Biological Journal of the Linnean Society, 2018, **XX**, 1–20. With 9 figures.

Genetic and climatic approaches reveal effects of Pleistocene refugia and climatic stability in an old giant of the Neotropical Dry Forest

GONZALO A. CAMPS[1,2], ENRIQUE MARTÍNEZ-MEYER[3,4], ANIBAL R. VERGA[1], ALICIA N SÉRSIC[2†] and ANDREA COSACOV[2†*]

[1]*Instituto de Fisiología y Recursos Genéticos Vegetales (IFRGV), CIAP, INTA, 5020 Córdoba, Argentina*
[2]*Laboratorio de Ecología Evolutiva—Biología Floral, Instituto Multidisciplinario de Biología Vegetal (IMBIV), CONICET-Universidad Nacional de Córdoba, 5000 Córdoba, Argentina*
[3]*Departamento de Zoología, Instituto de Biología, Universidad Nacional Autónoma de México, 04510 Ciudad de México, México*
[4]*Centro del Cambio Global y la Sustentabilidad en el Sureste, AC, 86080 Villahermosa, México*

Received 4 June 2018; revised 17 July 2018; accepted for publication 18 July 2018

Neotropical Dry Forests are important biodiversity hotspots characterized by intermediate to high levels of species richness and endemism. A possible explanation for these characteristics is that such forests have been less affected by drastic glacial impacts than other biomes. Using two approaches, geo-statistical phylogeography, based on two chloroplast markers, and multi-algorithm-based niche modelling, for the present and for the past, we explored if, during glacial periods, the geographical range of *Bulnesia sarmientoi* was stable or underwent expansions or retractions in space and time and if there is a relationship among past climatic refugia, the current climatic optimum and genetic diversity. We estimated that *B. sarmientoi* would have diverged from other *Bulnesia* at the beginning of the Pliocene (5 Mya), with diversification of the current lineages occurring in the Pleistocene (1.4–1.1 Mya). Our results suggest that Dry Forests underwent population expansion events during the glacial periods, whereas they would have undergone population stasis during interglacial periods. Furthermore, we identified a putative refugial area in the Dry Chaco that has been climatically stable through time, consistent with the area of highest genetic diversity and with the spatial location of the climatic optimum of the focal species.

ADDITIONAL KEYWORDS: arid environments – Bayesian diffusion phylogeography – climatic optimum – climatic centroid – Dry Chaco – landscape ecology – ecological niche modelling – endangered forest species – Quaternary climatic changes – Zygophyllaceae.

INTRODUCTION

Neotropical Dry Forests (NDFs) are important biodiversity hotspots characterized by intermediate to high levels of species richness and endemisms (Ceballos & Brown, 1995; Olson *et al.*, 2001; Sánchez-Azofeifa *et al.*, 2005; Miles *et al.*, 2006). A possible explanation fir these characteristics is that NDFs are considered to have been less affected by drastic glacial impacts than other biomes; however, recent reviews have shown a relative scarcity of studies exploring the impacts of past climatic oscillations on this region and particularly among tree species (Turchetto-Zolet *et al.*, 2013). The Gran Chaco of South America is the largest continuous NDF in the world (Olson *et al.*, 2001; Kuemmerle *et al.*, 2017), and includes two ecoregions, the Dry Chaco and the Humid Chaco; the Dry Chaco is located in central-northern Argentina, south-eastern Bolivia, western Paraguay and marginally in Brazil (Pennington *et al.*, 2000; Clark *et al.*, 2010). As in the entire Chaco region, the geology and climate of the Dry Chaco are linked to earlier volcanic and tectonic events affecting the Andean mountain range during the late Miocene and Pliocene (Ramos & Ghiglione, 2008). Iriondo (2010) postulated that the biogeographical identity of the

*Corresponding author. E-mail: acosacov@imbiv.unc.edu.ar
†A.C. and A.N.S. share senior authorship.

Chaco was moulded during the Pliocene, whereas Quaternary climatic changes influenced through variations in humidity rather than through changes in temperature (Argollo Bautista & Iriondo, 2008). It has also been proposed that there were two dry periods affecting South America in the last several thousand years, one during the late Pleistocene, probably linked to the Last Glacial Maximum (LGM), and the other, less severe, in the late Holocene (Iriondo, 2010). The Chaco, particularly the Dry Chaco, was an unstable and more extensive area during the glacial periods than at present (Ab'Saber, 2000; Pennington *et al.*, 2004; Speranza *et al.*, 2007; Iriondo, 2010; Iriondo & Brunetto, 2016). During these periods of drier and colder climates, xerophytic vegetation of north-northwestern Argentina advanced northwards very deep into the interior basins of central South America (Ab'Saber, 2000).

Phylogeography (Avise, 2000) stands as a powerful molecular tool to gain insights into the way these historical processes affected the geographical distribution of the Dry Chaco plant species (for reviews see: Taberlet *et al.*, 1998; Soltis *et al.*, 2006; Sérsic *et al.*, 2011; Hardy *et al.*, 2013; Turchetto-Zolet *et al.*, 2013; Bowen *et al.*, 2016). However, in several areas of the world, as in the Chaco biome, biogeographical patterns have been scarcely studied from a phylogeographical perspective (Werneck, 2011; Turchetto-Zolet *et al.*, 2013). In fact, in their review, Turchetto-Zolet *et al.* (2013) pointed out that the Pantanal–Chaco ecoregions, as a whole, stand as the least explored areas in phylogeographical terms, particularly for plants. Some phylogeographical studies have been conducted in animals (e.g. Eizirik *et al.*, 2001; Caparroz *et al.*, 2009; Campos-Krauer & Wisely, 2011; Werneck *et al.*, 2012), and some plant species (e.g. Speranza *et al.*, 2007; Caetano & Naciri, 2011) for the Dry Chaco, but few of them discuss the biogeographical history inferred from phylogeographical patterns (e.g. Speranza *et al.*, 2007; Caetano & Naciri, 2011; Werneck *et al.*, 2012; Bartoleti *et al.*, 2017). Phylogeography and ecological niche modelling (ENM; Peterson *et al.*, 1999) have been recently integrated to better understand biogeographical patterns and evolutionary history of species (e.g. Gao *et al.*, 2015; Peterson & Graves, 2016; Baranzelli *et al.*, 2017), even in the Chaco and surrounding ecoregions (e.g. Campos-Krauer & Wisely, 2011 in Paraguayan Chaco; Collevatti *et al.*, 2012 in Neotropical Seasonally Dry Forest; Werneck *et al.*, 2012 in Chaco, Cerrado and Caatinga; Vitorino *et al.*, 2016 in Seasonally Dry Tropical Forests; Bartoleti *et al.*, 2017 in Chaco and Cerrado). However, as described above, few are especially focused on the Dry Chaco.

Ongoing climate change increases the need to understand and forecast the responses of biodiversity to climate oscillations. Disentangling the processes and ecological factors that drove species expansion and persistence during past climatic changes would ultimately help to improve conservation efforts. In this context, identifying refugia, which might represent hotspots of genetic diversity (but see: Widmer & Lexer, 2001; Petit *et al.*, 2003), has been perhaps one of the main contributions of phylogeography to conservation planning. Many areas were proposed as refugia or relict in diverse regions of the world based only on phylogeographical evidence (reviewed by Keppel *et al.*, 2012), or from the joint evidence provided by phylogeography and ENM (e.g. Carstens & Richards, 2007; Waltari *et al.*, 2007; Collevatti *et al.*, 2012; Cosacov *et al.*, 2013; Baranzelli *et al.*, 2017), showing that in many species worldwide high genetic diversity is associated with climatically stable areas. However, whether these hotspots of genetic diversity (i.e. refugia) are concordant not only with areas of climatic stability but also with the spatial location of the climatic optimum of a species remains virtually unexplored (but see Lira-Noriega & Manthey, 2014). Understanding this aspect is of great relevance in the context of current climate change, and will allow us to improve our characterization and understanding of climatic refugia (Keppel *et al.*, 2012). All these considerations are of particular relevance for the studied area because, unfortunately, unsustainable exploitation of the Gran Chaco and agricultural expansion has had severe and accelerated consequences not only on the general biodiversity of this region (Hansen *et al.*, 2013; WWF, 2014; Vallejos *et al.*, 2015), but also on the exploited forest species, such as the emblematic *Bulnesia sarmientoi* Lorentz ex Griseb. (SAyDS, 2007; Mereles & Pérez de Molas, 2008; Waller *et al.*, 2012). The information generated here regarding the evolutionary history and genetic diversity of the species could be useful for future conservation programmes.

The general aim of this contribution was to infer the effects of historical climatic changes that occurred during the Quaternary glaciations on the demography, diversification and spatial distribution of *B. sarmientoi*, an emblematic tree species of the Dry Chaco, using geo-statistical phylogeography and spatial distribution modelling. In particular, our aims were: (1) to reconstruct the spatio-temporal dynamics of *B. sarmientoi* lineage diversification; (2) to explore if its geographical range was stable through space and time or, rather, if it underwent expansions or retractions; (3) if range expansions/retractions are inferred, to investigate if they were associated with glacial or interglacial periods; and (4) to explore if there is a relationship among climatic refugia, the current climatic optimum and genetic diversity. We hypothesized that *B. sarmientoi* underwent range expansions during the glacial periods and *in situ* persistence during the interglacials, because it is adapted to a dry climate that prevailed during glaciations,

whereas during interglacials, when the climate was more humid, it would not have been favoured.

MATERIAL AND METHODS

Study area and species

The Dry Chaco extends between 17°32′26″S and 33°52′7″S latitude and 67°43′12″W and 57°59′26″W longitude, covering approximately 800 000 km^2. Annual temperature has a north–south gradient from 21 to 18 °C, and rainfall declines east–west from 1000 to 500 mm/year, with a strong seasonality (Clark *et al.*, 2010).

Bulnesia sarmientoi (Fig. 1) is a large tree (up to 20 m in height) dominant on clay-rich, heavy soils with highly impeded drainage and temporary anaerobiosis (Adamoli *et al.*, 1972; Prado, 1993b). The distribution of the species is relatively well known, occurring in central-northern Argentina, south-eastern Bolivia, western Paraguay and marginally in Brazil (Zuloaga *et al.*, 2008; Waller *et al.*, 2012). The distribution area mainly includes the Dry Chaco ecoregion and marginally the Chiquitano Dry Forest and Pantanal (Fig. 1). The tree is pollinated by the Meliponini bee species *Geotrigona argentina* (Vossler, 2014) and the fruit is a three-winged capsule; each mericarp contains one seed that is dispersed by autochory (Abraham de Noir *et al.*, 2002).

Plant material, DNA isolation, amplification and sequencing

For the genetic study, fresh leaves of 144 *B. sarmientoi* individuals from 25 localities (six individuals per locality, each one 80 m distance from the next) were collected, covering the entire distribution range of *B. sarmientoi*. Samples of *B. bonariensis* Griseb. (Code 4143), *B. arborea* (Jacq.) Engler (Code 9041) and *B. schickendantzii* Hieron. ex Griseb (Code 15312) were obtained from herbarium (CORD, IS) material to be used as outgroups in the phylogenetic analyses.

DNA was extracted from dried leaves using a NucleoSpin Plant II kit (Macherey-Nagel, Düren, Germany). Two regions of the plastid genome were selected and amplified, the *trn*L-*trn*F region, using the primers *c* and *f* designed by Taberlet *et al.* (1991), and *rpl*32_F-*trn*L region (Shaw *et al.*, 2007). These regions were the most variable, with highest number of polymorphic sites, in contrast to the regions *trn*D-trnT, *trn*H-psbA, *rpl*32_R-*ndh*F, *rps*2_47F-*rps*2_661R and 8*for*-1525*rev*, which were also tested. The polymerase chain reaction (PCR) cycling conditions were template denaturation at 80 °C for 5 min followed by 30 cycles of denaturation at 95 °C for 1 min, primer annealing at 50 °C for 1 min, followed by a ramp of 0.3 °C/s to 65 °C, and primer extension at 65 °C for 4 min, followed by a final extension step of 5 min at 65 °C. PCR

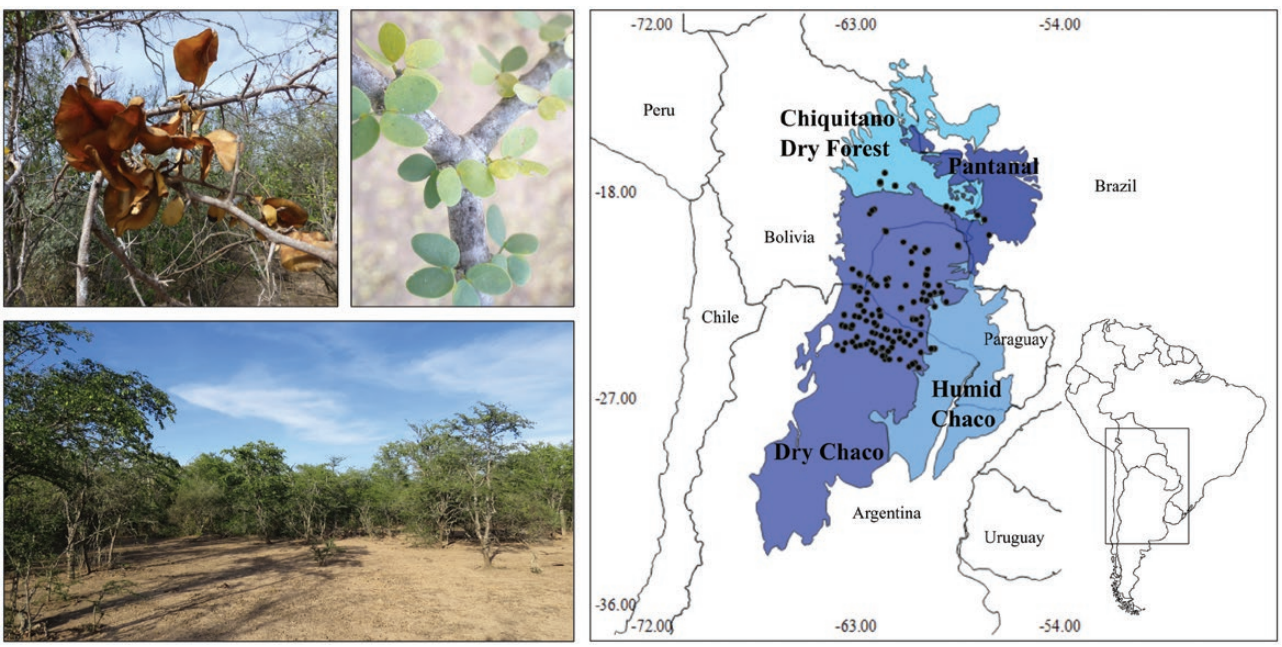

Figure 1. Geographical distribution area of *Bulnesia sarmientoi*. Left: fruits, leaves and trees of *B. sarmientoi*, in a particular vegetation formation called 'palosantal'. Right: partial map of South America showing the ecoregions (in blue hues) where the species is distributed. The Gran Chaco Americano comprises the Dry Chaco and Humid Chaco. Black points: presence points of *B. sarmientoi*. Ecoregions correspond to those of Olson *et al.* (2001).

was performed using an Eppendorf (Westbury, NY, USA) Mastercycler.

Purification and sequencing were performed by Macrogen Inc. (Seoul, Korea, http://dna.macrogen.com). Chromatograms were visualized and manually edited using Chromas Pro v.2.0.1. (Technelysium Pty. Ltd, Helensvale, Australia), aligned with ClustalX (Thompson *et al.*, 1997) and manually adjusted using BioEdit v.7.0.9.0 (Hall, 1999). Indels ($N = 7$) were coded as single binary characters using the simple indel coding method (Simmons & Ochoterena, 2000).

Haplotype network, genetic diversity and demographic indices

Haplotypes were inferred with DnaSP v.5 (Librado & Rozas, 2009) for the concatenated *trn*L-*trn*F/*rpl*32_F-*trn*L matrices. Genealogical relationships among haplotypes were estimated using the median-joining method (Bandelt *et al.*, 1999), implemented in Network 5.0 (Fluxus Technology Ltd, http://Fluxus-engineering.com).

Haplotype diversity (h; Nei, 1987), nucleotide diversity (π; Nei, 1987) and mean number of pairwise differences (p; Tajima, 1983) indices were calculated for the species and each locality using Arlequin v.3.1 (Excoffier *et al.*, 2005). To map genetic diversity on the range of the species we projected h and π of the 25 georeferenced localities into the modelled distribution of the species using the inverse distance weighting (IDW) spatial interpolation method (Watson, 1992). The distance coefficient (p) value in IDW was 2. For this method, and for the construction of the other maps and figures in this paper, we used Quantum GIS 2.18 (QGIS Development Team, http://www.qgis.org/) and Inkscape 0.91 (Inkscape Project, https://inkscape.org/).

For the species and the three phylogroups retrieved by genealogical inference we used two neutrality tests: (1) Tajima's D (Tajima, 1989) and (2) Fu's FS (Fu, 1997), using Arlequin 3.11 (Excoffier *et al.*, 2005). Mismatch distributions were simulated under the sudden-demographic expansion and the spatial-demographic expansion models using DnaSP v.5 (Librado & Rozas, 2009). For those groups that presented signals of demographic expansion we also performed Bayesian Skyline Plot (BSP) analyses implemented in BEAST v.1.7.5 (Drummond *et al.*, 2005, 2012). jModelTest 2.1.9 (Darriba *et al.*, 2012) was used to statistically select the best-fit model of nucleotide substitution according to the Akaike information criterion (AIC). For the species, we selected a GTR+I model, and the Markov chain Monte Carlo (MCMC) length setting was 5×10^8. For phylogroup 1 (see Results) we selected an HKY substitution model and the length of the chain was 1×10^8. Settings for the other parameters were shared for both analyses: lognormal relaxed clock, mean mutational rate of 0.005 (0.01–0.001, uniform distribution; Alsos *et al.*, 2005), constant population size, skyland.popsize was changed to uniform distribution, initial value 50, upper value 100 and lower value 10. Parameter convergence and effective sample size (ESS > 200) were verified using Tracer v.1.5 (Rambaut *et al.*, 2013). We also compared the two different skyline models (linear vs. constant) using the Bayes factor (BF) in Tracer.

Phylogenetic reconstruction and divergence time estimates

To reconstruct the phylogenetic relationships among haplotypes and to estimate the divergence time of the phylogroups, we performed a Bayesian phylogenetic reconstruction using the 18 sequences of the concatenated primers *rpl*32_F-*trn*L and *trn*L-*trn*F, and sequences of *B. bonariensis*, *B. arborea* and *B. schickendantzii* as outgroup. First, we obtained the divergence time of the *Bulnesia* clade, which was set at 13.8 Mya [95% highest posterior density (HPD) 21.86–6.37 Mya] by performing two indirect dating analyses; the first one was constructed from the literature downloading 24 *RuBisCO large subunit* sequences of the Order Zygophyllales (Supplementary Material S1) from GenBank (https://www.ncbi.nlm.nih.gov/genbank/). This analysis gave us information necessary for dating the second phylogenetic tree for which we used 41 *trnL-trnF* sequences of Zygophyllaceae downloaded from GenBank and selected sequences of *B. sarmientoi* (Supplementary Material S1). The substitution sites model selected was HKY+I; other parameters used were: lognormal relaxed clock and mutational rate of 0.005 substitutions per site per million years (0.01–0.001; Alsos *et al.*, 2005). The MCMC length was run for 8×10^8 generations, starting with a random tree, with parameters sampled every 80 000 steps. Parameter convergence and ESS values (> 200) were verified in Tracer v.1.5. Tree topologies were assessed using TreeAnnotator v.1.7.5 (Rambaut & Drummond, 2013) and FigTree v.1.6.1 (http://tree.bio.ed.ac.uk/software/figtree/).

Bayesian spatio-temporal diffusion analyses

As a complementary approach, we reconstructed the evolutionary history of lineages through time and space in BEAST assuming continuous spatial diffusion, using a time-heterogeneous random walk model ('Relaxed Random Walk', RRW; Lemey *et al.*, 2010). For this analysis we included one individual per haplotype per locality ($N = 74$). We used a normally distributed diffusion rate, a GTR substitution model, a lognormal relaxed clock with a uniform rate (0.01–0.001) and a coalescent Bayesian Skyride model. We manually modified .xml files to include geographical uncertainties in the sequences that have the same geographicalal

coordinates (i.e. those that belong to the same location) using the option jitter on statistical TraitLikelihood with a parameter of 0.01. We performed a run of 500 million generations sampled every 50 000 generations, to obtain a total of 10 000 trees from the posterior distribution. We inspected parameters with Tracer to check for stationarity. To summarize the posterior distribution of ancestral ranges using the RRW model, we annotated nodes in a maximum clade credibility (MCC) tree using the program TreeAnnotator v.1.7.5. This tree was then used as an input in SPREAD v.1.0.7 (Bielejec *et al.*, 2011) to reconstruct the pattern of spatial diffusion (Bielejec *et al.*, 2016) and was viewed using Google Earth (available from http://earth.google.com).

ENM IN GEOGRAPHICAL AND CLIMATIC SPACE

To perform ENM for the present we used 19 bioclimatic variables (2000s decade; 2.5 arcmin resolution) downloaded from the Dryad Digital Repository (http://datadryad.org/resource/doi:10.5061/dryad.s2v81), which are obtained from a global set of satellite-based bioclimatic variables (MERRAclim; Vega *et al.*, 2017), and Hydro-1K (HYDRO1k Elevation Derivative Database), a topographical database, downloaded from the US Geological Survey (2001; https://lta.cr.usgs.gov/HYDRO1K). Because it is not possible to reconstruct palaeoclimatic variables from satellite-based information, to obtain distribution models for past periods we performed a new ENM for the present with 19 climate variables for current conditions (~1950–2000; 2.5 arcmin) from the WorldClim dataset (http://worldclim.org/) and projected it to the mid-Holocene (~6000 years BP; 2.5 arcmin), LGM (~21 000 years BP; 2.5 arcmin) and Last Interglacial Maximum (LIG, ~116 000–130 000 years BP; 2.5 arcmin) climatic scenarios, also available in the WorldClim website. We used two Global Climate Models (GCMs), CCSM4 (cc) and MIROC-ESM (mr) for the mid-Holocene and LGM. The values of the MERRAclim and WorldClim variables are equivalent, differing only in that MERRAclim uses humidity instead of precipitation. We obtained 79 points of presence of *B. sarmientoi* in the field, 41 points from online databases (Global Biodiversity Information Facility, http://www.gbif.org/; SpeciesLink, http://splink.cria.org.br/), 22 points from the Herbarium of Chemistry Sciences of Asunción, Paraguay, and 363 points from the Ministry of Environment, Argentina. From this dataset, we kept a total of 138 presence points. The model was manually calibrated, and 57% of the clean occurrence dataset was used for calibration and 43% for validation. Supplementary Material S2 gives details on the cleaning of occurrence points, M estimate of the BAM diagram (Soberón & Peterson, 2005; Soberón, 2007) and the calibration process.

We tested six different modelling algorithms (Supplementary Material S2), which were validated in Niche ToolBox (Osorio-Olvera *et al.*, 2016) using partial-ROC (Peterson *et al.*, 2008), a binomial test (Anderson, Gómez-Laverde & Peterson, 2002), omission error (false negative of Fielding & Bell, 1997), and the omission/predicted surface relationship. For the current model with MERRAclim, we selected MaxEnt (Phillips *et al.*, 2006) (feature classes: LQHPT, regularization multiplier: 2), BIOCLIM (Busby, 1991) and Support Vector Machine (Vapnik, 1998) algorithms. For the models with WorldClim, only MaxEnt was selected (feature classes: LQH, regularization multiplier: 3). Binary maps of all models were obtained using the smallest suitability value from the species presence points as the threshold value. To test if past projected areas were detected as strict extrapolations or were areas with environmental similarity between the calibration (M) and projection (G) regions (Owens *et al.*, 2013), we used Multivariate Environmental Similarity Surface (MESS; Elith *et al.*, 2010) and Mobility-Oriented Parity (MOP; Owens *et al.*, 2013) analyses, implemented in MaxEnt and with the R-script available at https://github.com/luismurao, respectively.

The current ENM obtained with MERRAclim was complemented with an analysis of the ENM in ecological space; this calculates suitability values from Euclidean and Mahalanobis distances, and allows an approximation to the climatic optimum of the species (approximations of ecological niche centroid in ENM can be seen in Martínez-Meyer *et al.*, 2013; Tocchio *et al.*, 2014; Qiao *et al.*, 2015). The three climatic variables with the greatest contribution were used to reconstruct the model: 'Mean temperature of most humid quarter', 'Specific humidity mean of coldest quarter' and 'Annual mean specific humidity'. Suitability values greater than 0.8 were referenced on a map using Niche ToolBox.

GENETIC DIVERSITY, CLIMATIC STABILITY AND CLIMATIC OPTIMUM

To explore if there was spatial correspondence among climatically stable areas, current climate optimum and hotspots of genetic diversity, we generated a single map by combining the layers obtained in the previous results. For this, we transformed the following layers and thresholds into binary maps with Quantum GIS 2.18: genetic diversity (threshold = 0.733), LGM palaeodistribution (models *cc* + *mr*) (threshold = 0.063), LIG palaeodistribution (threshold = 0.063) and suitability of the ENM in ecological space (threshold = 0.8). Using the raster calculator, we added these layers and obtained a synthetic map showing the overlapped areas.

RESULTS

HAPLOTYPE NETWORK, GENETIC DIVERSITY AND DEMOGRAPHIC INDICES

Concatenated intergenic spacers, *trn*Lc-*trn*Ff and *rpl*32_F-*trn*L, generated a matrix of 1515 bp, which provides 17 polymorphic sites and 18 haplotypes. The most frequent and widespread haplotype (H9), which was present in all localities except NY, formed the core of the 'star-like' network topology (Fig. 2). The network shows three phylogroups sparsely distributed across the spatial range. Phylogroup 1 (P1) contained the largest number of individuals (N = 91), including H9 and its derivatives (H1, H8, H12, H14–H17), with five of them being exclusive to sites 52, NY, EE, LH and FE. Phylogroup 2 (P2) included five haplotypes, with H11 being the most frequent and only one (H18) being exclusive to locality SS. Phylogroup 3 (P3) contained five haplotypes: H2 was the most frequent and broadly distributed (being present in 58% of the localities); H3 was retrieved in eight individuals and six localities, and three exclusive haplotypes (H4, H5, H6) were present at LT, Mo and LH, respectively (Fig. 2). The three phylogroups were represented in localities 41, BP, EI, Es, EV, Fr, LT, Mo, NJ and NY, whereas localities 52, FE, Sa and SV all belonged to the same phylogroup, although they had more than one haplotype.

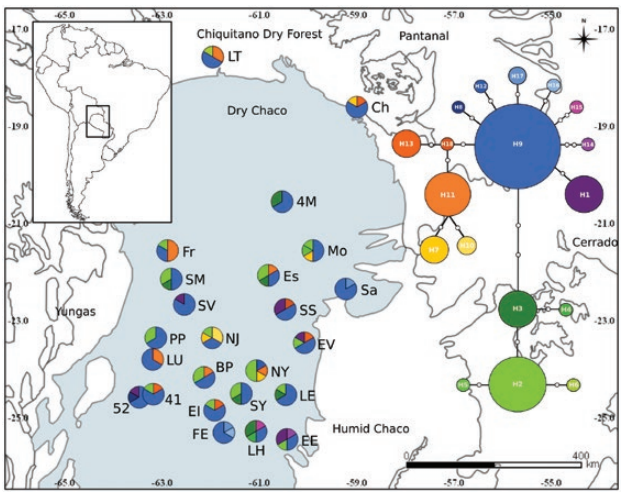

Figure 2. Geographical distribution and genealogical relationships among haplotypes recovered from 24 populations of *Bulnesia sarmientoi*. On the map, pie charts show the haplotype frequency in each population. Haplotype colours correspond to those shown in the network on the right. In the network, haplotypes are designated with numbers, and circle sizes are proportional to haplotype frequency. Codes are given in Table 1.

At the species level, haplotype diversity was h = 0.685 (SD: 0.039), nucleotide diversity was π = 0.0012 (SD: 0.0007), and the mean number of pairwise differences p = 1.798 (SD: 1.044). At the population level, the highest haplotype diversity was found for EE, ES, LH and NJ (h = 08667, SD 0.1291) and the highest nucleotide diversity was found for NY (π = 0.002332, SD 0.001591) (Fig. 3). The lowest haplotype diversity was found in Sa and SV (h = 0.3333, SD: 0.2152; Table 1).

For phylogroup P1 and for the whole species, demographic analyses showed evidence of recent demographic expansion, as indicated by the negative and significant values for Fu's neutrality test and by the sum of squared deviations (SSD) index, which indicates that the observed mismatch distribution did not differ from an expected sudden expansion model (Table 2). Evidence from demographic analyses supports BSPs

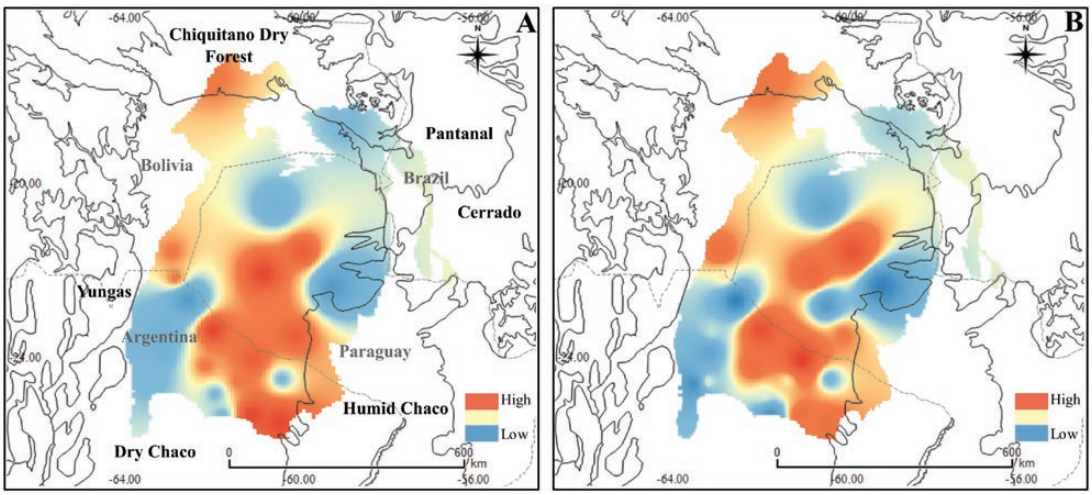

Figure 3. Spatial analysis of genetic diversity, showing the spatial interpolation of (A) gene diversity and (B) nucleotide diversity of chloroplast DNA from 24 localities of *B. sarmientoi*. Ecoregions and national boundaries are delimited with bold and grey dotted lines, respectively. In both indices, red indicates higher values and blue lower values.

Table 1. Collection localities, coordinates and number of individuals per haplotype, and gene diversity and nucleotide diversity (mean ± SD) for the sampled sites of *Bulnesia sarmientoi* in the Gran Chaco

Locality (code)	Latitude (S)	Longitude (W)	Haplotype (number of individuals)	Gene diversity	Nucleotide diversity
Santa Monica (SM)	22.1545387	62.8138602	2 (2) 3 (1) 9 (3)	0.7333 ± 0.1552	0.001144 ± 0.000888
Santa Victoria (SV)	22.6728865	62.5470401	1 (1) 9 (5)	0.3333 ± 0.2152	0.000220 ± 0.000290
Pluma de Pato (PP)	23.3766027	63.0904380	2 (2) 9 (4)	0.5333 ± 0.1721	0.001056 ± 0.000835
La Unión (LU)	23.8554779	63.1592984	9 (4) 11 (2)	0.5333 ± 0.1721	0.000704 ± 0.000618
RP41 (41)	24.5241833	63.1860494	2 (1) 9 (4) 13 (1)	0.6000 ± 0.2152	0.001100 ± 0.000862
RP52 (52)	24.5887598	63.4991095	1 (1) 8 (1) 9 (4)	0.6000 ± 0.2152	0.000440 ± 0.000447
Norte Juárez (NJ)	23.3576944	61.9628888	2 (1) 7 (1) 9 (2) 10 (2)	0.8667 ± 0.1291	0.002024 ± 0.001410
Bolsa Palomo (BP)	24.1740555	62.1414166	2 (2) 9 (3) 11 (1)	0.7333 ± 0.1552	0.001496 ± 0.001099
Norte Yema (NY)	24.0341666	61.0468055	2 (3) 7 (1) 11 (1) 12 (1)	0.8000 ± 0.1721	0.002332 ± 0.001591
Sur Yema (SY)	24.5164444	61.3787222	2 (2) 3 (1) 9 (3)	0.7333 ± 0.1552	0.001144 ± 0.000888
La Estrella (LE)	24.4668055	60.3936388	2 (1) 3 (1) 9 (4)	0.6000 ± 0.2152	0.000924 ± 0.000755
El Espinillo (EE)	25.4482562	60.4258352	1 (2) 2 (1) 9 (2) 14 (1)	0.8667 ± 0.1291	0.001452 ± 0.001072
Las Hacheras (LH)	25.2898319	61.0650607	3 (2) 4 (1) 9 (2) 15 (1)	0.8667 ± 0.1291	0.001452 ± 0.001072
El Impenetrable (EI)	24.8138040	61.8999646	2 (1) 9 (4) 13 (1)	0.6000 ± 0.2152	0.001100 ± 0.000862
Fuerte Esperanza (FE)	25.2970766	61.7390826	9 (4) 16 (1) 17 (1)	0.6000 ± 0.2152	0.000440 ± 0.000447
Selva Serena (SS)	22.7782830	60.4676000	1 (2) 9 (3) 18 (1)	0.7333 ± 0.1552	0.000572 ± 0.000534

Table 1. Continued

Locality (code)	Latitude (S)	Longitude (W)	Haplotype (number of individuals)	Gene diversity	Nucleotide diversity
Esperanza Viva (EV)	23.4631330	60.0743170	1 (1) 2 (1) 9 (3) 13 (1)	0.8000 ± 0.1721	0.001320 ± 0.000994
Salar (Sa)	22.3555830	59.2188330	9 (5) 17 (1)	0.3333 ± 0.2152	0.000220 ± 0.000290
Montanía (Mo)	21.5638500	59.8905000	2 (1) 5 (1) 7 (1) 9 (3)	0.8000 ± 0.1721	0.001936 ± 0.001359
4 de Mayo (4M)	20.5615170	60.5377000	3 (2) 9 (4)	0.5333 ± 0.1721	0.000704 ± 0.000618
Estigarribia (Es)	22.0747670	60.7946330	2 (2) 3 (1) 9 (2) 11 (1)	0.8667 ± 0.1291	0.001584 ± 0.001151
Frontera (Fr)	21.5701330	62.8896170	2 (1) 9 (2) 11 (3)	0.7333 ± 0.1552	0.001452 ± 0.001072
Chiquitanía (Ch)	18.6212830	58.9981830	7 (1) 9 (4) 13 (1)	0.6000 ± 0.2152	0.001012 ± 0.000808
Llanto de Tigre (LT)	17.5905830	61.9591330	6 (1) 9 (3) 11 (2)	0.7333 ± 0.1552	0.001364 ± 0.001020

(Fig. 4), which show a very wide confidence interval reducing the statistical power of the observed population growth pattern. The BSP analyses performed at the species level displayed a pattern consistent with population growth starting about 1.75 Mya and reaching the highest effective population size about 100 Mya; since then, there has been a slight decrease in population size. The BSP for P1 showed a very similar pattern, with population growth beginning about ~3 Mya, a constant increase until ~100 kya and a decrease since then.

Phylogenetic reconstruction and divergence time estimates

The Bayesian phylogenetic reconstruction is depicted in Figure 5. *Bulnesia sarmientoi* diverged from other *Bulnesia* species at about 5 Mya (12.06–1.04 Mya, 95% HPD) and about 2.2 Mya (5.26–0.34 Mya, 95% HPD) the species diversified into two clades: one is well supported [posterior probability (PP) = 1] and retrieved all haplotypes of P3; the other clade, weakly supported (PP = 0.32), nests two well-supported (PP = 1) subclades clustering the haplotypes of P1 and P2, respectively. Diversification times of phylogroups started about 1.4 Mya (3.53–0.2 Mya, 95% HPD) for P1, 1.1 Mya (2.94–0.1 Mya, 95% HPD) for P2, and 1.1 Mya (2.93–0.08 Mya, 95% HPD) for P3.

Bayesian spatio-temporal diffusion analyses

The spatial diffusion analysis (Fig. 6 and Supplementary Material Video) suggests that the expansion of lineages began in the central-southern area of the distribution ~200 kya. After 0.18 kya, the species expanded its range significantly towards the north, south and west, until 0.15 kya, when numerous diffusion events occurred over already colonized areas. Approximately 60–50 kya, a second significant expansion is observed, with lineages colonizing new areas mainly towards the west. From ~30 kya until about 16 kya particularly accelerated diversifications occurred. Since then, only diffusion of several lineages was observed, with no significant spatial expansion of the species range. The mean spatial diffusion rate was 7146.64 km/Myr (=7.15 m/year; 95% HPD: 15611.86–139.77).

ENM in geographical and climatic space

The results of the three algorithms selected to model the current potential distribution using the bioclimatic

Table 2. Results of demographic analyses

Index	All lineages	Phylogroup 1 (P1)	Phylogroup 2 (P2)	Phylogroup 3 (P3)
Ns	144	91	21	32
D	−1.11748	**−2.04397**	0.25046	−0.57853
	P = 0.12240	***P* = 0.00170**	*P* = 0.64370	*P* = 0.32230
Fs	**−7.49782**	**−7.32088**	−0.56463	−2.12025
	***P* = 0.00850**	***P* = 0.00000**	*P* = 0.33680	*P* = 0.04610
SSD	**0.02435**	**0.00073**	**0.00077**	**0.02502**
	***P* = 0.34080**	***P* = 0.59750**	***P* = 0.89090**	***P* = 0.07020**
rg	**0.07969**	**0.27756**	**0.04834**	**0.19821**
	***P* = 0.42970**	***P* = 0.57500**	***P* = 0.80580**	***P* = 0.06630**

Number of sequences (*Ns*), Tajima's *D* (significance level: $P < 0.05$), Fu's *Fs* (significance level: $P < 0.02$), sum of squared deviations (SSD) for a stepwise expansion model as null hypothesis, and the raggedness index (rg) for population expansion as null hypothesis are shown. In bold, results supporting demographic expansion.

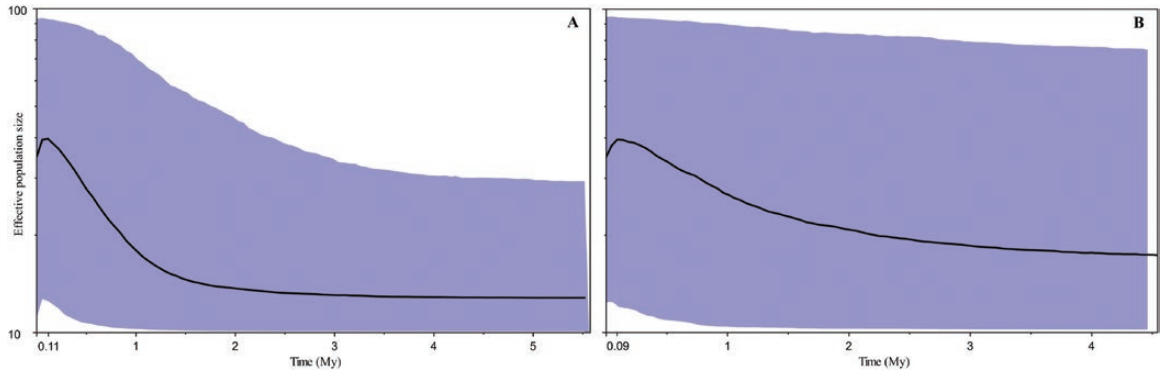

Figure 4. Temporal variation in effective population size in *Bulnesia sarmientoi*. Bayesian Skyline Plot for all species lineages (A) and for phylogroup P1(B). The *y*-axis represents effective population size expressed on a logarithmic scale. Bold lines correspond to the median values of the effective population size over time, and the shaded areas represent the 95% highest posterior densities over the median estimates along the coalescent history of the lineages.

variables of MERRAclim are summarized in Figure 7A. The validation values and thresholds of the models for each algorithm are shown in Table 3. In sum, the validation values of the chosen models indicate that our model is statistically different from random, predicting that percentages of presences and successes are higher than failures. The variables that most contributed to the current model for *B. sarmientoi* and their permutation importance values are summarized in Supplementary Material S2; notably, five of the nine selected variables contribute with 85.3% to the model. The current area of climatic favourability occurs mainly in the Dry Chaco, also occupying the ecoregions of Humid Chaco, Chiquitano Dry Forest and Pantanal, revealing clear limits for the species range, such as the Yungas ecoregion to the west of the distribution (which presents increasing altitude, marked thermal seasonality and greater precipitation than the Dry Chaco), and the extreme arid portion of the Dry Chaco to the south of the present distribution. The variables and their permutation importance values that most contributed in the current model and that were used to obtain the palaeodistribution models are summarized in Supplementary Material S2. Two variables contributed more than 83%: 'Mean temperature of warmest quarter' and 'Annual precipitation'. The models (Fig. 8) showed that the extent of the climatically favourable area would have been similar under the mid-Holocene and current climatic scenarios, whereas in the LGM and LIG, climatically favourable areas occurred strictly within the current Dry Chaco. The MESS and MOP analyses suggest that there are no strict extrapolations in the projection zone, meaning that in the palaeodistribution area there are no 'out of range' climatic combinations with respect to current environmental conditions of the calibration area (M).

The climatic optimum, which holds the greatest climatic suitability (> 0.8), is located in the Dry Chaco region, coincident with the climatic refugia of the LGM and LIG scenarios (Fig. 7B).

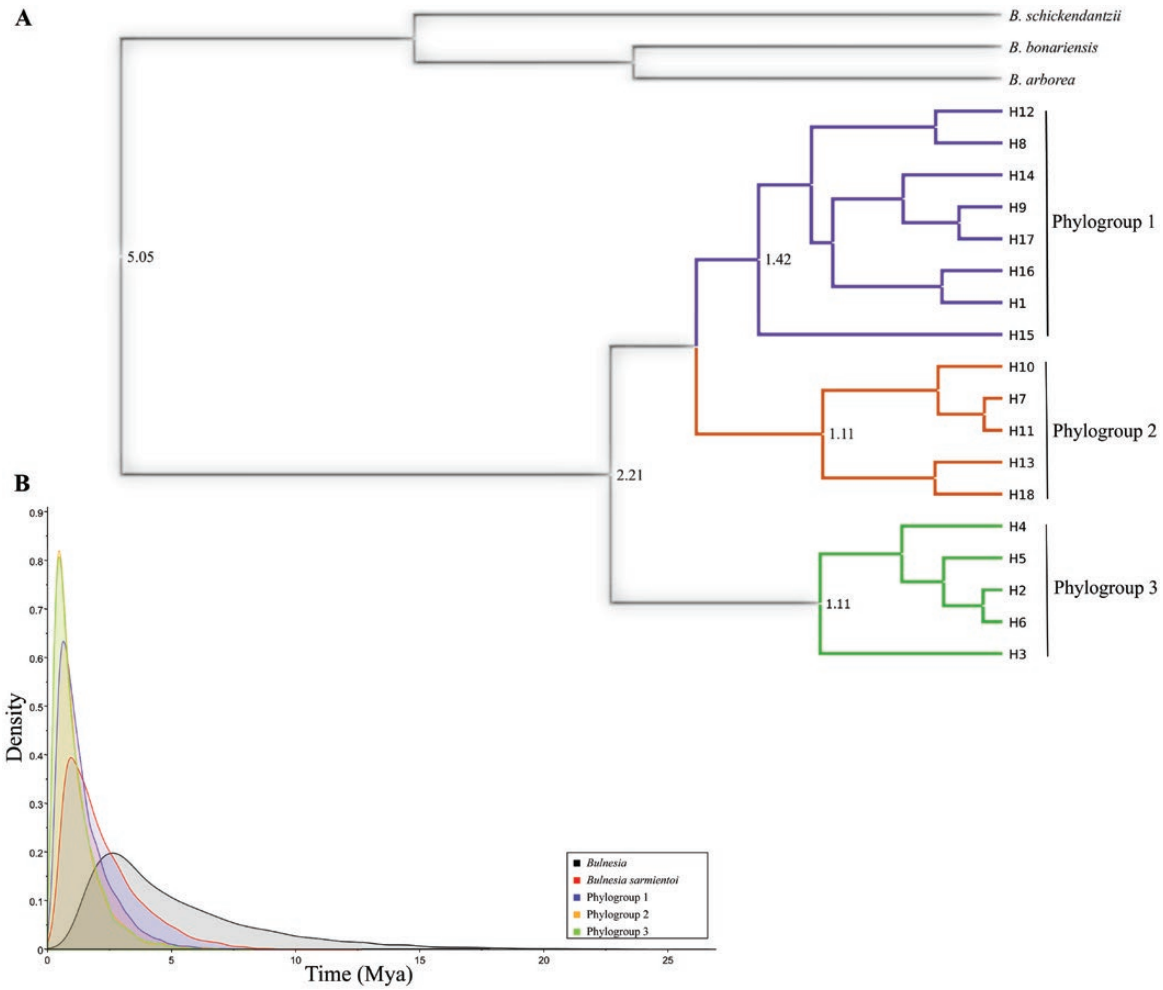

Figure 5. Phylogenetic reconstruction and divergence time estimates for the three phylogroups identified in *Bulnesia sarmientoi*. A, numbers in the tree indicate the mean dating value estimated for each supported node (PP > 0.95) expressed in million years ago. B, posterior density plots of divergence times for each node. Colours for each phylogroup correspond to the haplotype network shown in Figure 2.

GENETIC DIVERSITY, CLIMATIC STABILITY AND CLIMATIC OPTIMUM

We found areas of spatial agreement of all layers (climatic optimum, highest levels of genetic diversity, and LGM and/or LIG palaeodistribution) and areas where only two layers were coincident: (1) climatic optimum and highest levels of genetic diversity, (2) LGM palaeodistribution and climatic optimum, and (3) LGM palaeodistribution and highest levels of genetic diversity (Fig. 9). All layers coincided in two areas, one in the Paraguayan Chaco and the other near the triple borderline formed by Argentina, Bolivia and Paraguay. The areas where only the climatic optimum and high levels of genetic diversity coincide are found in the south-east of the distribution, whereas the LGM palaeodistribution and climatic optimum overlap in the central area (Paraguay).

DISCUSSION

ARIDIFICATION AND LINEAGE DIVERSIFICATION: FROM THE MIOCENE TO THE QUATERNARY

The Oligocene–Miocene origin of the xerophytic flora of the Chaco has been proposed to be associated with uplift of the Andes and the consequent drastic environmental changes (Spichiger *et al.*, 1995). Our results agree with this hypothesis, because our indirect dating analysis, with the *trn*L-*trn*F sequences, suggests that the clade Larreoideae would have diversified about 26 Mya; a similar dating was obtained by Wu *et al.* (2015) with *rbc*L sequences, and by Bellstedt *et al.* (2012), with the same molecular markers used in our study. In this line, fossil records found in the Valles Calchaquíes and Mesopotamia phytogeographical regions (Barreda *et al.*, 2007) support the existence of the family Zygophyllaceae during the Early–Middle Miocene. The

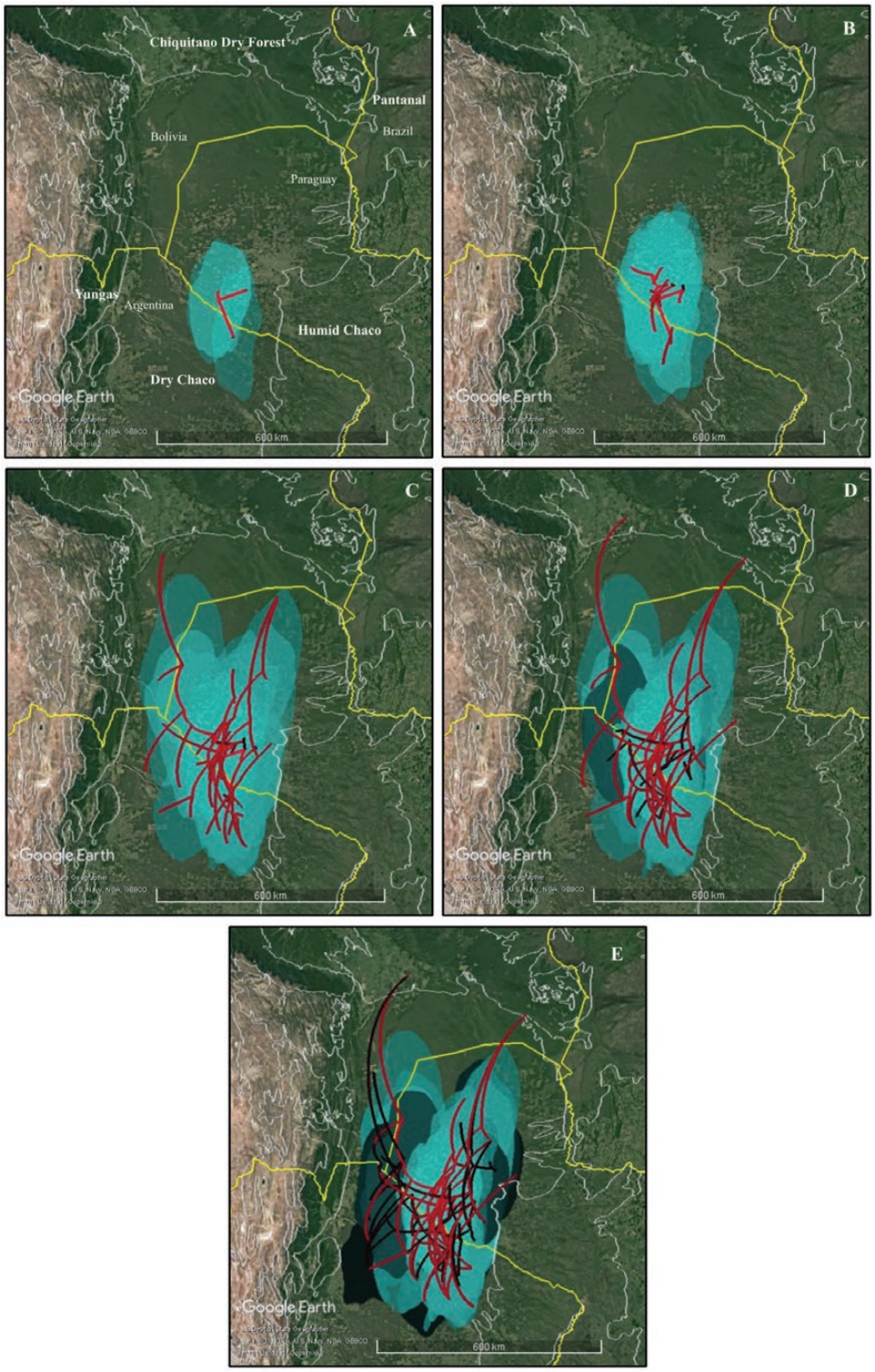

Figure 6. Bayesian spatio-temporal diffusion analysis of *Bulnesia sarmientoi* showing lineages at different time points based on the maximum clade credibility (MCC) tree. Time slices are: A, 212 000 years ago; B, 180 000 years ago; C, 150 000 years ago; D, 55 000 years ago; E, 16 000 years ago. Shaded areas correspond to the 80% HPD uncertainty in the location of ancestral branches; a gradient from light to dark shades indicates older vs. younger events.

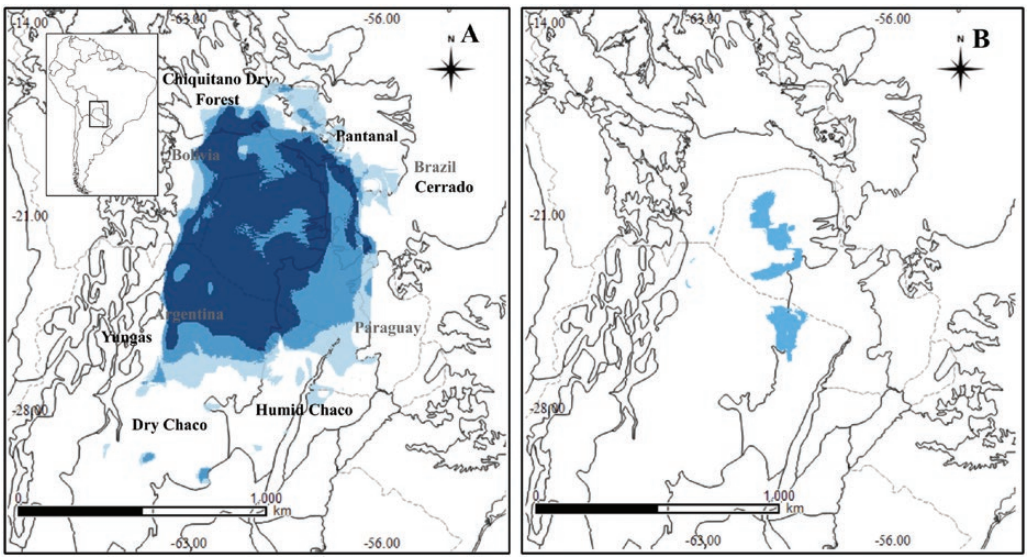

Figure 7. Current ENM performed with MERRAclim bioclimatic variables. A, binary map of the current ENM in geographical space; the darkest colour represents the area where the three algorithms join (MaxEnt, BIOCLIM and Support Vector Machine). B, Climatic centroid shown in geographical space; the blue area represent an approximation to the climatic optimum with suitability > 0.8. Ecoregions and national boundaries are delimited with bold and grey dotted lines, respectively.

ancestral lineages of *B. sarmientoi* diverged from the remaining *Bulnesia* species about 5 Mya, at the beginning of the Pliocene. This dating is coincident with the most important last Andean uplift (Ramos & Ghiglione, 2008); this geological event produced intense aridification that impacted on the climate and ecology of the arid regions of South America (Hoorn *et al.*, 2010).

The current lineages of *B. sarmientoi* diversified at the beginning of the Pleistocene, about 2.2 Mya. During that period the Chaco acquired its biogeographical identity (Iriondo, 2010), characterized by soils formed by the accumulation of Andean sediments (Prado, 1993a; Iriondo & Brunetto, 2016), an important aspect for the species because it grows especially on hard and poorly drained soils (Adamoli *et al.*, 1972; Prado, 1993b). Diversifications within phylogroups occurred more recently, estimated at ~1.4–1 Mya during the maximum of the Greatest South American Glaciation, when the arid conditions during this glacial period (Iriondo, 2010) probably favoured species diversification. Divergence times of *B. sarmientoi* lineages coincide with the dating reported for intraspecific phylogroups of mammals, birds, fishes and invertebrates in South America (references in Turchetto-Zolet *et al.*, 2013), supporting the idea that Pleistocene glaciation cycles greatly influenced speciation and diversification processes.

Range dynamics, demography and diffusion patterns during the Quaternary

The hypothesis that the Chaco, especially the Dry Chaco, was an unstable area during the glacial periods (Ab'Saber, 2000; Pennington *et al.*, 2004; Speranza *et al.*, 2007; Iriondo, 2010; Iriondo & Brunetto, 2016) is supported in the present work by the absence of population structure, the inferred spatial and demographic expansions, and changes in the location and extent of the climatically favourable areas over time. We suggest that the absence of population structure is a consequence of the environmental instability in the Dry Chaco. However, it should be noted that within the three phylogroups, exclusive haplotypes are found at sites located approximately 21°–25°S and 60°W, suggesting *in situ* persistence in this area. As reported for other phytogeographical regions (e.g. Sérsic *et al.*, 2011; Turchetto-Zolet *et al.*, 2013), in addition to climatic oscillations, river basin dynamics could have been an important factor shaping the retrieved genetic pattern. Because during the dry and cold glacial climate basins would have been smaller than today (Iriondo & Brunetto, 2016), terrestrial species dispersion would have been more feasible than today. Unfortunately, the lack of previous phylogeographical studies in the Chaco regions does not allow comprehensive comparisons among multiple species. Interestingly, however, similar genetic patterns were previously detected in a tree species using microsatellite and chloroplast DNA markers (Caetano *et al.*, 2008); the authors found a population unit composed of Chaco individuals, and this unit was ambiguously substructured and comprised individuals also from the Humid Chaco, Chiquitanía and the Chaco–Yungas transition. Caetano & Naciri (2011), Campos-Krauer & Wisely (2011) and Werneck *et al.* (2012) obtained

Table 3. Validation values and thresholds for each algorithm and MaxEnt models

	MerraC *Bioclim*	MerraC *Domain*	MerraC *GLM*	MerraC *Rforest*	MerraC *SVM*	MerraC *Mx LQHPT_3.5*	MerraC *Mx LQHPT_2*	MerraC *Mx LQH_3*	WorldC *Mx LQH_4*	WorlC *Mx LQH_3*	WorldC *Mx LQ_1*
AUC ratio 0.05	**1.613737**	1.686361	1.716084	1.969881	**1.721223**	1.790372	**1.798265**	1.756822	1.931816	**1.938229**	1.92623
partial AUC 0.05	**0.8068686**	0.8430497	0.8579649	0.9849398	**0.8606045**	0.895186	**0.8991327**	0.8784098	0.965908	**0.969115**	0.963117
partial AUC 0.5	**0.4999999**	0.4999208	0.4999541	0.4999997	**0.4999959**	0.4999999	**0.5**	0.4999994	0.499999	**0.5**	0.499999
p binomial	**1.52e-07**	2.2e-16	8.05e-016	1,16E-010	**2.2e-16**	2.2e-16	**2.2e-16**	2.2e-16	2.2e-16	**2.2e-16**	2.2e-16
Unpredicted points, threshold 100%	**1**	0	1	24	**1**	0	**1**	1	0	**0**	0
Percentage unpredicted points	**1.69**	0.00	1.69	40.68	**1.69**	0.00	**1.69**	1,69	0	**0**	0
Predicted surface, threshold 100%	**29244**	48235	78322	4067	**25069**	38940	**34158**	44756	47836	**43281**	46283
(Unpredicted points/ surface) × 100	**0.003**	0	0.001	0.59	**0.004**	0	**0.003**	0.002	0	**0**	0
Thresholds	**0.01266**	0.23674	0.01177	0.48233	**0.02734**	0.14747	**0.11521**	0.09714	0.09135	**0.06330**	0.05947

MerraC: models with MERRAclim bioclimatic variables; WorldC: models with WorldClim bioclimatic variables; GLM: Generalized Linear Model; RForest: Random Forest; SVM: Support Vector Machines; Mx: MaxEnt; for Mx models: Linear, Quadratic, Product, Threshold and/or Hince features_value of regularization multiplier. The selected algorithms and Maxent models are indicated in bold.

similar results for two woody species, for mammals and for reptiles, respectively. A study carried out in an herbaceous species complex (Speranza *et al.*, 2007) found population structure, but not specifically in the Chaco region.

Based on the spatial diffusion analysis, three significant range expansions occurred in *B. sarmientoi*, which are coincident with three glacial periods (at 180, 60 and 16 kya; Rabassa & Clapperton, 1990); during these glaciations the semi-arid forests of the Chaco would have occupied a larger area. Conversely, in the interglacial periods no range expansions were evidenced, but processes of diffusion and lineage diversification were inferred in the occupied areas. It seems that demographic expansion was not coupled with range expansion, because historical demography analyses show that the highest effective population size was estimated for the final stages of the Last Interglacial (~120 kya), followed by a population decline until the present, possibly as a consequence of the increase of humidity in the Chaco during interglacial periods compared to today (Markgraf, 1985; Iriondo, 2010).

The ENM obtained for the last 120 000 years (under the assumption of niche conservatism, e.g. Wiens *et al.*, 2010) shows a growing increase in the climatically suitable area for *B. sarmientoi* from the LIG to the mid-Holocene, followed by a period of stability until the present. Regarding the area in the model, the projection to the past does not strictly reflect the area of distribution of the species in the past, but rather areas of suitable climate (Wiens *et al.*, 2009; Peterson *et al.*, 2011). This stability found for the distribution area of the focal species, and indirectly for the Dry Chaco, since about 6000 years ago was also reported for the Chaco and surrounding ecoregions in previous studies (e.g. Werneck *et al.*, 2011; Bartoleti *et al.*, 2017; Trujillo-Arias *et al.*, 2017). As mentioned above, there are few works using ENM that focused particularly on the Chaco ecoregion, which hinders the comparison of our results mainly because of differences in the calibration areas for each model. The smaller potential distribution inferred for the LIG than for the LGM is consistent with results retrieved by diffusion analysis, which showed range expansions during glacial periods instead of interglacials. Argollo Bautista & Iriondo (2008) suggest that the climatic changes of the Quaternary were more extreme in relation to humidity than to temperature, which is in agreement with the observed contribution of the variables, and with the general results of the ENM.

Based on geomorphological and stratigraphic studies, Iriondo (1992) postulated that the Chaco Forest was an unstable area during Quaternary fluctuations, and that plant communities alternated between xerophytic and tropical–subtropical vegetation types. In particular, Iriondo postulated a spatial shift of

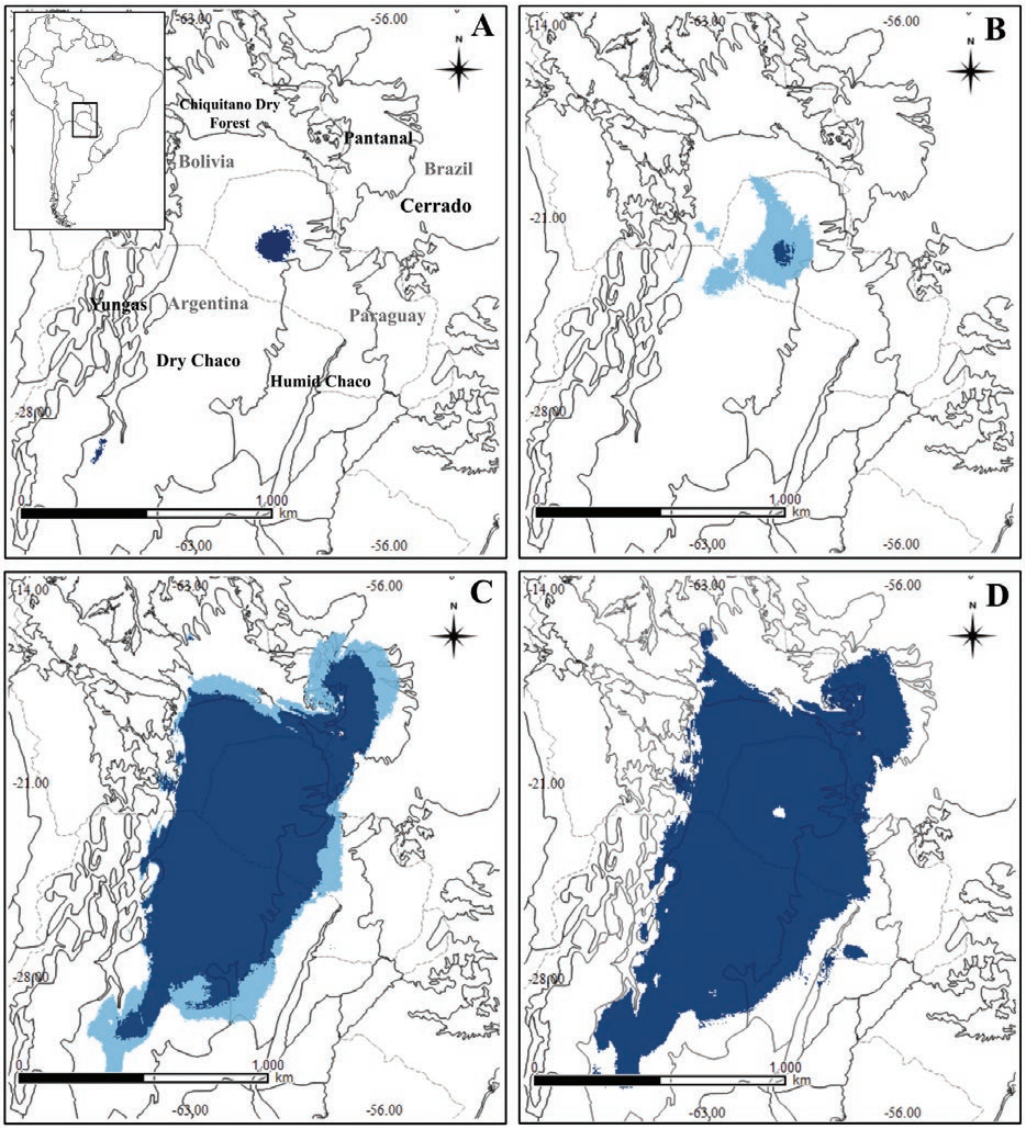

Figure 8. ENM over time in geographical space with WorldClim bioclimatic variables. A, LIG; B, LGM, cc and mr GCM models; C, mid-Holocene, cc and mr GCM models; D, present. Ecoregions and national boundaries are delimited with bold and grey dotted lines, respectively.

the Chaco region towards the north-east during dry and cold periods, the species contacting with tropical and subtropical elements, followed by a retraction to the south-west during humid and warm periods. This hypothesis is supported by findings reported by Speranza *et al.* (2007) and the present results. Similar vegetation movements bordering the Dry Chaco were suggested by Pennington *et al.* (2004) and Werneck *et al.* (2011) for the Seasonally Dry Tropical Forest (SDTF). Considering the dynamics of the distribution area of *B. sarmientoi*, the expansion of the arid climate towards the north-east during the glacial period would have promoted species shifts to areas currently corresponding to the Chiquitanía and Humid Chaco; these areas today remain marginal, and current species populations would represent relicts of a more extended ancient distribution. Accordingly, previous studies postulated that the Chiquitanía and Humid Chaco would have represented climatic refugia during the LGM (Werneck *et al.*, 2011; Vitorino *et al.*, 2016; Trujillo-Arias *et al.*, 2017). Interestingly, our diffusion analysis suggests that the Chiquitanía populations, which are the northernmost ones, would have been colonized during dry and cold glacial periods. Thus, the high genetic variability found in that geographical area coupled with the inferred processes of colonization and diffusion validates this hypothesis, suggesting that the highly diverse Chiquitania would have been

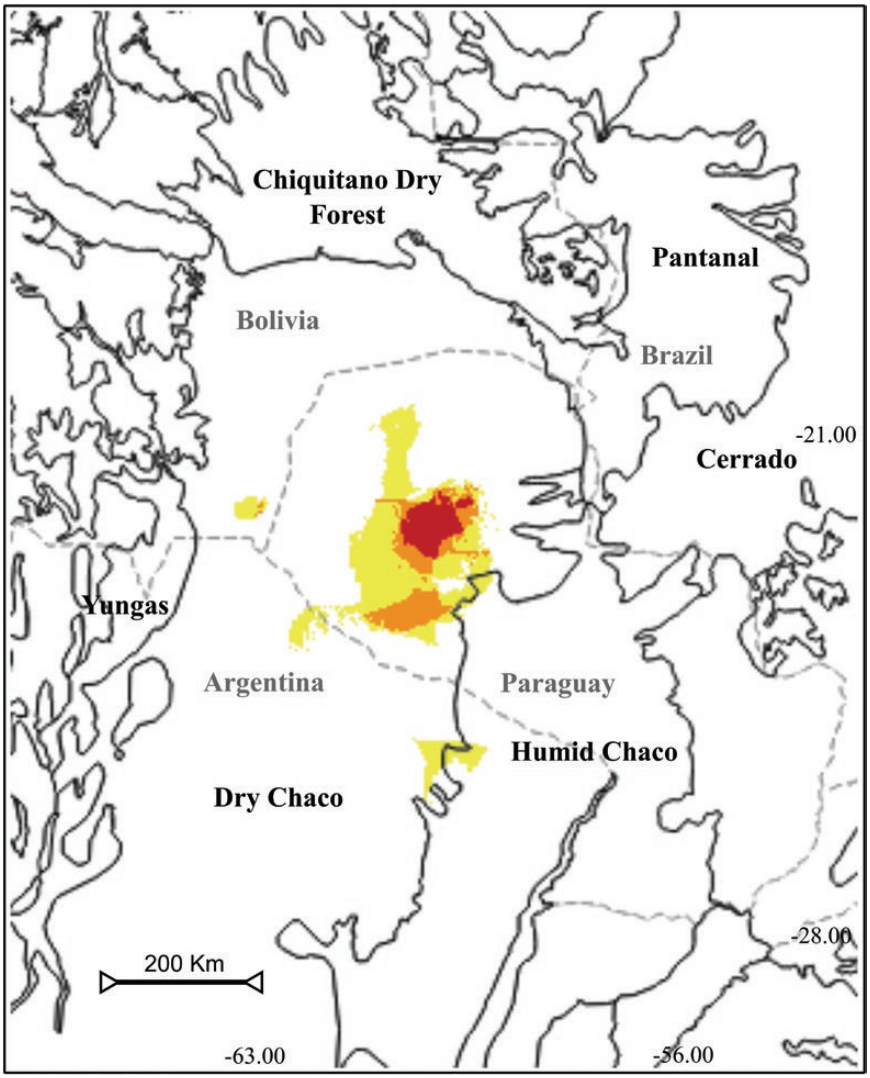

Figure 9. Climatic stability, climatic optimum and genetic diversity. Red: area of coincidence of LIG, LGM, climatic optimum and genetic diversity; orange: area of coincidence of LGM, climatic optimum and genetic diversity; yellow: area of coincidence of any of these three conventions – (1) climatic optimum and genetic diversity, or (2) LGM and climatic optimum, or (3) LGM and genetic diversity. Ecoregions and national boundaries are delimited with bold and grey dotted lines, respectively.

a refugium where flora from different phytogeographical origins met during glacial/interglacial periods.

LINKING HOTSPOTS OF GENETIC DIVERSITY WITH CLIMATIC STABILITY AND THE CLIMATIC OPTIMUM

Genetic evidence indicates that *B. sarmientoi* is a species of Pleistocene origin that would have suffered three range expansion events (i.e. spatial expansions) during glacial periods, whereas its effective population size would have increased progressively (i.e. demographic expansion) until ~120 kya, declining after the LIG. The spatial diffusion analysis suggests an origin of diversification in the mid-south areas of the distribution, which is consistent with the high levels of genetic variability areas found for the species. Interestingly, according to the ENM, this area would have been stable over time; therefore, we propose it as a possible climatic refugium (following Gavin *et al.*, 2014) for the Dry Chaco region. The location of the highest climatic suitability in ecological space – the current climatic optimum – is even more interesting because it coincides with the climatic refugium and with the highest levels of genetic variability. A similar combination is found marginally in the triple boundary between Argentina, Bolivia and Paraguay. Similarly, populations located in the Chiquitanía, an area of climatic stability during the LGM for the SDTF (Werneck *et al.*, 2011; Vitorino *et al.*, 2016; Trujillo-Arias

et al., 2017), are also genetically diverse, and it is therefore supported as a climatic refugium by our results. These populations are also valuable because they are located marginally in the climatic space of the species, away from the climatic centroid (typically semi-arid), and are relevant as a potential evolutionary source under new climatic scenarios (Parmesan, 2006; Macdonald *et al.*, 2017). Our next step is to obtain information about molecular markers indicative of rapid mutation rates, which would be very useful to understand the recent evolutionary history of the species, as well as to provide further evidence for the identification of evolutionarily significant units.

CONCLUSION

In this study we integrated geo-statistical phylogeography and multi-algorithm-based niche modelling to explore the spatio-temporal dynamics of an emblematic species of the NDFs. Our results suggest that *B. sarmientoi* forests underwent population range expansion events during the glacial periods of the Pleistocene, whereas during the interglacial periods they would have undergone population range stasis, although with demographic expansion and genetic diversification. We suggest that the Dry Chaco forest shifted in a north-east to south direction as climate conditions changed, being restricted during the LGM by the stable areas of SDTF to the north and east. Finally, we identified a putative refugium in climatically stable areas during the last 120 kyr in the Dry Chaco, which is coincident with the area of highest genetic diversity and with the spatial location of the climatic optimum of the focal species. Future demographic history and ENM studies including several species from the Chaco region will allow us to unravel the evolutionary history of this ecoregion and to explore if hotspots of genetic diversity are concordant with climatically stable areas and with the spatial location of the climatic optimum, as found in our study. This is of great relevance in the context of current climate change, allowing us to improve our characterization and understanding of climatic refugia in one of the largest dry forests of the world.

ACKNOWLEDGEMENTS

A.C. and A.N.S. acknowledge the National Research Council of Argentina (CONICET) and the Universidad Nacional de Córdoba (UNC) as researchers, and G.A.C. as doctoral fellowship holder. A.R.V. thanks the Instituto Nacional de Tecnología Agropecuaria (INTA) and E.M.M. the Universidad Nacional Autónoma de México (UNAM). We also thank Diego López Lauenstein, Nicolas Rocamundi, Paula Venier and Carmen Vega for the help in the collection of samples. Laboratory assistance was provided by Yohana Ruartes. We thank Jorgelina Brasca for reviewing the English text. The Ministry of Environment and Sustainable Development (Argentina) provided assistance and points of occurrence of the species. This work was supported by INTA (PNFOR-1104064) to SN Marcucci Poltri, FONCyT (PICT-2015-3089; PICTO-2014-0013) to A.N.S., and CONICET (PIP-11220150100690CO) to A.N.S. G.A.C. also thanks the Red de Macrouniversidades de América Latina y el Caribe for a research stay in LAE (UNAM). We are grateful to IMBIV (UNC-CONICET), IFRGV (CIAP, INTA) and LAE (UNAM). Finally, we thank two anonymous reviewers for their helpful comments.

REFERENCES

Abraham de Noir F, Bravo S, Abdala R. 2002. Dispersal mechanisms in some woody native species of Chaco Occidental and Serrano. *Quebracho* **9:** 140–150.

Adamoli J, Neumann R, De Colina ADR, Morello J. 1972. El Chaco aluvional salteño. *Revista de Investigaciones Agropecuarias* **9:**165–237.

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

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.

Anderson RP, Gómez-Laverde M, Peterson AT. 2002. Geographical distributions of spiny pocket mice in South America: insights from predictive models. *Global Ecology and Biogeography* **11:** 131–141.

Argollo Bautista J, Iriondo M. 2008. *El Cuaternario de Bolivia y regiones vecinas*. Santa Fe: Museo Provincial de Ciencias Naturales Florentino Ameghino.

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, Cosacov A, Ferreiro G, Johnson LA, Sérsic AN. 2017. Travelling to the south: phylogeographic spatial diffusion model in *Monttea aphylla* (Plantaginaceae), an endemic plant of the Monte Desert. *PLoS ONE* **12:** e0178827.

Barreda V, Anzótegui LM, Prieto AR, Aceñolaza P, Bianchi MM, Borromei AM, Brea M, Caccavari M, Cuadrado GA, Garralla S, Grill S, Guerstein GR, Lutz AI, Mancini MV, Mautino LR, Ottone EG, Quattrocchio ME, Romero EJ, Zamaloa MC, Zucol A. 2007. Diversificación y cambios de las angiospermas durante el Neógeno en Argentina. *Ameghiniana* **11:** 173–191.

Bartoleti LFM, Peres EA, Sobral-Souza T, Fontes FV, Silva MJ, Solferini VN. 2017. Phylogeography of the dry

vegetation endemic species *Nephila sexpunctata* (Araneae: Araneidae) suggests recent expansion of the Neotropical Dry Diagonal. *Journal of Biogeography* **44:** 2007–2020.

Bellstedt DU, Galley C, Pirie DP, Linder HP. 2012. The migration of the palaeotropical arid flora: Zygophylloideae as an example. *Systematic Botany* **37:** 951–959.

Bielejec F, Rambaut A, Suchard MA, Lemey P. 2011. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* **27:** 2910–2912.

Bielejec F, Baele G, Vrancken B, Suchard MA, Rambaut A, Lemey P. 2016. Spread3: Interactive visualization of spatiotemporal history and trait evolutionary processes. *Molecular Biology and Evolution* **33:** 2167–2169.

Bowen BW, Gaither MR, Di Battista JD, Iacchei M, Andrews KR, Grant WS, Toonen RJ, Briggs JC. 2016. Comparative phylogeography of the ocean planet. *Proceedings of the National Academy of Sciences USA* **113:** 7962–7969.

Busby JR. 1991. BIOCLIM A bioclimate analysis and prediction system. In: Margules CR, Austin MP, eds. *Nature conservation: cost effective biological surveys and data analysis*. Canberra: CSIRO, 64–68.

Caetano S, Prado D, Pennington RT, Beck S, Oliveira-Filho A, Spichiger R, Naciri Y. 2008. The history of Seasonally Dry Tropical Forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). *Molecular Ecology* **17:** 3147–3159.

Caetano S, Naciri Y. 2011. The biogeography of seasonally dry tropical forests in South America. In: Dirzo R, Young HS, Mooney HA, Ceballos G, eds. *Seasonally Dry Tropical Forests*. Washington: Island Press, 23–44.

Campos-Krauer JM, Wisely SM. 2011. Deforestation and cattle ranching drive rapid range expansion of capybara in the Gran Chaco ecosystem. *Global Change Biology* **17:** 206–218.

Caparroz R, Seixas GHF, Berkunsky I, Collevatti RG. 2009. The role of demography and climatic events in shaping the phylogeography of Amazona aestiva and definition of management units for conservation. *Diversity Distribution* **15:** 459–468.

Carstens BC, Richards CL. 2007. Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution* **61:** 1439–1454.

Ceballos G, Brown JH. 1995. Global patterns of mammalian diversity, endemism and endangerment. *Conservation Biology* **9:** 559–568.

Clark M, Aide TM, Grau R, Riner G. 2010. A scalable approach to mapping annual land cover at 250 m using MODIS time series data: a case study in the Dry Chaco ecoregion of South America. *Remote Sensing of Environment* **114:** 2816–2832.

Collevatti RG, Terribile LC, Lima-Ribeiro MS, Nabout JC, Oliveira G, Rangel TF, Rabelo SG, Diniz-Filho JA. 2012. A coupled phylogeographical and species distribution modelling approach recovers the demographical history of a Neotropical seasonally dry forest tree species. *Molecular Ecology* **21:** 5845–5863.

Cosacov A, Johnson LA, Paiaro V, Cocucci AA, Córdoba FE, Sérsic AN. 2013. Precipitation rather than temperature influenced the phylogeography of the endemic shrub *Anarthrophyllum desideratum* in the Patagonian steppe. *Journal of Biogeography* **40:** 168–182.

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

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.

Elith J, Kearney M, Phillips S. 2010. The art of modelling range-shifting species. *Methods in Ecology and Evolution* **1:** 330–342.

Eizirik E, Kim JH, Menotti-Raymond M, Crawshaw PG, O'Brien SJ, Johnson WE. 2001. Phylogeography, population history and conservation genetics of jaguars (*Panthera onca*, Mammalia, Felidae). *Molecular Ecology* **10:** 65–79.

Excoffier L, Laval G, Schneider S. 2005. Arlequin version 3.01: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47–60.

Fielding AH, Bell JF. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* **24:** 38–49.

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

Gao Y-D, Zhang Y, Gao X-F, Zhu Z. 2015. Pleistocene glaciations, demographic expansion and subsequent isolation promoted morphological heterogeneity: a phylogeographic study of the alpine *Rosa sericea* complex (Rosaceae). *Scientific Reports* **5:**11698.

Gavin DG, Fitzpatrick MC, Gugger PF, Heath KD, Rodrıguez-Sanchez F, Dobrowski SZ, Hampe A, Hu FS, Ashcroft MB, Bartlein PJ, Blois JL, Carstens BC, Davis EB, de Lafontaine G, Edwards ME, Fernandez M, Henne PD, Herring EM, Holden ZA, Kong WS, Liu J, Magri D, Matzke NJ, McGlone MS, Saltré F, Stigall AL, Tsai YH, Williams JW. 2014. Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytologist* **204:** 37–54.

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.

Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman SV, Goetz SJ, Loveland TR, Kommareddy A, Egorov A, Chini L, Justice CO, Townshend JR. 2013. High-resolution global maps of 21st-century forest cover change. *Science* **342:** 850–853.

Hardy OJ, Born C, Budde K, Daïnou K, Dauby G, Duminil J, Ewédjé EBK, Gomez C, Heuertz M, Koffi GK, Lowe AJ, Micheneau C, Ndiade-Bourobou D, Piñeiro R, Poncet V. 2013. Comparative phylogeography of African rain forest trees: a review of genetic signatures of vegetation history in the Guineo-Congolian region. *Comptes Rendus Geoscience* **345:** 284–296.

Hoorn C, Wesselingh FP, ter Steege H, Bermudez MA, Mora A, Sevink J, Sanmartín I, Sanchez-Meseguer A, Anderson CL, Figueiredo JP, Jaramillo C, Riff D, Negri FR, Hooghiemstra H, Lundberg J, Stadler T, Särkinen T, Antonelli A. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution and biodiversity. *Science* **330:** 927–931.

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

Iriondo MH. 2010. *Geología del Cuaternario en la Argentina*. Santa Fe: Museo Provincial de Ciencias Naturales Florentino Ameghino.

Iriondo MH, Brunetto E. 2016. *Cuaternario de Brasil, Paraguay y Uruguay*. Santa Fe: Museo Provincial de Ciencias Naturales Florentino Ameghino.

Keppel G, Van Niel KP, Wardell-Johnson GW, Yates CJ, Byrne M, Mucina L, Schut AG, Hopper SD, Franklin SE. 2012. Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global Ecology and Biogeography* **21:** 393–404.

Kuemmerle T, Altrichter M, Baldi G, Cabido M, Camino M, Cuellar E, Cuellar RL, Decarre J, Díaz S, Gasparri I, Gavier-Pizarro G, Ginzburg R, Giordano AJ, Grau HR, Jobbágy E, Leynaud G, Macchi L, Mastrangelo M, Matteucci SD, Noss A, Paruelo J, Piquer-Rodríguez M, Romero-Muñoz A, Semper-Pascual A, Thompson J, Torrella S, Torres R, Volante JN, Yanosky A, Zak M. 2017. Forest conservation: remember gran chaco. *Science* **355:** 465.

Lemey P, Rambaut A, Welch JJ, Suchard MA. 2010. Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology and Evolution* **27:** 1877–1885.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25:** 1451–1452.

Lira-Noriega A, Manthey JD. 2014. Relationship of genetic diversity and niche centrality: a survey and analysis. *Evolution* **68:** 1082–1093.

Macdonald SL, Llewelyn J, Moritz C, Phillips BL. 2017. Peripheral isolates as sources of adaptive diversity under climate change. *Frontiers in Ecology and Evolution* **5:** 88.

Markgraf V. 1985. Paleoenvironmental history of the last 10 000 years in northwestern Argentina. *Zentralblatt für Geologie und Paläontolgie*, **11:** 1739–1749.

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

Mereles F, Pérez de Molas L. 2008. Bulnesia sarmientoi *Lorentz ex Griseb. (Zygophyllaceae): estudio de base para su inclusión en el Apéndice II de la Convención CITES*. Lambaré, Paraguay: WWF Paraguay, **15**.

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.

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.

Osorio-Olvera L, Barve V, Barve N, Soberón J. 2016. *nichetoolbox: From getting biodiversity data to evaluating species distribution models in a friendly GUI environment. R package version 0.1.6.0*. Available at: http://shiny.conabio.gob.mx:3838/nichetoolb2/.

Owens HL, Campbell LP, Dornak LL, Saupe EE, Barve N, Soberón J, Ingenloff K, Lira-Noriega A, Hensz CM, Myers CE, Townsend Peterson A. 2013. Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecological Modelling* **263:** 10–18.

Parmesan C. 2006. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics* **37:** 637–669.

Pennington RT, Prado DE, Pendry CA. 2000. Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography* **27:** 261–273.

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, Sánchez-Cordero V. 1999. Conservatism of ecological niches in evolutionary time. *Science* **285:** 1265–1267.

Peterson AT, Papes M, Soberon J. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling* **213:** 63–72.

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

Peterson BJ, Graves WR. 2016. Chloroplast phylogeography of *Dirca palustris* L. indicates populations near the glacial boundary at the last glacial maximum in eastern North America. *Journal of Biogeography* **43:**314–327.

Petit RJ, Aguinagalde I, De Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Müller-Starck G, Demesure-Musch B, Palmé A, Martín JP, Rendell S, Vendramin GG. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300:** 1563–1565.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190:** 231–259.

Prado DE. 1993a. What is the Gran Chaco vegetation in South America?. I. A review. Contribution to the study of flora and vegetation of the Chaco. V. *Candollea* **48:** 145–172.

Prado DE. 1993b. What is the Gran Chaco vegetation in South America?. II. A redefinition. Contribution to the study of flora and vegetation of the Chaco. VII. *Candollea* **48:** 615–629.

Qiao H, Soberon J, Peterson AT. 2015. No silver bullets in correlative ecological niche modelling: insights fromtesting among many potential algorithms for niche estimation. *Methods in Ecology and Evolution* **6:** 1126–1136.

Rabassa J, Clapperton CM. 1990. Quaternary glaciations of the southern Andes. *Quaternary Science Reviews* **9:** 153–174.

Rambaut A, Drummond AJ. 2013. *TreeAnnotator v1.7.0*. Available at: http://beast.bio.ed.ac.uk

Rambaut A, Suchard MA, Xie D, Drummond AJ. 2013. *Tracer v1.5*. Available at: http://beast.bio.ed.ac.uk/Tracer.

Ramos VA, Ghiglione MC. 2008. Tectonic evolution of the Patagonian Andes. *Developments in Quaternary Science* **11:** 57–71.

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

SAyDS. 2007. *Primer Inventario Nacional de Bosques Nativos. Informe Regional Parque Chaqueño.* Buenos Aires: Secretaría de Ambiente y Desarrollo Sustentable de la Nación.

Sánchez-Azofeifa GA, Kalacska M, Quesada M, Calvo-Alvarado JC, Nassar JM, Rodríguez JP. 2005. Need for integrated research for a sustainable future in tropical dry forests. *Conservation Biology* **19:** 285–286.

Sérsic AN, Cosacov A, Cocucci AA, Johnson LA, Pozner R, Avila LJ, Sites Jr JW, Morando M. 2011. Emerging phylogeographical patterns of plants and terrestrial vertebrates from Patagonia. *Biological Journal of the Linnean Society* **103:** 475–494.

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.

Soberón J, Peterson AT. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics* **2:** 1–10.

Soberón J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters* **10:** 1115–1123.

Soltis DE, Morris AB, Mclachlan JS, Manos PS, Soltis PS. 2006. Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology* **15:** 4261–4293.

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.

Spichiger R, Palese R, Chautems A, Ramella L. 1995. Origin, affinities and diversity hot spots of the Paraguayan dendrofloras. *Candollea* **50:** 515–537.

Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of plastid DNA. *Plant Molecular Biology* **17:** 1105–1109.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7:** 453–464.

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.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25:** 4876–4882.

Tocchio LJ, Gurgel-Goncalves R, Escobar LE, Peterson AT. 2014. Niche similarities among white-eared opossums (Mammalia, Didelphidae): is ecological niche modelling relevant to setting species limits? *Royal Swedish Academy of Sciences* **44:** 1–10.

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, Pinheiro F, Salgueiro F, Palma-Silva C. 2013. Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology* **22:** 1193–1213.

US Geological Survey. 2001. *HYDRO 1k elevation derivative database*. Sioux Falls: U.S. Geological Survey.

Vallejos M, Volante JN, Mosciaro MJ, Vale LM, Bustamante ML, Paruelo JM. 2015. Transformation dynamics of the natural cover in the Dry Chaco ecoregion: a plot level geo-database from 1976 to 2012. *Journal of Arid Environments* **123:** 3–11.

Vapnik V. 1998. *Statistical learning theory*. New York: Wiley.

Vega GC, Pertierra LR, Olalla-Tárraga MÁ. 2017. MERRAclim, a high-resolution global dataset of remotely sensed bioclimatic variables for ecological modelling. *Scientific Data* **4:** 170078.

Vitorino LC, Lima-Ribeiro MS, Terribile LC, Collevatti RG. 2016. Demographical history and palaeodistribution modelling show range shift towards Amazon Basin for a Neotropical tree species in the LGM. *BMC Evolutionary Biology* **16:** 213.

Vossler FG. 2014. Small pollen grain volumes and sizes dominate the diet composition of three South American subtropical stingless bees. *Grana* **54:** 68–81.

Waller T, Barros M, Draque J, Micucci P. 2012. Conservation of the Palo Santo tree, Bulnesia sarmientoi Lorentz ex Griseb, in the South America Chaco Region. *Medicinal Plant Conservation* **15:** 4–9.

Waltari E, Hijmans RJ, Peterson AT, Nyári ÁS, Perkins SL, Guralnick RP. 2007. Locating pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS ONE* **2:** e563.

Watson DF. 1992. *Contouring: a guide to the analysis and display of spatial data*. New York: Pergamon Press.

Werneck FR. 2011. The diversification of eastern South American open vegetation biomes: historical biogeography and perspectives. *Quaternary Science Reviews* **30:** 1630–1648.

Werneck FP, Costa GC, Colli GR, Prado DE, Sites Jr JW. 2011. Revisiting the historical distribution of Seasonally Dry

Tropical Forests: new insights based on palaeodistribution modelling and palynological evidence. *Global Ecology and Biogeography* **20:** 272–288.

Werneck FP, Gamble T, Colli GR, Rodrigues MT, Sites JW Jr. 2012. Deep diversification and long-term persistence in the South American 'Dry Diagonal': integrating continent-wide phylogeography and distribution modeling of geckos. *Evolution* **66:** 3014–3034.

Widmer A, Lexer C. 2001. Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. *Trends in Ecology & Evolution* **16:** 267–269.

Wiens JA, Stralberg D, Jongsomjit D, Howell CA, Snyder MA. 2009. Niches, models, and climate change: Assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences USA* **106**(Suppl 2): 19729–19736.

Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV, Damschen EI, Jonathan Davies T, Grytnes JA, Harrison SP, Hawkins BA, Holt RD, McCain CM, Stephens PR. 2010. Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters* **13:** 1310–1324.

Wu SD, Lin L, Li HL, Yu SX, Zhang LJ, Wang W. 2015. Evolution of Asian interior arid-zone biota: evidence from the diversification of Asian *Zygophyllum* (Zygophyllaceae). *PLoS ONE* **10:** e0138697.

WWF—World Wide Fund For Nature. 2014. *The growth of soy: impacts and solutions*. Gland: WWF International.

Zuloaga OF, Morrone O, Belgrano MJ, eds. 2008. *Catálogo de Las Plantas Vasculares Del Cono Sur: Argentina, Sur de Brasil, Chile, Paraguay y Uruguay*. St. Louis: Missouri Botanical Garden Press.

SUPPORTING INFORMATION

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

Table S1. List of species, molecular markers and GenBank accession numbers.
Table S2. List of bioclimatic variables used for each ENM marked with permutation importance value (%). The names of the variables (1 to 19) correspond to the MERRAclim and WorldClim databases.
Figure S1. MOP and MESS analysis. (A) Mid-Holocene, cc model; (B) mid-Holocene, mr model; (C) LGM, cc model; (D) LGM, mr model; (E) LIG. 1 to 3, MOP analysis, percentiles 0.1, 0.5 and 1; 4, MESS analysis. MOP: areas in white are of strict interpolation risk (i.e. lack of similar environmental combinations) in past climatic scenarios; in red, areas distant from region M; the blue areas are more similar to M (see Owens *et al.*, 2013). MESS: areas in red have one or more environmental variables outside the range presented in the training data, so predictions in those areas should be treated with strong caution (see Elith *et al.*, 2010). The shape is the M in the past and the green polygon is the binary layer of the ENM for each time period.