Biological Journal of the Linnean Society, 2020, **129**, 793–809. With 5 figures.

Ecophylogeography of the disjunct South American xerophytic tree species *Prosopis chilensis* (Fabaceae)

DANA LUCÍA AGUILAR[1], MARÍA CRISTINA ACOSTA[1,*], MATÍAS CRISTIAN BARANZELLI[1], ALICIA NOEMÍ SÉRSIC[1], JOSE DELATORRE-HERRERA[2], ANIBAL VERGA[3] and ANDREA COSACOV[1,*]

[1]*Laboratorio de Ecología Evolutiva y Biología Floral, Instituto Multidisciplinario de Biología Vegetal (IMBIV), CONICET-Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba, Argentina*
[2]*Facultad de Recursos Naturales Renovables, Universidad Arturo Prat, Iquique, Chile 4 Instituto de Fisiología y Recursos Genéticos Vegetales (IFRGV), CIAP, INTA, Córdoba, Argentina*
[3]*Instituto de Fisiología y Recursos Genéticos Vegetales (IFRGV), CIAP, INTA, Córdoba, Argentina*

Received 27 September 2019; revised 8 January 2020; accepted for publication 8 January 2020

The intraspecific evolutionary history of South American xerophytic plant species has been poorly explored. The tree species *Prosopis chilensis* has a disjunct distribution in four South American regions: southern Peru, southern Bolivia, central–western Argentina and central Chile. Here, we combined phylogeographical (based on chloroplast and nuclear markers), morphological and climatic data to evaluate the relative contribution of historical demo-stochastic and adaptive processes in differentiating the disjunct areas of distribution. The results obtained with the two molecular markers revealed two closely related phylogroups (Northern and Southern, predominating in Bolivian Chaco and in Argentine Chaco/Monte, respectively), which would have diverged at ~5 Mya, probably associated with transgression of the Paranaense Sea. Bolivia and Argentina have a larger number of exclusive haplotypes/alleles and higher molecular diversity than Chile, suggesting a long-lasting *in situ* persistence in the former and a relatively recent colonization in the latter, from the Bolivian and Argentinian lineages. The two main lineages differ in morphology and climatic niche, revealing two significant, independent evolutionary units within *P. chilensis* promoted by local adaptation and geographical isolation.

ADDITIONAL KEYWORDS: Chaco – Chilean Matorral – climatic centroid – dry forests – ecological niche modelling – Monte – leaf morphology – Miocene marine transgression – phylogeography – Pleistocene refugia.

INTRODUCTION

Xerophytic forests constitute fundamental wooded masses in South America (SA; Naumann, 2006) and are the second largest forests after the Amazon rainforest. Given the characteristic low precipitation and severe climate, these regions harbour a unique biota adapted to arid and semi-arid conditions with complex geographical distributions, being important biodiversity hotspots with intermediate to high levels of species richness and endemism (Miles *et al.*, 2006).

Although xerophytic forests in SA have great environmental (and cultural) value, the intraspecific evolutionary history of arid SA plant species has been poorly explored (but see Speranza *et al.*, 2007; Baranzelli *et al.*, 2014; Moreno *et al.*, 2018; Ossa *et al.*, 2019) and, surprisingly, only recently have tree species been the focus of phylogeographical studies in these areas (Camps *et al.*, 2018). The arid and semi-arid ecoregions, as currently known, are the result of the Andean, Sierras Pampeanas and Sierras Subandinas uplifts (Gregory-Wodzicki, 2000), Miocene marine transgressions (Hernández *et al.*, 2005), and several alluvial systems (Iriondo, 1992), which have continuously influenced the distribution and the climatic conditions of this biome from the Palaeogene until today. Other geographically important historical events in southern SA are glaciations, many of which contributed to aridization at these latitudes (Ortiz-Jaureguizar & Cladera, 2006).

*Corresponding author. E-mail: acosacov@imbiv.unc.edu.ar

Prosopis (Fabaceae) is an emblematic genus of SA dry forests. Of the 44 described species of trees and shrubs in the genus (Burkart, 1976), 15 are distributed in this ecoregion. All these species are associated with semi-arid or arid lands. Moreover, diversification and expansion of *Prosopis* in SA would have occurred in the Neogene (Catalano *et al.*, 2008), during the last pulses of the Andean range uplift (about 5–10 Mya) (Gregory-Wodzicki, 2000; Ortiz-Jaureguizar & Cladera, 2006), when extremely xeric conditions on the south-eastern and north-western side of the Andes developed. It was also suggested that the spread of the genus in SA may have been favoured in the Quaternary, when arid and semi-arid lands reached their current extent (Bessega *et al.*, 2006).

Prosopis chilensis (Molina) Stuntz emend. Burkart, locally known as 'algarrobo blanco' or 'algarrobo chileno', is a tree species with a disjunct geographical distribution in four regions (Burkart, 1976). The most extensive of these areas occurs in Argentina from 26°S to 34°S, with the species occurring in the north-eastern part of the Monte Desert and in the southernmost areas of the Dry Chaco (Olson *et al.*, 2001), whereas the second largest distribution area is the northern Dry Chaco in Bolivia, between *c.* 16°S and 22°S. The species is also distributed in central Chile, across Chilean Matorral, between *c.* 29°S and 34°S, and in Peru, between *c.* 13°S and 16°S (Burkart, 1976; Burghardt *et al.*, 2010). Archaeobotanical and palaeoecological data suggest that the species would have appeared to the west of the Andes in the late Holocene, probably associated with human activity (McRostie *et al.*, 2017). However, no phylogeographical study has allowed us to unravel the evolutionary history of this species and of its disjunct distribution. Phylogeographical studies of disjunctly distributed species are particularly important because they may provide information about the past relationships within the current allopatric group; such information may help to identify barriers to gene flow that allow us to reconstruct, at least in part, the biogeographical history of the region (Turchetto-Zolet *et al.*, 2012).

According to its wide distribution, *P. chilensis* exhibits considerable variation in phenotypic characters, mainly in leaf morphology (Burkart, 1976). In addition, geographical isolation could have resulted in morphological differentiation of disjunct populations via genetic drift (i.e. non-adaptive divergence) or, alternatively, via local adaptation (i.e. adaptive divergence). A previous genetic study based on nuclear markers [i.e. isozymes and random amplification of polymorphic DNAs (RAPDss)] performed in Argentine *P. chilensis* populations suggested that genetic differentiation among populations would be better explained by adaptive rather than neutral (i.e. geographical isolation) processes (Ferreyra *et al.*, 2010). Non-coding fragments of the chloroplast genome are the most appropriate markers for phylogeographical studies due to their uniparental inheritance and their capacity to detect neutral processes of evolution (Avise, 2000). In addition, ecological niche modelling (ENM), commonly used to characterize the climatic niche of allopatrically distributed populations, combined with phylogeographical and morphological data, i.e. an ecophylogeographical perspective (Avise, 2000), can provide a powerful means for evaluating the relative contribution of adaptive factors in differentiating populations with marked discontinuities in their geographical ranges (e.g. Baranzelli *et al.*, 2014, 2018).

Here we performed a comprehensive morphological, climatic and genetic study of the tree species *P. chilensis*. We characterized the leaf morphology and climatic niche of each disjunct area of distribution and then analysed sequence variation of both chloroplast (cpDNA) and nuclear (nDNA) markers to elucidate the evolutionary history of this species. We expect a disruptive pattern between ecological and phylogeographical characterization of *P. chilensis*, suggesting that morphological characteristics respond to local adaptations, regardless of genealogical history.

Thus, our aims were to: (1) evaluate the extent of differentiation among disjunct areas in leaf morphological traits and climatic conditions, as evidence of local adaptation; (2) infer the phylogeographical history underlying the distributional pattern of *P. chilensis*; and (3) propose a possible biogeographical scenario for the distribution and diversification of *P. chilensis*.

MATERIAL AND METHODS

STUDY SPECIES

Prosopis chilensis, a species belonging to Section *Algarobia*, is a flat-topped tree that can reach 12 m in height. It has bipinnate leaves and distant leaflets on the rachis. Flowers are small, densely gathered in cylindrical, spike-like racemes 7–12 cm in length. Fruits are indehiscent pods that are compressed, coriaceous to subligneous, elongated straight to slightly curved, and light yellow (Burkart, 1976). The species is pollinated by insects, being mostly outcrossing (although selfing can occur), while seed dispersal is endozoic (Burkart, 1976; Bessega *et al.*, 2000). Before the introduction of livestock by Europeans, the main vectors for seed dispersal of *Prosopis* species would have been small or sloth herbivorous mammals, some of them now extinct (Mooney *et al.*, 1977). Both these pollen and seed dispersal strategies are usually associated with limited dispersal (Burkart, 1976; Bessega *et al.*, 2000).

Sample collection

Samples were collected from four localities in Chile, 14 in Argentina and 17 in Bolivia, representative of most of the entire native range of the species (no samples from Peru were available; Supporting Information, Table S1). Within each sampled locality, 2–7 individuals were sampled at least 100 m apart from one another. Geographical coordinates were recorded with a handheld GPS. Taxonomic identification of the specimens was based on Burkart (1976). Despite the sampling effort, no specimens of this species were found at intermediate localities between the distribution areas in Bolivia and Argentina, which confirms the disjunct distribution of the species mentioned by Burkart (1976).

Morphological traits

From each sampled tree, five totally expanded leaves were collected from different places of the canopy and photographed (N_{trees} = 112; N_{leaves} = 560; Supporting Information, Table S1). Eight leaf characters usually used for species identification (Burkart, 1976; Bessega *et al.*, 2009; Ferreyra *et al.*, 2013) were measured: pinna length (PIL), number of leaflet pairs per pinna (NLeP), leaflet length (LeL), leaflet width (LeW), leaflet area (LeA), leaflet shape (LeL/LeW), leaflet apex area (LeAPX) and distance between leaflets (DBLe) calculated as PIL/NLeP (Fig. 1A). Measurements were taken on digital images using the software HOJA v.3.4 (software available from the author upon request: A. Verga, INTA-IFRGV).

Morphological analyses

The mean and standard errors of each of the eight leaf variables were calculated. The ordination of individuals according to their morphological variation was visualized using a principal components analysis (PCA; Hotelling, 1933) performed on the 112 sampled trees and the eight leaf morphometric variables, which were previously $\log_{10}$-transformed. Biplots were obtained and individuals were identified by their geographical provenance (i.e. Argentine Chaco/Monte, Bolivian Chaco, Chilean Matorral). Then, to analyse possible significant phenotypic differences among geographical groups, we performed a multivariate analysis of variance (MANOVA) using all the measured morphological traits. Additionally, with the aim of finding the subset of morphological variables that best explains the differences among disjunct groups, we performed a linear discriminant analysis (LDA). LDA is a supervised statistical learning technique which, unlike unsupervised methods such as cluster analysis, uses a priori knowledge about which observation belongs to which group. LDA tests the hypothesis that groups are clustered based on the available set of predictors. In the present study, the predictors were leaf morphometric variables and each tree was re-classified according to its geographical provenance (i.e. Argentine Chaco/Monte, Bolivian Chaco, Chilean Matorral). The classification obtained with the LDA was compared by using a cross-classification table and a chi-square test. Analyses were performed using Infostat v.2018 (Di Rienzo *et al.*, 2016).

Environmental variation and niche comparisons

We obtained 109 trustable presence points for *P. chilensis* from field trips performed between 2015 and 2017. Additional points were obtained from the Global Biodiversity Information Facility (GBIF, 2019; https://www.gbif.org/, N_{loc} = 122). Current climatic data with a grid cell resolution of 0.25° (~5-km² cell) for all these georeferenced locations were downloaded from the WorldClim database v.1.4 (http://www.worldclim.org/; Hijmans *et al.*, 2005); following Braunisch *et al.* (2013), we used all 19 bioclimatic variables derived from the monthly temperature and rainfall values. One-way ANOVAs and MANOVAs were performed using the obtained environmental values to test whether the observed environmental features significantly differed among geographical provenances. Then, to find the climatic conditions that best explain the differences among geographical groups, an LDA was performed using all bioclimatic variables as discriminant variables, and geographical groups as the grouping variable.

The climatic niches for each geographical group were then characterized by combining environmental variables with model prediction, following Martínez-Meyer *et al.* (2013). First, a PCA was performed to reduce the climatic data to three axes explaining most of the climatic variance (93.7%). Second, for each geographical group, the climatic centroid in the environmental space was calculated as the mean value of suitable pixels in each climatic dimension. Then, the Euclidean distances from all pixels to the ecological niche centroid were calculated, and the climatic niche ellipsoid that included 95% of the occurrence data for each geographical group (i.e. fundamental niche; Peterson *et al.*, 2011) was constructed. These ellipsoids defined the internal structure of tolerable climatic conditions for each group, with the optimal climatic conditions being placed at the centre of this hypervolume, i.e. the centroid (Maguire, 1973). Thus, if disjunct areas of the distribution of *P. chilensis* are not climatically different, we expect the centroid and the hypervolume of each group to overlap; if

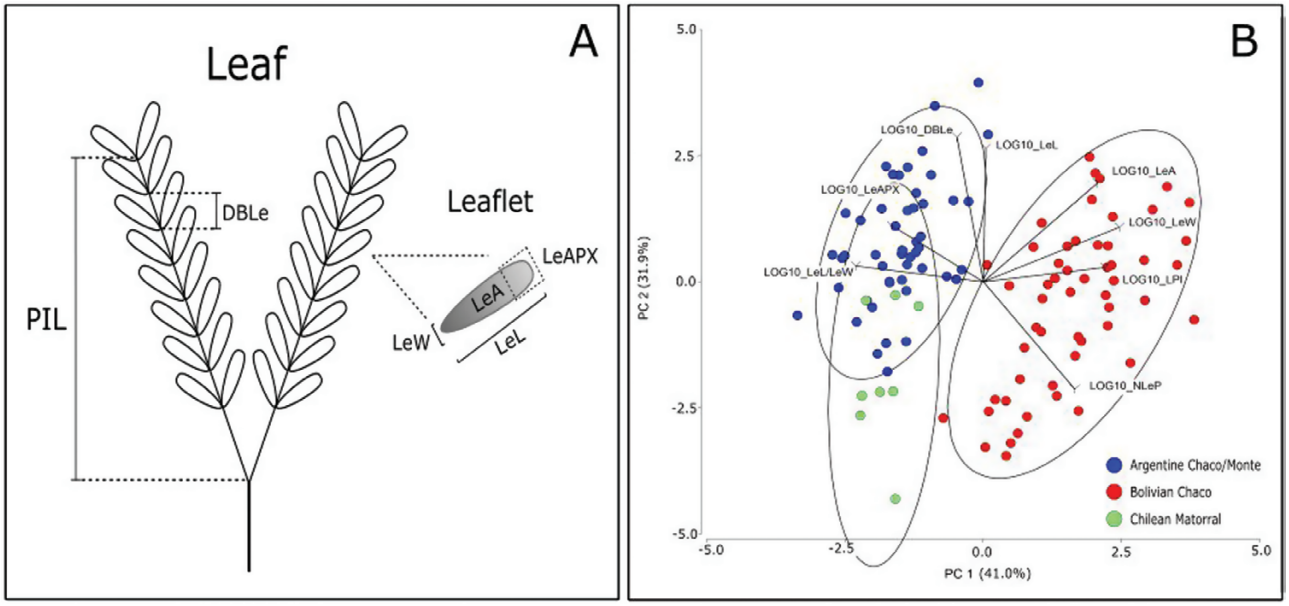

Figure 1. A, schematic representation of leaf morphological traits measured in this study: pinna length (PIL), leaflet length (LeL), leaflet width (LeW), leaflet area (LeA), leaflet apex area (LeAPX) and distance between leaflets (DBLe). B, biplot of the first two axes of the principal components analysis performed with morphometric variables. Blue dots: Argentine Chaco/Monte individuals; red dots: Bolivian Chaco; green dots: individuals from the Chilean Matorral. The amount of variance explained by each axis is provided in parentheses. For morphometric variable abbreviations, see panel A and the Material and Methods.

they are divergent groups, not only spatially but also climatically, we also expect a divergent pattern of their niche ellipsoids. The ellipsoid of each geographical group was visualized in an environmental space using three-dimensional plots and *c.* 18 000 background points from SA. Finally, the ellipsoids were spatially projected onto a map of SA, and favourability values were assigned to each pixel of the map according to its distance to the climatic conditions at the centroid. All these analyses were performed in R v.3.2.3 (R Core Team, 2017). For environmental variation analyses we used the *multcomp* (Hothorn *et al.*, 2008), *lsmeans* (Lenth, 2016), *sciplot* (Murdoch, 2017), *candisc* (Friendly & Fox, 2017) and *Lattice* (Sarkar, 2008) packages, and for niche modelling analyses we used the *NicheToolBox* package (Osorio-Olvera *et al.*, 2018).

DNA extraction, amplification and sequencing

A subset of sampled individuals (2–6) were analysed genetically (Supporting Information, Table S1). Total DNA was extracted from 40 mg of leaf material per tree using the NucleoSpin Plant II kit (Macherey-Nagel, 2014). Nuclear intergenic spacers ITS1, 5.8S and ITS2 (primers *ITS5-ITS4*; White *et al.*, 1990) and the non-coding chloroplastic region *ndhF-rpL32* (primers *ndhF-rpL32*; Shaw *et al.*, 2007) were sequenced in 29 and 61 *P. chilensis* individuals, respectively (Table S1). The *ndhF-rpL32* region showed the greatest variation among several surveyed loci (*trnQ-rps16*, *trnH-psbA*, *rpl32R-ndhF*, *rpl32F-trnL*, *trnD-trnT*; Shaw *et al.*, 2007).

The PCR mix contained 1 µL of template DNA (100 ng/µL), 0.15 µL GoTaq DNA polymerase (Promega, Madison, WI, USA), 6 µL Green GoTaq reaction buffer (Promega), 0.6 µL of dNTPs, 1.2 µL of each primer and 19.85 µL of water in a total volume of 30 µL. Chloroplast and nuclear DNA amplifications consisted of an initial denaturation step at 80 °C for 5 min, followed by 30 cycles of denaturation at 95 °C for 1 min, annealing at 48 °C for 1 min, a ramp of 0.3 °C/s to 65 °C, extension primer at 65 °C for 4 min, a final extension at 65 °C for 5 min and a final hold at 13 °C (Shaw *et al.*, 2007). Sequencing was performed by Macrogen Inc. (Seoul, South Korea, http://dna.macrogen.com). Electropherograms were visualized and edited using ChromasPro v.1.7.7. Sequences were aligned with BioEdit (Hall, 1999) and MAFFT v.6 (Katoh *et al.*, 2002) using default parameters and were adjusted manually. Gaps were coded as binary characters using simple indel coding (Simmons & Ochoterena, 2000). Accessions representing all sequences were deposited in GenBank.

Genetic analyses

Haplotype network and geographical distribution

Haplotypes were inferred with DnaSP v.5 (Rozas *et al.*, 2003) and haplotype genealogy was reconstructed for both cpDNA and nDNA matrices using the

median-joining algorithm implemented in Network v.5.0.0.1 (Bandelt *et al.*, 1999). Ambiguous connections (loops) in the networks were resolved using predictions from coalescent theory (Crandall & Templeton, 1993), according to the geographical criterion. Haplotype frequencies of each site were depicted on a map of the study region with Quantum GIS v.2.18 (Quantum GIS Development Team, 2016).

Population genetic structure and genetic diversity

To identify genetic populations and infer the presence of past historical barriers among groups of populations, a spatial analysis of the molecular variance with and without constraint for the geographical composition of the groups was implemented in SAMOVA v.2.0 (Dupanloup *et al.*, 2002). Analyses were run to evaluate different numbers of groups (K) ranging from $K = 2$ to 5 based on 10 000 simulated annealing steps. The best clustering for each K value was selected based on the among-group component (FCT) of the overall genetic variance. For each retrieved genetic group, the following genetic diversity parameters were estimated: haplotype number, number of polymorphic sites (S), haplotype diversity (h; Nei, 1987), nucleotide diversity (π; Nei, 1987) and mean number of pairwise differences (p; Tajima, 1983) using Arlequin v.3.11.

Demographic history analyses

To assess the population demographic history of the groups obtained by SAMOVA, Tajima's D and Fu's F_S (Tajima, 1989; Fu, 1997) neutrality tests were performed. Significance for both values was calculated from 10 000 simulated samples using a coalescent algorithm. Additionally, a mismatch distribution of pairwise differences among individuals was calculated to test the hypothesis of a sudden demographic expansion model (Excoffier & Lischer, 2010). The goodness of fit of the observed mismatch distribution to that expected under a sudden expansion model was evaluated with the sum of squared deviations (SSD) using parametric bootstrapping (10 000 replicates). Neutrality tests and mismatch distribution analyses were performed using Arlequin v.3.11 (Excoffier & Lischer, 2010) and DnaSP. Finally, to visualize changes in effective population size (N_e) over time in the groups that presented signals of demographic expansion (see Results) we used a Bayesian Skyline Plot (BSP) implemented in BEAST v.1.7.5 (Drummond *et al.*, 2005, 2012), under the HKY model selected with j-ModelTest v.2.1.9 (Darriba *et al.*, 2012). For these analyses, we used a strict molecular clock with a uniformly distributed prior based on the mutation rate of cpDNA (0.001–0.01 substitutions/site/Myr; Alsos *et al.*, 2005). The number of grouped intervals was set to 5, with an initial popsize value of 200, a minimum of 0.1, a maximum of 1000, and a piecewise constant model. Two Monte Carlo Markov chains (MCMCs) starting with a random tree were run for 1×10^7 generations with parameters sampled every 1000 steps. Convergence of estimated parameters [effective sample size (ESS) > 200) was verified using Tracer v.1.6 (Rambaut *et al.*, 2014). After a 10% burn-in, log files were combined and demographic plots were visualized. All these demographic analyses were performed using only the cpDNA sequence matrix, because the sampling size was suitable for this analysis (Hall *et al.*, 2016).

Molecular dating

To reconstruct phylogenetic relationships between cpDNA and nDNA haplotypes and to estimate divergence times, a Bayesian phylogenetic reconstruction was performed in BEAST. For those intraspecific nodes (lineages) with high posterior probability values (PP > 0.90), time to the most recent common ancestor (tMRCA) and high-density posterior intervals (HDP) were calculated.

In all cases, Bayesian analyses were performed using one sequence for each retrieved haplotype/allele. Outgroups consisted of sequences of several species of the genus *Prosopis* (*P. affinis*, *P. alba*, *P. alpataco*, *P. argentina*, *P. caldenia*, *P. flexuosa*, *P. glandulosa*, *P. hassleri*, *P. juliflora*, *P. kuntzei*, *P. laevigata*, *P. nigra*, *P. pallida*, *P. palmeri*, *P. pubescens*, *P. reptans*, *P. ruscifolia*, *P. sericantha*, *P. velutina* and *P. vinalillo*), species of the subfamily *Mimosoideae*, and the genera *Acacia*, *Inga* and *Leucaena*, which were downloaded from GenBank or sequenced specifically for this study (Supporting Information, Table S2).

The best-fit models of nucleotide substitution selected in J-ModelTest were HKY for the cpDNA matrix and GTR+I+G for nDNA. Three independent analyses for each marker were run using different calibration settings (see Supporting Information, Table S3). The MRCA of the four genera (i.e. *Prosopis*, *Acacia*, *Inga* and *Leucaena*) was set to 42.4 ± 10 Mya, and the divergence time between *Acacia* and *Inga* to 30.5 ± 3.1 Mya; calibration was based on the mean ages derived from previous analyses (nodes 10 and 6 of Lavin *et al.*, 2005, respectively). The age of the genus *Prosopis* was set at 33.9 ± 1.0 Mya, based on fossil pollen grain of *Prosopis quesnelli*; this value stands as the first record of the genus, and dates from the Early Oligocene (Piel, 1971; Caccavari, 1996). A similar age was estimated for a fruit of *Prosopis lazarii* found in the palaeoflora of Puebla, Mexico (Magallón-Puebla & Cevallos-Ferriz, 1994). Finally, we also included mutation rate priors for non-coding regions of cpDNA (0.001–0.01 substitutions/site/Myr; Alsos *et al.*, 2005), and for the nuclear intergenic

spacer region (0.0031 ± 0.0003 substitutions/site/Myr) in Fabaceae (Kay *et al.*, 2006).

For all analyses, we used a lognormal relaxed clock model and a Yule tree prior. MCMCs were run with 1×10^8 iterations with samples taken every 10 000 steps. Convergence of estimated parameters (ESS > 200) was evaluated using Tracer v.1.6. After a 10% burn-in, a Maximum Clade Credibility (MCC) tree was obtained in TreeAnnotator (Drummond *et al.*, 2012) and visualized in FigTree v.1.4.2 (Rambaut & Drummond, 2012). Although all our estimates are tentative and should be interpreted with caution, they provide approximations that allow us to hypothesize scenarios of possible lineage divergence.

RESULTS

Morphological traits

Leaf morphological characters varied among the three provenances (Table 1). The first two principal components (PC1 and PC2) obtained from the analysis of leaf morphological traits explained 72.9% of the total variability among individuals (Fig. 1B). High values of PC1 (which accounted for 41.0% of the total variability) were positively correlated with the number of leaflet pairs per pinna (NLeP), pinna length (PIL), leaflet width (LeW) and leaflet area (LeA), and negatively correlated with leaflet shape (LeL/LeW) (Fig. 1). PC2 was positively related to the distance between leaflets (DBLe), leaflet length (LeL) and leaflet area (LeA), and negatively associated with the number of leaflets per pinna (NLeP), accounting for 31.9% of the total variability. The first axis differentiated *P. chilensis* from Bolivia from those from Chile and Argentina, showing that Bolivian trees have leaves with larger pinnae, wider leaflets and higher number of leaflets per pinnae than the trees from Argentina and Chile. The second axis revealed a gradual differentiation between trees from Argentina and Chile, showing that the latter have a smaller distance between leaflets, higher number of leaflets per pinna and shorter leaflets (Fig. 1). The MANOVA performed using all measured traits and the post-hoc Hotelling test with a Bonferroni correction indicated significant differences among the three geographical groups ($F_{112;109} = 63.74$; $P < 0.0001$; Table 1). The total error rate of the LDA performed with the foliar traits was 1.85%. All individuals from Argentina and Chile were classified correctly, whereas one Bolivian tree was assigned to the Argentine group (see Supporting Information, Table S4). The variables that best explained differences among geographical groups were leaflet width, leaflet length and leaflet shape.

Table 1. Population mean (±SD) values of leaf traits measured for the three disjunct areas of distribution of *Prosopis chilensis* in southern South America

Distribution area	*N/n*	PIL		NLeP		LeL		LeW		LeA		LeAPX/LeA		LeL/W		DBLe	
Argentine Chaco/Monte	14/50	11.17	±1.83	17.41	±3.15	2.71	±0.54	0.18	±0.03	0.39	±0.13	0.18	±0.01	14.88	±2.62	0.65	±0.12
Bolivian Chaco	18/54	14.45	±2.11	31.47	±9.68	2.42	±0.34	0.30	±0.11	0.59	±0.23	0.13	±0.03	8.83	±2.78	0.50	±0.16
Chilean Matorral	4/8	9.11	±1.66	18.56	±2.83	1.89	±0.45	0.14	±0.02	0.27	±0.06	0.11	±0.01	14.07	±4.03	0.51	±0.17

The number of sampled localities (*N*) and individuals (*n*) are indicated. Codes for the different morphological variables are provided in the Material and Methods and Figure 1.

Environmental variation and niche comparisons

All 19 bioclimatic variables differed significantly among the three geographical areas of *P. chilensis* distribution ($P < 0.001$; Supporting Information, Fig. S1 and Table S5) and posterior comparisons showed that the Bolivian Chaco locations are the most different of the three regions (Table S5). Bolivian Chaco locations occur in subtropical environments, with the highest mean values but lowest variance in annual and seasonal rainfall, and temperatures, whereas Chilean Matorral locations occur in the driest environment, with the lowest annual and seasonal precipitation. Argentine Chaco/Monte locations occur in a semi-arid environment, with intermediate rainfall but higher annual and seasonal temperature than Chilean locations (Fig. S1). LDA canonical axis 1, which separated the three geographical groups, explained 72.84% of the total variation; it was positively correlated with annual precipitation and with temperature of the coldest quarter, and negatively associated with annual temperature, annual temperature range and precipitation of the driest quarter. Axis 2, which separated locations of Bolivia from those of Argentina and Chile, was positively correlated to precipitation seasonality and temperature of the warmest quarter, and negatively correlated to precipitation of the warmest quarter (Fig. 2A). In the cross-validation procedures, all individuals were correctly assigned to their respective geographical groups (Table S6).

Ellipsoids constructed from the climatic centroid of each geographical group based on the climatic

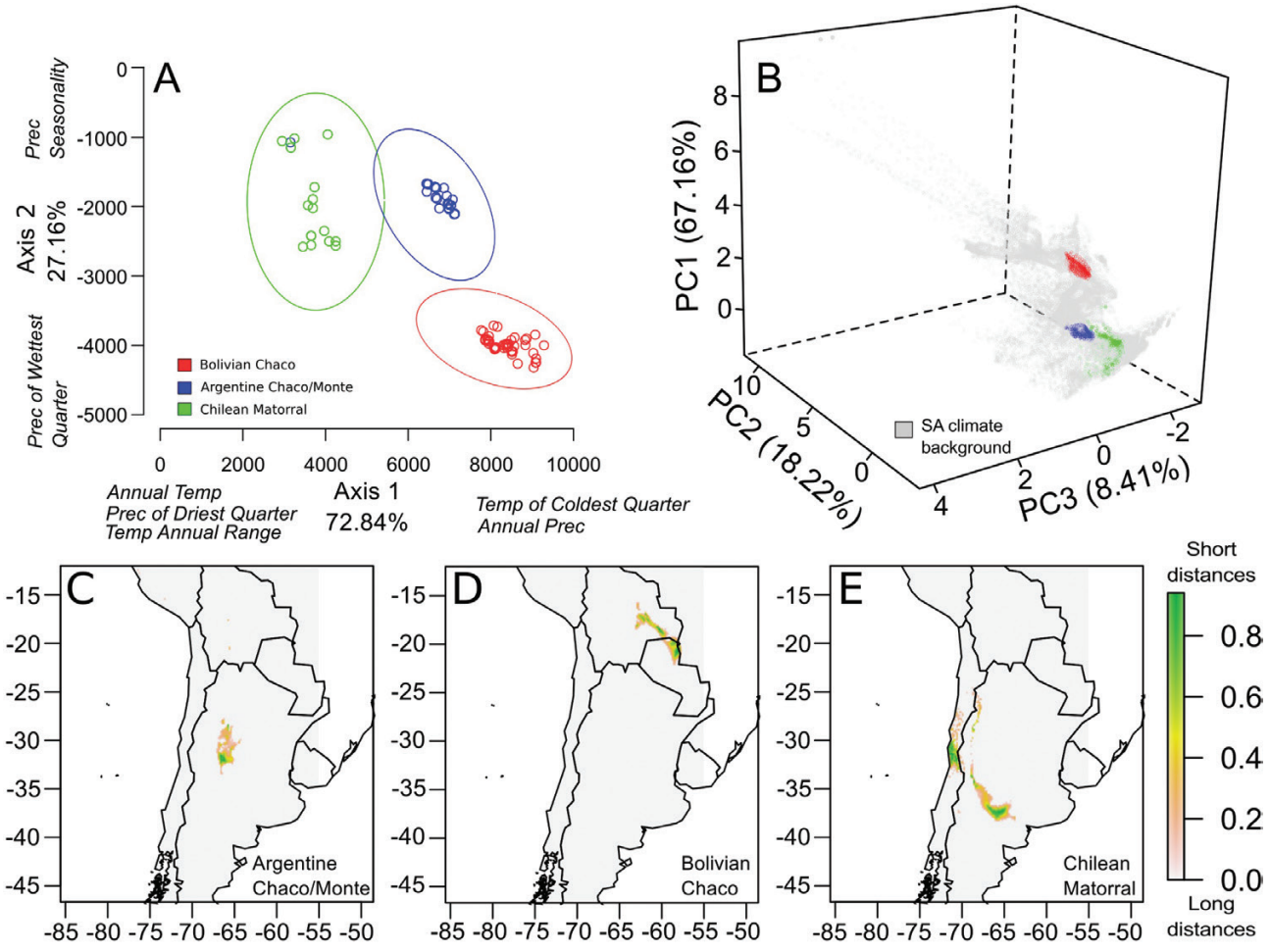

Figure 2. Environmental variation and niche comparisons among three geographical provenances of *Prosopis chilensis* (N_{loc} = 121; Argentine Chaco/Monte in blue, Bolivian Chaco in red and Chilean Matorral in green). A, ordination of the three groups along the first two axes of the linear discriminant analysis. The variables that most contribute to the axes are indicated. B, triplot of the first three principal components of climatic space showing the climatic background points and the climatic points corresponding to each group. C–E, geographical projection of ellipsoids in geographical space. The colour gradient represents favourability values according to distance to the climatic centroid.

background of SA did not overlap in ecological space (Fig. 2B). The ellipsoids for Bolivian Chaco and Argentine Chaco/Monte regions were smaller than that for the Chilean Matorral, indicating more restrictive climatic conditions in the two former regions; the Chilean ellipsoid appeared nearly to overlap with the Argentine one, indicating some relationship in the fundamental niche between these groups. The geographical projections of the ellipsoids (Fig. 2C–E) for Bolivian Chaco and Argentine Chaco/Monte were highly coincident with their current distribution, whereas for the Chilean Matorral, the projection of the ellipsoid covered its current distribution but was also projected onto the Argentine Chaco/Monte region.

Genetic analyses

DNA sequence matrices

The matrix obtained for the *ndhF-rpL32* region of the chloroplast genome was 754 bp in size. Eight polymorphic sites were retrieved: four corresponded to point mutations, three to deletions/insertions of one nucleotide, and one to a 41-bp fragment identified as a duplication. This duplication was eliminated and coded as a binary character to be considered as a single evolutionary event. After this editing, eight cpDNA haplotypes were identified among the 60 amplified sequences. Regarding the amplification of the ITS1, 5.8S and ITS2 regions of the nuclear genome, a 640-bp matrix was obtained with 33 polymorphic sites. Of these, 32 corresponded to point mutations and one to a deletion/insertion. Thirteen alleles were identified among the 29 obtained sequences. The aligned databases were deposited in TreeBase.

Haplotype networks and geographical distribution

Genealogical relationships and the geographical distribution of the eight haplotypes (cpDNA), and the 13 alleles (nDNA) are shown in Figure 3. The cpDNA network shows two most frequent and widespread haplotypes (H1 and H6; Fig. 3A). H1 predominates in the northernmost area of the Bolivian Chaco (retrieved in 17 of the 24 sampled individuals) but is

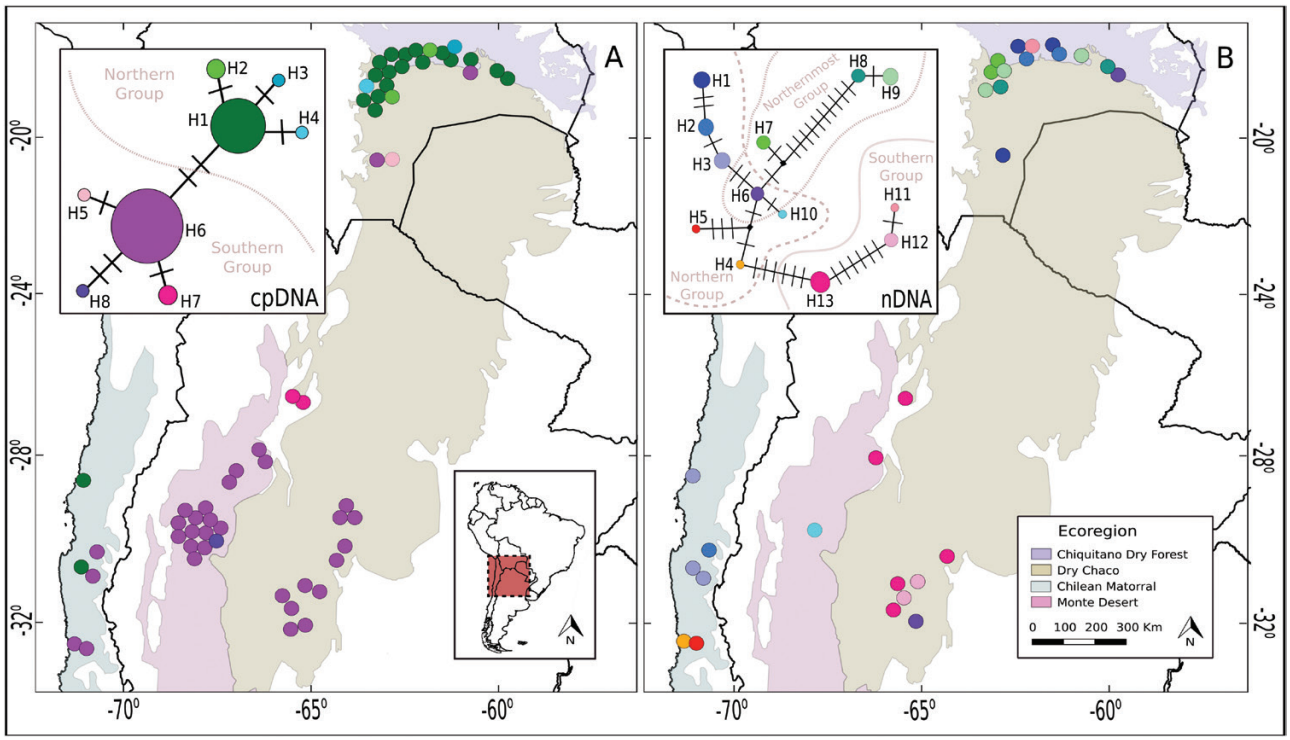

Figure 3. Geographical distribution and genealogical relationships of (A) the eight chloroplast haplotypes (cpDNA) and (B) the 13 nuclear alleles (nDNA), recovered from sampled *Prosopis chilensis* individuals. On the maps, the sampled individuals with their respective haplotype/alleles are shown. Haplotype colours correspond to those shown in the networks. Shaded areas indicate different South American ecoregions based on Olson *et al.* (2001). In the networks, haplotypes/alleles are designated with numbers, and circle sizes are proportional to their frequency. Black squares represent missing intermediate haplotypes/alleles not observed in the analysed individuals and cross hatches represent mutational steps. Dashed and filled grey lines indicate each detected genetic group retrieved by SAMOVA.

also distributed in the Chilean Matorral (2/6), whereas H6 prevails in Argentine Chaco/Monte (29/32), but is also found in the Chilean Matorral (4/6) and in Bolivia (2/24). Thus, in the Chilean individuals only the two most frequent and most widely distributed haplotypes were retrieved. H2, H3 and H4 are low-frequency haplotypes exclusive to Bolivia, derived from H1 by one mutational step; H5 is exclusive to southern Bolivia, and H7 and H8 are exclusive to Argentina, derived from H6 by a few mutational steps (Fig. 3A). The nDNA network shows low-frequency alleles separated by numerous mutational steps (Fig. 3B); in general, they are restricted to specific geographical areas, found in a single locality or shared with nearby locations. Only H6, which is a low-frequency haplotype with a central position in the network, is shared by Bolivia and Argentina, whereas Chile shares only one haplotype with Bolivia (H2). The remaining alleles are exclusive to each disjunct area: Bolivia has five exclusive haplotypes (H1, H7, H8, H9 and H11), whereas three are exclusive to Argentina (H10, H12 and H13) and three to Chile (H3, H4 and H5; Fig. 3B).

Population genetic structure and genetic diversity

The optimal partitioning of cpDNA genetic diversity inferred from SAMOVA was obtained with $K = 2$ (FCT = 0.84; $P < 0.0001$; Supporting Information, Table S7) with and without constraint for the geographical composition of the groups. Analyses for the remaining K values generated single population groupings although the differences among the groups were maximized (Table S7). Thus, two genealogically well-defined genetic populations were recognized: a group that included all the individuals present in Argentina, southern Bolivia and three Chilean individuals (hereafter termed the Southern group), and the Northern group including the remaining individuals located in the northernmost area of the Bolivian Chaco, and one individual from Chile (Fig. 3A). For nDNA, optimal genetic diversity partitioning was obtained with $K = 3$ (FCT = 0.73; $P < 0.0001$; Table S7) with and without constraint for the geographical composition of the groups. Analyses for the remaining K values generated population groupings with one or two individuals, although the difference among the groups was maximized (Table S7). The three retrieved genetic groups consisted of a genetic group clustering mainly individuals from Bolivia and Chile, but also from Argentina (the Northern group), a group including individuals from Argentina and one individual from Bolivia (the Southern group) and a third group exclusively located in the northernmost area of the Bolivian Chaco (the Northernmost group; Fig. 3B). Molecular diversity indices for the Northern groups were higher than those retrieved for the Southern groups (Table 2).

Demographic history analyses

Southern and Northern chloroplast groups showed evidence of recent demographic expansion based on the SSD index, because the observed mismatch distribution did not differ from an expected sudden expansion model. This scenario was supported by the negative and significant values of Tajima's D and Fu's F_s parameters obtained for each group, confirming that the two genetic populations are not in mutation–drift balance (Table 2). The Bayesian skyline analyses performed independently for the two groups showed a slight increase of population size that would have started at about 200 kya (Fig. 4).

Molecular dating

The MCC tree based on cpDNA (Fig. 5A) recovered *P. chilensis* as a monophyletic group with high support (PP = 1) and the Northern and Southern groups as reciprocal monophyletic (although the Southern clade has low support) groups that diverged ~5.22 Mya [95% highest posterior density (HPD): 1.13–11.91 Mya]. Within the Northern phylogroup, diversification would have begun at ~1.92 Mya (95% HPD: 0.30–3.98 Mya) and the Southern phylogroup at ~2.71 Mya (95% HPD: 0.37–7.02 Mya). The MCC tree based on nDNA (Fig. 5B) also recovered *P. chilensis* as a monophyletic group with high support (PP = 1). Two main clades were retrieved within *P. chilensis*, which would have diverged at ~8.03 Mya (95% HPD: 3.42–15.08 Mya): one of them exclusively nesting haplotypes located in the northernmost area of Bolivia (called the Northernmost phylogroup, NMP), and the other one nesting two subclades, called the Northern and Southern phylogroups (NP and SP). The NP includes haplotypes distributed mainly in Bolivia and Chile, but also in Argentina, while the SP also nests haplotypes distributed in the three disjunct areas but mainly in Argentina and Chile (Figs 3B,5B). The split among the three main clades would have occurred ~5.35 Mya, whereas diversification within these clades would have occurred between 0.3 and 1.56 Mya. The results of each calibration scheme are reported in Supporting Information, Table S3.

DISCUSSION

ECOGEOGRAPHICAL PATTERNS

Previous studies have suggested considerable morphological variation in *P. chilensis* across its geographical range (Burkart, 1976;

Table 2. Sample size (N), haplotypes/alleles (H), polymorphic sites (S), molecular diversity indices and demographic analyses of the retrieved genetic groups (based on chloroplast and nuclear markers) of *Prosopis chilensis*; haplotype diversity (h), nucleotide diversity (π), mean number of pairwise differences (p), Tajima's D, Fu's F_s and sum of squared deviations (SSD) of the mismatch distribution analyses are shown

Marker	Group	N	H	S	Diversity indices			Demographic analyses		
					h (SD)	π (SD)	p (SD)	D	F_s	SSD
cpDNA	Northern	23	4	3	0.32 (0.12)	0.0005 (0.001)	0.34 (0.35)	−1.48	**−2.32****	**0.002**
	Southern	37	4	4	0.20 (0.08)	0.0004 (0.0004)	0.27 (0.3)	**−1.75****	**−2.50***	**0.001**
nDNA	Northernmost	6	3	12	0.73 (0.16)	0.006 (0.004)	4.26 (2.45)	–	–	–
	Northern	15	7	11	0.90 (0.05)	0.005 (0.003)	3.49 (1.88)	–	–	–
	Southern	8	3	7	0.60 (0.16)	0.005 (0.003)	3.46 (1.97)	–	–	–

Results consistent with demographic expansion are shown in bold. *$P < 0.05$; **$P < 0.01$.

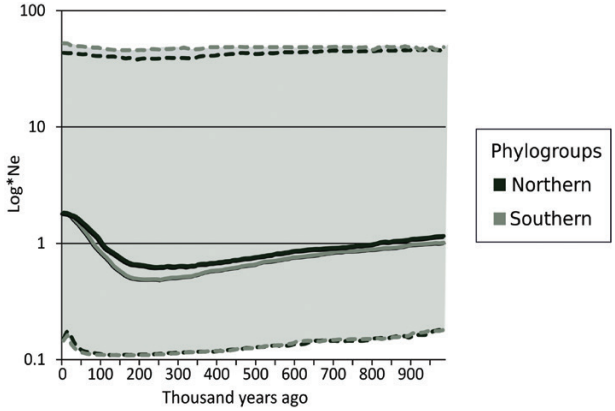

Figure 4. Bayesian Skyline Plot analyses showing variation through time in the effective population size (N_e) of the Northern and Southern genetic groups based on cpDNA. The y-axis represents effective population size expressed on a logarithmic scale. Bold lines correspond to the median values of the effective population size over time, and the shaded areas represent the 95% highest posterior densities over the median estimates throughout the coalescent history of the lineages.

Vázquez-Garcidueñas *et al.*, 2003; Verzino *et al.*, 2003). However, until now, morphological variation had been evaluated only in populations from Argentina (Vázquez-Garcidueñas *et al.*, 2003; Verzino *et al.*, 2003), so the extent to which disjunct areas differed in phenotypic characters was previously unknown. The present study is the first to explore the evolutionary process underlying the disjunct distribution of *P. chilensis*.

Our results confirm a significant leaf trait divergence among the disjunct areas of distribution of the focal species. Moreover, the climatic niche differs significantly among these areas, suggesting that environmental factors might be promoting differentiation among *P. chilensis* populations. The overall pattern obtained with leaf trait analyses indicates that Bolivian trees have the largest leaves, suggesting higher photosynthetic capacity and evapotranspiration, whereas the Chilean trees have the smallest leaves. Consistent with this pattern, Bolivian Chaco locations occur in subtropical environments, with the highest mean values but lowest variance in annual and seasonal rainfall and temperatures, whereas Chilean Matorral locations occur in the driest environment, with the lowest annual and seasonal precipitation. Thus, as rainfall decreases, leaves become smaller, with a higher number of but smaller leaflets, a well-documented pattern in several groups of plants (e.g. Fonseca *et al.*, 2000). In agreement with this pattern, climatic ellipsoids of the most similar morphological groups (i.e. Argentina and Chile) nearly overlapped,

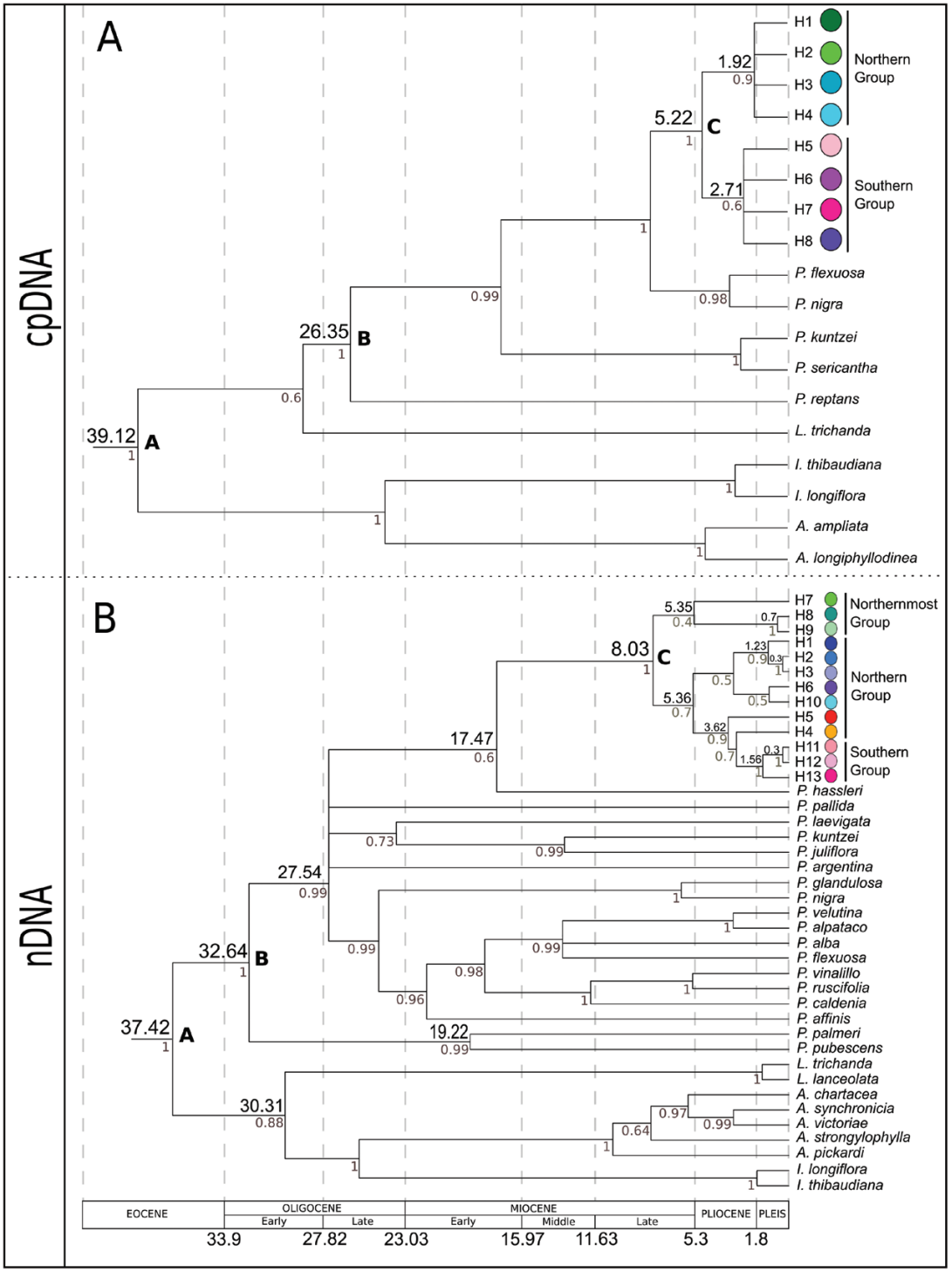

Figure 5. Inferred evolutionary relationships among *Prosopis chilensis* lineages using Bayesian inference. Maximum clade credibility tree of *P. chilensis* based on (A) cpDNA (*ndhF_rpL32* region) and (B) nDNA (ITS region). Numbers given at each node are the average estimated (median) divergence time expressed in million years ago, based on three different calibration settings (Supporting Information, Table S3). Nodes A and B indicate the prior calibration points and correspond to the estimated age of the tree root, and to the age of divergence of the genus *Prosopis*, respectively. Node C represents the estimated age of *P. chilensis*. The estimated age of the nodes and posterior probability values are indicated above and below branches, respectively. Haplotype/allele colours are match those in Figure 3.

indicating not only morphological proximity but also a relationship in the fundamental niche. Moreover, the spatial projection of the Chilean Matorral ellipsoid not only covered its current distribution but also was projected onto the Chaco/Monte region, the ecoregion of the Argentine group.

These morphological and climatic results suggest a putative role of local adaptation in shaping divergent phenotypes. In fact, all sampled trees located in the Bolivian Chaco are more similar to each other than to trees of the other two regions, even though some are genealogically more closely related to the southern phylogroup (located in the Argentinian Chaco). In addition, Chilean samples make up a different morphological and climatic group, regardless of the genetic grouping to which they belong. A previous genetic study based on isozyme and RAPDs performed in Argentine populations of *P. chilensis* suggested that genetic differentiation among populations would be better explained by adaptive rather than neutral (i.e. geographical isolation) processes (Ferreyra *et al.*, 2010). The results obtained here with morphological data agree with that hypothesis, because leaf morphology is coupled with climatic niches of the species in each disjunct area and is uncoupled from the neutral genetic variation patterns (i.e. chloroplast data set).

Phylogeographical patterns

The retrieved phylogeographical patterns based on cpDNA and nDNA were different but compatible. Overall, the two markers revealed geographically structured phylogroups, which is expected for mammal-mediated endozoic fruit dispersion, moving over relatively short distances (Burkart, 1976; Bessega *et al.*, 2000). cpDNA revealed two divergent phylogroups (differing in two mutational steps), one of them predominating in Bolivia but also found in Chile, and the other one predominating in Argentina, but also found in Chile and Bolivia. Because the two phylogroups are geographically structured, the SAMOVA spatial structuring analysis identified two genetic populations that are genealogically well defined in the haplotype network. The retrieved pattern suggests that the two genetic groups differentiated due to drift in the recent past (Avise, 2000). While Bolivia and Argentina have exclusive haplotypes (and higher molecular diversity than Chile), suggesting long-lasting *in situ* persistence, in Chile only the two most frequent and most widely distributed haplotypes from Bolivia and Argentina are present, suggesting a possible recent colonization (or introduction) of *P. chilensis* from the Bolivian Chaco and Argentine Chaco/Monte to the Chilean Matorral. Archaeobotanical and palaeoecological data suggest that the species would have appeared in desert areas of Chile in the late Holocene, probably associated with human activity (McRostie *et al.*, 2017). However, our dating analyses suggest an earlier appearance of the focal species west of the Andes (see below).

As expected, the nuclear region, which generally has a higher mutation rate than the chloroplast genome (Wolfe *et al.*, 1987), showed a higher genetic variability and higher differentiation among the geographical groups. In fact, three genetic groups (instead of two) were retrieved based on ITS sequences. As occurred with the chloroplast haplotypes, the nuclear alleles also evidenced two closely related phylogroups, one predominant in Bolivia and the other in Argentina, but present in the three disjunct areas. A third phylogroup was evidenced with the nuclear marker, being exclusive to the northernmost area of the Bolivian Chaco and limited to the Chiquitania phytogeographical province. Additionally, the nuclear marker revealed exclusive variants, not only in Bolivia and Argentina (as retrieved by cpDNA), but also in Chile.

The results obtained with the chloroplast and nuclear relaxed molecular clock estimated the divergence between the two closely related Northern and Southern phylogroups at around 5 Mya, at the Miocene–Pliocene boundary. In addition, the exclusive lineage from the Bolivian Chaco retrieved with the nuclear marker would have diverged from the Northern and Southern phylogroups at around 8.03 Mya in the Late Miocene. Thus, the estimated divergence for the main lineages within *P. chilensis*, dated to ~5–8 Mya, could be associated with the known 'Paranean Sea' (Ortiz-Jaureguizar & Cladera, 2006, and references therein). During the Middle and Late Miocene (5–11 Mya), three successive Atlantic marine transgressions were recorded in southern SA, all of them informally known as the 'Paranean Sea', separating terrestrial environments (Ortiz-Jaureguizar & Cladera, 2006). This geological event probably fragmented an ancestral distribution of the focal species, favouring lineage differentiation into two main unflooded ancestral areas located in the Bolivian Chaco and Argentine Chaco/Monte, which would explain the origin of our Northern and Southern phylogroups. In this regard, previous studies have proposed that Miocene marine incursions would have promoted lineage divergence in some plant genera, such as *Jaborosa* (Moré *et al.*, 2015) and *Nierembergia* (Acosta *et al.*, 2016), and in animal species, such as geckos (*Homonota*, Morando *et al.*, 2014) and frogs (*Lepidobatrachus*, Brusquetti *et al.*, 2018). Moreover, a similar disjunct distribution as observed in *P. chilensis* was registered in fairy armadillos, proposing a vicariant origin due to the disruption of their ancestral range by the Paranean Ssea (Delsuc *et al.*, 2012). Finally, the highly divergent lineage retrieved with the nuclear marker in *P. chilensis*, exclusively distributed in the limit between the Bolivian Chaco and the Chiquitania,

supports the idea that the Bolivian Chaco would have been a fragmented ancestral area.

The extended surface flooded by the Paranean Sea was followed by the spread of plains generating new terrestrial niches, and allowing species to extend their geographical ranges (Gregory-Wodzicki, 2000; Ortiz-Jaureguizar & Cladera, 2006). The appearance of these new habitats occurred along with the aridization process associated with progressive Andean uplift (from *c.* 14 to 5 Mya). Interestingly, the nDNA retrieved two exclusive nuclear alleles in Chile (H4, H5) that would have diverged from Argentinean alleles at ~3.62 Mya, which suggest a possible vicariant event promoted by the last pulse of the Andean uplift, as was suggested for several plant species (reviewed by Luebert & Weigend, 2014). Thus, the last uplift of the Andes and the rain-shadow effect generated to the east probably influenced the diversification of *P. chilensis*, as was also proposed in *Monttea* (Baranzelli *et al.*, 2014). More recent diversifications within each of the main clades of *P. chilensis* were estimated between 300 kya and 2 Mya, and the detected demographic expansion signals were dated to ~200 kya. Thus, these diversifications and demographic processes seem to have occurred during the Pleistocene.

The Chaco, particularly the Dry Chaco, was a climatically unstable and more extensive area during the Pleistocene glacial/interglacial periods compared to the present time (Ab'Saber, 2000; Pennington *et al.*, 2004). In particular, geological (Iriondo, 1992; Ab'Saber, 2000) and phylogeographical studies (Speranza *et al.*, 2007; Camps *et al.*, 2018; Moreno *et al.*, 2018; Scaldaferro *et al.*, 2018) suggest that during these alternating periods of dry and cold climate, xerophytic vegetation of north-northwestern Argentina advanced northwards very deep into the interior basins of central SA with the species contacting with tropical and subtropical elements, followed by a retraction to the south and south-west during humid and warm periods. In this regard, the Chiquitania phytogeographical region (limiting with the northern Chaco) has been postulated as a climatic refugium during Pleistocene glaciations (Werneck *et al.*, 2011; Trujillo-Arias *et al.*, 2017; Camps *et al.*, 2018, Scaldaferro *et al.*, 2018). Our results support these postulated scenarios because the highest chloroplast and nuclear genetic variability, and exclusive haplotypes/alleles, were detected in the northernmost area of the Bolivian Chaco (limited within the Chiquitania phytogeographic province), supporting the hypothesis that these areas could have been refugia during glacial periods. Additionally, the presence of the southern lineage of *P. chilensis* in the northern areas of the Chaco regions is consistent with the postulated vegetation movements promoting a secondary contact between lineages in the Bolivian Chaco. Interestingly, this secondary contact area was also detected in the arid Chacoan tree species *Bulnesia sarmientoi* (Camps *et al.*, 2018) and in the chili pepper *Capsicum baccatum* (Scaldaferro *et al.*, 2018). The Bolivian northern Chaco/Chiquitania are therefore probably recent secondary contact areas where the previously differentiated lineages meet.

Supporting the inferred lineage movements in *P. chilensis*, demographic analyses revealed an expansion signal in both Northern and Southern phylogroups, and the BSP reconstructions revealed an increase in effective population size starting at ~200 kya. Interestingly, the nuclear marker retrieved an allele exclusive to Bolivia (H11), but genealogically related to Argentina, and an allele exclusive to Chile (H3) but genealogically related to Bolivia, dated at ~300 kya, which agrees with the inferred demographic expansion at ~200 kya. These results suggest relatively recent gene flow from Argentina to Bolivia and from Bolivia to Chile, respectively.

CONCLUSIONS

Phylogeographical patterns suggest that the evolutionary history of *Prosopis chilensis* is linked to ancient late Miocene processes, in particular the Paranean Sea, but also to more recent processes of aridization that occurred during the Pleistocene. Our results show two main regions that harbour high genetic variability, exclusive haplotypes/alleles and ancient lineages, namely the Bolivian Chaco and the Argentine Chaco/Monte, although the Bolivian Chaco seems to be a more ancient (or stable) area for the species than the Argentine ecoregions. The small sampling in Chile does not allow us to elucidate the history of this species in that region, but it appears that the Chilean Matorral is a more recent area of distribution that was colonized at a different timescale from both Bolivia and Argentina. Future studies with increased sample size in Chile, and including samples from Peru, will be necessary to elucidate the evolutionary history of *P. chilensis* west of the Andes.

Regarding the climatic and morphological characterization, our results showed the Chilean and Argentine locations to be morphologically and climatically most similar. Overall, a disruptive ecological and phylogeographical pattern in *P. chilensis* was revealed, suggesting that morphological characteristics respond to local adaptations, regardless of the genealogical history revealed by the phylogeographical patterns. However, from the genetic and morphological standpoints, it can be concluded that within *P. chilensis* there are two clearly differentiated groups of trees: those from the northern Bolivian Chaco and those from Argentine Chaco/Monte and southern Bolivia, suggesting the presence of two significant

different evolutionary units within *P. chilensis*. Finally, it is important to note that between the disjunct areas of *P. chilensis* in Argentina and Bolivia, other *Prosopis* species are found (Burkart, 1976). The current disjunction between *P. chilensis* lineages may therefore be maintained by ecological processes such as niche separation and interspecific competition. This is the first ecophylogeographical approach to the evolutionary history of the emblematic genus *Prosopis* in SA. Ongoing studies from this perspective, including several *Prosopis* species, will contribute to our understanding of the eco-evolutionary processes underlying the diversity and distribution patterns of the South American *Prosopis*.

ACKNOWLEDGMENTS

D.A. is a PhD student at the FCEFyN, Universidad Nacional de Córdoba, and fellowship holder of CONICET. A.V. is a research at INTA and M.C.A., M.C.B., A.N.S. and A.C. are researchers at CONICET. We thank Dr Eduardo Ruiz and two anonymous reviewers for their helpful comments and to Jorgelina Brasca for reviewing the English text. This work was supported by INTA funds PNFOR-1104064 and PNFOR-1104063, BIRF Project LN 7520 AR, and by MINCyT-FONCyT PICT 2015-3089. The authors declare that they have no conflicts of interest.

REFERENCES

Ab'Saber AN. 2000. Spaces occupied by the expansion of dry climates in South America during the Quaternary ice ages. *Revista do Instituto Geológico* **21:** 71–78.

Acosta MC, **Moscone EA**, **Cocucci AA. 2016.** Using chromosomal data in the phylogenetic and molecular dating framework: karyotype evolution and diversification in *Nierembergia* (Solanaceae) influenced by historical changes in sea level. *Plant Biology* **18:** 514–526.

Alsos IG, **Engelskjøn T**, **Gielly L**, **Taberlet P**, **Brochmann C. 2005.** Impact of ice ages on circumpolar molecular diversity: insights from an ecological key species. *Molecular Ecology* **14:** 2739–2753.

Avise JC. 2000. *Phylogeography. The history and formation of species*. Cambridge: Harvard University Press.

Bandelt HJ, **Forster P**, **Rohl A. 1999.** Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16:** 37–48.

Baranzelli MC, **Johnson LA**, **Cosacov A**, **Sérsic AN. 2014.** Historical and ecological divergence among populations of *Monttea chilensis* (Plantaginaceae), an endemic endangered shrub bordering the Atacama Desert, Chile. *Evolutionary Ecology* **28:** 751–774.

Baranzelli MC, **Cosacov A**, **Espíndola A**, **Iglesias MR**, **Chan LM**, **Johnson LA**, **Sérsic AN. 2018.** Echoes of the whispering land: interacting roles of vicariance and selection in shaping the evolutionary divergence of two *Calceolaria* (Calceolariaceae) species from Patagonia and Malvinas/Falkland Islands. *Evolutionary Ecology* **32:** 287–314.

Bessega C, **Ferreyra L**, **Julio N**, **Montoya S**, **Saidman B**, **Vilardi JC. 2000.** Mating system parameters in species of genus *Prosopis* (Leguminosae). *Hereditas* **132:** 19–27.

Bessega C, **Vilardi JC**, **Saidman BO. 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.

Bessega C, **Saidman BO**, **Darquier MR**, **Ewens M**, **Sánchez L**, **Rozenberg P**, **Vilardi JC. 2009.** Consistency between marker and genealogy-based heritability estimates in an experimental stand of *Prosopis alba* (Leguminosae). *American Journal of Botany* **96:** 458–465.

Braunisch V, **Coppes J**, **Arlettaz R**, **Suchant R**, **Schmid H**, **Bollmann K. 2013.** Selecting from correlated climate variables: a major source of uncertainty for predicting species distributions under climate change. *Ecography* **36:** 971–983.

Brusquetti F, **Netto F**, **Baldo D**, **Haddad CF. 2018.** What happened in the South American Gran Chaco? Diversification of the endemic frog genus *Lepidobatrachus* Budgett, 1899 (Anura: Ceratophryidae). *Molecular Phylogenetics and Evolution* **123:** 123–136.

Burghardt AD, **Brizuela MM**, **Mom MP**, **Albán L**, **Palacios RA. 2010.** Análisis numérico de las especies de *Prosopis* L. (Fabaceae) de las costas de Perú y Ecuador. *Revista Peruana de Biología* **17:** 317–324.

Burkart A. 1976. A monograph of the genus *Prosopis* (Leguminosae subfam. Mimosoideae). *Journal of the Arnold Arboretum* **57:** 219–249; 495–498.

Caccavari MA. 1996. Analysis of the South American fossil pollen record of Mimosoideae (Leguminosae). *Review of Palaeobotany and Palynology* **94:** 123–135.

Camps GA, **Martinez-Meyer E**, **Verga A**, **Sérsic AN**, **Cosacov A. 2018.** Genetic and climatic approaches reveal effects of Pleistocene refugia and climatic stability in an old giant of the Neotropical Dry Forest. *Biological Journal of the Linnean Society* **125:** 401–420.

Catalano SA, **Vilardi JC**, **Tosto D**, **Saidman BO. 2008.** Molecular phylogeny and diversification history of *Prosopis* (Fabaceae: Mimosoideae). *Biological Journal of the Linnean Society* **93:** 621–640.

Crandall KA, **Templeton AR. 1993.** Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* **134:** 959–969.

Darriba D, **Taboada GL**, **Doallo R**, **Posada D. 2012.** jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9:** 772.

Delsuc F, **Superina M**, **Tilak MK**, **Douzery EJ**, **Hassanin A. 2012.** Molecular phylogenetics unveils the ancient evolutionary origins of the enigmatic fairy armadillos. *Molecular Phylogenetics and Evolution* **62:** 673–680.

Di Rienzo JA, **Casanoves F**, **Balzarini MG**, **Gonzalez L**, **Tablada M**, **Robledo CW. 2016.** *InfoStat versión 2016*. Córdoba: Grupo InfoStat, FCA, Universidad Nacional de Córdoba. Available at: http://www.infostat.com.ar.

Drummond AJ, **Rambaut A**, **Shapiro B**, **Pybus OG. 2005.** Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* **22:** 1185–1192.

Drummond AJ, **Suchard MA**, **Xie D**, **Rambaut A. 2012.** Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29:** 1969–1973.

Dupanloup I, **Schneider S**, **Excoffier L. 2002.** A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* **11:** 2571–2581.

Excoffier L, **Lischer HEL. 2010.** Arlequin Suite ver 3.5, a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10:** 564–567.

Ferreyra LI, **Vilardi JC**, **Tosto DS**, **Julio NB**, **Saidman BO. 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* **6:** 1011–1025.

Ferreyra LI, **Vilardi JC**, **Verga A**, **López V**, **Saidman BO. 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.

Fonseca CR, **Overton JM**, **Collins B**, **Westoby M. 2000.** Shifts in trait-combinations along rainfall and phosphorus gradients. *Journal of Ecology* **88:** 964–977.

Friendly M, **Fox J. 2017.** *candisc: Visualizing generalized canonical discriminant and canonical correlation analysis. R package version 0.8-0*. Available at: https://CRAN.R-project.org/package=candisc.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147:** 915–925.

GBIF.org. **2019.** *GBIF Occurrence Download*. Available at: https://doi.org/10.15468/dl.di5egc (accessed 13 February 2019).

Gregory-Wodzicki KM. 2000. Uplift history of the Central and Northern Andes: a review. *GSA Bulletin* **112:** 1091–1105.

Hall MD, **Woolhouse MEJ**, **Rambaut A. 2016.** The effects of sampling strategy on the quality of reconstruction of viral population dynamics using Bayesian skyline family coalescent methods: a simulation study. *Virus Evolution* **2:** vew003.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41:** 95–98.

Hernández RM, **Jordan TE**, **Farjat AD**, **Echavarría L**, **Idleman BD**, **Reynolds JH. 2005.** Age, distribution, tectonics, and eustatic controls of the Paranense and Caribbean marine transgressions in southern Bolivia and Argentina. *Journal of South American Earth Sciences* **19:** 495–512.

Hijmans RJ, **Cameron SE**, **Parra JL**, **Jones PG**, **Jarvis A. 2005.** Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25:** 1965–1978.

Hotelling H. 1933. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* **24:** 417–441, 498–520.

Hothorn T, **Bretz F**, **Westfall P. 2008.** Simultaneous inference in general parametric models. *Biometrical Journal* **50:** 346–363. Available at: http://www.worldclim.org/.

Iriondo MH. 1992. El Chaco. *Holoceno* **1:** 50–63.

Katoh K, **Misawa K**, **Kuma KI**, **Miyata T. 2002.** *MAFFT*: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30:** 3059–3066.

Kay KM, **Whittall JB**, **Hodges SA. 2006.** A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. *BMC Evolutionary Biology* **6:** 36.

Lavin M, **Herendeen PS**, **Wojciechowski MF. 2005.** Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Systematic Biology* **54:** 575–594.

Lenth RV. 2016. Least-squares means: the R Package lsmeans. *Journal of Statistical Software* **69:** 1–33.

Luebert F, **Weigend M. 2014.** Phylogenetic insights into Andean plant diversification. *Frontiers in Ecology and Evolution* **2:** 27.

Maguire Jr, B. 1973. Niche response structure and the analytical potentials of its relationship to the habitat. *American Naturalist* **107:** 213–246.

Macherey-Nagel. **2014.** *Genomic DNA from plant. User manual. NucleoSpin Plant II. Rev.09*. Düren: Macherey-Nagel.

Magallón-Puebla S, **Cevallos-Ferriz SRS. 1994.** Fossil legume fruits from Tertiary strata of Puebla, Mexico. *Canadian Journal of Botany* **72:** 1027–1038.

Martínez-Meyer E, **Díaz-Porras D**, **Peterson AT**, **Yáñez-Arenas C. 2013.** Ecological niche structure and rangewide abundance patterns of species. *Biology Letters* **9:** 20120637.

McRostie VB, **Gayo EM**, **Santoro CM**, **De Pol-Holz R**, **Latorre C. 2017.** The pre-Columbian introduction and dispersal of Algarrobo (*Prosopis*, Section Algarobia) in the Atacama Desert of northern Chile. *PLoS One* **12:** e0181759.

Miles L, **Newton AC**, **DeFries RS**, **Ravilious C**, **May I**, **Blyth S**, **Kapos V**, **Gordon JE. 2006.** A global overview of the conservation status of tropical dry forests. *Journal of Biogeography* **33:** 491–505.

Morando M, **Medina CD**, **Avila LJ**, **Perez CH**, **Buxton A**, **Sites JW Jr**. **2014.** Molecular phylogeny of the New World gecko genus *Homonota* (Squamata: Phyllodactylidae). *Zoologica Scripta* **43:** 249–260.

Moré M, **Cocucci AA**, **Sérsic AN**, **Barboza GE. 2015.** Phylogeny and floral trait evolution in *Jaborosa* (Solanaceae). *Taxon* **64:** 523–534.

Moreno SEM, **Brandão de Freitas L**, **Speranza PR**, **Solís Neffa VG. 2018.** Impact of Pleistocene geoclimatic events on the genetic structure in mid-latitude South American plants: insights from the phylogeography of *Turnera sidoides* complex (Passifloraceae, Turneroideae). *Botanical Journal of the Linnean Society* **188:** 377–390.

Mooney HA, **Simpson BB**, **Solbrig OT. 1977**. Phenology, Morphology, Physiology. In: Simpson BB, ed. *Mesquite, its biology in two desert ecosystems*. Stroudsburg: Dowden, Hutchinson & Ross, Inc, 26–43.

Murdoch D. 2017. *sciplot: Scientific Graphing Functions for Factorial Designs. R package v.1.1-1*. Available at: https://CRAN.R-project.org/package=sciplot.

Naumann M. 2006. *Atlas del Gran Chaco Sudamericano. Sociedad Alemana de Cooperación Técnica (GTZ)*. Buenos Aires: ErreGé & Asoc., 92 pp.

Nei M. 1987. *Molecular Evolutionary Genetics*. New York: Columbia University Press.

Olson DM, **Dinerstein E**, **Wikramanayake ED**, **Burgess ND**, **Powell GVN**, **Underwood EC**, **D'amico JA**, **Itoua I**, **Strand HE**, **Morrison JC**, **Loucks CJ**, **Allnutt TF**, **Ricketts TH**, **Kura Y**, **Lamoreux JF**, **Wettengel WW**, **Hedao P**, **Kassem KR. 2001.** Terrestrial ecoregions of the world: a new map of life on Earth. *Bioscience* **51:** 933–938.

Ortiz-Jaureguizar E, **Cladera GA. 2006.** Paleoenvironmental evolution of southern South America during the Cenozoic. *Journal of Arid Environments* **66:** 498–532.

Osorio-Olvera L, **Barve V**, **Barve N**, **Soberón J**, **Falcomi M. 2018.** *ntbox: From getting biodiversity data to evaluating species distribution models in a friendly GUI environment. R package v. 0.2. 5.4*. Available at: http://shiny.conabio.gob.mx:3838/nichetoolb2/.

Ossa GC, **Montenegro P**, **Larridon I**, **Pérez F. 2019.** Response of xerophytic plants to glacial cycles in southern South America. *Annals of Botany* **124:** 15–25.

Pennington RT, **Lavin M**, **Prado DE**, **Pendry CA**, **Pell SK**, **Butterworth CA. 2004.** Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both tertiary and quaternary diversification. *Philosophical Transactions of the Royal Society B: Biological Sciences* **359:** 515–537.

Peterson AT, **Soberón J**, **Pearson RG**, **Anderson RP**, **Martínez-Meyer E**, **Nakamura M**, **Araújo M. 2011.** *Ecological niches and geographic distributions*. Princeton: Princeton University Press.

Piel KM. 1971. Palynology of Oligocene sediments from central British Columbia. *Canadian Journal of Botany* **49:** 1885–1920.

Quantum GIS Development Team. **2016.** *QGIS geographic information system*. Open source geospatial foundation project. Available at: http://www.qgis.org/.

R Development Core Team. **2017.** *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available at: http://www.R-project.org/.

Rambaut A, **Drummond AJ. 2012.** *FigTree: tree figure drawing tool, v1.4.2*. Institute of Evolutionary Biology, University of Edinburgh. Available at: http://tree.bio.ed.ac.uk/software/figtree.

Rambaut A, **Suchard MA**, **Xie D**, **Drummond AJ. 2014.** *Tracer v1.6*. Available at: http://beast.bio.ed.ac.uk/Tracer.

Rozas J, **Sánchez-DelBarrio JC**, **Messeguer X**, **Rozas R. 2003.** DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19:** 2496–7.

Sarkar D. 2008. *Lattice: multivariate data visualization with R*. New York: Springer.

Scaldaferro MA, **Barboza GE**, **Acosta MC. 2018.** Evolutionary history of the chili pepper *Capsicum baccatum* L. (Solanaceae): domestication in South America and natural diversification in the Seasonally Dry Tropical Forests. *Biological Journal of the Linnean Society* **124:** 466–478.

Shaw J, **Edgar B**, **Lickey EB**, **Schilling EE**, **Small R. 2007.** Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in Angiosperms: the tortoise and the hare III. *American Journal of Botany* **94:** 275–288.

Simmons MP, **Ochoterena H. 2000.** Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* **49:** 369–381.

Speranza PR, **Seijo JG**, **Grela IA**, **Solís Neffa VG. 2007.** Chloroplast DNA variation in the *Turnera sidoides* L. complex (Turneraceae): biogeographical implications. *Journal of Biogeography* **34:** 427–436.

Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105:** 437–460.

Tajima F. 1989. The effect of change in population size on DNA polymorphism. *Genetics* **123:** 598–601.

Trujillo-Arias N, **Dantas GPM**, **Arbeláez-Cortés E**, **Naoki K**, **Gómez MI**, **Santos FR**, **Miyaki CY**, **Aleixo A**, **Tubaro PL**, **Cabanne GS. 2017.** The niche and phylogeography of a passerine reveal the history of biological diversification between the Andean and the Atlantic forests. *Molecular Phylogenetics and Evolution* **112:** 107–121.

Turchetto-Zolet AC, **Cruz F**, **Vendramin GG**, **Simon MF**, **Salgueiro F**, **Margis-Pinheiro M**, **Margis R. 2012.** Large-scale phylogeography of the disjunct Neotropical tree species *Schizolobium parahyba* (Fabaceae-Caesalpinioideae). *Molecular Phylogenetics and Evolution* **65:** 174–182.

Vázquez-Garcidueñas S, **Palacios RA**, **Segovia-Quiroz J**, **Frías-Hernández JT**, **Olalde-Portugal V**, **Martínez-de la Vega O**, **Mollard FPO**, **Vázquez-Marrufo G. 2003.** Morphological and molecular data to determine the origin and taxonomic status of *Prosopis chilensis var. riojana* (Fabaceae, Mimosoideae). *Canadian Journal of Botany* **81:** 905–917.

Verzino G, **Carranza C**, **Ledesma M**, **Joseau J**, **Di Rienzo J. 2003.** Adaptive genetic variation of *Prosopis chilensis* (Mol) Stuntz: preliminary results from one test-site. *Forest Ecology and Management* **175:** 119–129.

Werneck FP, **Costa GC**, **Colli GR**, **Prado DE**, **Sites JW Jr**. **2011.** Revisiting the historical distribution of Seasonally Dry Tropical Forests: new insights based on paleodistribution modelling and palynological evidence. *Global Ecology and Biogeography* **20:** 272–288.

White TJ, **Bruns T**, **Lee S**, **Taylor J. 1990.** Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, eds. *PCR Protocols: a guide to methods and applications*. New York: Academic Press, 315–322.

Wolfe KH, **Li WH**, **Sharp PM. 1987.** Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the USA* **84:** 9054–8.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website.

Table S1. Sampled localities within each geographical area, coordinates, chloroplast haplotypes, nuclear alleles and number of measured individuals for the sampling sites of *Prosopis chilensis* across its natural distribution in southern South America.
Table S2. Accession numbers of the species used as outgroups for calibration analyses available in GenBank or obtained in this study. *Prosopis* species with two accession numbers indicate that the amplified region was divided into two fragments (ITS1 and ITS2).
Table S3. Results of calibration analyses based on chloroplast and nuclear sequences. The 95% HPD is indicated between square brackets. Node A corresponds to the estimated age of the tree root. Node B shows the age of divergence of the genus *Prosopis*. Node C represents the estimated age of *P. chilensis* (see Fig. 5).
Table S4. Cross-classification table of the 112 sampled *P. chilensis* individuals derived from linear discriminant analyses using leaf morphological characters.
Table S5. Results of one-way ANOVA and posterior Tukey contrasts to compare means of each of the 19 bioclimatic variables among the three geographical areas. Df, degrees of freedom; Sq, squared errors.
Table S6. Cross-classification table of the 121 sampled *P. chilensis* locations derived from the linear discriminant analysis using 19 bioclimatic variables.
Table S7. Results of the spatial analysis of molecular variance (SAMOVA) for *Prosopis chilensis*, based on cpDNA and nDNA. The percentage of the total variance explained by each hierarchical level of variation is shown: AG, among groups of populations; AP, among populations within groups; WP, within populations.
Figure S1. Comparison of environmental variation among three geographical provenances of *Prosopis chilensis* (N_{loc} = 121; Argentine Chaco/Monte in blue, Bolivian Chaco in red and Chilean Matorral in green). Vertical grey bars represent the maximum and minimum for each bioclimatic variable. Dots represent mean value and vertical lines represent the standard deviation for each bioclimatic variable.