Regional climate oscillations and local topography shape genetic polymorphisms and distribution of the giant columnar cactus *Echinopsis terscheckii* in drylands of the tropical Andes

Vilma B. Quipildor¹,² | Thomas Kitzberger¹ | Pablo Ortega-Baes² | Maria P. Quiroga¹ | Andrea C. Premoli¹

¹Instituto de Investigaciones en Biodiversidad y Medioambiente, CONICET, Universidad Nacional del Comahue CRUB, Bariloche, Argentina
²Laboratorio de Investigaciones Botánicas (LABIBO), CONICET, Facultad de Ciencias Naturales, Universidad Nacional de Salta, Salta, Argentina

**Abstract**

**Aim:** We sought to determine if the present fragmentary distribution of the giant columnar cactus *Echinopsis terscheckii* in tropical drylands is a relict of a previously more widespread range during cold and dry phases of the Last Glacial Maximum (LGM).

**Location:** Tropical and subtropical dry ecotonal areas of northern and central Andes of Argentina.

**Methods:** We combined ecological niche models (ENM) with molecular polymorphisms of isozymes and DNA sequences. We collected samples from 30 individuals at 24 locations for genetic analysis covering a wide range of environmental conditions. We sequenced the nuclear ITS and three non-coding regions of the chloroplast DNA and we resolved 15 isozyme loci. Potential distribution was modelled using 88 *E. terscheckii* presence training records and a reduced set of 10 modern bioclimatic variables. LGM and the Mid-Holocene distributions were derived by projecting bioclimatic data under present to past environmental conditions according to CCSM4 and MIROC-ESM Global Climate Models.

**Results:** We detected high isozyme diversity towards the south. The multivariate cluster analysis yielded two groups of populations that were geographically concordant with the DNA haplotypes located north and south of a divide at 27°S. Distribution models show range expansion during the LGM in two north and south areas separated by a gap of low suitability at 27°S. Suitable areas in the south were close to current populations, while in the north, populations survived in more disjunct locations that probably suffered from founder effects. In contrast, Mid-Holocene bioclimatic conditions were relatively unsuitable in the south.

**Main conclusions:** Our results suggest that the divergence of north and south groups of *E. terscheckii* populations reflect long-lasting persistence through climatic cycles that were reinforced by the presence of an orogenic divide at mid-latitudes. Latitudinally divergent groups of populations should be treated as distinct evolutionary significant units that deserve independent conservation actions. Increased genetic diversity and inbreeding towards the south may guide setting up priorities for the long-term protection of a dominant element of drylands as *E. terscheckii.*
1 | INTRODUCTION

Slightly less than half of the world’s subtropical and tropical forests is comprised of dry forest (Murphy & Lugo, 1986). Nonetheless, the evolutionary ecology of species from tropical and subtropical dry forests and woodlands has received less attention than that of species from rain forests. Studies on arid and semi-arid regions are important because they have been subjected to agricultural conversion with devastating ecological consequences. For example in Argentina, one of the world’s leading producers of soybeans, deforestation in the lowlands of the semi-arid Chaco has been accelerating, reaching >28,000 ha/year since 1997 (Grau, Gasparri, & Aide, 2005).

The southern central Andes of northern Argentina, Bolivia and Chile are extremely arid between c. 15 and 27°S. Despite this aridity, dry tropical forests consist of diverse vegetation types ranging from forests of tall trees on the moist sites and cactus scrubs on the driest sites (Murphy & Lugo, 1995). Parallel mountain ranges block moisture-bearing winds that originate in the Amazon basin and the Atlantic, leading to humid eastern flanks and aridity within the Puna-Altiplano Plateau in the west (Strecker et al., 2007). Arid climates of the western Andes are recorded from the late Triassic (Clarke, 2006). Thus, aridity-inducing factors during the Miocene uplift intensified the Andean rain shadow, which already existed 15 Ma ago, rather than changing it (Hartley, 2003). While arid conditions have prevailed in this area throughout geologic time, intense climate oscillations, such as the northward displacement in weather patterns during the Pleistocene and increased westerly circulation at 28°S (Clarke, 2006), probably had local effects on the biota. Climate variation in this area is of interest not only because of its impact on water resources in adjacent lowlands but also because of its links with continental and presumably hemispheric-scale atmospheric circulation (e.g. Lenters & Cook, 1999). Furthermore, species’ responses to climatic cycles may be complex in the tropics, where areas with significant precipitation might be located adjacent to those with low precipitation due to the orographic effects of the Andes (Cook & Vizy, 2006). Consequently, the extent and timing of climate change in tropical areas is still a matter of debate.

Ice ages are associated with global-scale cooling and drying (CLIMAP Project Members, 1976). However, the relative effects of precipitation (e.g. Clapperton, 1995) and temperature fluctuation (e.g. Colinvaux et al., 1996) on changes in the distribution of vegetation in the Neotropics is controversial, partially because fossil records are fragmentary (Pennington, Prado, & Pendry, 2000). Distribution maps of patchy dry woodland formations in South America support the Pleistocene arc theory that aims to explain current disjunct fragments of dry forests in central tropical South America. These isolated drylands would constitute modern-day refugia resulting from vicariance of a formerly widespread distribution during the more arid climatic conditions associated with the Last Glacial Maximum (LGM). It postulates that with the onset of the last glaciation, a dry-cool climate resulted in the expansion of semi-arid vegetation, and that subsequent contraction occurred due to more humid climates during the past 12,000 Ka (Prado & Gibbs, 1993). Analysis of the ‘dry forest refugia’ in relation to biogeographical and floristic patterns in the Neotropics using a cladistic approach suggested that diversification patterns of woody dry forest elements in South America might be better explained by pre-Pleistocene (i.e. orogeny) factors than climate cycles associated to the Pleistocene age (Pennington et al., 2004). This lack of evidence for Quaternary allopatric speciation raises the possibility that vicariance might not have been the underlying cause for the disjunct distributions of dry forests. In support with the latter view, a recent study showed that the Bolivian Chiquitano Dry Forest probably developed by long-distance dispersal from southern Amazonia during the Holocene (Mayle, 2004). This evidence suggests that historical biogeographical processes in the Neotropics involved a combination of long-distance dispersal–vicariance. Moreover, the work of Mayle (2004) highlights the relevance of performing population genetic and phylogeographic studies to disentangle potential responses to past dry-cold and humid-warm climatic cycles.

In the present study, we analysed present and past distribution patterns of Echinopsis terscheckii, a dominant columnar cactus of tropical and subtropical dry ecotones of northern and central eastern slopes of the Andes in Argentina. The area inhabited by E. terscheckii has been subjected to substantial climatic variation on time-scales of 10^7 to 10^5 years (Strecker et al., 2007). We aimed to test whether the current fragmented distribution of E. terscheckii is a relic of a previously widespread range during cold and dry phases of the LGM, using a combination of ecological niche models and distribution patterns of molecular polymorphisms. We predict a wider range during the LGM, and a contraction during warmer phases of the Holocene, and a correlation between the extent of area occupation and genetic diversity as an indication of long-lasting persistence along the range. In addition, given the complexity of the area inhabited by E. terscheckii, we expected a direct orographic barrier for gene flow and an indirect effect of precipitation regime along its current and historical distributions.

2 | MATERIALS AND METHODS

2.1 | Study species

Echinopsis terscheckii (Parmentier ex Pfeiff.) H.Friedrich & G.D. Rowley (= Trichocereus terscheckii) is a giant columnar cactus...
endemic to north-western Argentina (Hunt, 2006; Kiesling, 1978). It is distributed in the provinces of Jujuy, Salta, Tucumán, Catamarca, La Rioja and San Juan at the western limit of the Chaco and Monte phytogeographic provinces (Cabrera, 1976). It is found between 300 and 2,200 m a.s.l. with high-elevation populations located towards the northern range. It inhabits dry mountain slopes mixed with xerophytic vegetation of Central Andean Puna and in the lowlands at the ecotone with transitional forests as the dry Chaco, High Monte and Yungas (Figure 1). The eastern plateau margin in NW Argentina receives 1,000–3,000 mm/year rainfall (Strecker et al., 2007), 80% of which falls in summer from November to February. Flowering occurs at various peaks from late spring to summer, although it may vary throughout its distribution. Flowers have nocturnal anthesis with an extended floral cycle until next morning. The reproductive system is self-incompatible and floral visitors are mainly nocturnal moths and bees (Ortega-Baes, Saravia, Suhring, Godinez, & Zamar, 2010).

2.2 Sampling of natural populations

We selected 24 sampling locations for genetic analysis covering a wide environmental range of conditions inhabited by the species (Table S1.1 in Appendix S1). Study sites were chosen within areas with suitability ≥0.55 based on the current potential distribution model (Sosa, 2016). At each location, we randomly sampled 30 individuals. From each individual, we collected a 5 cm³ cube of fresh tissue from the main trunk or secondary branches. Samples were labelled and stored in a portable cooler for posterior transport to Laboratorio Ecotono, Universidad Nacional del Comahue in Bariloche, Argentina.

2.3 Isozyme electrophoresis

Proteins were extracted by crushing 2 g of fresh parenchymatic tissue with liquid nitrogen using mortar and pestle and adding 1.5 ml of grinding buffer (Mitton, Linhart, Sturgeon, & Hamrick, 1979). Protein homogenates were stored at −70°C as soaked wicks of Whatman No. 3 chromatography paper and in 0.5 ml Eppendorf tubes as backups.

Three buffer systems were used to resolve 15 isozyme loci. These were aldolase (Ald), glyceraldehyde-3-phosphate dehydrogenase (G3PDH), isocitrate dehydrogenase (Idh), and malate dehydrogenase (Mdh-1, Mdh-2, and Mdh-3) that were run using the histidine system by King and Dancik (1983); menadione reductase (Mnr-1 and Mnr-2), peroxidase (Per), phosphoglucomutase (PgM-1 and PgM-2), and shikimate dehydrogenase (Skdh) using the morpholine–citrate system (Ranker, Haufier, Soltis, & Soltis, 1989); and phosphoglucoisomerase (PgI-1, PgI-2 and PgI-3) using the B system by Conkle, Hodgskiss, Nunally, and Hunter (1982). Electrophoresis was performed for 6 hr using 12% P/V starch gels (Starch Art, Starch Art Corporation, Smithville, TX, USA). Stains for specific enzymes were suspended in 1% agar individually poured on gel slices. Banding patterns were visualized using a light table (Porta-Trace, Gagne Inc. NY USA).

Loci and alleles were scored by relative mobility assigning the most anodal (fast) the lowest numeral. Loci are considered putative given that no genetic control was performed on the analysed enzymes. Nonetheless, banding patterns are similar for the same enzymes analysed in other plant species for which formal analyses were performed (Murphy, Sites, Buth, & Hufler, 1996).

2.4 Sampling, DNA extraction, amplification and sequencing

We randomly chose 2–4 individuals from each population for DNA analysis. Approximately 0.025 g of fresh parenchymatic tissue was ground with liquid nitrogen. Subsequent DNA extraction followed the ATMB protocol (adapted from Doyle & Doyle, 1990; Dumolin, Demesure, & Petit, 1995). We amplified three non-coding regions of the chloroplast DNA (cpDNA) using universal primers trnH-psbA (HA; Hamilton, 1999), psbM-trnD (MD; Shaw, Lickey, Schilling, & Small, 2007) and trnS-trnG (SG; Hamilton, 1999) and one nuclear region (nDNA) ITS1-2 (White, Bruns, Lee, & Taylor, 1990), following PCR conditions described in Appendix S2. DNA extracts and PCR products were verified on 1% agarose gels with TAE buffer. Amplified products were sequenced by Macrogen Inc., South Korea. Sequences were edited manually using the software ALIVEVIEW (Larsson, 2014) and aligned using MEGA 5.05 (Tamura et al., 2011).

2.5 Species distribution modelling

Potential E. terscheckii distribution was modelled using 88 widely distributed presence training records. This occurrence data consisted of the sampled populations for genetic analyses (see above) and available information that were downloaded by October 2015 from public databases such as Documenta Florae Australis (wwwdarwin.edu.ar/iris/), Flora Argentina (www.floraargentina.edu.ar), Global Biodiversity Information Facility (http://www.gbif.org/) and the Instituto Miguel Lillo Herbarium (http://www.lillo.org.ar). In addition, from March to September 2014 and 2015, we used the Street View application of Google Earth to locate the actual presence of the specimens throughout main routes across the documented distribution of the species.

For the modern bioclimatic dataset, we downloaded the WorldClim tiled 1960–1990 30 arc-s data (Bio 1–19) (www.worldclim.org) and windowed it to cover a study area extending c. 60°–76°W and 17°–37°S. Bioclimatic data of past climatic conditions were downloaded for two Global Climate Models (GCMs) CCSM4 and MIROC-ESM from the WorldClim database (30 arc-s resolution) for LGM (c. 22,000 yr bp) and Mid-Holocene (c. 6,000 yr bp) time slices.

To model E. terscheckii distribution, we used the maximum entropy approach implemented in Maxent 3.0 (Phillip, Anderson, & Schapire, 2006). This program works well with few presence data, does not require absence data, combines continuous and categorical variables, provides the contribution of each variable to the model, controls the excessive adjustment of distributions and the output is a continuous occurrence probability. Maxent generates distribution models based on present occurrences and extracts actual climatic variables and then projects the relationship of such variables using
present and past climatic scenarios to develop potential present and past ranges respectively. Therefore, each model used in this study has current and past potential ranges of *E. terscheckii*.

To obtain a better fit and a measure of the model dispersion, we performed 100 simulations using Maxent standard parameters. Thirty per cent of presence data were used for internal validation and the remaining data were used for model construction. The overall efficiency evaluation of the model was made by the ROC operator, from which the indicator of the area under the curve (AUC) is derived (Fielding & Bell, 1997). A value above 0.75 is considered adequate in studies oriented to management and conservation (Pearcy & Ferrier, 2000). To get a model with a reduced set of variables, we used heuristic relative contributions and jackknifing (Phillip et al., 2006; Phillips & Dudík, 2008) to eliminate highly correlated redundant variables. A reduced final set of 10 modern bioclimatic variables that maximized training gain consisted of: mean annual temperature, seasonal temperature, minimum temperature of the coldest month, annual temperature range, average temperature of the wettest quarter, average temperature of the coldest quarter, annual rainfall, precipitation of the driest month, seasonality of precipitation, and precipitation of the coldest quarter. The final model was re-projected onto past bioclimatic conditions estimated by the two GCMs for LGM and Mid-Holocene. Past *E. terscheckii* suitability values for each time slice were averaged between the two models. Absolute and relative change in suitability between each time slice and the present were calculated based on averaged suitability values in the 88 training site locations.

**FIGURE 1** Heterogeneous physical conditions inhabited by *Echinopsis terscheckii* in north-western Argentina. Background colours depicted in grey scale (see online version for full colours) correspond to distinct ecoregions that also follow a precipitation gradient, i.e. annual rainfall decreasing from north-eastern flanks to south-western Andean slopes. Black lines represent mountain ranges. Triangles and circles correspond to different chloroplast haplotypes found at each side of a divide at c. 27°S latitude. Triangles and circles depict high-elevation northern and low-elevation southern sampled populations respectively. Population names are those as Table S1.1 in Appendix S1.
2.6 Genetic analysis

Isozyme genotypes were used to calculate: the mean number of alleles per locus ($N_a$); the mean effective number of alleles per locus that measures number and frequency distribution of alleles ($N_e$); the number of unique alleles, defined here as those that are present in just one population ($A_{un}$) and/or groups of populations ($A_{uc}$); the proportion of polymorphic loci sensu stricto criterion ($P$); and observed ($H_o$) and expected heterozygosities ($H_e$) under Hardy–Weinberg equilibrium conditions. Observed and expected heterozygosities for each polymorphic locus were compared by calculating Wright’s fixation index (where $F = 1–H_o/H_e$) to determine deviations from random-mating expectations, i.e., from zero values, which were analysed using $\chi^2$ tests. Variation of within-population genetic diversity parameters with latitude, longitude and elevation were analysed by multiple lineal regressions using forward stepwise models with Statistica v.7.0 (StatSoft, Inc, 2004).

Among-population differences were analysed by heterogeneity in allelic frequencies by $\chi^2$ tests and by AMOVA using Genalex (Peakall & Smouse, 2006). Hierarchical $F$ statistics as mean within- ($F_{st}$) and among-population inbreeding ($F_{is}$) were estimated by resampling methods that estimate confidence intervals [CI] in Fstat v2.9.3 (Goudet, 2001). Isolation by distance was analysed by Mantel tests that correlate linearized $F_{st}$ values as $F_{st}/(1–F_{st})$ with corresponding geographic distances between all population pairs. Genetic relations among populations were explored by multivariate cluster analysis using Nei (1978) unbiased genetic distance following Unweighted Pair Group Methods with Arithmetic mean (UPGMA).

Mean isozyme metrics of diversity ($N_a$, $N_e$, $P$, $H_o$, $H_e$ and $F_{is}$) and divergence ($F_{st}$) of north and south groups of populations defined by cpDNA sequences were compared by nonparametric Mann–Whitney U tests. Population bottlenecks were compared within each north and south groups, respectively, were analysed by the excess of heterozygosity relative to that expected under mutation-drift equilibrium under the infinite allele model (IAM) using sign tests by the software Bottleneck (Cornuet & Luikart, 1996). Divergence time between north and south groups of populations was estimated using isozymes by approximate Bayesian computation (ABC) using the Diyabc software (Cornuet et al., 2014).

Concatenated ITS and cpDNA sequences were used to estimate the number of polymorphic sites ($S$), haplotype diversity (HD) and nucleotide diversity ($\pi$). Demographic analyses were conducted by means of Tajima’s D, Fu and Li’s F, and observed and expected mismatch distribution graphics, using the program DnaSP 5 (Librado & Rozas, 2009).

3 RESULTS

3.1 Genetic variation and structure by isozymes

At the species level, 100% of the 15 analysed isozyme loci were polymorphic (Table S1.1 in Appendix S1). Allelic frequencies were heterogeneous across populations for all analysed loci ($\chi^2$ tests $p < .01$ data not shown). Linear regressions of genetic diversity parameters, $N_a$, $N_e$, $H_e$ and $P$, yielded a significant effect of latitude increasing towards the south (Figure 2a-d).

Out of a total of 325 by-locus tests for departure from equilibrium conditions, 41.54% were not significant, whereas 40.61% yielded positive (heterozygous deficit) and 17.85% negative (heterozygous excess) fixation indices ($\chi^2$, $p < .01$) (Table S1.2 in Appendix S1). The average within-population inbreeding was not significantly different from zero $F_{st} = 0.11$ [CI –0.06 to 0.34], while the mean among-population divergence was significantly positive $F_{st} = 0.12$ [CI 0.09–0.15]. These parameters indicate that, on average, no significant inbreeding was measured within populations and that a moderate divergence exists among populations of E. terscheckii. A substantial genetic differentiation among distinct populations is also evident by AMOVA analysis that yielded a $\Phi_{PT} = 0.199$ ($p < .001$). The significant correlation between genetic and geographic distances indicated by the Mantel test $r = 0.27$, $p < .05$ (Figure 3) suggests that the genetic structure may be partially due to differences in gene flow rates relative to distance.

The multivariate cluster analysis yielded two latitudinally structured groups that contained most analysed populations, located north and south of a divide at 27°S (Figure 4). These groups also differed in the number of unique alleles, two in the north and 12 in the south (data not shown). On average, southern populations attained significantly greater values for diversity ($N_a$, $N_e$, $H_e$ and $P$) and within-population inbreeding $F_{is}$ than northern ones. The divergence ($F_{st}$) among southern populations was significantly lower than among northern ones (nonparametric Mann–Whitney U test $p < .05$) (Table 1). Tests for heterozygous excess derived from a genetic bottleneck rejected the equilibrium hypothesis of mutation-drift equilibrium in 15 of the 24 studied populations. More than half of possible tests (62%) yielded bottleneck effects in southern populations, whereas only a quarter of those (26%) did so in northern ones (Table S1.3 in Appendix S1). The latitudinal divergence between these groups as dated by coalescence inference and using a generation time of 30 years was $c.$ 10.740 Ka $\pm$ CI 1.536–28.080.

3.2 Diversity at DNA sequences

Chloroplast sequences yielded a total of 11 polymorphic sites (Table S2.1 in Appendix S2). Main sources of diversity were nucleotide substitutions and also indels whose length varied at distinct cpDNA regions: 5 base pairs (bp) long at MD, 3 bp at SG and 72 bp at HA. The nuclear and chloroplast DNA sequences yielded two geographically structured haplotypes concordant with the divide at mid-latitudes (Figure 1). Reduced nucleotide diversity indicates that few mutations occurred in the complete dataset (Tables S2.2 in Appendix S2). Neutrality tests yielded significant and positive Tajima D ($3.67$, $p < .001$) and Fu & Li’s F (2.56, $p < .02$) and thus low levels of both low- and high-frequency polymorphisms, indicating a decrease in population size. Mismatch distribution yielded a bimodal pattern revealing demographic equilibrium or stable populations (Figure S2.3 in Appendix S2).
Species distribution models

A final model was reached based on 10 bioclimatic variables that yielded an AUC of 0.989 and a regularized training gain of 2.769. Suitability values in the model above 0.420 included at least 90% of the *E. terscheckii* training presences. Projections onto LGM and Mid-Holocene conditions led to zero or near zero clamping throughout the study region, thus suggesting that past climate conditions were within the range of values represented in the training data (Figure 4).

Modern *E. terscheckii* potential distribution consists of two areas of high suitability: a smaller northern area in the Argentinean provinces of Salta, Tucumán and southern Jujuy, and a more widespread N–S-extended area encompassing Catamarca, La Rioja and northern San Juan. Both areas are clearly separated by a gap of low suitability located at c. 26°42′–27°18′S (Figure 5).

### TABLE 1

Average genetic diversity parameters of north (11 populations) and south (13 populations) of *Echinopsis terscheckii* by means of 15 isozyme loci. *N*<sub>A</sub>, mean number of alleles per locus; *N*<sub>E</sub>, effective number of alleles; *P* (%), percent polymorphism *sensu stricto*; *H*<sub>O</sub> and *H*<sub>E</sub>, observed and expected heterozygosity respectively. Standard deviations are in parenthesis. *F*<sub>IS</sub>, within-population inbreeding; *F*<sub>ST</sub>, among-population divergence.

<table>
<thead>
<tr>
<th>Sector</th>
<th><em>N</em>&lt;sub&gt;A&lt;/sub&gt;</th>
<th><em>N</em>&lt;sub&gt;E&lt;/sub&gt;</th>
<th><em>P</em> (%)</th>
<th><em>H</em>&lt;sub&gt;O&lt;/sub&gt;</th>
<th><em>H</em>&lt;sub&gt;E&lt;/sub&gt;</th>
<th><em>F</em>&lt;sub&gt;IS&lt;/sub&gt;</th>
<th><em>F</em>&lt;sub&gt;ST&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td>2.67 (0.82)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.90 (0.06)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>85.45 (2.51)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.37 (0.02)</td>
<td>0.39 (0.01)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.018&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.116&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>South</td>
<td>3.07 (0.07)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.23 (0.06)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>94.36 (1.66)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.40 (0.02)</td>
<td>0.49 (0.01)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.218&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.088&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Distinct letters indicate significant differences between north and south groups of populations (nonparametric Mann–Whitney U test *p* < .05).
Past climatic conditions during the LGM and Holocene generated strong and spatially heterogeneous changes in *E. terscheckii* suitability. During the LGM, species distribution models show a general eastern (lowland) expansion of suitable areas (Figure 5a) and a general worsening of conditions in higher elevation western locations. While the southern portion of *E. terscheckii* distribution shows increased suitability during LGM in areas close (a few kilometres) to extant populations, the northern portion of the distribution shows a core area of increased LGM suitability located some 100 kilometres east of extant populations (Figure 5d). Locally, extant population location witnessed overall decreases in suitability during the LGM (Figure 5a); however, conditions were far worse for northern than southern present population locations (Figure 5e). Mid-Holocene conditions were also generally worse than present-day suitability with little or no area of improved conditions (Figure 5b, f). Generally, the southern portion of *E. terscheckii* distribution showed poorer conditions compared to the northern portion which remained more stable (Figure 5f, g).

### DISCUSSION

Populations of *E. terscheckii* inhabiting arid and semi-arid environments of north-western Argentina are genetically diverse yet geographically structured possibly due to cycles of contraction/expansion driven by climate oscillations of the Quaternary as well as orogenic influences. Our distribution models suggested that climate conditions of the LGM probably favoured range expansion of elements inhabiting drylands, as exemplified by *E. terscheckii*. Nonetheless, range expansion was far from homogeneous throughout different latitudes. Niche models show that climatic conditions during the LGM in the south allowed local persistence as well as in nearby lower elevation sites. In contrast, northern suitable areas were overall reduced and possibly located faraway in the eastern lowlands of the Chaco Salteño. Therefore, current northern populations are the result of local survival in small populations that withstood LGM conditions and/or long-distance dispersal. In either case, populations in the north have suffered from genetic drift and founder effects that have eroded genetic diversity. Environmental conditions during the Mid-Holocene improved in the north, while they deteriorated in the south. However, surviving southern populations have maintained genetic variation and cohesiveness, as reflected by significantly lower genetic divergence by means of $F_{ST}$ values than those among northern ones.

The presence of the Sierra del Aconquija located at ~27°S maybe in part responsible for the differences between northern and southern groups of *E. terscheckii* populations. The Sierra del Aconquija is an orographic barrier that was built up over the past 6 Myr. It currently forms a prominent landscape element above its adjacent foreland plains, reaching elevations greater than 5 km (Bonnet, 2009). It represents a major topographic barrier to the moisture flux coming from the Atlantic Ocean: the eastern flank receives much more precipitation (>2 Myr$^{-1}$) than the western flank (Sobel & Strecker, 2003), which highlights the relevant role of tectonics in the evolution of climate in this part of the Andes (Coutand et al., 2006). Fixed (i.e. monomorphic) chloroplast and nuclear DNA sequences show a significant genetic structure north and south of c. 26–27°S that correlates with the presence of the Sierra del Aconquija and its climatic influence. Although the lack of intra-group polymorphism impeded the DNA-based molecular dating of this divergence, we may speculate that these two groups diverged from a widespread common ancestor due to an orographic barrier that reduced precipitation particularly on northern populations. Thus, under this hypothesis, populations of the southern lineage had a more continuous input of precipitation through time, which resulted in more stable populations. This hypothesized stability is also reflected in the higher isozyme variation in the south which indicates that southern populations are a long-term reservoir of genetic diversity. The presence of unique genetic variants of chloroplast DNA was also present in southern populations of the montane *Podocarpus parlatorei* which were interpreted as long-lasting refugia (Quiroga, Pacheco, Malizia, & Premoli, 2012). Thus, genetic results highlight the need of conservation actions towards the southern range of these endemic elements of the tropical Andes inhabiting montane Yungas (Quiroga & Premoli, 2013) and dry lowlands (this study).

In contrast, the area occupied by northern populations of *E. terscheckii* is located on the leeward side of the Sierra del Aconquija and they probably suffered from hyperaridity as a consequence. However, the retention of ancestral variants at cpDNA and ITS sequences, the formation of distinctive north and south groups by means of cluster analysis and the presence of unique isozyme alleles in both groups of populations support the idea of long-lasting persistence of *E. terscheckii* along its entire range. Palynological evidence indicates that most taxa survived locally throughout the last
Glacial–interglacial cycle, even if at much reduced levels, and/or shifting towards more favourable conditions (Gosling, Bush, Hanselman, & Chepstow-Lusty, 2008). Hence, reduced isozyme diversity in the north is not necessarily a consequence of migration from southern areas but probably a reflection of the effects of drift in such populations. Therefore, our results provide evidence of multiple refugia during the LGM as suggested for microthermic (i.e. cold hardy) taxa inhabiting temperate areas of the southern-most Andes that persisted locally along its range during cold periods (sensu Premoli, 1998; Premoli, Kitzberger, & Veblen, 2000). The moderate divergence measured by $F_{ST}$ which is also associated with the geographic distance among populations (as shown by significant Mantel tests) suggests that population expansions occurred from local populations within north and south sectors separately. The novelty of our results is that glacial–interglacial fluctuations affected not only vegetation composition of the tropical Andes (Gosling et al., 2008) but also intraspecific genetic patterns as well which in turn were differently impacted along latitude.

Modelling of Mid-Holocene conditions yielded a significant reduction of suitability throughout *E. terscheckii* range. Our Mid-Holocene modelling yielded more optimal conditions that improve occupancy especially in the northern range where *E. terscheckii* reaches higher elevations (Figure S3 in Appendix S3). Climates of the tropical Andes are the result of a balance between the easterlies that foster the Amazon moisture input during the austral summer (December–March) and the northward shift of westerlies that moderate eastern moisture input (Tchilinguirian & Morales, 2013). Therefore, Mid-Holocene conditions might have fostered the expansion of

**FIGURE 5** Distribution models of *Echinopsis terscheckii* showing potential suitability during (a) LGM (c. 22,000 yr BP), (b) Mid-Holocene (c. 6,000 yr BP) and (c) modern (pre-industrial). Suitability values for past scenarios represent averages of two Global Climate Models (CCSM4 and MIROC-ESM, see text). Colours depicted in grey scale (see online version for full colours). Red-orange values represent high suitability, whereas green represents low suitability. Dots represent training points. Absolute change in suitability are compared between modern pre-industrial conditions with (d) LGM and (f) Mid-Holocene where red-orange colours represent suitability losses and blue colours gains in suitability. Boxplots depict per cent change in suitability (compared to modern pre-industrial) in training site locations during (e) LGM and (g) Mid-Holocene. Upper and lower plots in (e) and (g) represent the distribution of % suitability change in northern (<27°S) and southern (>27°S) *E. terscheckii* training site locations.
northern windward populations while favouring other species in the south, e.g. probably transitional forests, that out-competed E. terscheckii lowland populations and thus reducing its southern natural range. This scenario is consistent with our results that provide evidence of genetic bottlenecks and increased inbreeding in southern populations due to relatively recent population decline. In contrast, northern populations have suffered from a contraction–expansion dynamics that resulted in low genetic diversity due to drift and reduced inbreeding occurring in historically small populations where selection have purged homozygote deleterious alleles. Hence, range-retraction at early to Mid-Holocene has enhanced the genetic divergence between northern and southern groups of populations dated by isozymes as occurring c. 10,000 yr BP.

Cactaceae, especially South American columnar cacti, occupy spatially isolated fragmented arid habitats yet they retain high genetic variation (Moraes, Abreu, Andrade, Sene, & Solferini, 2005; Nassar, Hamrick, & Fleming, 2002). Our results also show high within-population variation exceeding 0.35 mean gene diversity ($H_e$) in all studied populations and retaining values as high as 0.573 in some (Table S1.4 in Appendix S1). These values of $H_e$ are as high as those found in other columnar cacti such as Stenocereus pruinosus, which has the highest diversity reported so far ($H_e = 0.583$; Parra, Pérez-Nasser, Lira, Pérez-Salicrup, & Casas, 2008). Columnar cacti are species with a long life span, usually with high outcrossing rates, and long-distance, animal-mediated dispersal of seeds and pollen. These life history traits have been generally associated with high genetic diversity in vascular plants and columnar cacti in particular (Moraes et al., 2005; Nassar, Hamrick, & Fleming, 2003). Vegetative propagation, such as resprouting species (Premoli & Steinke, 2008), may also be a mechanism that maintains genetic diversity over time in Cactaceae (Moraes et al., 2005). Historical biogeographical factors such as former widespread distributions may explain the particularly high polymorphism found in currently fragmented cactus species (Moraes et al., 2005; Pennington et al., 2000). In addition, vegetation cover in drylands may also impact on the fine-scale spatial configuration of genotypes. Under desert climates as those consisting of open vegetation similar genotypes of E. terscheckii cluster at small-scales probably due to the establishment underneath nurse plants, while, in xerophilous forests under moister climates and higher resource competition for light, E. terscheckii populations consist of sparse individuals that produce coarse-grained neighbourhoods (Quipildor, Mathiasen, & Premoli, 2017). Thus, fragmented landscapes may hold spatially complex and genetically diverse cacti populations.

In the tropics, the last glacial period was primarily reflected in altered patterns of precipitation (Mosblech et al., 2012) which was the main driver in vegetation change (Gosling et al., 2008). Superimposed on long glacial–interglacial cycles, millennial and submillennial variability in moisture have rapidly, i.e. 1,000 years, impacted on vegetation as reflected by high-resolution pollen records (Gosling et al., 2008). Therefore, more information is needed on plant responses under rapidly changing climates. For example, it is important to determine if poleward expansion predicted under increasing temperatures fits for tropical plants. Recent modelling evidence suggests that montane tropical taxa may shift towards the equator, which prompts questions on the velocity of change and potential expansion into areas far away from current sources (M. Paula Quirorga, Universidad Nacional del Comahue, pers. comm.).

Fixed variants of nuclear and chloroplast DNA sequences together with multivariate cluster analysis of isozyme data yielded two latitudinally concordant groups of populations strongly controlled by orogeny. In addition, the presence of unique, i.e. distinctive, isozyme variants in northern and southern clusters suggests that latitudinally distinct groups of E. terscheckii populations should be considered different evolutionary significant units that deserve independent conservation actions. Also, the increased diversity and inbreeding with latitude calls for the need to prioritize conservation efforts of southern populations to protect their evolutionary potential. This is particularly critical in areas where the expansion of land for agriculture is at present threatening E. terscheckii persistence.

ACKNOWLEDGEMENTS

The study was funded by CONICET AVG 2011/2014 Grant No. 01149 and FONCYT Grant No. 1492. We thank Secretarías de Ambiente from Salta, Catamarca, and San Juan provinces for allowing sample collection. We are grateful to M. Chanampa, M. González, P. Gorostiague, A. Gutierrez, N. López, E. Quiroga, J. Sajama, and C. Sosa for field assistance. For laboratory work, we thank M. Fasanella, G Ignazi, P. Mathiasen, and C. Souto. R. Vidal Russell and P. Mathiasen provided valuable help with DNA sequence analysis. We also thank the valuable comments on our manuscript provided by A. Jump.

ORCID

Andrea C. Premoli http://orcid.org/0000-0002-4632-1497

REFERENCES


**BIOSKETCH**

*Vilma Quipildor* is a doctorate fellow from Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) of Argentina. Her interests are population genetics of native plant species from arid areas of tropical South America applying genetic markers to infer their biogeographical history and to define conservation actions.

**How to cite this article:** Quipildor VB, Kitzberger T, Ortega-Baes P, Quiroga MP, Premoli AC. Regional climate oscillations and local topography shape genetic polymorphisms and distribution of the giant columnar cactus Echinopsis terscheckii in drylands of the tropical Andes. *J Biogeogr*. 2017;00:1–11. https://doi.org/10.1111/jbi.13106