Journal of Biogeography (J. Biogeogr.) (2012)

ORIGINAL ARTICLE

Precipitation rather than temperature influenced the phylogeography of the endemic shrub *Anarthrophyllum desideratum* in the Patagonian steppe

Andrea Cosacov[1], Leigh A. Johnson[2], Valeria Paiaro[1], Andrea A. Cocucci[1], Francisco E. Córdoba[3] and Alicia N. Sérsic[1]*

[1]*Laboratorio de Ecología Evolutiva y Biología Floral, Instituto Multidisciplinario de Biología Vegetal (IMBIV), CONICET–Universidad Nacional de Córdoba, 5000 Córdoba, Argentina,* [2]*Department of Biology and M. L. Bean Life Science Museum, Brigham Young University, Provo, UT 84602, USA,* [3]*Centro de Investigaciones en Ciencias de la Tierra (CICTERRA), CONICET–Universidad Nacional de Córdoba, 5000 Córdoba, Argentina*

*Correspondence: Alicia N. Sérsic, Laboratorio de Ecología Evolutiva y Biología Floral, Instituto Multidisciplinario de Biología Vegetal (IMBIV), CONICET–Universidad Nacional de Córdoba, Casilla de Correo 495, 5000 Córdoba, Argentina.
E-mail: asersic@com.uncor.edu

ABSTRACT

Aim In order to assess the impact of precipitation changes during Pleistocene glaciations on plant species of the Patagonian steppe, a phylogeographical study of the endemic shrub *Anarthrophyllum desideratum* was performed.

Location Southern Patagonia: Argentina and Chile.

Methods Chloroplast intergenic spacers *trn*S–*trn*G and *rpo*B–*trn*C were sequenced for 264 individuals from 33 localities spanning the entire distribution of *A. desideratum*. Phylogenetic (statistical parsimony, maximum likelihood and Bayesian inference) and population genetic analyses (spatial analyses of molecular variance, mismatch distributions, neutrality tests and Bayesian skyline plot) were performed. Divergence time estimates using a calibrated molecular clock were also conducted. Niche modelling was used to reconstruct the palaeodistribution to validate phylogeographical patterns.

Results Thirty haplotypes were identified that clustered into two main lineages, revealing a significant latitudinal phylogeographical break north and south of the Deseado River (*c.* 47° S). Infra-specific diversification began in the late Miocene, with northern and southern lineages separating *c.* 3 Ma, after the eastern Patagonian lowlands started to become increasingly arid. Three areas of high molecular diversity were identified: one in southern and two in northern Patagonia where niche modelling indicates that the species may have survived during the Last Glacial Maximum. These putative refugia received more moisture than much of the steppe during glaciation-associated aridization. The south-western refugium is the more likely source for eastward range expansion during post-glacial humidification.

Main conclusions *Anarthrophyllum desideratum* responded differently to historical processes north and south of the Deseado River. In the north this species survived *in situ* in fragmented populations, whereas in the south it survived in localized refugia that presumably avoided extreme aridization, and from which it expanded eastwards. For southern Patagonia, our results support a new historical scenario affected more by precipitation regimes than by temperature changes associated with glacial cycles. This hypothesis should be considered in future plant phylogeographical studies from the Patagonian steppe.

Keywords

Aridization, cold–warm hypothesis, ecological niche modelling, Fabaceae, historical water balance, *in situ* survival, plant phylogeography, Pleistocene glaciations, refugia, westerlies.

Journal of Biogeography

INTRODUCTION

During the Quaternary period, beginning *c.* 2.58 Ma (Gradstein & Ogg, 2004), climatic fluctuations between glacial and interglacial periods affected the geographical distribution and genetic structure of many species (e.g. Hewitt, 2004; Buckley *et al.*, 2010; Jacquemin & Pyron, 2011). Outside the tropics, the effects of glacial and interglacial periods have been viewed primarily in terms of temperature changes and ice sheet extension/retraction (e.g. Hewitt, 2004). The major signature in the genetic structure of Northern Hemisphere organisms indicates that they were pushed to the south and subsequently recolonized northwards (e.g. Hewitt, 2004). Assuming a similar scenario in the Southern Hemisphere, Pleistocene glaciations are thought to have forced many plant species and communities to the north and/or east of the ice sheet in the Patagonian steppe during cold cycles, with dispersal occurring southwards and westwards as the climate warmed and the ice sheet retreated (e.g. Markgraf, 1983; Heusser, 1987). Emerging phylogeographical patterns confirm that several plant taxa followed this pattern (Sérsic *et al.*, 2011), especially for those distributed along the Andes. For the vast Patagonian steppe, only three plant phylogeographical studies have been published (Jakob *et al.*, 2009; Cosacov *et al.*, 2010; Sede *et al.*, 2012). Two of these focused on species occurring exclusively in the steppe (Jakob *et al.*, 2009; Sede *et al.*, 2012) and both support an *in situ* survival hypothesis. The third focused on a species also distributed along the Andean forest, with evidence for a marked impact of Pleistocene cold cycles (Cosacov *et al.*, 2010). All three studies analysed and discussed data in a framework based on Pleistocene cold cycles and ice sheet extension (reviewed in Sérsic *et al.*, 2011).

Precipitation regimes also fluctuated greatly during the Pleistocene because atmospheric circulation patterns varied substantially between glacial and interglacial periods, particularly in southern Patagonia (south of *c.* 47° S, *sensu* Ramos, 1999; Compagnucci, 2011). Because critical biological processes and ecosystem dynamics in arid and semi-arid environments are mainly regulated by seasonal to inter-annual variation in precipitation (Paruelo *et al.*, 1998; Schwinning & Sala, 2004), changes in precipitation regimes during Pleistocene glaciations are likely to have been influential for the distribution of Patagonian plant species. Although temperature and precipitation effects are not mutually exclusive, it is possible to evaluate which had the dominant effect. For example, dendrochronological studies of shrubs from the Patagonian steppe or from dry areas of the Andes indicate a prevalent influence of water availability over temperature on growth (Srur & Villalba, 2009). Thus, we propose that scenarios and genetic patterns different from those traditionally proposed under the cold–warm hypothesis should be expected in this region if historical water balance had a predominant effect on the distribution range of the study species.

Although a matter of debate (Compagnucci, 2011), palaeoecological records indicate a northward shift of the westerlies during the Last Glacial Maximum (LGM) – and probably during the Greatest Patagonian Glaciation (GPG) – resulting in extreme aridization of southern Patagonia (Paez *et al.*, 1999; Glasser *et al.*, 2004; Cusminsky *et al.*, 2011). Because the weakened influence of the westerlies was restricted to the area west of *c.* 71° W and south-eastern Patagonia was influenced by easterly humid air masses from the Atlantic during this period (Compagnucci, 2011), the region most affected by aridization would have been the central area of southern Patagonia (*c.* 68–70° W). With the exposure of the Atlantic submarine shelf as sea levels receded and the coastline moved eastwards during glacial periods (Rabassa *et al.*, 2011), aridity was particularly acute in the southernmost region of these inland areas of the Patagonian steppe, due to increased continentalization (Ponce *et al.*, 2011; Rabassa *et al.*, 2011).

Shrubs are one of the dominant functional groups in the Patagonian steppe (Golluscio & Sala, 1993), yet phylogeographical information on woody perennials in this environment is limited at present to a single species (Sede *et al.*, 2012). To better understand the processes influencing geographical patterns of genetic variation in the Patagonian steppe, we studied an endemic shrub characteristic of these communities, *Anarthrophyllum desideratum* (DC.) Benth. (Fabaceae). We performed phylogeographical analyses and palaeoclimatic niche modelling to evaluate the possible effect of Pleistocene precipitation regimes on genetic patterns, considering three alternative scenarios. If southern populations persisted *in situ* during the arid and cold Pleistocene glacial periods (‘*in situ* survival hypothesis’), then genetic diversity and the number of unique haplotypes within populations should be high and comparable to populations located in northern Patagonia. Furthermore, demographic analyses should be consistent with historical population stasis, and palaeoecological niche modelling should show comparable areas of predicted presence in northern and southern regions. In contrast, if populations were affected mainly by extreme aridization of southern Patagonia (‘historical water balance hypothesis’), populations inhabiting central and eastern areas (i.e. inland Patagonia during glacial periods) of southern Patagonia should exhibit a markedly different pattern: first, there should be low to no haplotype diversity compared with populations in south-western and northern Patagonia. Second, a signal of eastward population expansion accompanying the increase of effective moisture during interglacial periods should be evident. Third, palaeoecological niche modelling should show low probabilities of occurrence in central and eastern areas of southern Patagonia. Alternatively, if populations were affected mainly by low temperature associated with glacial cycles (‘cold–warm hypothesis’) the following is expected: first, a northward range shift should be evident through palaeoecological niche modelling. Second, demographic analyses should show a signal of southward population expansion accompanying rising temperatures during interglacial periods and, most importantly, areas

of high genetic diversity and exclusive haplotypes (i.e. evidence of *in situ* long-lasting persistence) should be present only in northern areas of the species' range.

MATERIALS AND METHODS

Study species

Anarthrophyllum desideratum (Fabaceae, Papilionoideae–Genisteae) is endemic to southern Patagonia down to the Strait of Magellan (Sorarú, 1974). In Argentina, this species is distributed from 43°10′ to 51°60′ S, while in Chile it is restricted to the southernmost continental region, in the Torres del Paine National Park area, 51°00′ S, 72°50′ W (Fig. 1). It forms isolated, dense populations mostly in the arid and semi-arid steppe of the Patagonian phytogeographical province, although some populations also occur in the sub-Andean grasslands in the Subantarctic province – a narrow strip extending northwards along the Andean Cordillera (Roig, 1998). This landscape is heterogeneous, with elevations ranging from 0 to 900 m a.s.l., marked climatic gradients and a variety of topographic and soil conditions (Paiaro, 2011).

Anarthrophyllum desideratum is a spiny dwarf shrub with an extensive pivoting root system (Oliva *et al.*, 2001). It produces numerous showy flowers pollinated by birds – an unusual plant–pollinator mutualism within the Patagonian steppe (Paiaro, 2011). Although its fruits are explosive, seed dispersal is limited to a few metres, contributing to the patchy distribution of populations.

Sample collection

Geographical coordinates from 66 localities spanning the entire distribution of *A. desideratum* were recorded using a hand-held GPS unit. Fresh leaves were collected from 264 individuals of *A. desideratum* representing 33 localities in Argentina and one in Chile that span the entire distribution of the species (Fig. 1, Table 1). Samples were taken from seven to ten individuals selected haphazardly at each site, separated by a minimum distance of 50 m to minimize the sampling of closely related individuals, and dehydrated in silica gel. Four *Anarthrophyllum* species, *Anarthrophyllum rigidum* (Gillies ex Hook. and Arn.) Hieron., *Anarthrophyllum strigulipetalum* Sorarú, *Anarthrophyllum subandinum* Speg. and *Anarthrophyllum ornithopodum* Sandw., and the monotypic sister genus *Sellocharis* (Lewis *et al.*, 2005; Conterato *et al.*, 2007) were used as outgroups (see Appendix S1 in Supporting Information).

DNA extraction, amplification and sequencing

DNA was isolated using a modified cetyl trimethyl ammonium bromide (CTAB) protocol (Doyle & Doyle, 1987; Cullings, 1992). Chloroplast intergenic spacers *trn*S–*trn*G (primers trnS^{GCU} and trnG^{UCC}; Hamilton, 1999) and *rpo*B–*trn*C (primers rpoB and trnC^{GCAR}; Shaw *et al.*, 2005) were sequenced because they showed the greatest variation among several surveyed loci. Amplification consisted of 94°C for 3 min followed by 30 cycles of 94°C for 1 min, 52°C for 1 min and 72°C for 1 min. Products were purified using Millipore PCR96 plates (Millipore Corp, Billerica, MA, USA), sequenced with BigDye v.3 (Applied Biosystems, Foster City, CA, USA), and purified with Sephadex (GE Healthcare, Piscataway, NJ, USA) before electrophoresis on an AB 3730xl automated sequencer. Electropherograms were edited using Sequencher 4.6 (Gene Codes, Ann Arbor, MI, USA). Sequences were aligned with mafft v.6 (Katoh *et al.*, 2002) and adjusted by hand. Seven gaps were coded as binary characters using simple indel coding (Simmons & Ochoterena, 2000). All sequences were deposited in GenBank (*trn*S–*trn*G, JN712264–JN712297; *rpo*B–*trn*C, JN709816–JN709849).

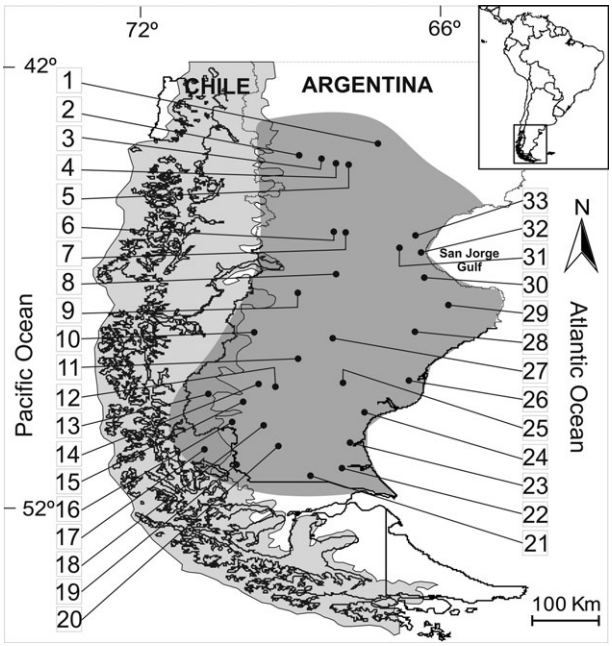

Figure 1 The inset depicts a map of South America, with the study area indicated by a box. The expanded map shows the extent of the study area (in grey) and the locations of the 33 sampling sites of *Anarthrophyllum desideratum*. The extent of the ice cap during the Last Glacial Maximum is indicated by the light grey shaded area. Locality numbers correspond to those in Table 1.

Haplotype relationships and divergence time estimates

Haplotype relationships were inferred using statistical parsimony implemented in tcs 1.21 (Clement *et al.*, 2000) with the default 0.95 probability connection limit. Two ambiguous connections (loops) in the network were resolved using frequency and topological criteria (Crandall & Templeton, 1993).

Relationships among ingroup and outgroup haplotypes were reconstructed using Bayesian inference (BI) and maximum

Table 1 Sampling site, geographical coordinates, elevation (metres), sample size (N_{ind}), haplotypes (H) and molecular diversity indices of sampled *Anarthrophyllum desideratum* populations: haplotype diversity (h), nucleotide diversity (π) and mean number of pairwise differences (p). Voucher numbers deposited at CORD (herbarium of the Museo Botánico de Córdoba) are indicated in the final column.

No.	Sampling site	Latitude (S)	Longitude (W)	Elevation	N_{ind}	H	h (± SD)	π (± SD)	p (± SD)	Voucher no. CORD
1	El Escorial	−43°13′55.39″	−68°26′44.22″	721	8	H26	0	0	0	AAC 4115
2	Tecka	−43°34′10.91″	−70°33′55.08″	899	8	H1	0	0	0	AAC 3514
3	Mata Grande	−43°42′51.83″	−70°03′04.68″	819	8	H13	0	0	0	AAC 3510
	Cajón ginebra					H27	0.5238	0.0004	0.5714	
4	Chico	−43°45′53.59″	−69°37′27.20″	760	7	H28	(± 0.2086)	(± 0.0005)	(± 0.5208)	AAC 3521
						H29				
5	El Pajarito	−43°47′49.66″	−69°18′40.09″	592	7	H29	0	0	0	AAC 3528
6	L. Manantiales	−45°27′39.96″	−69°35′26.16″	548	8	H14	0.2500	0.0008	1	AAC 3934
						H15	(± 0.1802)	(± 0.0006)	(± 0.7481)	
7	Boca del Diablo	−45°28′24.60″	−69°30′30.96″	671	8	H16	0	0	0	AAC 4557
8	Pampa Verdún	−46°36′50.76″	−69°36′16.20″	359	8	H4	0.2500	0.0002	0.2500	AAC 3623
						H5	(± 0.1802)	(± 0.0003)	(± 0.3112)	
9	Sumich	−46°58′41.15″	−70°41′19.68″	760	8	H6	0	0	0	AAC 3967
10	PN P Moreno	−48°04′38.64″	−71°40′27.84″	870	8	H7	0	0	0	AAC 3661b
11	La Primitiva	−48°43′54.47″	−70°31′06.20″	341	8	H8	0	0	0	AAC 3670
12	Tres Lagos	−49°29′39.09″	−71°31′05.16″	359	8	H6	0.5357	0.0008	1	AAC 3975
						H20	(± 0.1232)	(± 0.00071)	(± 0.7856)	
13	Helsingfors	−49°39′41.99″	−72°51′48.60″	303	8	H6	0	0	0	AAC 3981
14	Cerro Bagual	−49°24′02.40″	−71°30′53.28″	467	8	H6	0.5357	0.0004	0.5357	AAC 3970
						H21	(± 0.1232)	(± 0.0004)	(± 0.4927)	
15	La Leona	−50°04′05.04″	−72°07′32.39″	235	8	H30	0	0	0	AAC 4196
16	Huyliche	−50°21′57.99″	−72°16′39.00″	337	8	H25	0	0	0	AAC 4193
17	Torres del Paine	−51°02′33.36″	−72°53′52.08″	315	8	H22	0	0	0	AAC 4188
						H6	0.6071	0.0007	0.9643	
18	Escarchados	−50°27′10.80″	−71°28′00.12″	700	8	H9	(± 0.1640)	(± 0.0006)	(± 0.7292)	AAC 3991
						H11				
19	Río Turbio	−51°27′16.39″	−72°13′20.64″	416	8	H22	0	0	0	AAC 3992
20	La Vanguardia	−50°59′03.48″	−71°06′11.52″	312	8	H12	0	0	0	AAC 4630
21	Potrok Aike	−51°56′01.68″	−70°23′59.28″	200	8	H6	0	0	0	AAC 4551
22	Güer Aike	−51°37′45.11″	−69°37′26.04″	46	8	H9	0	0	0	AAC 3692
23	Cdon Boliche	−50°57′20.08″	−69°16′55.09″	11	8	H9	0	0	0	AAC 4553
24	Monte León	−50°12′12.96″	−68°56′56.40″	247	8	H10	0	0	0	AAC 3949
25	La Julia	−49°38′33.19″	−69°23′23.28″	90	8	H10	0.25	0.0025	3.25	AAC 3946
						H19	(± 0.1802)	(± 0.0016)	(± 1.8698)	
26	San Julián	−49°18′56.51″	−67°46′16.68″	7	8	H10	0	0	0	AAC 4588
27	D. Manantiales	−48°15′07.89″	−69°46′59.52″	726	8	H6	0	0	0	AAC 3953
28	Tres Cerros	−48°03′41.76″	−67°36′49.32″	172	8	H18	0	0	0	AAC 3945
29	Pto. Deseado	−47°23′23.28″	−66°46′02.28″	145	8	H10	0	0	0	AAC 4174
30	Caleta Olivia	−46°41′27.60″	−67°23′00.60″	190	8	H17	0	0	0	AAC 4587
31	El Trébol	−45°49′55.20″	−67°56′00.24″	557	8	H3	0	0	0	AAC 3591
32	Pta. Marqués	−45°57′33.01″	−67°32′11.69″	152	10	H23	0.3556	0.0003	0.3556	AAC 4586
						H24	(± 0.1591)	(± 0.0003)	(± 0.3753)	
33	Rïo Chico	−45°30′25.56″	−67°37′24.24″	524	8	H2	0	0	0	AAC 3570
	TOTAL	–		–	264		0.8985	0.0037	4,674	
						–	(± 0.0137)	(± 0.0020)	(± 2.2977)	

likelihood (ML). BI used BEAST 1.6.1 (Drummond & Rambaut, 2007) with nucleotides and coded indels partitioned separately and a Yule tree prior. The GTR+G model was selected using the Bayesian information criterion (BIC) as implemented in JMODELTEST 0.1.1 (Posada, 2008), while coded indels used the stochastic Dollo model ('simple model' in BEAST; Alekseyenko *et al.*, 2008). Two Monte Carlo Markov chains (MCMC) starting with a random tree were run for 60 million generations (burn-in fraction of 0.25), with parameters sampled every 1000 steps. Convergence of estimated parameters was verified using TRACER 1.6.1 (Rambaut & Drummond, 2009), log and tree files were combined using LOGCOMBINER 1.6.1 (Drummond & Rambaut, 2007), and topologies were assessed using TREEANNOTATOR 1.6.1 (Drummond & Rambaut, 2007) and FIGTREE 1.6.1 (Rambaut, 2008). The ML analyses used PHYML 3.0 (Guindon *et al.*, 2005) and the GTR+G substitution model with four categories. The robustness of the phylogenetic relationships was evaluated through 1000 bootstrap replications.

With no known fossils of *Anarthrophyllum*, previously published values were used to approximate divergence times for the main phylogroups in our study. Lavin *et al.* (2005) date the split between *Anarthrophyllum* and *Spartium* at *c.* 19.2 Ma (*Spartium* nests in a clade sister to *Anarthrophyllum* + *Sellocaris* within Genisteae). Here, the split of the stem lineage *Anarthrophyllum* from *Sellocharis* was considered the root of the tree. Using BEAST, a hard maximum bound for the age of the root node was set to 19 Ma (SD = 2.5 Ma). The best molecular clock model (strict or relaxed) was evaluated by comparing Bayes factors (Suchard *et al.*, 2001), with the strict molecular clock model fitting the data best ($\log_{10}$ Bayes factor = 0.5).

Molecular diversity and population genetic structure

Historical barriers among localities or geographical zones were inferred by a spatial analysis of molecular variance (SAMOVA) with SAMOVA 1.0 (Dupanloup *et al.*, 2002). Analyses were run for *K* values ranging from 2 to 15, using 10,000 independent annealing processes. The best grouping option for each *K* value, based on the among-group component (F_{CT}) of the overall genetic variance, was used to conduct independent analyses of molecular variance (AMOVA) in ARLEQUIN 3.5.1.3 (Excoffier & Lischer, 2010). For these analyses, three localities (1, 2 and 28) were removed where one of four highly divergent haplotypes were fixed (H26, H1 and H18, respectively) because subsequent increase of *K* values separated these localities individually. Haplotype diversity (*h*; Nei, 1987), nucleotide diversity (π; Nei, 1987) and mean number of pairwise differences (*p*; Tajima, 1983) were calculated for each sampling site and population group using ARLEQUIN.

Inferring historical events

To disentangle processes and/or historical events involved in shaping the geographical patterns of genetic variation, an exploratory nested clade phylogenetic analysis (NCPA; Templeton, 1998) was performed, followed by other methods to investigate alternative assessments of demographic history. The haplotype network derived from statistical parsimony analysis (see above) was converted into a hierarchical nested design. Clade (D_c) and nested clade (D_n) distances were estimated to assess the association between nested clades and geographical distances among sampled localities using GEODIS (Posada *et al.*, 2000). Null distributions for permutational contingency table test comparisons were generated from 10,000 Monte Carlo replications, with a 95% confidence level. For significant associations, Templeton's (2011) key was used to infer demographic processes and/or historical events.

Tajima's *D* and Fu's F_S (Tajima, 1989; Fu, 1997) neutrality tests were used to further assess population demographic history. Both assume that populations have been in mutation–drift balance for long periods. When this is not true due to sudden expansion, these indices usually have negative values. Significance for both values was determined from 10,000 simulated samples using a coalescent algorithm. Additionally, a mismatch distribution of pairwise differences among individuals was calculated (Rogers & Harpending, 1992). A multimodal distribution is typical of samples drawn from populations at demographic equilibrium, whereas a unimodal distribution is typical of populations having experienced recent demographic expansion (Excoffier, 2004). Goodness of fit of the observed mismatch distribution to that expected under a sudden expansion model was evaluated with the sum of squared deviations (SSD) using parametric bootstrapping (10,000 replicates). Neutrality tests and mismatch distribution analyses were performed with ARLEQUIN and DNASP 5.0 (Librado & Rozas, 2009). Finally, in the two main population groups (see SAMOVA results), past population dynamics through time were estimated using a Bayesian skyline plot implemented in BEAST (Drummond *et al.*, 2005). Separate analyses were run for each group under the model of nucleotide substitution selected with JMODELTEST (in both cases, HKY), using a strict molecular clock with a uniformly distributed prior for clock rate (0.0001–0.01 substitution per site per million years; based on Alsos *et al.*, 2005; Lavin *et al.*, 2005). The number of grouped intervals was set to 10 and the Bayesian skyline analysis performed in the piecewise constant model. Two MCMC starting with a random tree were run for 70 million generations, with parameters sampled every 1000 steps. Convergence of estimated parameters was verified using TRACER (Rambaut & Drummond, 2009). After 25% burn-in, log files were combined and demographic plots visualized (Drummond & Rambaut, 2007).

Niche modelling

To validate and complement phylogeographical analyses and identify areas where the species would have survived during the Pleistocene, the present ecological niche of *A. desideratum* was modelled and projected onto the LGM using environmental data and point locality information in DIVA-GIS 7.4.0.1 (Hijmans *et al.*, 2005) and MAXENT 3.3.2 (Phillips *et al.*, 2006).

Climatic layers at 2.5 arcmin resolution for current and past conditions were obtained from the WorldClim database (http://www.worldclim.org/; Hijmans *et al.*, 2005). To model the potential past distribution, the 19 available climatic layers for present and past conditions were used. MAXENT was run using the following settings: duplicate records removed; random test percentage = 25; regularization multiplier = 1; convergence threshold = 0.00001; maximum iterations = 1000 and averaged across 10 cross-validation runs. Ultimately, a jackknife test was conducted to assess the importance of each variable. Area under the receiver operating characteristic curve (AUC) statistics, which are independent of a threshold and consider both presence and absence records, were used to assess predictive performance of the test data, namely the probability that a random positive instance and a random negative instance are correctly ordered by the classifier. An AUC of 0.5 indicates that model performance is equal to

random prediction, while an AUC of 0.8 means that in places where a species is present, in 80% of cases the predicted values will be higher than where the species has not been recorded (Phillips *et al.*, 2006). To identify areas where the presence of the species was predicted with the highest probability for the past and present scenarios, past × present ecological niche models were multiplied using the overlay option in DIVA-GIS. Through this arithmetical operation, potential stable areas between time periods are identified.

RESULTS

DNA sequence matrices

The entire *trn*S–*trn*G intergenic spacer measured 713–739 nucleotides in length. For the *rpo*B–*trn*C region, a series of poly-A, poly-G and poly-T strings at the 3′ end (*c.* 100 nucleotides) were trimmed, resulting in fragments of 625–640 nucleotides in length. Because the chloroplast is typically inherited uniparentally without recombination in angiosperms, these two regions were combined for analyses a priori.

Haplotype distribution, phylogenetic and divergence time analyses

Thirty chloroplast DNA (cpDNA) haplotypes were identified that form a single network with two primary groups (Fig. 2). The 'southern group' (SG) consists of 10 haplotypes distributed exclusively south of 47° S. This haploclade includes the three most frequent and widespread haplotypes: H6, H9 and H10 (found in 16, 8 and 12% of sampled individuals, and in 21, 9 and 12% of sampled localities, respectively). Haplotype

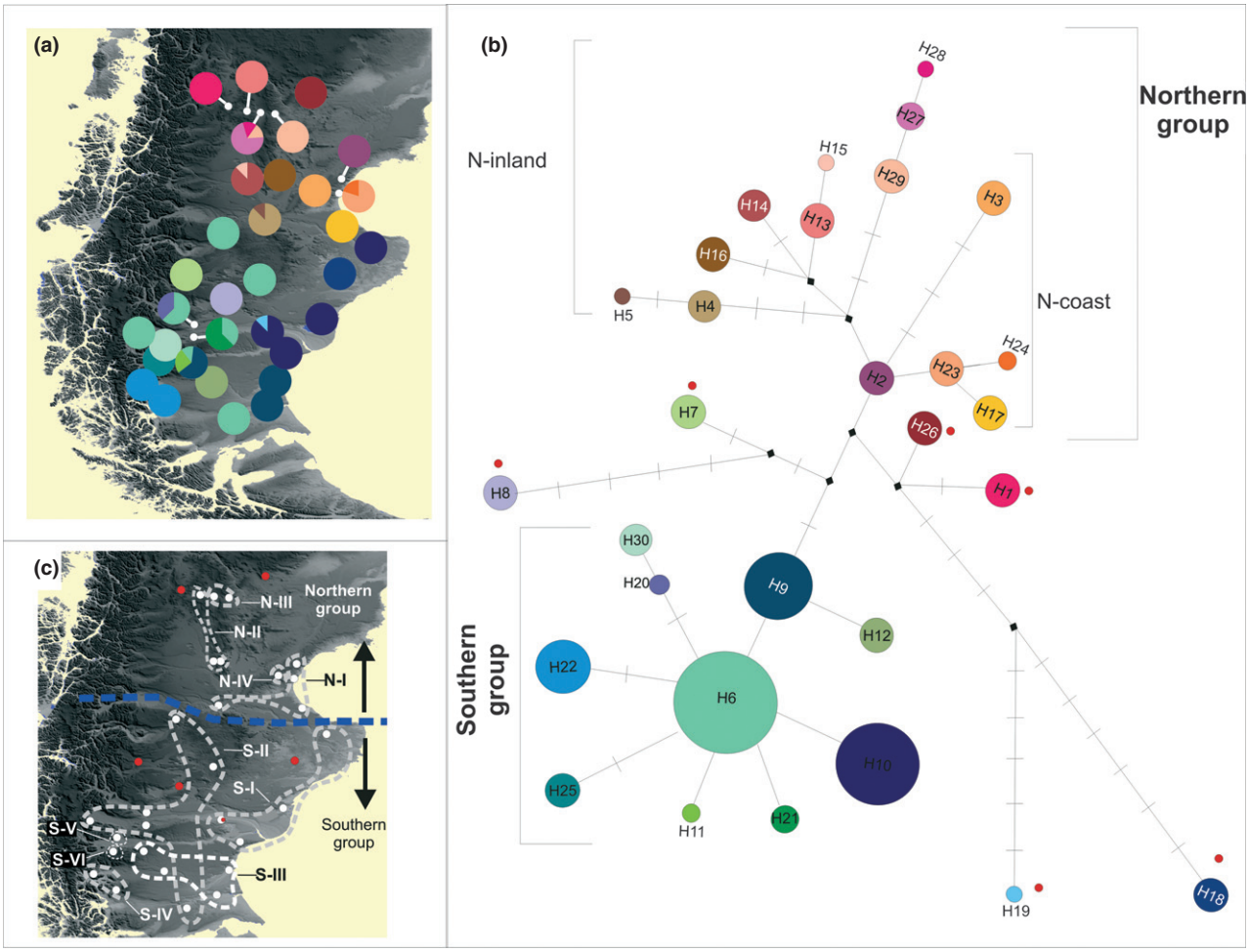

Figure 2 Geographical distribution and genealogical relationships of the 30 chloroplast DNA haplotypes recovered from 33 populations of *Anarthrophyllum desideratum* in southern South America. (a) Pie charts reflect the frequency of occurrence of each haplotype in each population. Haplotype colours correspond to those shown in panel (b). (b) Statistical parsimony network linking the 30 haplotypes, which are designated by numbers. Circle sizes are proportional to haplotype frequencies. Cross-hatching and diamonds represent inferred intermediate haplotypes not observed in the analysed individuals. (c) Dashed delineated areas in the map correspond to the four northern groups (N-I to N-IV) and the six southern groups (S-I to S-VI) retrieved by spatial analysis of molecular variance. The blue dashed line indicates the boundary between northern and southern groups, which approximates the course of the Deseado River in Santa Cruz province, Argentina. Red dots indicate the geographical location of the highly divergent haplotypes shown in panel (b).

9 appears at an internal node connected through one step to H6, which formed the core of a star-like network topology (Fig. 2). Of seven haplotypes connected by one or two steps to H6, four are found in a single locality, and three (H9, H10 and H22) are shared by two or more sampling sites restricted to specific geographical areas (Fig. 2). Overall, the geographical structure of SG follows a longitudinal pattern divided into an eastern area (east of 70° W) where only H10 and H9 are present, and a western area, where H6 and the remaining seven haplotypes are found (Fig. 2). North of 47° S, haplotypes form a 'northern group' (NG) separated by four steps from SG (Table 1, Fig. 2). Haplotype 2 appears as an internal node and is connected by one to four steps to haplotypes restricted to the coast of San Jorge Gulf (H17, H23, H24 and H3; 'N-coast'), and by three to six steps to the remaining haplotypes located in inland Patagonia (68–70° W, 'N-inland') that are all derived from a common hypothetical haplotype. Except haplotype 29, which is shared by two localities, all remaining haplotypes of the NG are exclusive to single localities (Table 1, Fig. 2). Two additional haplotype groups diverged from two of the hypothetical ancestors separating SG from NG. One includes two pairs of highly genetically and geographically divergent haplotypes observed in the two northernmost localities (H1 and H26; sampling sites 1 and 2, respectively) and in two south-eastern localities (H18 and 19, sampling sites 25 and 28, respectively; Fig. 2). The second group contains two divergent haplotypes from two south-western localities (H7 and H8, sampling sites 10 and 11, respectively; Fig. 2).

BI and ML analyses recovered topologies very similar to the parsimony network (Figs 2 & 3). All *A. desideratum* haplotypes appear in a strongly supported clade (Fig. 3). The split separating *Sellocharis* from *Anarthrophyllum* was estimated at *c.* 17 Ma [95% highest posterior density (HPD) = 11.45–23.85], while the divergence between *A. desideratum* and the remaining *Anarthrophyllum* species was *c.* 14 Ma (95% HPD = 8.32–20.76; Fig. 3). Within *A. desideratum,* the first split was estimated at 6.7 Ma (95% HPD = 3.62–10.85) separating most *A. desideratum* haplotypes from a small, weakly supported clade formed by haplotypes H1, H26 and H18 plus H19 (Fig. 3). Diversification within this clade was estimated at *c.* 3.6 Ma (95% HPD = 1.41–6.75), and between H18 and H19 at 2.74 Ma (95% HPD = 0.92–5.43). All remaining haplotypes are nested in a weakly supported clade whose diversification was estimated at *c.* 5.02 Ma (95% HPD = 2.52–7.86). This clade forms a trichotomy with two well-supported subclades

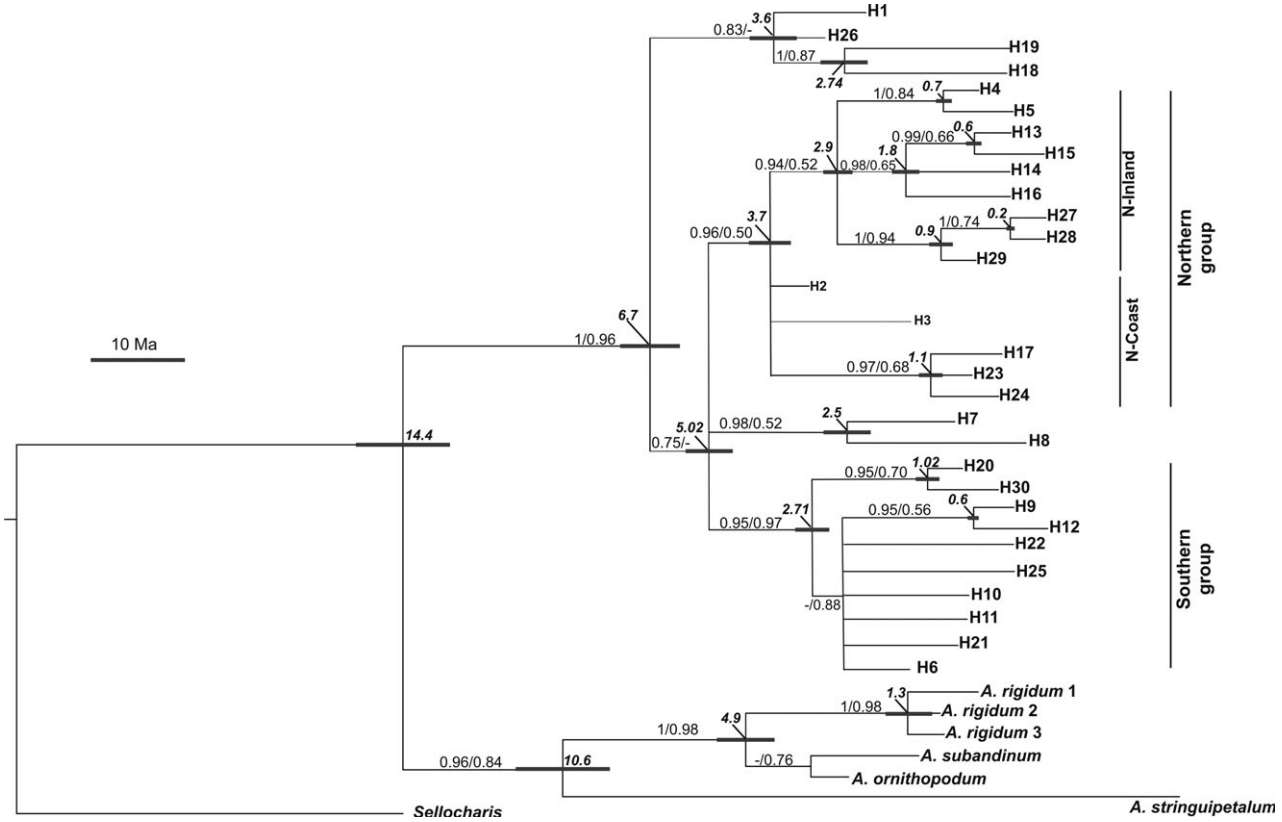

Figure 3 Evolutionary relationships of *Anarthrophyllum desideratum* reconstructed using maximum likelihood (ML) (ln*L* = −2450.0219) based on single haplotypes. Numbers above the branches are support values from Bayesian posterior probabilities/bootstrap resampling for ML analyses, shown when > 0.5. Numbers given for each node are the estimated (median) divergence time in million years, and grey bars indicate the 95% highest posterior densities over the median value. Vertical lines on the right indicate geographical groups. Haplotypes of *A. desideratum* correspond to those in Fig. 2.

corresponding to NG and SG, and a minor subclade composed of H7 and H8 (Fig. 3). NG diversified *c.* 3.7 Ma (95% HPD = 1.88–6.13), with a distinctive N-inland subclade forming a polytomy with N-coast populations. Diversification within these groups occurred during the Pleistocene (0.2–1.8 Ma; Fig. 3). SG haplotypes show little phylogenetic structure, with diversification starting *c.* 2.71 Ma (95% HPD = 1.15–4.82); the split between H9 and H12, and between H20 and H30 were dated to *c.* 0.6 Ma (95% HPD = 0.01–1.71) and 1.02 Ma (95% HPD = 0.03–2.82), respectively.

Molecular diversity and population genetic structure

Genetic diversity within populations was highest for three south-western sites (localities 12, 14 and 18), each containing private haplotypes (Figs 1 & 2). Sites with no haplotype diversity ($h = 0$) occur throughout the distribution, with 25 of 33 localities monomorphic. Six of eight (75%) monomorphic sites located north of 47° S possess exclusive haplotypes, whereas only 4 of 17 (36%) monomorphic localities south of this latitude possess exclusive haplotypes. These four sites are restricted to the west; the remaining monomorphic southern localities, predominantly eastern, possess one of three common haplotypes (H6, H9 or H10). Nucleotide diversity ranged from 0 to 0.0025, with locality 25 possessing the highest value (Table 1).

Optimal partitioning of genetic diversity by SAMOVA was obtained with $K = 2$ ($F_{CT} = 0.52$, $P < 0.0001$; Appendix S2), recovering northern and southern groups concordant with the geographical division at 47° S latitude and haploclades SG and NG identified above (Figs 2 & 3). Further increase of K removed single localities from NG, rather than forming informative multisite clusters. Therefore, to explore spatial genetic structure within each group, independent analyses for the northern and southern groups were run. Optimal partitioning of genetic diversity for NG was obtained with $K = 4$ ($F_{CT} = 0.63$, $P < 0.0001$; Fig. 2, Appendix S2). This analysis recovered a group of localities ('N-I') distributed in the central steppe and the coast (67–69° W, 45°30′–46°41′ S), a group of inland sampling sites ('N-II') located at the western flanks of a complex of plateau and hills (Sierras and Mesetas Occidentales), a third group of two sites located at the top and eastern flanks of the plateau ('N-III'), and a fourth group containing a single locality ('N-IV'; this group appears in the initial partition at $K = 2$; Fig. 2). SG localities clustered in six groups ($F_{CT} = 0.78$, $P < 0.0001$; Appendix S2) coincident with the geographical distribution of the most common haplotypes of the southern clade: an eastern group ('S-I'), clustered the four localities where H10 is fixed or predominates; group S-II, a western group clustering all localities where H6 is fixed or predominant; a southern group S-III clustering all localities where H9 is fixed or predominant; a south-western group S-IV clustering two monomorphic localities for H22; and two groups (S-V and S-VI) each formed by a single locality with an exclusive haplotype (H25 and H30, respectively; Fig. 2).

Levels of gene and nucleotide diversity were higher for NG ($h = 0.9193 \pm 0.0067$, $\pi = 0.035 \pm 0.0019$) than SG ($h = 0.7402 \pm 0.0316$, $\pi = 0.0025 \pm 0.0014$; Table 2). Within NG, values of gene and nucleotide diversity were in the ranges 0.582–0.811 and 0.0005–0.001, respectively, with the highest values from the coastal group N-I. Within SG, gene and nucleotide diversity ranged from 0.00 to 0.518, with the highest values found within the western group S-II and the southern

Table 2 Diversity indices and results of demographic analyses used to test range expansion in *Anarthrophyllum desideratum* population groups derived from spatial analysis of molecular variance. Haplotype (h) and nucleotide (π) diversity, mean number of pairwise differences (p), Tajima's D, Fu's F_S and sum of squared deviations (SSD).

	Diversity indices			Demographic analyses		
Groups	$h \pm$ SD	$\pi \pm$ SD	$P \pm$ SD	D	F_S	SSD
Northern Patagonia	0.925 ± 0.01	0.003 ± 0.002	4.570 ± 2.269	0.181	0.126	0.014*
Southern Patagonia (3.2 NCA)	0.830 ± 0.01	0.001 ± 0.001	1.875 ± 1.078	−0.341	−0.591	**0.002**
N-I	0.811 ± 0.025	0.001 ± 0.080	2,121 ± 1.212	1.236	0.776	0.200**
N-II	0.721 ± 0.036	0.002 ± 0.001	2,362 ± 1.334	1.422	2663	0.143*
N-III	0.582 ± 0.092	0.0005 ± 0.0005	0.670 ± 0.545	1.434	0.055	**0.02**
N-IV[a]	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	—	—	—
S-I	0.063 ± 0.058	0.0006 ± 0.0005	0.812 ± 0.600	**−2.375****	2.823	**0.005**
S-II	0.297 ± 0.080	0.0003 ± 0.0003	0.430 ± 0.395	−0.759	0.044	**0.002**
S-III	0.518 ± 0.081	0.0005 ± 0.0004	0.684 ± 0.534	−0.190	−0.444	**0.001**
S-IV	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	—	—	—
S-V[a]	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	—	—	—
S-VI[a]	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	—	—	—
(S-I)+(SII)+(SIII)	0.752 ± 0.022	0.001 ± 0.001	1.356 ± 0.845	**−1.457****	−0.759	**0.006**
(S-I)+(SII)+(SIII)+(IV)	0.796 ± 0.018	0.001 ± 0.0008	1.660 ± 0.983	**−1.315****	−0.566	**0.003**

NCA, nested clade analysis.

Results consistent with range expansion are indicated in bold.

[a]Groups composed of a single population.

*$P < 0.05$; **$P \leq 0.001$.

group S-III, while the lowest values were found in the eastern (S-I) and south-western groups (S-IV, S-V and S-VI; Table 2).

Inferring historical events

Nested clade phylogeographical analysis (Fig. 4) shows a significant relationship between genetic and geographical distributions in *A. desideratum* (Table 3). The most frequent processes inferred were past fragmentation or allopatric fragmentation, with these processes being inferred for 9 of 16 clades, including all nesting levels, and being particularly frequent in NG (Fig. 4). Restricted gene flow or restricted gene flow with isolation by distance was inferred for four clades, three of them southern. Finally, contiguous range expansion was inferred for the northern clade 4–1 (which includes NG and clade 3–1), and the southern clade 3–2 (coincident with SG; Table 3, Fig. 4).

The southern group (clade 3–2), all southern groups retrieved by SAMOVA (S-I to S-III) and the northern group N-III showed evidence of recent demographic expansion based on the SSD index, because the observed mismatch distribution did not differ from an expected sudden expansion model (Table 2, Fig. 5). Negative and significant values for Tajima's *D* were obtained for group S-I, and when groups S-I, S-II, S-III and S-IV were analysed together (Table 2). Bayesian skyline analyses showed a constant population size through time for NG and a pattern consistent with a recent demographic expansion for SG, indicating that population growth would have started *c.* 500 ka (Fig. 5).

Niche modelling

Niche modelling identified suitable areas for occupation (Fig. 6) and 10-fold cross-validation runs yielded similar average AUC values for past (0.784, SD 0.049) and current (0.776, SD 0.074) simulations. Precipitation of the coldest quarter, the wettest month, annual precipitation and temperature seasonality were generally the most important variables contributing to the past and current models (see Appendix S3).

The current model (Fig. 6a) accurately reflects the extant distribution of this species, although some inconsistencies between predicted and actual distributions were found, such as Tierra del Fuego Island and the Falkland Islands (Islas Malvinas), where *A. desideratum* does not occur, although the occurrence probability is lowest in these areas.

Considering the species' overall range, the distribution model for the LGM (Fig. 6b) did not show shifts to lower latitudes compared with the present-day model (Fig. 6a). However, comparing just the areas with the highest probability of occurrence between the current and past distribution models, a shift to lower latitudes during the LGM is evident. In southern inland Patagonia, a considerable area appeared unsuitable for the species (Fig. 6b). An eastward shift is also evident under the LGM model towards the exposed Patagonian shelf of the Atlantic coast and to the Falkland Islands (Islas Malvinas), although with the lowest probability. No projection of the species was predicted at the centre of the steppe

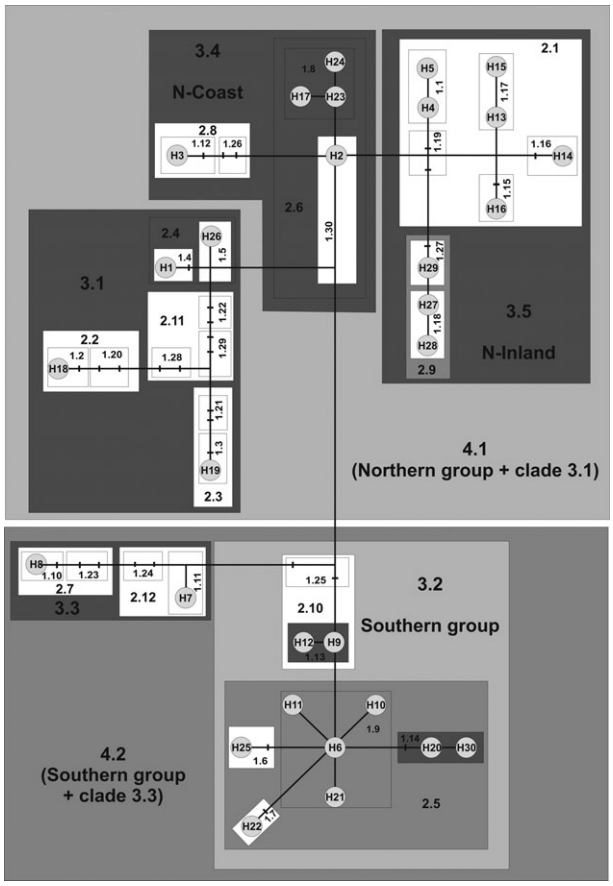

Figure 4 Statistical parsimony network and the resulting nested clade design of the 30 haplotypes found in *Anarthrophyllum desideratum*. Haplotypes correspond to those in Fig. 2; cross hatches represent nucleotide differences between haplotypes. Haplotypes belonging to the same clade level are boxed. The corresponding inference for each haploclade is shown: grey, restricted gene flow/restricted gene flow with isolation by distance; light grey, contiguous range expansion; dark grey, allopatric fragmentation or past fragmentation.

in Santa Cruz province during the LGM, yet this area is currently occupied by the species. Niche modelling also revealed areas where the presence of the species was predicted with the highest probability for the past and present scenarios (Fig. 6c).

DISCUSSION

From our estimates, diversification of *A. desideratum* occurred before the onset of the major Quaternary glaciations, during the late Miocene (*c.* 6.7 Ma). However, the origin of the major clades, including the northern and southern groups, occurred *c.* 5 Ma (middle Pliocene), and diversification within these groups would have started *c.* 2.7–3.7 Ma. This estimate agrees with estimates for the initial diversification of extra-Andean lineages in *Calceolaria polyrhiza* (Cosacov *et al.*, 2010) and in Patagonian *Hordeum* species (Jakob *et al.*, 2009), reinforcing the hypothesis of a recent origin for steppe vegetation. Diversification began as

Table 3 Inferences of historical processes affecting genetic structure in *Anarthrophyllum desideratum* populations based on nested clade phylogenetic analysis (NCPA). Geographical groups, hierarchically nested clades, results of permutational contingency tests, NCPA inference events, and inferred chain are shown.

Geographical region	Clade	χ^2 statistic	Probability	Inferred event	Inference chain
N-coast	1.8	18.0	< 0.0001	Past fragmentation	1–2–11–12–13–14 NO
S	1.9	176.1	< 0.0001	Restricted gene flow with isolation by distance	1–2–3–4 NO
S	1.13	29.0	< 0.0001	Allopatric fragmentation	1–19 NO
S	1.14	10.0	< 0.0001	Allopatric fragmentation	1–19 NO
N-inland	2.1	88.9	< 0.0001	Inconclusive outcome	1–2
N-inland	2.4	16.0	< 0.0001	Allopatric fragmentation	1–19 NO
S	2.5	319.9	< 0.0001	Restricted gene flow/dispersal but with some long-distance dispersal	1–2–3–5–6–7 YES
N-coast	2.6	25.0	< 0.0001	Allopatric fragmentation	1–19 NO
N-inland	2.9	10.0	< 0.0001	Restricted gene flow	1–2–3–5–6–7 YES
N-S	3.1	49.0	< 0.0001	Past fragmentation	1–19–20–2–11–12–13–14 NO
S	3.2	131.4	< 0.0001	Contiguous range expansion	1–2–11–12 NO
S	3.3	16.0	< 0.0001	Allopatric fragmentation	1–19 NO
N-coast	3.4	34.0	< 0.0001	Allopatric fragmentation	1–19 NO
N-inland	3.5	46.0	< 0.0001	Allopatric fragmentation	1–19 NO
N	4.1	209.0	< 0.0001	Contiguous range expansion	1–2–11–12 NO
S	4.2	159.0	< 0.0001	Restricted gene flow with isolation by distance	1–2–3–4 NO
Total		243.6	< 0.0001	Inconclusive outcome	1–19–20–2

N, northern group; S, southern group; N-coast, coastal populations of the northern group, N-inland, steppe populations of the northern group.

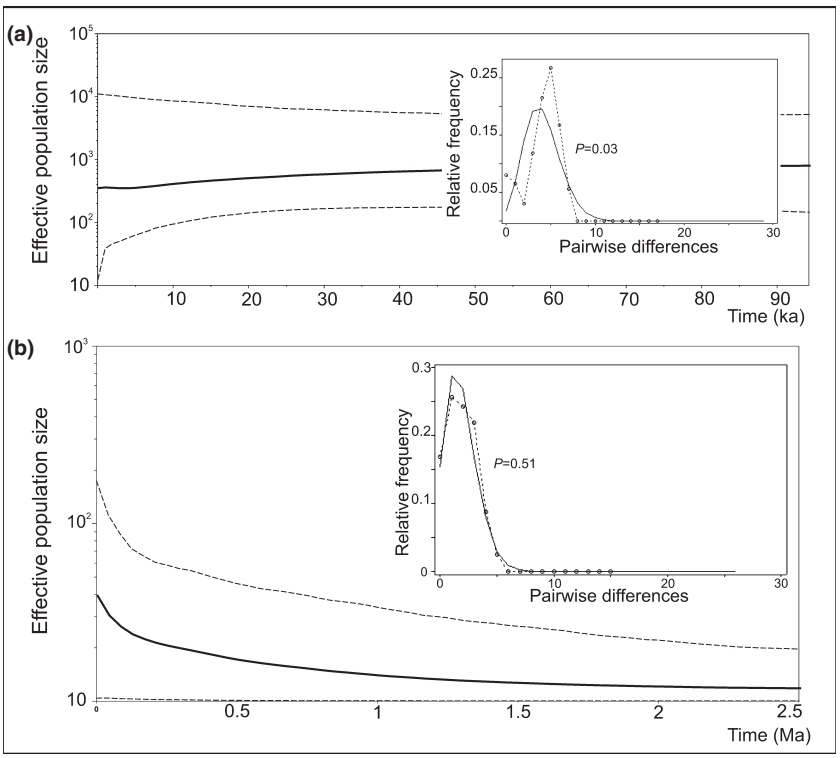

Figure 5 Bayesian skyline plot depicting population size in relation to time for the northern (a) and southern (b) lineages of *Anarthrophyllum desideratum*. The *y*-axis represents effective population size expressed on a logarithmic scale. Dotted lines give the 95% highest posterior densities over the median (bold line) estimates. Insets in (a) and (b) show the observed (closed circles, solid lines) and expected (open circles, dashed lines) chloroplast DNA mismatch distributions under a sudden expansion model; significance level for these analyses are shown.

the Andean range achieved its current elevation (*c.* 5 Ma; Ramos & Ghiglione, 2008), the humid Pacific winds decreased considerably, and aridization started, favouring the establishment of the Patagonian steppe at the eastern lowlands of the Andes (Barreda *et al.*, 2008).

Phylogeographical analyses identified two main phylogroups for *A. desideratum*, located to the north and south of *c.* 47° S, respectively. This phylogeographical break approximates the present course of the Deseado River in Santa Cruz province, Argentina. Latitudinal genetic discontinuities have been previously reported for other Patagonian organisms (Sérsic *et al.*, 2011), along with proposals that such discontinuities may be related to river basins that drain (currently or in the past) the Andean east watershed

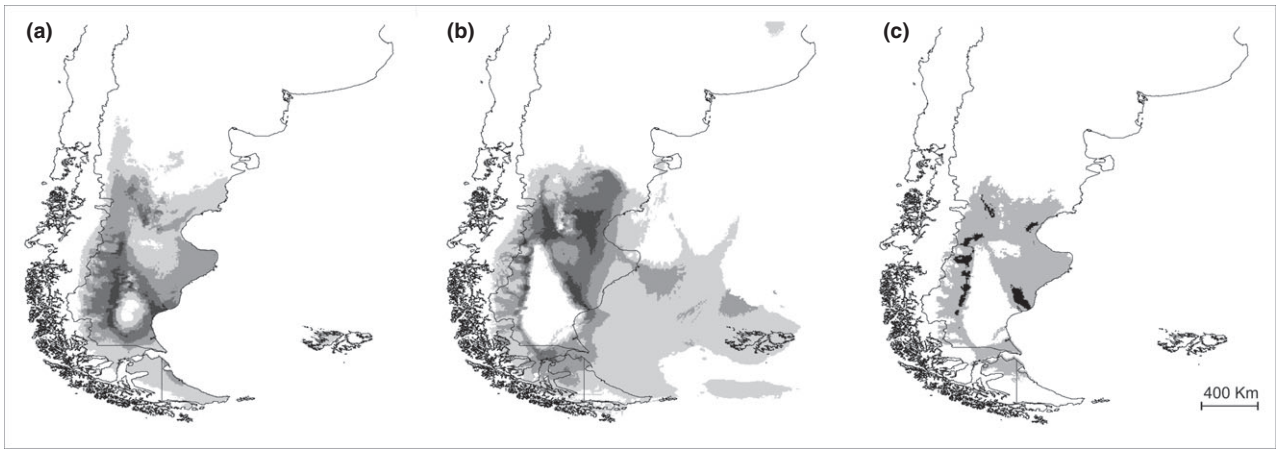

Figure 6 Ecological niche models for *Anarthrophyllum desideratum* in southern South America: (a) current distribution, (b) Last Glacial Maximum distribution, (c) potential stable areas between past and current distributions. Dark areas indicate more suitable habitat, averaged across 10 cross-validation runs.

by crossing the steppe to the Atlantic Ocean (e.g. Morando *et al.*, 2007; Cosacov *et al.*, 2010). In particular, the Deseado River was an important historical drainage (Martínez & Coronato, 2008) and a phylogeographical study of the shrub *Mulinum spinosum* (Sede *et al.*, 2012) found the same phylogeographical break. Together, these inferences suggest that the Deseado River has been important in isolating plant populations of northern Patagonia from those of southern Patagonia.

Our estimates indicate that diversification within NG and SG occurred primarily during the Pleistocene. Consequently, associated changes in climate and landscape dynamics, such as shifts in river basins, Quaternary palaeobasins and coastlines (Martínez & Kutschker, 2011; Ponce *et al.*, 2011), constitute factors that may contribute to the differentiation of Patagonian lineages (Sérsic *et al.*, 2011). Based on genetic diversity and inferred historical processes for NG and SG haplotypes, different evolutionary pathways occurred in these respective areas.

The high proportion of NG localities with private and fixed haplotypes suggests that small isolated populations persisted in this area. Combined with inferences from NCPA, it is evident that diversification within NG was particularly strongly affected by past fragmentation processes, which were detected not only in more ancestral haploclades (deeper levels) but also in the most recent (the shallowest). Enlargement of Quaternary river basins converted Patagonia into a highly fragmented landscape (Martínez & Coronato, 2008; Martínez & Kutschker, 2011) that probably promoted the isolation of populations. In particular, differentiation within NG into inland and coastal lineages may have been associated with the Chubut Basin. Other Quaternary river basins have been associated with fragmentation patterns in other terrestrial taxa (e.g. Kim *et al.*, 1998; Morando *et al.*, 2007; Cosacov *et al.*, 2010; Sede *et al.*, 2012) as have coastline shifts (e.g. San Jorge Gulf zone; Sérsic *et al.*, 2011). From niche modelling it appears that, climatically, the whole northern area was suitable for this species' survival during the Pleistocene. In particular, two areas concordant with the distribution of the N-inland and N-coast groups showed the highest probability of occurrence.

In contrast, genetic patterns in SG seem to be imprinted by successive bottlenecks and colonization processes, given a prevalence of monomorphic localities possessing any one of the three most common haplotypes (H6, H9 and H10), which are also found in putative source populations containing high genetic diversity and/or exclusive haplotypes (for example, localities 12, 14, 18 and 25). In fact, demographic analysis and NCPA inferences indicate demographic expansion for SG. Considering haplotype diversity and distribution patterns of shared haplotypes (i.e. H6, H9, H10), the most likely range expansion was from western to northern, eastern and southern Patagonia. Under a model of post-glacial range expansion associated with glacial retreat and warmer conditions (i.e. a cold–warm hypothesis), however, we should observe a westward and/or southward expansion (e.g. Cosacov *et al.*, 2010; Lessa *et al.*, 2010). Consequently, we propose that phylogeographical patterns of *A. desideratum* in southern Patagonia are related more to precipitation regimes and aridization processes (i.e. a historical water balance hypothesis) during Pleistocene glaciations than to the traditional cold–warm hypothesis. This is supported, first, by niche-modelling analyses that identify precipitation as the most important variable predicting the distribution of *A. desideratum.* This agrees with the suggestion that water is a critically limiting resource affecting the composition and structure of vegetation in Patagonia (Paruelo *et al.*, 1998; Schwinning & Sala, 2004). Of particular relevance to this study, a dendrochronological examination of the co-distributed *A. rigidum* suggested a dominant influence of water availability on the growth of this shrub throughout its range (Srur & Villalba, 2009). Second, palaeodistribution modelling shows an extensive area in southern inland Patagonia that was unsuitable for this species, while areas with the highest probability of occurrence were restricted to western and

north-eastern parts of southern Patagonia. Again, this inland area was probably extremely arid during the LGM and GPG, while western and eastern areas received some influence of humid air masses from the weakened westerlies and from the Atlantic Ocean, respectively (Cusminsky *et al.*, 2011; Ponce *et al.*, 2011; Rabassa *et al.*, 2011). Third, good agreement exists between areas with a high probability of occurrence predicted by the palaeoecological niche model, and areas with the highest genetic diversity and/or areas with exclusive haplotypes. This reinforces the proposal that western and north-eastern areas of the southern Patagonian region represent areas where species persisted *in situ* (e.g. Hewitt, 2004; Mráz *et al.*, 2007). Moreover, differentiation and isolation between south-eastern monomorphic localities containing H10 (group S-I derived by SAMOVA) and the remaining southern localities are probably the consequence of the large unsuitable inland area during the Pleistocene. Fourth, as mentioned above, demographic analyses provide evidence of range expansion in southern Patagonia. Because expansion from refugia leaves genetic signatures of high diversity in refugial areas and low diversity in colonized areas attributed to successive bottlenecks (Hewitt, 2004), we conclude that the most likely range expansion was eastwards, northwards and southwards from western Patagonia (H6 and H9). Nevertheless, there is also evidence of westward expansion into Helsingfors and Torres del Paine, the only two sampled localities in formerly glaciated areas. In particular, all surveyed individuals of Helsingfors possess the most frequent haplotype (H6), probably indicative of dispersal from the non-glaciated western Patagonian steppe.

Together, this evidence suggests that a historical water balance hypothesis can better explain genetic patterns than the traditional cold–warm hypothesis, particularly in the south-eastern area of the distribution range. Relevant to this conclusion, previous plant studies covering the Patagonian steppe (i.e. Jakob *et al.*, 2009; Cosacov *et al.*, 2010; Sede *et al.*, 2012) found some genetic patterns that cannot be explained by the cold–warm hypothesis or by *in situ* survival during Pleistocene glaciations. For example, in the southern geographical range of *Hordeum* species (Jakob *et al.*, 2009), *Calceolaria* (Cosacov *et al.*, 2010), and most evident in *Mulinum* (Sede *et al.*, 2012), eastern localities have low to no haplotype diversity and western localities possess the highest molecular diversity and numbers of exclusive haplotypes. Further paralleling our results, the palaeodistribution of *Mulinum* shows the southern inland area with a low probability of occurrence (Sede *et al.*, 2012). Although few studies for the Patagonian steppe exist, patterns observed for these three taxa support the water balance hypothesis, suggesting this could be a shared factor of influence among Patagonian steppe plant species.

CONCLUSIONS

The effect of glacial and interglacial periods in species range dynamics outside the tropics has been primarily interpreted in terms of cyclical warming and cooling, and/or the associated ice sheet extension/retraction cycles rather than precipitation regimes (Hewitt, 2004; Keppel *et al.*, 2012). However, water balance is a principal environmental factor determining the distribution of terrestrial vegetation (Stephenson, 1990; Neilson, 1995) and, in Patagonia, water is a critically limiting resource affecting the composition and structure of vegetation (Paruelo *et al.*, 1998; Schwinning & Sala, 2004). As suggested by some recent world-wide studies that identified factors such as increasing aridity associated with glacial periods providing evidence for mesic refugia (Keppel *et al.*, 2012), temperature and ice sheet coverage alone are insufficient or inadequate for explaining patterns of genetic variation in *A. desideratum*. The recent eastward demographic expansion of this taxon, the location of areas of high genetic diversity and the dynamic geographical range for this species reconstructed through time via niche modelling, suggest a phylogeographical hypothesis with an important precipitation-related component that is different from those previously proposed for Patagonian organisms. We suggest that future studies should explicitly consider the historical water balance hypothesis as a complement to the traditional cold–warm hypothesis, to better understand the evolutionary history and diversification patterns of Patagonian plants.

ACKNOWLEDGEMENTS

We thank G. Oliva, M. Nicola, L. Ávila and staff of INTA-Santa Cruz for field assistance; A. Wright for laboratory assistance; E. Castro for help with BEAST; and M. Morando and three anonymous referees for valuable comments on previous versions of the manuscript. A.A.C. and A.N.S. acknowledge the National Research Council of Argentina (CONICET) as researchers and A.C. and V.P. as their post-doctoral fellowship holders. Financial support was provided by the following grants: NSF DBI-0520978, NSF-PIRE (OISE-0530267) for support of collaborative research on Patagonian Biodiversity awarded to the following institutions (listed alphabetically): Brigham Young University (USA), Centro Nacional Patagónico (Argentina), Dalhousie University, Instituto Botánico Darwinion (Argentina), Universidad Austral de Chile (Chile), Universidad de Concepción (Chile), Universidad Nacional del Comahue (Argentina), Universidad Nacional de Córdoba (Argentina), and University of Nebraska (USA), CONICET (PIP 11220080101264), FONCyT (PICT 01-10952 and 01-33755) and SeCyT–UNC (to A.N.S. and A.A.C.).

REFERENCES

Alekseyenko, A.V., Lee, C.J. & Suchard, M.A. (2008) Wagner and Dollo: a stochastic duet by composing two parsimonious solos. *Systematic Biology*, **57**, 772–784.

Alsos, I.G., 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.

Barreda, V., Guler, V. & Palazzesi, L. (2008) Late Miocene continental and marine palynological assemblages from Patagonia. *The Late Cenozoic of Patagonia and Tierra del Fuego* (ed. by J. Rabassa), pp. 343–350. Elsevier, Oxford.

Buckley, T.R., Marske, K. & Attanayake, D. (2010) Phylogeography and ecological niche modelling of the New Zealand stick insect *Clitarchus hookeri* (White) support survival in multiple coastal refugia. *Journal of Biogeography*, **37**, 682–695.

Clement, M., Posada, D. & Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1999.

Compagnucci, R.H. (2011) Atmospheric circulation over Patagonia from the Jurassic to present: a review through proxy data and climatic modelling scenarios. *Biological Journal of the Linnean Society*, **103**, 229–249.

Conterato, I.F., Miotto, S.T.S. & Schifino-Wittmann, M.T. (2007) Chromosome number, karyotype, and taxonomic considerations on the enigmatic *Sellocharis paradoxa* Taubert (Leguminosae, Papilionoideae, Genisteae). *Botanical Journal of the Linnean Society*, **155**, 223–226.

Cosacov, A., Sérsic, A.N., Sosa, V., Johnson, L.A. & Cocucci, A.A. (2010) Multiple periglacial refugia in the Patagonia steppe and post-glacial colonization of the Andes: the phylogeography of *Calceolaria polyrhiza*. *Journal of Biogeography*, **37**, 1463–1477.

Crandall, K.A. & Templeton, A.R. (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics*, **134**, 959–969.

Cullings, K.W. (1992) Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. *Molecular Ecology*, **1**, 233–240.

Cusminsky, G., Schwalb, A., Pérez, A.P., Pineda, D., Viehberg, F., Whatley, R., Markgraf, V., Gilli, A., Ariztegui, D. & Anselmetti, F.S. (2011) Late Quaternary environmental changes in Patagonia as inferred from lacustrine fossil and extant ostracods. *Biological Journal of the Linnean Society*, **103**, 397–408.

Doyle, J.J. & Doyle, J.L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, **19**, 11–15.

Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Drummond, A.J., Rambaut, A., Shapiro, B. & Pybus, O.G. (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, **22**, 1185–1192.

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

Excoffier, L. (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology*, **13**, 853–864.

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

Fu, Y.X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Glasser, N.F., Harrison, S., Winchester, V. & Aniya, M. (2004) Late Pleistocene and Holocene palaeoclimate and glacier fluctuations in Patagonia. *Global and Planetary Change*, **43**, 79–101.

Golluscio, R.A. & Sala, O.E. (1993) Plant functional types and ecological strategies in Patagonian forbs. *Journal of Vegetation Sciences*, **4**, 839–846.

Gradstein, F.M. & Ogg, J.G. (2004) Geologic Time Scale 2004 – why, how, and where next! *Lethaia*, **37**, 175–181.

Guindon, S., Lethiec, F., Duroux, P. & Gascuel, O. (2005) PHYML online – a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Research*, **33**, W557–W559.

Hamilton, M.B. (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology*, **8**, 521–523.

Heusser, C.J. (1987) Quaternary vegetation of southern South America. *Quaternary of South America and Antarctic Peninsula*, Vol. 5 (ed. by J. Rabassa), pp. 197–221. A. A. Balkema, Leiden.

Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 183–195.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Jacquemin, S.J. & Pyron, M. (2011) Impacts of past glaciation events on contemporary fish assemblages of the Ohio River basin. *Journal of Biogeography*, **38**, 982–991.

Jakob, S.S., Martinez-Meyer, E. & Blattner, F.R. (2009) Phylogeographic analyses and paleodistribution modeling indicate Pleistocene *in situ* survival of *Hordeum* species (Poaceae) in southern Patagonia without genetic or spatial restriction. *Molecular Biology and Evolution*, **26**, 907–923.

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

Keppel, G., Van Niel, K.P., Wardell-Johnson, G.W., Yates, C. J., Byrne, M., Mucina, L., Schut, A.G.T., Hopper, S.D. & Franklin, S.E. (2012) Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global Ecology and Biogeography*, **21**, 393–404.

Kim, I., Phillips, C.J., Monjeau, J.A., Birney, E.C., Noack, K., Pumo, D.E., Sikes, R.S. & Dole, J.A. (1998) Habitat islands, genetic diversity, and gene flow in a Patagonian rodent. *Molecular Ecology*, **7**, 667–678.

Lavin, M., Herendeen, P.S. & Wojciechowski, M.F. (2005) Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Systematic Biology*, **54**, 575–594.

Lessa, E.P., D'Elía, G. & Pardiñas, U.F.J. (2010) Genetic footprints of late Quaternary climate change in the diversity of Patagonian–Fueguian rodents. *Molecular Ecology*, **19**, 3031–3037.

Lewis, G., Schrire, B., Mackinder, B. & Lock, M. (2005) *Legumes of the world*. Royal Botanic Gardens, Kew.

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

Markgraf, V. (1983) Late and postglacial vegetational and paleoclimatic changes in subantarctic, temperate and arid environments in Argentina. *Palynology*, **7**, 43–70.

Martínez, O.A. & Coronato, A.M.J. (2008) The Late Cenozoic fluvial deposits of Argentine Patagonia. *The Late Cenozoic of Patagonia and Tierra del Fuego* (ed. by J. Rabassa), pp. 205–226. Elsevier, Oxford.

Martínez, O.A. & Kutschker, A. (2011) The 'Rodados Patagónicos' (Patagonian Shingle Formation) of eastern Patagonia: environmental conditions of gravel sedimentation. *Biological Journal of the Linnean Society*, **103**, 336–345.

Morando, M., Avila, L.J., Turner, C.R. & Sites, J.W., Jr (2007) Molecular evidence for a species complex in the Patagonian lizard *Liolaemus bibronii* and phylogeography of the closely related *Liolaemus gracilis* (Squamata: Liolaemini). *Molecular Phylogenetics and Evolution*, **43**, 952–973.

Mráz, P., Gaudeul, M., Rioux, D., Gielly, L., Choler, P., Taberlet, P. & the IntraBioDiv Consortium (2007) Genetic structure of *Hypochaeris uniflora* (Asteraceae) suggests vicariance in the Carpathians and rapid post-glacial colonization of the Alps from an eastern Alpine refugium. *Journal of Biogeography*, **34**, 2100–2114.

Nei, M. (1987) *Molecular evolutionary genetics*. Columbia University Press, New York.

Neilson, R.P. (1995) A model for predicting continental-scale vegetation distribution and water-balance. *Ecological Applications*, **5**, 362–385.

Oliva, G.E., Noy-Meir, I. & Cibils, A. (2001) Fundamentos de ecología de pastizales. *Ganadería sustentable en la Patagonia austral* (ed. by P. Borrelli and G.E. Oliva), pp. 81–98. INTA, Río Gallegos, Argentina.

Paez, M.M., Prieto, A.R. & Mancini, M.V. (1999) Fossil pollen from Los Toldos locality: a record of the Late-glacial transition in the Extra-Andean Patagonia. *Quaternary International*, **53–54**, 69–75.

Paiaro, V. (2011) *Gradientes ambientales y márgenes de distribución: patrones espaciales de variabilidad fenotípica, atributos poblacionales y caracteres reproductivos en* Anarthrophyllum desideratum *(DC.) Benth*. PhD Thesis, Universidad Nacional de Córdoba, Córdoba, Argentina.

Paruelo, J.M., Beltrán, A., Jobbágy, E., Sala, O.E. & Golluscio, R.A. (1998) The climate of Patagonia: general patterns and controls on biotic processes. *Ecología Austral*, **8**, 85–101.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Ponce, J.F., Rabassa, J., Coronato, A. & Borromei, A.M. (2011) Palaeogeographical evolution of the Atlantic coast of Pampa and Patagonia since the last glacial maximum to the Middle Holocene. *Biological Journal of the Linnean Society*, **103**, 363–379.

Posada, D. (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Posada, D., Crandall, K.A. & Templeton, A.R. (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488.

Rabassa, J., Coronato, A. & Martínez, O. (2011) Late Cenozoic glaciations in Patagonia and Tierra del Fuego: an updated review. *Biological Journal of the Linnean Society*, **103**, 316–335.

Rambaut, A. (2008) *FigTree v1.1.1*. Available at: http://tree.bio.ed.ac.uk/software/figtree.

Rambaut, A. & Drummond, A.J. (2009) *Tracer v1.5.0*. Available at: http://beast.bio.ed.ac.uk/Tracer (accessed 20 August 2011).

Ramos, V. (1999) Las provincias geológicas del territorio argentino. *Anales del Servicio Geológico Minero Argentino*, **29**, 41–96.

Ramos, V.A. & Ghiglione, M.C. (2008) Tectonic evolution of the Patagonian Andes. *The Late Cenozoic of Patagonia and Tierra del Fuego* (ed. by J. Rabassa), pp. 205–226. Elsevier, Oxford.

Rogers, A.R. & Harpending, H. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Roig, F.A. (1998) La vegetación de la Patagonia. *Flora Patagónica*, Vol. VIII, Part I (ed. by M.N. Correa), pp. 48–166. INTA, Buenos Aires, Argentina.

Schwinning, S. & Sala, O.E. (2004) Hierarchy of responses to resource pulses in arid and semi-arid ecosystems. *Oecologia*, **141**, 211–220.

Sede, S.M., Nicola, M.V., Pozner, R. & Johnson, L.A. (2012) Phylogeography and palaeodistribution modelling in the Patagonian steppe: the case of *Mulinum spinosum* (Apiaceae). *Journal of Biogeography*, **39**, 1041–1057.

Sérsic, A.N., Cosacov, A., Cocucci, A.A., Johnson, L.A., Pozner, R., Avila, L.J., Sites, J.W., Jr & Morando, M. (2011) Emerging phylogeographical patterns of plants and terrestrial vertebrates from Patagonia. *Biological Journal of the Linnean Society*, **103**, 475–494.

Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E. & Small, R.L. (2005) The tortoise and the hare. II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, **92**, 142–166.

Simmons, M.P. & Ochoterena, H. (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology*, **49**, 369–381.

Sorarú, S.B. (1974) Revisión de *Anarthrophyllum*, género Argentino-Chileno de Leguminosas. *Darwiniana*, **18**, 453–488.

Srur, A.M. & Villalba, R. (2009) Annual growth rings of the shrub *Anarthrophyllum rigidum* across Patagonia: interannual variations and relationships with climate. *Journal of Arid Environments*, **73**, 1074–1083.

Stephenson, N.L. (1990) Climatic control of vegetation distribution: the role of the water balance. *The American Naturalist*, **135**, 649–670.

Suchard, M.A., Weiss, R.E. & Sinsheimer, J.S. (2001) Bayesian selection of continuous-time Markov chain evolutionary models. *Molecular Biology and Evolution*, **18**, 1001–1013.

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**, 597–601.

Templeton, A.R. (1998) Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381–397.

Templeton, A.R. (2011) *Inference key for the nested haplotype tree analysis of geographical distances*. Available at: http://darwin.uvigo.es/download/geodisKey_06Jan11.pdf (accessed 20 August 2011).

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Outgroup species, locality, coordinates and voucher numbers.

Appendix S2 Spatial analysis of molecular variance results for the southern and northern groups of *Anarthrophyllum desideratum* populations.

Appendix S3 Percentage contribution of climatic variables to the past and current simulations of the distribution of *Anarthrophyllum desideratum*.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCH

The authors of this paper constitute an interdisciplinary team of molecular and field researchers interested in ecology, evolution, systematics and conservation of plants with special interest in Patagonian species; http://patagonia.byu.edu/Home.aspx, http://www.efn.uncor.edu/departamentos/divbioeco/otras/bioflor/.

Author contributions: A.C. and A.N.S. conceived the idea; A.C., A.N.S., V.P. and A.A.C., collected the samples; L.A.J. performed the molecular analyses; A.C. and A.N.S. analysed the data statistically; F.E.C. contributed the geological and climatic framework of Patagonia; A.C. and A.N.S. led the writing; L.A.J., V.P. and A.A.C. improved the final version of the manuscript.

Editor: Jorge Crisci