Spatial structure of phenotypic traits in seven provenances of *Neltuma alba* (Fabaceae)

Estructura espacial de caracteres fenotípicos en siete procedencias de *Neltuma alba* (Fabaceae)

María V. Vega[1]* , Beatriz O. Saidman[2] & Juan C. Vilardi[2]

1. Laboratory of Plants Biotechnology, FRN, SECyT, UNaF, Formosa, Argentina
2. Department of Ecology, Genetics and Evolution. FCEN. UBA, IEGEBA, CONICET, Buenos Aires, Argentina

*mavivega@yahoo.es

Summary

Backgrounds: *Neltuma alba* is one of the most important native woody species in Argentina in the Dry Chaco, part of the Argentine Mesopotamia and the Paraguayan Chaco. It shows great variability due to interspecific hybridization and introgression associated with protogyny and self-incompatibility systems. This species is adapted to arid and semi-arid climates with saline and degraded soils. Environmental heterogeneity and wide distribution may result in local adaptation and autocorrelated spatial patterns in genetic and quantitative variables.

Aim: To analyze the spatial structure in provenances of *N. alba* influenced by isolation by distance in the Gran Chaco Region.

M&M: This work studied spatial structure in seven provenances of *N. alba* from the Dry and Humid Chaco regions, based on fifteen foliar, fruit and germination traits in 68 individuals, together with five environmental variables.

Results and Conclusion: univariate statistical analyses showed significant or highly significant differences among provenances. According to Moran's *I* index phenotypic and geographical distances are significantly autocorrelated for the first distance class (0-0.643 km). Partial Mantel test showed significant correlation for the first two distance classes. The overall analysis showed that 11 of the analyzed traits showed significant spatial autocorrelation. The local spatial analysis indicated that for several traits their hot spots of high similarity between neighboring individuals and cold spots where nearby individuals are highly differentiated.

Keys words

Local adaptation, morphological traits, *Neltuma*, origin, *Prosopis*, spatial autocorrelation analyses.

Resumen

Introducción: *Neltuma alba* es una de las leñosas nativas más importantes de Argentina. Ocupa la ecorregión de Gran Chaco y parte de la Mesopotamia Argentina. Muestra gran variabilidad parcialmente atribuida a hibridación e introgresión con otros algarrobos, asociados a la protoginia y sistemas de autoincompatibilidad. Se adapta a climas áridos y semiáridos con suelos salinos y degradados. La heterogeneidad ambiental y su amplia distribución pueden asociarse a adaptación local y autocorrelacionado espacial en variables genéticas y cuantitativas.

Objetivo: Analizar la estructura espacial en procedencias de *N. alba* influenciada por el aislamiento por distancia en la Región del Gran Chaco.

M&M: Se evaluó la autocorrelación en siete procedencias de *N. alba* de las ecorregiones del Chaco Seco y el Chaco Húmedo, basado en quince rasgos foliares, de fruto y germinación en 68 individuos adultos, incluyendo información de cinco variables ambientales.

Resultados y conclusión: Los análisis estadísticos univariados demostraron diferencias significativas o altamente significativas entre procedencias. La asociación entre la similitud fenotípica y la distancia geográfica mostró valores de autocorrelación significativa para la primera clase de distancia (0-0.643 km). El test de Mantel parcial indicó que la correlación entre la distancia fenotípica y geográfica se pierde a partir de 40 - 50 km. El análisis global mostró que 11 de los rasgos analizados presentaron una autocorrelación significativa. El análisis espacial local indicó que para varios rasgos existen puntos de alta similitud entre individuos vecinos (*hotspots*) y puntos donde los individuos cercanos están muy diferenciados (*coldspots*).

Palabras clave

Análisis de autocorrelación espacial, caracteres adaptativos, *Neltuma*, origen, *Prosopis*, rasgos morfológicos.

INTRODUCTION

In 1943, Wright introduced the term isolation by distance to describe a pattern in which genetic differentiation increases with geographic distance. In this context, genetic differentiation among populations would be the result of the interaction between drift and gene flow. Wang & Bradburd (2014) incorporate the concept of isolation by environment as a pattern in which genetic differentiation increases with environmental differences, regardless the geographic distance. Together, processes that reduce the rate of dispersal and environmental differences among populations will generate patterns of increased genetic differentiation. These distribution patterns of variation can be generated by climatic and edaphic variations and geographic isolation (Siabato & Guzman Manrique, 2019).

One way to assess the association between phenotypic similarity and geographic distance is spatial autocorrelation (SA) analysis, defined as the property of pairs of spatial objects of being more (positive SA) or less similar (negative SA) to each other than randomly expected (Getis & Ord, 1992). The first index to measure SA, proposed by Moran (Moran's Global I index) in 1950, was used to evaluate whether the values of one (phenotypic) variable studied tended to cluster spatially. In the 90s, measurements of local autocorrelation were developed, that allowed to capture local spatial autocorrelation indicators such as the coefficients Gi and Gi^* (Ord & Getis, 1995; Garcia, 2019), which described spatial clustering around individual sites, to discover local "packages" (hot spots) of autocorrelated points. In the early 2000s, a new approach defined as landscape genetics emerged, oriented to the analysis of interactions between landscape features and evolutionary processes such as gene flow and selection (Bessega et al., 2015).

In the arid and semiarid forest regions of the Humid Chaco and Dry Chaco in Argentina, the woody species of the *Neltuma* Raf. (formerly *Prosopis* L., Hughes *et al.,* 2022) genus stand out, constituting an important multipurpose natural resource. The fruits of these species and their derivatives are used for human food and fodder and their wood is of high quality for furniture and charcoal production. They show high morphological variability and ecological amplitude, offering the possibility of selecting trees adapted to diverse environments. Morphometric studies carried out in *Neltuma* species, account for phenotypic differentiation among populations, which can be attributed to genetic differences caused by adaptation to various environmental conditions (Bessega *et al.*, 2015; Fontana *et al.*, 2018), observable at different levels: species, populations, individuals and even within the same individual (Fontana *et al.*, 2018; Castillo *et al.*, 2021).

The factors capable of generating SA patterns and affecting the genetic coherence among provenances include the mechanisms of pollen and seed dispersal and the environmental heterogeneity (Roser, 2017). Based on studies conducted by researchers on *Neltuma* species, the dispersal distances of pollen and seeds are short, which is why in local populations a marked reduction in kinship and phenotypic resemblance among individuals was observed depending on the distance that separates them (Bessega *et al.*, 2012; Roser, 2017).

MATERIALS AND METHODS

Sampling

Relatively homogeneous natural areas within the distribution of *Neltuma alba* (Griseb.) C.E.Hughes & G.P.Lewis were identified, and individuals from seven provenances representing three complexes of two ecoregions within the Argentine Chaco region were collected covering a latitudinal and longitudinal geographic gradient (Table 1). Five provenances correspond to the Dry Chaco Ecoregion: Ibarreta (ib), Laguna Yema (ly), Isla Cuba (ic), Las Breñas (lb) and Charata (ch). The other two provenances are located in the Humid Chaco Ecoregion: Formosa (fo) and Villa Dos Trece (vi) (Table 1, Fig. 1). The provenances of ib, ly and ic are located in the Pilcomayo-Bermejo Interfluvial Complex, while lb and ch are located in the Antiguos Cauces of Juramento-Salado Complex (Table 1). These complexes belong to the Dry

Table 1. Location of the different provenances sampled. Ecoregions, sub-regions and complexes according to Morello *et al.* (2012). Geographical coordinates, meanl rainfall, mean, maximum and minimum temperatures and water vapor pressure for each provenance are provided.

Ecoregions	Subregions	Complex	Provenances	long	lat	mr (mm)	mean T (°C)	max T (°C)	min T (°C)	wvp (kPa)
Dry Chaco	Semiarid Chaco	Bermejo-Pilcomayo Interfluvium	Ibarreta (Formosa)	-59.97	-25.44	1002.8	21.2	28.8	15.4	131.6
Dry Chaco	Semiarid Chaco	Bermejo-Pilcomayo Interfluvium	Laguna Yema (Formosa)	-61.25	-24.26	795.2	23.2	29.8	16.6	131.6
Dry Chaco	Semiarid Chaco	Bermejo-Pilcomayo Interfluvium	Isla Cuba (Formosa)	-61.87		673.2	23.2	30.2	15.9	133.9
Dry Chaco	Semiarid Chaco	Old Channel of Juramento Salado	Las Breñas (Chaco)	-61.07	-27.10	906.8	20.4	29.3	14.5	117
Dry Chaco	Semiarid Chaco	Old Channel of Juramento Salado	Charata (Chaco)	-61.19	-27.24	939	20.4	29.3	14.5	117
Humid Chaco	Chaco of Forests and Glens	Eastern Lower Paraguay River	Formosa (Formosa)	-58.25	-26.25	1279	21.9	28.5	14.6	103.3
Humid Chaco	Chaco of Forests and Glens	Eastern Lower Paraguay River	Villa Dos Trece (Formosa)	-59.35	-26.16	1204	21.9	28.5	14.6	117
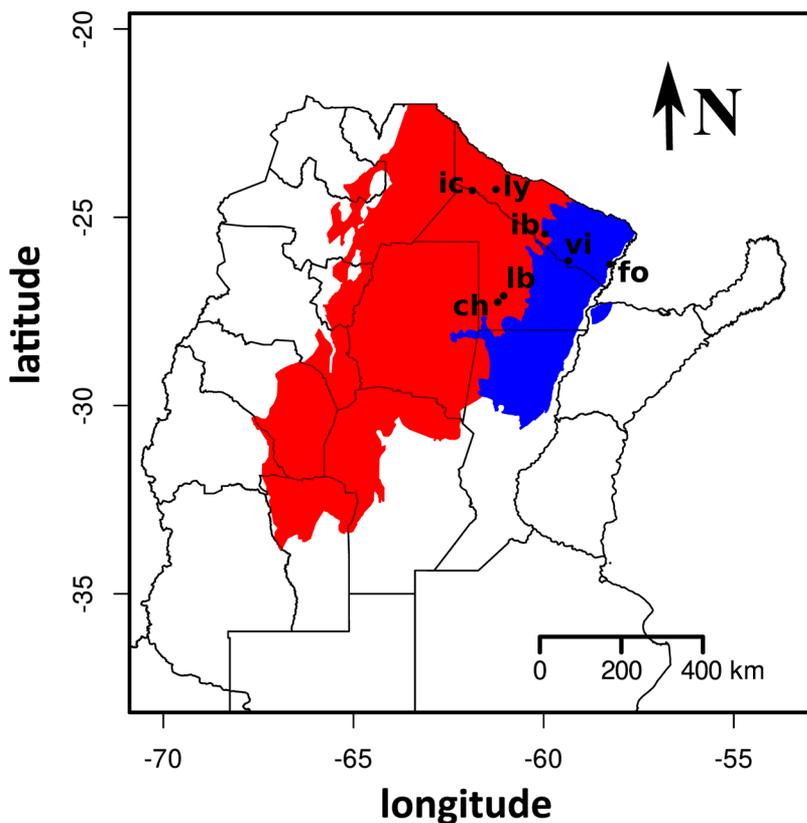
Fig. 1. Distribution of the Dry Chaco (red) and Humid Chaco (blue) ecoregions in Argentina with the geographic location of the sampled sites (provenances). The distribution of the ecoregions was obtained from Morello *et al.* (2012). Abbreviations: fo: Formosa; ch: Charata, ib: Ibarreta, ic: Isla Cuba, lb: Las Breñas, ly: Laguna Yema, vi: Villa Dos Trece.

Chaco Ecoregion, where the warm continental subtropical climate predominates, with absolute maximum temperatures exceeding 47°C. Absolute minimum temperatures range between -6° and -7°C in the plains and foothills, and between -12° and -16°C in the foothills of Chaco Serrano (Morello *et al*., 2012). Precipitation ranges from 700 mm (isohyet between Santiago and Northern Santa Fe and Central Southern Chaco) to 400 mm in the valleys of Güemes, Tapia-Trancas and Catamarca hills (Morello *et al*., 2012).

The provenances of fo and vi are located in the Eastern Complex of the Lower Paraguay River (Table 1), where annual precipitation decreases from West to East with more than 1300 mm in

the area of the large river collectors and about 750 mm in the border with the Dry Chaco. The Eastern Complex of the Lower Paraguay River is located approximately 150-400 km from the Bermejo-Pilcomayo Interfluvial Complex, and 430 km from the Antiguos Cauces of Juramento-Salado Complex. While the Bermejo-Pilcomayo Interfluvial Complex and the Antiguos Cauces del Juramento-Salado Complex are 496 km apart (Table 1; Fig. 1).

To capture a good representation of the variation, 10 individuals (mother plants) per provenance were randomly chosen, with typical morphological characteristics of *N. alba*, according to the description of Burkart (1976). Seeds were collected from each individual to set

up a progeny trial in a nursery in the Formosa National University. The set of seeds collected from each mother plant was considered a family made up of half-sibs, following the criterion that trees should be at least 50 m apart (Saidman & Vilardi, 1993; Bessega *et al.*, 2000) to reduce the probability of duplications in the sampling due to the possible relationship among neighboring trees.

The geographic location of the harvested individuals was recorded using a GPS (Garmin Etrex 10). Harvesting data sheets were filled in using the passport information described by the International Plant Genetic Resources Institute, adapted to the species under study, and documenting the topographic, edaphic and climatic characteristics of each provenance. For each sampled site, five environmental variables were analyzed: mean rainfall (mr), mean temperature (mean T), maximum temperature (max T), minimum temperature (min T) and water vapor pressure (wvp). These data were collected from: https://worldclim.org, with a 30 sec resolution. The information corresponds to the 1970-2000 period. For each variable, 12 GeoTiff files were obtained, one for each month of the year. The environmental variables of each sampling site were extracted from the GeoTiff files using the *extract* function of the *raster* package (Hijmans, 2019) of the R program (R Core Team, 2020). In each case, the records of all the months were averaged to obtain an annual mean (Table 1).

Analyzed traits

Herbarium vouchers and fruit samples were obtained from each sampled individual in order to measure the following morphological traits: leaflet length (LL) and width (LW) (cm), petiole length (PEL, cm), length of the longest pinnae of each leaf (PIL, cm), number of pinnae per leaf (NPL), interleaf distance (ID, cm), number of leaflet pairs per pinnae (LP), pod length (PL) and width (PW, cm), seed length (SL) and width (SW, cm), number of developed seeds per pod (SN), germination rate (GR) (%), germination power (GP) and mean germination time (GT, days) (Table 2). Measurements were carried out on fresh samples, using the methodology proposed by Palacios & Bravo (1981), for which leaves, seeds and fruits were randomly collected from branches located in the basal part of each individual in each provenance. Ten measurements were made of each of the organs mentioned in each of the 68 families sampled/provenance. To make up the herbarium, two duplicates were prepared, all from the same plant and with the same collection number.

Individual records of leaf, pod and seed traits, were based on the averages of 10 measurements per individual. To assess the germination traits, 25 seeds were used for each provenance and family. They were planted in plastic trays (260 mm x 190 mm x 60 mm) containing 400 g of sterilized sand, treated with Kaptan 50 % WP (CAS N° 133-06-2). Later, they were placed in an incubation chamber at 27°C±2°C, under continuous fluorescent white light of 150 mmol.m^{-2}.s^{-1} with a photoperiod of 8 h of light, for 20 days, recording the number of germinated seeds daily. Seeds were considered germinated when their coats were broken by the radicles.

The germination rate (GR) was measured as follows:

$$GR = \frac{g}{s} \times 100$$

where *g* is the number of germinated seeds at the end of the trial and *s* is the total number of planted seeds.

Germination power (*GP*) was quantified using the Djavanshir & Pourbeik (1976) expression, as follows:

$$VG = GDM \times VP$$

where *GDM* represents the average daily germination and is calculated as the total percentage of germinated seeds at the end of the trial divided by the number of days the trial lasted, that is,

$$GDM = \frac{100G}{ST}$$

where *G* is the number of germinated seeds, *S* is the total number of seeds sown and *T* is the total length of the trial in days. In addition, *VP* is the peak or maximum value obtained by dividing the germination rate accumulated day

Table 2. Basic statistics of the traits analyzed and significant differences among provenances (P) obtained by the variance analysis. Provenances: ch = Charata, fo = Formosa, ic = Isla Cuba, ib = Ibarreta, lb = Las Breñas, ly = Laguna Yema, vi = Villa Dos Trece. Abbreviations: GR = germination rate, GP = germination power, GT = mean germination time, LL = length of leaflet, LW = leaflet width, PEL = length of petiole, PIL = length of pinna, NPL = number of pinna/leaf, ID = inter-leaf distance, LP = number of pairs of leaflets/pinna, PL = pod length, PW = pod width, SL = seed length, SN = number of seeds, SW = seed width. SD = standard deviation, Adjusted p-value = P-values adjusted for multiple comparisons by the method described in Benjamini and Hochberg (1995).

Provenances	fo		ib		ly		ic		vi		lb		ch		P	Adjusted p-value
Traits	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD		
LL	0.7	0.1	0.8	0.2	0.7	0.1	0.7	0.09	0.6	0.07	0.6	0.1	0.6	0.09	1.4×10^{-5}	3.1×10^{-5}
LW	0.1	0.03	0.7	0.6	0.2	0.04	0.1	0.03	0.1	0.03	0.1	0.03	0.1	0.02	2.5×10^{-7}	1.3×10^{-6}
PEL	2.3	0.7	1.8	0.7	3.6	1.2	2.7	0.6	2.3	0.6	2.0	0.6	2.5	0.3	1.1×10^{-5}	2.9×10^{-5}
PIL	9.8	2.0	8.5	0.8	9.6	1.4	8.7	1.1	7.9	2.0	8.4	1.7	8.7	0.8	9.1×10^{-2}	1×10^{-1}
NLP	4.7	0.4	6.2	0.8	4.2	0.3	4.5	0.3	5.2	0.7	5.3	0.7	5.0	1.2	6×10^{-7}	2×10^{-6}
ID	0.3	0.3	0.1	0.0	0.2	0.08	0.1	0.0	0.2	0.0	0.2	0.0	0.2	0.0	4.7×10^{-1}	4.7×10^{-1}
LP	40.2	4.1	73.0	6.0	33.6	2.5	63.3	7.2	38.1	3.6	36.6	5.2	37.4	3.5	2.2×10^{-16}	3.5×10^{-15}
PL	15.9	2.1	17.6	3.8	14.3	1.7	15.9	2.4	16.3	1.4	18.5	5.6	16.4	3.3	2.3×10^{-3}	4.7×10^{-3}
PW	1.3	0.1	1.2	0.1	1.4	0.4	1.4	0.09	1.2	0.1	1.3	0.1	1.3	0.2	1.5×10^{-1}	1.6×10^{-1}
SN	24.2	6.9	13.6	4.9	18.7	3.3	25.5	6.9	25.0	3.8	28.1	3.4	29.9	1.9	4.5×10^{-11}	3.6×10^{-10}
SL	0.6	0.1	0.5	0.0	0.5	0.04	0.6	0.3	0.7	0.3	0.6	0.04	0.7	0.2	4×10^{-2}	5.3×10^{-2}
SW	0.3	0.0	0.3	0.0	0.3	0.01	0.4	0.3	0.4	0.03	0.5	0.2	0.4	0.03	5.2×10^{-2}	6.5×10^{-2}
H	10.7	2.4	8.4	1.7	9.1	1.4	9.4	1.5	11.2	4.0	11.1	1.5	9.1	1.7	2.2×10^{-2}	3.2×10^{-2}
GR	99.2	1.7	84.7	14.6	84.8	27.6	83.1	10.8	86	15.5	99.6	1.3	97.9	6.6	7.6×10^{-3}	1.2×10^{-2}
GP	22.6	7.4	11.6	4.9	19.7	6.0	15.6	4.6	19.2	7.2	20.2	8.9	25.7	10.8	4.7×10^{-3}	8.4×10^{-3}
GT	3.1	0.4	4.4	0.9	4.5	1.3	2.9	0.7	3.3	0.6	3.3	0.3	3.0	0.7	5.4×10^{-6}	1.7×10^{-5}

by day, by the number of days elapsed. Finally, mean germination time (*GT*) was quantified as proposed by Silva & Nakagawa (1995):

$$TMG = \sum(T_i N_i)/N$$

where T_i = number of days elapsed since the beginning of germination, N_i = number of germinated seeds on day *i*, and N = total number of germinated seeds at the end of the trial.

Statistical Methods

To characterize the phenotypic variation among the provenances, the following linear model was used:

$$Z_{ij} = \mu + P_i + e_{ij}$$

where Z_{ij} is the trait observation of the individual *j* from provenance *i*, μ is the general mean, P_i is the provenance effect and e_{ij} is the residual component. For continuous traits, a conventional linear model was applied, while in the case of GP, a generalized model was applied considering that the response is a binomial variable. All the analyses were carried out using R, version 4.2.1 (R Core Team, 2020). To control the Type 1 error, the significance *P* values of the individual tests were corrected for multiple tests using the *FDR* (*false discovery rate*) method (R Core Team, 2020).

Spatial autocorrelation analyses were carried out with the *EcoGenetics* package (Roser *et al*., 2017) of the R program. Spatial autocorrelation was analyzed using four methods. The first one was based on correlograms using Moran's *I* coefficient (Moran, 1950), applied to each trait analyzed and multivariate correlogram (average) using Moran's *I* coefficient. The analysis was performed using the *eco.correlog* function and significance was obtained by bootstrap (1000 simulations). The second method was based on the comparison of morphological distance and geographical distance matrices by means of Mantel correlograms using the *eco.cormantel* function, including the environmental distances among sampling sites as a third matrix. The third method consisted of a global spatial analysis, using the *eco.gsa* function, which produced a single global statistic for each trait across the

entire area analyzed. Significance was obtained using a permutation test with 1000 simulations. The fourth was a local spatial analysis by the *eco. lsa* function, using the *Gi** coefficient (Getis & Ord, 1992, 1995). This analysis identified points of high similarity or high dissimilarity between neighbors and significance was obtained by a permutation test (1000 simulations). In all the cases, the distance classes were defined so that they all included the same number of individuals pairs, representing with intense red coloration sites where neighbors tended to be very similar to each other (*hot spots*), and with intense blue coloration sites where there was high dissimilarity between neighbors (*cold spots*).

RESULTS

The measurements obtained for leaf, fruit, see, and germination traits are specified in Table 2. Highly significant differences were recorded for most of the variables, even after correction for multiple comparisons. The only variables that did not show significant differences were ID and PW. The multivariate analysis of variance showed that the differences among provenances were highly significant considering the total set of traits (*Pillai* = 3.3 $P<10^{-16}$).

Moran's Index

Univariate correlogram analyses were performed for each trait considered individually and an average correlogram based on Moran's *I* coefficient (Fig. 2). The first point represents the autocorrelation among individuals separated from each other by a distance of 0 to 0.643 km, which would correspond to the closest neighbors, located within the same provenance. The mean distance between pairs of individuals within this distance class is 0.192 km. The second point is the autocorrelation among individuals separated from each other by 0.643 to 15.445 km (mean = 7.736 km); in this case the individuals of each pair belong to the same provenance with the exception of some pairs involving individuals from lb and ch. In all remaining classes pairs involve individuals from different provenances (distances > 15.45 km) (Table 3).

Table 3. Classes of morphological-geographic distances among the individuals analyzed and Mantel correlogram significance for the set of traits analyzed. Observed correlation, p-values and p-values adjusted for multiple tests are reported. Black numbers indicate significant differences.

Traits	obs	p-value	Adjusted p-value
LL	0.4	0.001	**0.002**
LW	0.3	0.02	**0.03**
PEL	0.5	0.001	**0.002**
PIL	0.2	0.1	0.1
NLP	0.6	0.001	**0.002**
ID	0.2	0.03	**0.04**
LP	0.9	0.001	**0.002**
PL	0.2	0.1	0.1
PW	0.1	0.2	0.2
SN	0.7	0.001	**0.002**
SL	0.2	0.1	0.1
SW	0.3	0.03	**0.04**
GR	0.4	0.001	**0.002**
GP	0.5	0.001	**0.002**
GT	0.4	0.002	**0.004**

In 10 of the traits analyzed, autocorrelation was significant at least for the first class of distance (Fig 2A-C, E, G-H, J, M-O). Three traits (LL, LP and GP) showed significant autocorrelation in the third classes of distance (15.445 – 90.674 km). In the case of LL, the similarity is reduced to approximately 75-80 km, since the line that joins the points passes through 0 between points 3 and 4. From that distance, the similarity among individuals would be random (Table 3, Fig. 2A). For LW, LP, SN, GP and GT, autocorrelation is significant for the first and second interval, and autocorrelation is lost starting at 40 km (Fig. 2B, G, J, N-O). For PEL and PL, autocorrelation is significant only for the first interval (Fig. 2C, H). The average correlogram indicates that the phenotypic similarity is significant for the first two classes of distance and the corresponding chart (Fig. 2P) cuts the ordinate zero approximately at 50 km.

Mantel Correlograms

Mantel correlograms of phenotypic distances as a function of geographic distances, using environmental distances as a third matrix, gave highly significant results for the first two distance classes ($p = 0.0001$ and 0.0002 respectively) (Table 3, Fig. 3). The Mantel correlogram graph cuts the 0 ordinate at approximately 50 km and that from distances between 40 and 80 km the differentiation among individuals does not depend on geographic distance suggesting that individuals would be virtually isolated.

Global Autocorrelation

Global autocorrelation is significant for 11 of the traits analyzed after correction for multiple comparisons (Table 4). These results are consistent with the univariate Moran's correlograms: eight of the traits with higher global autocorrelation (LL, LW, NPL, LP, SN, GR, GP y GT) showed significant autocorrelation in the first two classes (Fig. 2 A-B, E, G, J, M-O), while PIL, ID, PW, SL, and SW did not show significant global autocorrelation (Table 4) not even for the first class of distance (Fig. 2D, F, I, K-L).

Local Spatial Analysis Getis Ord Statistics

In contrast to the global correlation, a marked variation among features and sites was observed for local autocorrelation (Fig. 4A-O). The $Gi*$ coefficient allowed the visualization of provenances groupings with similar values according to geographic area. Within each provenance, traits were observed with non-significant autocorrelation (yellow circles), negative (cold spots) represented by intense blue dots, which corresponded to values of dissimilar individuals in relation to their neighbors, and positive (hot spots), with intense red coloration corresponding to individuals with similar values (Fig. 4). No trend associated with the type of trait has been observed, since the leaf, fruit and germination traits presented hot and cold spots in different provenances. For traits such as PIL, PL, PW and SL the presence of families with non-significant $Gi*$ values was observed in all or almost all provenances (Table 4, Fig. 4D, H-I, K).
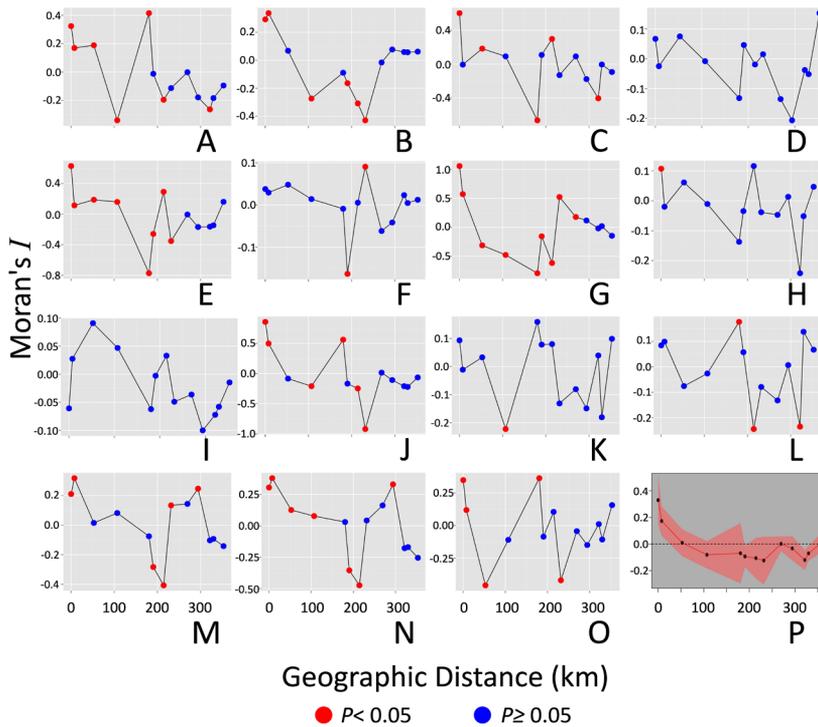
Fig. 2. Spatial autocorrelation quantified by the Moran's I index for the phenotypic traits analyzed. **A:** Length of leaflet, **B:** Leaflet width, **C:** Length of petiole, **D:** Length of pinna, **E:** Number of pinna/leave, **F:** Inter-leaf distance, **G:** Number of pairs of leaflets/pinna, **H:** Pod length, **I:** Pod width, **J:** Number of seeds, **K:** Seed length, **L:** Seed width, **M:** Germination rate, **N:** Germination power, **O:** Mean germination time. **P:** Average. Red dots represent p-values <0.05 and blue dots represent p-values ≥0.05. In all plots the x-axis represents geographic distance expressed in kilometers (km).
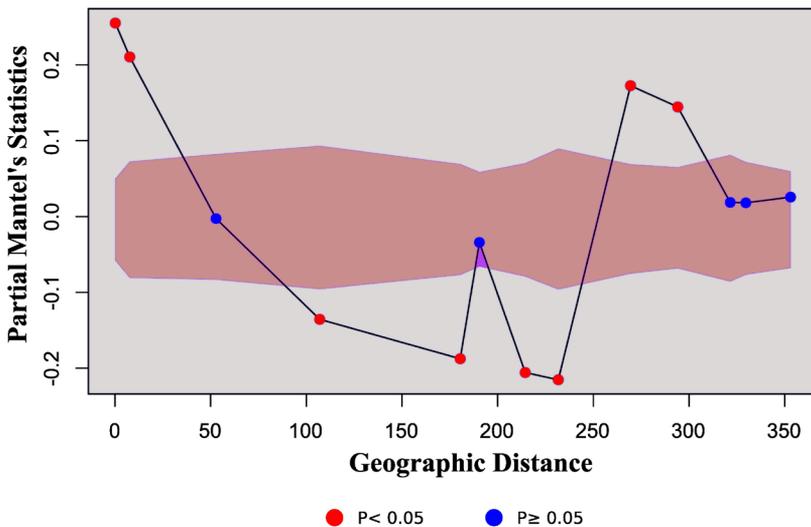
Fig. 3. Correlogram based on the partial Mantel statistic correlogram. It compares the morphological distance matrices with the geographical ones, with correction for environmental differences. The x-axis represents geographic distance expressed in kilometers (km). Red dots represent p-values <0.05 and blue dots represent p-values ≥0.05.

Table 4. Results of the global autocorrelation analysis. Provenances: ch = Charata, fo = Formosa, ic = Isla Cuba, ib = Ibarreta, lb = Las Breñas, ly = Laguna Yema, vi = Villa Dos Trece. Traits abbreviations: GR = germination rate, GP = germination power, GT = mean germination time, ID = inter-leaf distance, LL = leaflet length, LW = leaflet width, NPL = number of pinna/leave, LP = number of pairs of leaflets/pinna, PEL = length of petiole, PIL = length of pinna, PL = pod length, PW = pod width, SN = number of seeds, SL = seed length, SW = seed width. Values are presented for hot spots (hot) and cold spots (cold) for each provenance.

Provenances	ch		fo		ib		ic		lb		ly		vi	
Traits	hot	cold	hot	cold	hot	cold	hot	cold	hot	cold	hot	cold	hot	cold
LL	0	10 (2)	0	10 (5)	10 (3)	0	8	0	0	10 (5)	10	0	0	10
LW	0	10 (8)	0	10	10 (6)	0	0	8	0	10 (1)	10	0	0	10
PEL	2	8	0	10 (2)	0	10	8	0	0	10 (1)	10	0	6	4
PIL	0	10	9 (2)	1	3	7	0	8	5	5 (1)	10	0	0	10
NLP	8	2	0	10	10 (10)	0	0	8 (8)	10	0	0	10 (10)	4	6
ID	10	0	10 (8)	0	0	10 (3)	0	8 (8)	5	5	3	7	10	0
LP	0	10	0	10	10 (10)	0	8 (8)	0	0	10 (5)	0	10 (10)	0	10
PL	3	7	0	10	9 (3)	1	0	8	10 (5)	0	0	10 (9)	10	0
PW	4	6	10	0	0	10 (4)	8	0	0	10	7	3	0	10
SN	10 (10)	0	10 (5)	0	0	10 (10)	8	0	10 (5)	0	0	10 (9)	4	6
SL	10	0	8	2	0	10 (10)	4	4 (4)	10	0	0	10 (10)	10 (4)	0
SW	10	0	0	10	0	10 (10)	4	4 (4)	10 (10)	0	0	10 (10)	10	0
GR	10 (6)	0	10 (8)	0	0	10 (7)	0	8 (8)	10 (9)	0	7	3	0	10
GP	10 (6)	0	10 (10)	0	0	10 (9)	0	8 (4)	10 (10)	0	7	3	0	10
GT	0	10	0	10	10	0	0	8 (8)	9	1	10 (10)	0	6	4
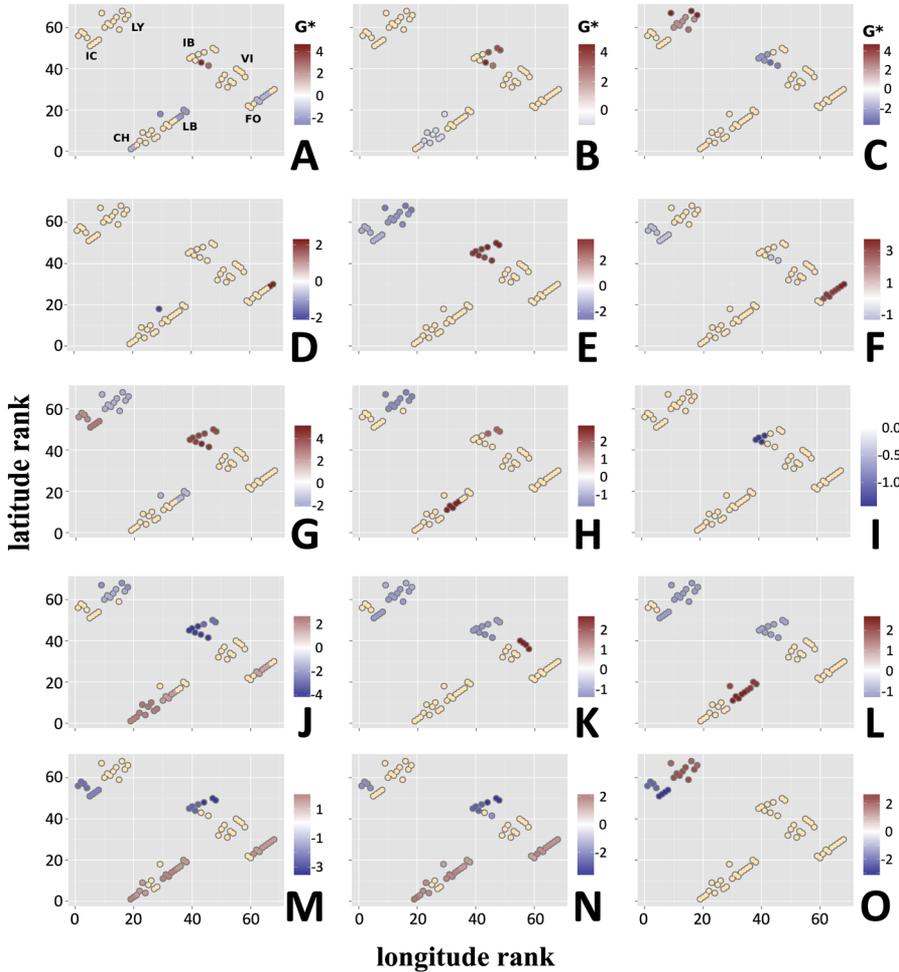
Fig. 4. Local spatial analysis for the 15 morphological traits analyzed, based on the Getis and Ord statistic Gi* (1992). The color scale to the right of each graph represents the Gi* value at each sampled point. On the left upper chart, the location of each sampled provenance is shown. Abbreviations: ch: Charata, fo: Formosa, ib: Ibarreta, ic: Isla Cuba, lb: Las Breñas, ly: Laguna Yema, vi: Villa Dos Trece. **A**: Length of leaflet, **B**: Leaflet width, **C**: Length of petiole, **D**: Length of pinna, **E**: Number of pinna/leave, **F**: Inter-leaf distance, **G**: Number of pairs of leaflets/pinna, **H**: Pod length, **I**: Pod width, **J**: Number of seeds, **K**: Seed length, **L**: Seed width, **M**: Germination rate, **N**: Germination power, **O**: Mean germination time. In all plots the x-axis and y-axis represent geographic distance expressed in kilometers (km).

DISCUSSION AND CONCLUSIONS

One of the forest species that adapts very well to a wide environmental heterogeneity is *N. alba*, that may yield local adaptation processes and autocorrelated spatial patterns. Learning about these patterns would be useful for the development of genetic improvement and conservation programs for this species. In general, it is of interest to quantify the association between phenotypic variability and genetic variability, since the morphological traits present high variability with the environment (Teich *et al.*, 2015). Unpublished preliminary results developed by our laboratory in nurseries suggest that most of the traits are associated with environmental variables of the parent's origin site.

The global autocorrelation of morphological variability analysis detected spatial structure for most of the analyzed traits, which could indicate the existence of a joint environment-foliar phenotype variation consistent with that observed by Roser (2017) at a smaller spatial scale in the same species. The Mantel correlogram revealed the existence of distribution patterns of phenotypic variation shown by a positive autocorrelation at geographic distances with the same geographical location. The disappearance of this pattern at longer distances may be explained by limited gene flow associated with short pollen and seed dispersal as described by Bessega *et al*. (2017). This fact also occurs in other species in addition to the local action of genetic drift and a strong family structure (Goncalvez, 2019).

In *N. alba*, the low dispersal rate determines that genetic differentiation increases rapidly over short distances and the results obtained with the Mantel correlogram are consistent with the analysis based on *Moran's I* index, which indicates a marked spatial structure as observed in other studies on *N. alba* (Teich *et al*., 2015; Goncalvez, 2019) and other tree species (Villareal, 2018; Ortiz *et al*., 2018). Our results indicate that phenotypic correlation among pairs of individuals would be lost on average at a distance of approximately 50 km. This result might mostly depend on the region analyzed since in *N. chilensis*, a species related to *N. alba*, Contreras Díaz *et al*. (2021) indicate that provenances separated by approximately 40 km do not show evidence of isolation.

The distribution patterns of variation may be due to geographic or environmental isolation. In the former case individuals mate randomly within a neighborhood, but are restricted from mating with more distant members (Wright, 1943). This reduction in the phenotypic similarity would be due to the fact that the dispersion of pollen and seeds in the species of the Algarobia section is generally reduced (between 5 and 31 meters), resulting in a significant decrease in the relatedness of individuals with increasing spatial distance separating them (Bessega *et al*., 2012), and with each mother plant receiving pollen from approximately seven different male parents (Bessega *et al*., 2017). In other cases, lack of genetic structuring associated with geographic origin has been observed, due to the dispersal process, and not to large distances (>500 meters), which would not act as a barrier to pollen flow (Aguirre Morales, 2017).

The provenances studied represent three complexes from two ecoregions with different environmental characteristics. In addition, some of these provenances are several hundred kilometers apart, so an important geographic isolation can be expected considering that pollen dispersal in *Neltuma* species (Algarobia section) was estimated by Bessega *et al*. (2000) in approximately 31 meters.

Likewise, previous studies in other mesquite species such as *Neltuma flexuosa* (DC.) C.E.Hughes & G.P.Lewis (Darquier *et al.,* 2013) and *Neltuma chilensis* (Molina) C.E.Hughes & G.P.Lewis (Bessega *et al.,* 2022) show that several of the leaf traits analyzed in this work would have adaptive value, so that they could be affected by the different environmental conditions of local populations, which may contribute to the observed diversity.

Neighboring families, separated by short distances, were highly related, showing high levels of kinship. However, local analysis with the *Gi** statistic revealed the existence of both hot spots and cold spots within each provenance, suggesting a complex internal structure that could reduce inbreeding by presenting hot spots with high phenotypic diversity at specific sites within populations.

The results suggest that traits characterizing leaflet size and number, seed size and number and germination are spatially structured. The analysis revealed the existence of local phenotypic patterns, which could be associated with limited gene flow. The observed spatial autocorrelation patterns and the high phenotypic diversity among the provenances could be explained on the basis of the isolation model by distance mediated by restricted gene flow proposed in the present and other analyses in *N. alba* and related species (Bessega *et al*., 2012; Roser, 2017).

AUTHOR CONTRIBUTIONS

MVV collected the material, carried out the measurements, trials in the nursery, and wrote the manuscript, BS carried out the design and collaborated with the manuscript drafting, JV carried out the experimental design, statistical analysis and final manuscript review.

Bibliography

AGUIRRE MOLARES, A. 2017. *Distribution patterns and genetic structure of Cedrela odorata and Albizia saman for the ecological conservation and restoration for the tropical dry forest in Colombia*. Thesis of Master in Biological Sciences. University National of Colombia, Colombia [online]. Available in: https://repositorio.unal.edu.co/bitstream/handle/unal/59023/1060651111.2017.pdf?sequence=1&isAllowed=y [Access: 20 september 2023].

BENJAMINI, Y. & Y. HOCHBERG. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. A Sta.* 57: 289-300.
https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

BESSEGA, C., L. FERREYRA, N. JULIO, S. MONTOYA, B. SAIDMAN & J. VILARDI. 2000. Mating system parameters in species of genus *Prosopis* (Leguminosae). *Hereditas* 132: 19-27. https://doi.org/10.1111/j.1601-5223.2000.00019.x

BESSEGA, C., C. POMETTI, M. EWENS, B. SAIDMAN & J. VILARDI. Strategies for conservation for disturbed *Prosopis alba* (Leguminosae, Mimosoidae) forests based on mating system and pollen dispersal parameters. 2012. *Tree Genet. Genomes* 8: 277-288. https://doi.org/10.1007/s11295-011-0439-6

BESSEGA, C., C. POMETTI, M. EWENS, B. SAIDMAN & J. VILARDI. 2015. Evidences of local adaptation in quantitative traits in *Prosopis alba* (Leguminosae). *Genetic* 143: 31-44.
https://doi.org/10.1007/s1070 9-014-9810-5

BESSEGA, C., C. POMETTI, B. SAIDMAN & J. VILARDI. 2017. Contribución de estudios genético poblacionales a la conservación de especies nativas de argentina de interés forestal. *Ciencia e Investigación* 67: 25-35.

BESSEGA, C., J. VILARDI, M. CONNY, B. SAIDMAN & C. POMETTI. 2022. Low genetic variation of foliar traits among *Prosopis chilensis* (Leguminosae) provenances. *J. Plant Res.* 135: 221-234. https://doi.org/10.1007/s10265-022-01378-9

BURKART, A. 1976. A monograph of the genus *Prosopis* (Leguminosae Subfam. Mimosoideae). *J. Arnold Arb.* 57: 219-249, 450-525.

CASTILLO, M., U. SCHAFFENER, B. van WILGEN, N. MONTAÑO, R. BUSTAMANTE, A. COSACOV, M. MATHESE & J. LE ROUX. 2021. Genetic insights into the globally invasive and taxonomically problematic tree genus *Prosopis*. *AoBP Plants* 13: 1-13. https://doi.org/10.1093/aobpla/plaa069

CONTRERA DIAZ, R., M. GONZALEZ GONZALEZ, F. AGUAYO CRUCES, S. GACITUA ARIAS & V. PORCILE SAAVEDRA. 2021. Analysis of the diversity and genetic structure of *Prosopis chilensis* populations between the provinces of Chacabuco and San Felipe de Aconcagua, using microsatellite molecular markers. *Bosque (Valdivia)* 42: 371-382. http://dx.doi.org/10.4067/S0717-92002021000300371

DARQUIER, M., C. BESSEGA, M. CONNY, J. VILARDI & B. SAIDMAN. 2013. Evidence of heterogeneous selection on quantitative traits of *Prosopis flexuosa* (Leguminosae) from multivariate QST–FST test. *Tree Genet. Genomes* 9: 307-320. https://doi.org/10.1007/s11295-012-0556-x.

DJAVANSHI, K. & H. POURBEIK. 1976. Germination Value - a new formula. *Silvae Genetica,* 25: 79-83 [online]. Available in: https://www.thuenen.de/media/institute/fg/PDF/Silvae_Genetica/1976/Vol._25_Heft_2/25_2_79.pdf. [Access: 20 september 2023].

FONTANA, M., V. PEREZ & C. LUNA. 2018. Effect of the geographic origin in the morphological quality of *Prosopis alba* (Fabaceae) plants. *Rev. Biol. Trop.* 66: 593-604.
http://dx.doi.org/10.15517/rbt.v66i2.33383.

GARCIA, C. 2019. Integrating the demographics and genetics of populations through the landscape genetics. *Ecosistemas* 28: 75-90. https://doi.org/10.7818/ECOS.1694.

GETIS, A. & J. ORD. 1992. The analysis of spatial association by use of distance statistics. *Geogr. Anal.* 24: 189-206.
https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

GONCALVES, L. 2019. *Special genetic structure in fragmented landscapes: A study of natural populations of curupay (Leguminosae): Anadenanthera colubrina var. cebil)*. Doctoral Thesis. University National of La Plata, Argentina. [online]. Available in: http://sedici.unlp.edu.ar/bitstream/handle/10915/73663/Documento_completo.pdf-PDFA2u.pdf?sequence=1&isAllowed=y. [Access: 20 september 2023]

HIJMANS, R. 2019. Raster: Geographic Data Analysis and Modeling. R package versión 2.9-23. https://CRAN.R-project.org/package=raster.

HUGHES, C., J. RINGELBERG, G. LEWIS & S. CATALANO. 2022. Desintegration of the genus *Prosopis* L. (Leguminosae, Caesalpinioideae, mimosoid clade). *PhytoKeys* 205: 147-189. https://doi.org/10.3897/phytokeys.205.75379

MORAN, P. 1950. Notes on continuous stochastic phenomena. *Biometrika* 37: 17-23. https://doi.org/10.2307/2332142.

MORELLO, J., S. MATTEUCCI & A. RODRIGEZ. 2012. *Ecorregiones y complejos ecosistemicos argentinos.* 1st ed. Orientación Gráfica Editora. Bs. As.

ORD, J. & A. GETIS. 1995. Local spatial autocorrelation statistics: distributional issues and an application. *Geogr. Anal.* 27: 286-306. https://doi.org/10.1111/j.1538-4632.1995.tb00912.x.

ORTIZ, N., S. GIANNONI & R. PAZ. 2018. Spatial genetic structure and genetic diversity of a natural population of *Ramorinoa girolae* in San Juan province (Argentina): An exploratory analysis. *Ecol. Austral* 28: 513–524. https://doi.org/10.25260/EA.18.28.3.0.604.

PALACIOS, R. & L. BRAVO. 1981. Natural *Prosopis* (Leguminosae) hybridization in Argentina's Chaco Region. Morphological and chromatographic evidences. *Darwiniana* 23: 3-35.

R CORE TEAM. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org (Versión 4.0.3).

ROSER, L. 2017. *Genetics of the landscape in Prosopis alba populations from the province of Santiago del Estero.* Doctoral Thesis. University of Buenos Aires.

ROSER, L., L. FERREYRA, B. SAIDMAN & J. VILARDI. 2017. EcoGenetics: An R package for the management and exploratory analysis of spatial data in landscape genetics. *Mol. Ecol. Resour.* 17: e241-e250. https://doi.org/10.1111/1755-0998.12697

SAIDMAN, B. & J. VILARDI. 1993. Genetic variability and germplasm conservation in the genus *Prosopis.* In: PURI, S., P.K. KHOSLA (eds.), *Nursery Technology for Agroforestry: Applications in Arid and Semiarid Regions*. pp. 187-198. Winrock-Oxford & IBH Publishing Co., New Delhi.

SIABATO, W. & J. GUZMAN MANRIQUE. 2019. Spatial autocorrelation and development of quantitative geography. Cuad. Geogr. *Rev. Colomb. Geogr*. 28: 1-22. https://doi.org/10.15446/rcdg.v28n1.76919.

SILVA, J. & J. NAKAGAWA. 1995. Estudo de Fórmulas para cálculo da velocidade de germinação. *Informativo Abrates* 5: 62-73.

TEICH, I., M. MOTTURA, M. BALZARINI & A. VERGA. 2015. Association between genetic and phenotypic variability with adjustment for spatial autocorrelation in *Prosopis. BAG* 26: 63-74.

VILLAREAL, M. 2018. *Effect of the Cedrela odorata L genetic variability management in the State of Hidalgo*. Thesis of Master in Biodiversity and Conservation Sciences. México.

WANG, I. & G. BRADBURG. 2014. Isolation by environment. *Mol. Ecol.* 23: 5649-5662. https://doi.org/10.1111/mec.12938.

WRIGHT, S. 1943. Isolation by distance. *Genetics* 28: 114-138. https://doi.org/10.1093/genetics/28.2.114