambas.

Es importante notar que el ACC asume correlación del tipo lineal y que otras correlaciones pueden pasar desapercibidas con este análisis. El APG, en cambio, permite detectar asociaciones no necesariamente lineales (Bramardi *et al.*, 2005). La ordenación de consenso de los individuos del enjambre híbrido bajo estudio permitió diferenciar principalmente a *P. flexuosa* de *P. chilensis*. Los híbridos, si bien se diferencian morfológicamente tanto de *P. flexuosa* como de *P. chilensis*, genéticamente no se pueden diferenciar de *P. chilensis*, al menos con los seis SSR utilizados en este trabajo (Teich, 2012). Los resultados de Mottura (2006) sugieren que si bien existe un fuerte flujo génico entre los diferentes grupos morfológicos, éste es asimétrico. La baja diferenciación de los híbridos en el espacio de consenso probablemente se deba a las mayores tasas de apareamiento de los mismos con individuos de *P. chilensis* y por lo tanto a los mayores niveles de introgresión hacia este grupo. Asimismo, nuestros resultados indican que la proporción de consenso entre datos genotípicos y fenotípicos de estos dos grupos es mayor que la observada en *P. flexuosa*.

La alta asociación detectada entre la variabilidad genética y fenotípica confiere valor taxonómico a los caracteres morfológicos propuestos para clasificación de híbridos de *P. flexuosa* y *P. chilensis*. Si bien los tres grupos de algarrobos analizados presentan identidad a nivel morfológico, respecto al consenso entre la información fenotípica y genética sólo se diferencian *P. flexuosa* y *P. chilensis*. Los híbridos no se reconocen como una unidad independiente respecto al nivel de asociación fenotipo-genotipo, al menos para los marcadores moleculares y los caracteres morfológicos analizados. Respecto a la incorporación de la autocorrelación espacial de los datos en estudios de asociación, concluimos que su omisión puede afectar los resultados de los estudios de asociación entre variabilidad genética y morfológica, aún cuando la estructura espacial en una u otra sea débil.

BIBLIOGRAFÍA

- Balzarini M.G., Di Rienzo J.A. (2011) InfoGen versión 2011. FCA, Universidad Nacional de Córdoba, Argentina. URL <http://www.info-gen.com.ar>.
- Balzarini M.G., Teich I., Bruno C., Peña A. (2011) Making genetic biodiversity measurable: a review of statistical multivariate methods to study variability at gene level. *Rev. Fac. Cienc. Agrar., Univ. Nac. Cuyo* 43: 261-275.
- Bates D., Mächler M., Bolker B., Walker S. (2015) lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-9.
- Besega C.F., Ferreyra L., Vilardi J., Saidman B.O. (2000) Unexpected low genetic differentiation among allopatric species of section *Algarobia* of *Prosopis* (Leguminosae). *Genética* 109: 255-266.
- Bivand R., Altman M., Anselin L., Assunção R., Berke O., Bernat A., Blanchet G., Blankmeyer E., Carvalho M., Christensen B., Chun Y., Dormann C., Dray S., Halbersma R., Krainski E., Legendre P., Lewin-Koh N., Hongfei Li J.M., Millo G., Mueller W., Ono H., Peres-Neto P., Piras G., Reder M., Tiefelsdorf M., Yu D. (2011) spdep: Spatial dependence: weighting schemes, statistics and models. R package version 0.5-31.
- Bonferroni C. E. (1936) Teoria statistica delle classi e calcolo delle probabilità. *Pub. Ist. Sup. Sci. Econom. Com. di Firenze* 8: 3-62.
- Brachi B., Faure N., Horton M., Flahauw E., Vazquez A., Nordborg M., Bergelson J., Cuguen J., Roux F. (2010) Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genet.* 6: e1000940.
- Bramardi S.J., Bernet G.P., Asíns M.J., Carbonell E.A. (2005) Simultaneous agronomic and molecular characterization of genotypes via the Generalised Procrustes Analysis. *Crop Sci.* 45: 1603-1609.
- Carrari F., Schauer N., Willmitzer L., Fernie A. (2006) Systems biology: a renaissance of the top-down approach for plant analysis. In: Saito K., Dixon R.A., Willmitzer L. (Eds) *Plant Metabolomics*. Springer Berlin Heidelberg, pp 185-198.
- Castellano S., Balletto E. (2002) Is the partial Mantel test inadequate? *Evolution* 56: 1871-1873.
- Cockram J., White J., Zuluaga D.L., Smith D., Comadran J., Macaulay M., Luo Z., Kearsey M.J., Werner P., Harrap D., Tapsell C., Liu H., Hedley P.E., Stein N., Schulte D., Steuernagel B., Marshall D.F., Thomas W.T.B., Ramsay L., Mackay I., Balding D.J., Consortium T.A., Waugh R., O'Sullivan D.M. (2010) Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. *Proc. Natl. Acad. Sci. USA* 107: 21611-21616.
- Degen B., Caron H., Bandou E., Maggia L., Chevallier M.H., Leveau A., Kremer A. (2001) Fine-scale spatial genetic structure of eight tropical tree species as analysed by RAPDs. *Heredity* 87: 497-507.
- Ding B., Gentleman R., Carey V. (2008) bioDist: Different distance measures. R package version 1.40.0.
- Dutilleul P., Jason D.S., Frigon D., Legendre P. (2000) The Mantel test versus Pearson's correlation analysis: assessment of the differences for biological and environmental studies. *J. Agric. Biol. Env. S* 5: 131-150.
- Gabriel K.R., Sokal R.R. (1969) A new statistical approach to geographic variation analysis. *Syst. Biol.* 18: 259-278.
- Gower J. (1975) Generalized procrustes analysis. *Psychometrika* 40: 33-51.
- Guillot G., Rousset F. (2013) Dismantling the Mantel tests. *Methods Ecol. Evol.* 4: 336-344.
- Hamrick J.L., Murawski D.A., Nason J.D. (1993) The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. *Plant Ecol.* 107-108: 281-297.
- Hottelling H. (1933) Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 24: 417-411.
- Houle D., Govindaraju D.R., Omholt S. (2010) Phenomics: the next challenge. *Nat. Rev. Genet.* 11: 855-866.

- Hunziker J.H., Saidman B.O., Naranjo C.A., Palacios R.A., Poggio L., Burghardt A.D. (1986) Hybridization and genetic variation of Argentine species of *Prosopis*. *Forest Ecol. Manag.* 16: 301-315.
- Jombart T. (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403-1405.
- Jombart T., Pontier D., Dufour A.B. (2009) Genetic markers in the playground of multivariate analysis. *Heredity* 102: 330-341.
- Keitt T.H., Bjørnstad O.N., Dixon P.M., Citron-Pousty S. (2002) Accounting for spatial pattern when modeling organism-environment interactions. *Ecography* 25: 616-625.
- Legendre P., Fortin M.J. (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Mol. Ecol. Resour.* 10: 831-844.
- Littell R.C., Milliken G.A., Stroup W.W., Wolfinger R.D. (2006) SAS System for Mixed Models. SAS Institute Inc., Cary, NC, USA.
- Mantel N. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209-220.
- Moran P.A.P. (1950) Notes on continuous stochastic phenomena. *Biometrika* 37 (1/2): 17-23.
- Mottura M.C. (2006) Development of microsatellites in *Prosopis* spp. and their application to study the reproduction system. Tesis Doctoral Georg-August University of Göttingen, Göttingen.
- Mottura M.C., Finkeldey R., Verga A.R., Gailing O. (2005) Development and characterization of microsatellite markers for *Prosopis chilensis* and *Prosopis flexuosa* and cross-species amplification. *Mol. Ecol. Notes* 5: 487-489.
- Novembre J., Johnson T., Bryc K., Kutalik Z., Boyko A.R., Auton A., Indap A., King K.S., Bergmann S., Nelson M.R., Stephens M., Bustamante C.D. (2008) Genes mirror geography within Europe. *Nature* 456: 98-101.
- Oden N., Sokal R. (1992) An investigation of three-matrix permutation tests. *J. Classif.* 9: 275-290.
- Oksanen J., Blanchet F.G., Kindt R., Legendre P., O'Hara R.B., Simpson G.L., Solymos P., Stevens M.H.H., Wagner H. (2011) vegan: Community Ecology R Package.
- Pasiecznik N.M., Felker P, Harris P.J.C., Harsh L.N., Cruz G., Tewari J.C., Cadoret K., Maldonado L.J. (2001) The *Prosopis juliflora* - *Prosopis pallida* complex: a monograph. HDRA, Coventry, United Kingdom.
- R Development Core Team R (2011) R: A language and environment for statistical computing. Vienna, Austria.
- Raufaste N., Rousset F (2001) Are partial mantel tests adequate? *Evolution* 55: 1703-1705.
- Rousset F (2002) Partial Mantel Tests: Reply to Castellano and Balletto. *Evolution* 56: 1874-1875.
- Saidman B.O. (1990) Isoenzyme studies on hybrid swarms of *Prosopis caldenia* and sympatric species. *Silvae Genet.* 39: 5-8.
- Saidman B.O., Bessega C.F., Ferreira L.I., Julio N., Vilardi J. (2000) The use of genetic markers to assess population structure and relationships among species of the genus *Prosopis* (Leguminosae). *Bol. Soc. Argent. Bot.* 35: 315-324.
- Smouse P.E., Long J.C., Sokal R.R. (1986) Multiple regression and correlation extensions of the Mantel Test of matrix correspondence. *Syst. Zool.* 35: 627-632.
- Sork V.L., Davis F.W., Westfall R., Flint A., Ikegami M., Wang H., Grivet D. (2010) Gene movement and genetic association with regional climate gradients in California valley oak (*Quercus lobata* Nee) in the face of climate change. *Mol. Ecol.* 19: 3806-3823.
- Teich I. (2012) Análisis de la estructura genética espacial de especies arbóreas y su asociación con la variabilidad fenotípica y ambiental. Tesis Doctoral, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina.

- Teich I., Planchuelo A.M., Balzarini M. (2011) Análisis de asociaciones en escenarios de datos masivos. 40JAIIO-Congreso Argentino de AgroInformática 166-177.
- Teich I., Verga A., Balzarini M. (2014) Assessing spatial genetic structure from molecular marker data via principal component analyses: A case study in a *Prosopis* sp. forest. *Adv. Biosci. Biotechnol.* 5: 89-99.
- Vega M.V., Hernández P. (2005) Molecular evidence for natural interspecific hybridization in *Prosopis*. *Agroforest Syst.* 64: 197-202.
- Vekemans X., Hardy O.J. (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol. Ecol.* 13: 921-935.
- Verga A. (2000) Clave para la identificación de híbridos entre *Prosopis chilensis* y *P. flexuosa* sobre la base de caracteres cuantitativos. *Multequina* 9: 17-22.
- Verga A.R. (1995) Genetische Untersuchungen an *Prosopis chilensis* und *P. flexuosa* (Mimosaceae) im trockenen Chaco Argentinien. Göttingen Research Notes in Forest Genetics 19:1-96. PhD tesis, Department of Forest Genetics and Tree Breeding, Faculty of Forestry, University of Göttingen.
- Verga A.R., López Lauenstein D., López C., Navall M., Joseau J., Gómez C., Royo O., Degano W., Marcó M. (2009) Caracterización morfológica de los algarrobos (*Prosopis* sp.) en las regiones fitogeográficas Chaqueña y Espinal norte de Argentina. *Quebracho* 17: 31-40.