

Quaternary range and demographic expansion of *Liolaemus darwini* (Squamata: Liolaemidae) in the Monte Desert of Central Argentina using Bayesian phylogeography and ecological niche modelling

ARLEY CAMARGO,* FERNANDA P. WERNECK,†‡ MARIANA MORANDO,* JACK W. SITES JR† and LUCIANO J. AVILA*

*Unidad de Diversidad, Sistemática y Evolución, Centro Nacional Patagónico, Consejo Nacional de Investigaciones Científicas y Técnicas, Boulevard Almirante Brown 2915, U9120ACD, Puerto Madryn, Chubut, Argentina, †Department of Biology & Bean Life Science Museum, Brigham Young University, Provo, UT 84602, USA, ‡Departamento de Zoologia, Universidade de Brasília, Brasília, DF, CEP 70910-900, Brazil

Abstract

Until recently, most phylogeographic approaches have been unable to distinguish between demographic and range expansion processes, making it difficult to test for the possibility of range expansion without population growth and vice versa. In this study, we applied a Bayesian phylogeographic approach to reconstruct both demographic and range expansion in the lizard *Liolaemus darwini* of the Monte Desert in Central Argentina, during the Late Quaternary. Based on analysis of 14 anonymous nuclear loci and the cytochrome *b* mitochondrial DNA gene, we detected signals of demographic expansion starting at ~55 ka based on Bayesian Skyline and Skyride Plots. In contrast, Bayesian relaxed models of spatial diffusion suggested that range expansion occurred only between ~95 and 55 ka, and more recently, diffusion rates were very low during demographic expansion. The possibility of population growth without substantial range expansion could account for the shared patterns of demographic expansion during the Last Glacial Maxima (OIS 2 and 4) in fish, small mammals and other lizards of the Monte Desert. We found substantial variation in diffusion rates over time, and very high rates during the range expansion phase, consistent with a rapidly advancing expansion front towards the southeast shown by palaeo-distribution models. Furthermore, the estimated diffusion rates are congruent with observed dispersal rates of lizards in field conditions and therefore provide additional confidence to the temporal scale of inferred phylogeographic patterns. Our study highlights how the integration of phylogeography with palaeo-distribution models can shed light on both demographic and range expansion processes and their potential causes.

Keywords: Bayesian phylogeography, demographic expansion, dispersal, *Liolaemus*, Monte Desert, range expansion

Received 10 October 2012; revision received 25 April 2013; accepted 25 April 2013

Introduction

A synthesis of phylogeographic studies of multiple taxa from the Northern Hemisphere has revealed repeated cycles of range retraction towards southern refugia

during glacial periods and range expansion towards higher (northern) latitudes during interglacial periods (Hewitt 2000, 2004; Hewitt & Ibrahim 2001). Comparative studies in the Southern Hemisphere have found similar, but reversed, cyclical patterns of persistence in low-latitude refugia and postglacial expansion to higher (southern) latitudes in multiple taxa (Lessa *et al.* 2010; Sersic *et al.* 2011; Fraser *et al.* 2012). In Patagonia, other

Correspondence: Arley Camargo, Fax: (+54) 2965 451024; E-mail: arley.camargo@gmail.com

taxa also show patterns of displacements towards the east and persistence in situ in the west (Breitman *et al.* 2012; Cosacov *et al.* 2012; Sede *et al.* 2012). These generalizations about shared phylogeographic patterns among species worldwide have been derived from multiple inferences of demographic history using genetic data and geographic information to elucidate, qualitatively or quantitatively, changes in population size (demographic expansion) and geographic range (spatial expansion). However, until recently, phylogeographic methods have failed to distinguish between demographic and spatial processes within a single statistical framework, making it difficult to test for the possibility of demographic expansion without range expansion, and vice versa.

Traditionally, inferences of population growth have been obtained with test statistics such as Tajima's D (Tajima 1989), Fu's F_s (Fu 1997), Harpending's raggedness index (Harpending 1994) and Ramos-Onsins and Rozas' R_2 (Ramos-Onsins & Rozas 2002). In addition, several Bayesian Skyline Plot (BSP) methods are able to infer changes in effective population size over time using flexible demographic models (Ho & Shapiro 2011). In contrast, assessments of range expansion have relied upon model-free inferences about the association between haplotype clades and their geographic distribution using Nested Clade Phylogeographic Analysis (NCPA) (Templeton 2004, 2008). NCPA has been used routinely to detect range expansions in empirical studies although inferences based on this approach are not statistically robust due to its inability to distinguish between alternative scenarios and high error rates (Beaumont & Panchal 2008; Garrick *et al.* 2008; Knowles 2008; Petit 2008; but see Garrick *et al.* 2008 and Templeton 2008). Other methods have generally lacked a spatially explicit component such as the spatial expansion test implemented in Arlequin 3.5 (Excoffier & Lischer 2010) based on Excoffier's range expansion model (Excoffier 2004). A fully integrated analysis of genetic and geographic information was achieved in the maximum-likelihood-based program Phylomapper (Lemmon & Lemmon 2008), which estimates ancestral distributions and assumes a constant rate of spatial diffusion, but does not incorporate either genealogical uncertainty (i.e. it uses a fixed tree topology) or a model of demographic expansion. More recently, a Bayesian method developed for area biogeography could potentially be applied to reconstruct ancestral distributions and dispersal events among discrete locations (Yu *et al.* 2011).

A Bayesian approach for spatio-temporal phylogeographic reconstruction has been implemented recently in the software BEAST 1.7 (Drummond *et al.* 2012). This approach allows inference about the geographic location of ancestors while accommodating genealogical

uncertainty, and using models of either discrete or continuous diffusion over space and time. While models of homogenous diffusion assume a Brownian constant rate among discrete locations (Lemey *et al.* 2009a), an extension of this approach that models diffusion on a continuous space allows for variation in the diffusion rate across branches of the genealogy (Lemey *et al.* 2010), based on analogous models to those implemented for relaxed-molecular clocks (Drummond *et al.* 2006). Because the analysis is embedded within the BEAST program, the spatial diffusion process can be coupled with the estimation of demographic models (e.g. constant size, exponential growth, BSP, etc.), and model-selection based on marginal likelihoods can be used to test for alternative diffusion and demographic models (Baele *et al.* 2012). This flexible phylogeographic approach that accomplishes both pattern inference and model testing has resulted in novel insights about the spread of several viral epidemics (Lemey *et al.* 2009b; Allicock *et al.* 2012; Pybus *et al.* 2012), and even the origin of human languages using vocabulary instead of molecular data (Walker & Ribeiro 2011; Bouckaert *et al.* 2012).

Ecological niche modelling (ENM) has become a popular tool in phylogeography, evolutionary biology and conservation biology for inference of potential geographic distributions of species in past, present and future climatic conditions (Chan *et al.* 2011). Among other applications, ENM has been used for delimiting species (Raxworthy *et al.* 2007; Rissler & Apodaca 2007), locating Pleistocene refugia in the Northern (Waltari *et al.* 2007) and Southern (Carnaval & Moritz 2008; Werneck *et al.* 2011, 2012b; Fontanella *et al.* 2012) Hemispheres, developing models of population history (Carstens & Richards 2007; Richards *et al.* 2007), and studying niche evolution (Warren *et al.* 2008). Despite several drawbacks of species-distribution models in general, and of ENM in particular (Pearson *et al.* 2006; Elith & Graham 2009; Elith & Leathwick 2009; Araujo & Peterson 2012), this approach offers the possibility to build a finite set of phylogeographic hypotheses for evaluation with independent genetic data using model-based statistical approaches (Knowles 2009; Bloomquist *et al.* 2010; Hickerson *et al.* 2010). Expanded phylogeographic approaches that integrate geospatial methods, such as ENM, can help elucidate the ecological and evolutionary mechanisms underlying population history both when spatio-temporal changes in climate and habitat suitability match demographic inferences (Chan *et al.* 2011), or when they do not explicitly agree (Thomé *et al.* 2010; Werneck *et al.* 2012a).

In this study, we reconstruct an explicit spatio-temporal phylogeographic model for *Liolaemus darwini* (Squamata: Liolaemidae), a lizard that occupies sandy habitats in the southern and central portions of the Monte Desert in Argentina (Etheridge 1993). A recent species-delimitation

study based on coalescent methods and multilocus data, found that a southern lineage of *L. darwinii* is sister to a new species distributed further north (Camargo *et al.* 2012). The two lineages have a distributional gap in central Mendoza Province and differ in scale counts and colouration (Etheridge 2001; Abdala 2007). A previous study found significant signal for demographic and range expansion in southern *L. darwinii* based on cytochrome *b* (cyt *b*) analyses using NCPA and several test statistics, and proposed that Pleistocene climatic cycles could have been associated with this population history (Morando *et al.* 2004).

Herein, we increased samples sizes and included nuclear loci in order to test the hypothesis of associated population growth and range expansion of *L. darwinii* during the Late Quaternary, using spatio-temporal phylogeographic reconstruction and demographic analyses. Further, we investigated the possible role of the Quaternary climate change in this expansion event via

palaeo-distribution modelling since the Last Interglacial (120 ka). We predicted that if population expansion was due to recent climatic change, we would be able to find: (i) a temporal association between the population growth and the range expansion; and (ii) a spatial association between the range expansion and palaeo-distribution models. Specifically, we would expect that the geographic range of the species inferred with both phylogeographic and palaeo-distribution reconstructions should match spatially during the progression of the expansion event.

Material and methods

We sampled 404 individuals of *L. darwinii* from 105 localities including the distribution of southern lineages (Fig. 1, Table S1, Supporting information). Tissue samples from liver and/or tail muscle were preserved in absolute ethanol and stored at -20°C . Specimens were

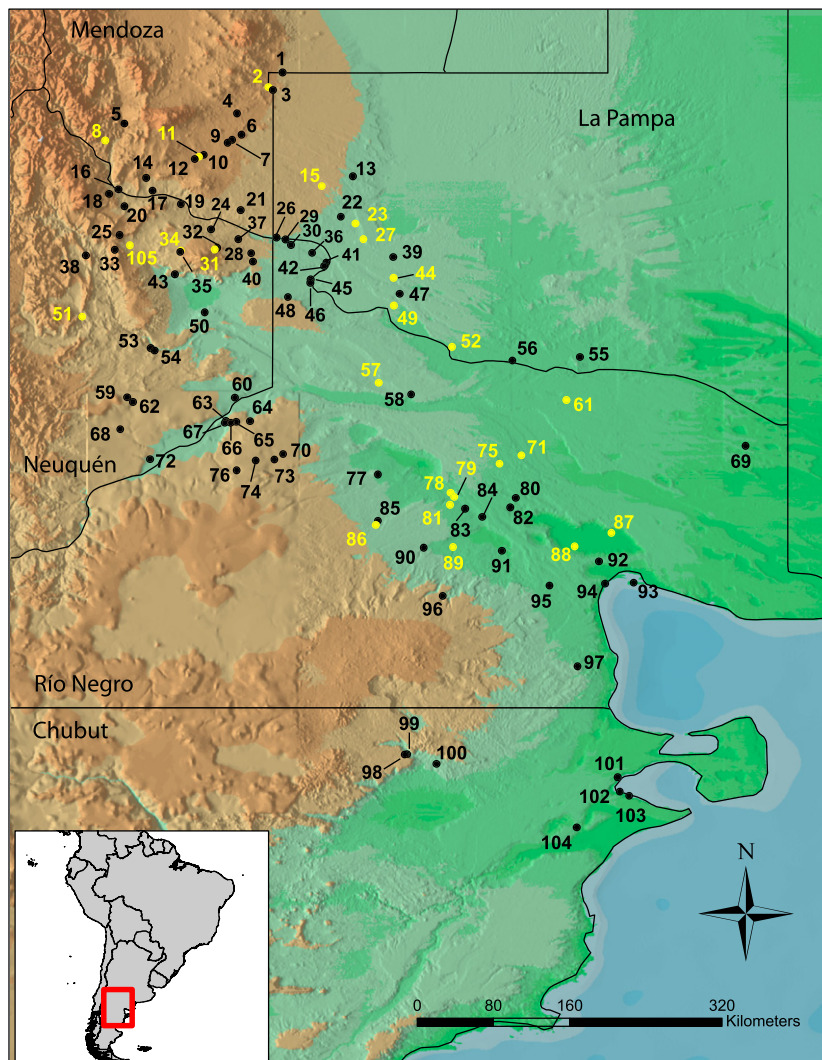


Fig. 1 Map of central Argentina showing the sampled localities of *L. darwinii* plotted against a digital elevation model for South America (green represents lower altitudes and brown represents higher altitudes). Numbers are the codes to identify the sampled localities in Table S1. Black dots represent the locations sampled for cyt *b* only and yellow dots identify the locations sampled for both cyt *b* and ANL loci.

fixed in 10–20% formalin, later transferred to 70% ethanol and deposited in the herpetological collection of the Centro Nacional Patagónico (LJAMM-CNP, CENPAT-CONICET; <http://www.cenpat.edu.ar/colecciones03.html>) and the Bean Life Science Museum, Brigham Young University (<http://mlbean.byu.edu/ResearchCollections/Collections/ReptilesandAmphibians.aspx>). Genomic DNA was extracted with the DNAeasy Qiagen kit (Qiagen).

We used the Green Go Taq PCR kit (Promega) for all PCRs in PTC-200 DNA Engine (MJ Research) or GeneAmp PCR 9700 thermal cyclers (Applied Biosystems, Inc.). Sequencing reactions used the Big-Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Inc.) in a GeneAmp PCR 9700 thermal cycler (Applied Biosystems, Inc.). Sequencing products were cleaned with Sephadex G-50 Fine (GE Healthcare Bio-Sciences AB) and sequenced in an ABI 3730xl DNA Analyzer (Applied Biosystems, Inc.). We sequenced the *cyt b* mtDNA gene for all individuals following protocols in Morando *et al.* (2004) and subsampled 29 individuals across the geographic range representing the variation found in mtDNA haplotypes for screening anonymous nuclear loci (ANL). We sequenced 14 ANL that were developed from the genomic DNA of a *L. darwini* individual (LJAMM-CNP 7097) following protocols in Noonan & Yoder (2009) (Table S2, Supporting information). We generated ~200 random fragments, cloned, sequenced and BLAST searched these to confirm they were anonymous, then designed primers for fragments with confirmed anonymity and used the PCR temperature profile of Noonan & Yoder (2009) to screen ANL in all subsampled individuals.

Chromatograms were checked by eye, and ambiguity codes were used to represent polymorphisms of heterozygote individuals in Sequencher 4.10.1 (Gene Codes Corporation). Gametic phase of heterozygotes was resolved with the program PHASE 2.1.1 (Stephens *et al.* 2001) based on the input files prepared with SeqPHASE (<http://www.mnhn.fr/jfflot/seqphase/>) (Flot 2010). The heterozygous indels (alleles with length variation) found in 8 ANL were phased along with base polymorphisms by specifying them as unknown phases in a separate input file (.known). We ran two independent runs consisting of 10 000 iterations, a thinning interval of 1 and a burn-in of 10 000. To ensure convergence between runs, we checked for identical best haplotype pairs estimated for each individual between runs. Input and output files of PHASE analyses are provided as supplemental files in Dryad (see Data Accessibility below). Subsequently, the output files from PHASE were converted back into FASTA alignments of phased sequences using SeqPHASE. Sequences were inspected by eye to check that there were no fixed heterozygotes

at any polymorphic site and to ensure we were not using multiple-copy markers (Thomson *et al.* 2010). Alignments were tested for recombination with the program RDP3 beta35 (Martin *et al.* 2010), and best-substitution models for each locus were selected with jModeltest 2 (Darriba *et al.* 2012).

Population structure

We estimated haplotype networks for all loci using TCS 1.21 (Clement *et al.* 2000) using a 95%-connection limit and considering gaps as a fifth state for the ANL. In the case of *cyt b*, we first eliminated positions with missing nucleotides to minimize ambiguous connections among haplotypes, reaching a final alignment length of 357 bp. We also estimated the optimal number of populations using the 15-locus data set representing 29 individuals with the Bayesian program Geneland 4.0.3 (Guillot *et al.* 2005a) under the assumption of Hardy–Weinberg equilibrium among populations and linkage equilibrium between loci (Guillot *et al.* 2005b). We ran five independent replicate runs with the number of populations ranging between 1 and 10, and assuming a correlated allele frequencies model and a spatial model without uncertainty on coordinates. Each run consisted of one million iterations, a thinning interval of 1000 and a burn-in phase of 200 iterations. To assess convergence among runs, we plotted burn-in plots of the posterior probability parameter and compared the posterior estimates of the number of populations. We created probability maps of population membership on a spatial domain defined by 100 pixels in both X- and Y-axes following the authors' recommendations.

Demographic expansion

We reconstructed the change in effective population size over time with a multilocus Extended Bayesian Skyride Plot (EBSP, Heled & Drummond 2008) in BEAST 1.7.4 (Drummond *et al.* 2012). We unlinked substitution, clock and tree models for all loci, and specified a linear model of population size, instead of the less realistic stepwise model. We applied a strict clock model for *cyt b* with the substitution rate sampled from a normal distribution with mean of 1.94×10^{-2} (standard deviation = 3.46×10^{-5}) to match the calibrated values of Olave *et al.* (in review). Olave *et al.* used a recently dated fossil of *Liolaemus* to calibrate a divergence event within the genus and to obtain an estimate of the substitution rate of multiple loci (including *cyt b*) based on a representative sample of species that were analysed with the *BEAST species-tree method. We applied a strict clock model because a likelihood ratio test could not reject the hypothesis of a molecular clock based on the

topology estimated with BEAST, and the optimized log-likelihood values calculated with PAUP* 4b10 (Swofford 2002) (ln-likelihood with clock: -2910.7, ln-likelihood without clock: -2787.4, χ^2 value = 246.6, critical χ^2 at $P < 0.05$ and 299 degrees of freedom = 340.3). Following the recommendations of the tutorial in the BEAST website, the weights for EBSF operators and the initial value for the population size mean were adjusted to improve MCMC mixing. We performed two independent analyses each consisting of 250 million generations with samples taken every 25 000 generations to obtain a total of 20 000 trees in both runs. Parameter traces were inspected with Tracer v1.5 (Rambaut & Drummond 2009) to check for stationarity and high effective sample sizes (ESS), and the EBSF plot was based on the output file in csv format produced by the combined results.

We also reconstructed the change in effective population size through time using a Bayesian Skyride model (Minin *et al.* 2008), which was implemented as the demographic tree prior of the phylogeographic diffusion analyses (see below). In this case, we analysed the complete cyt *b* data set with a strict clock model for the substitution rate using the calibration of Olave *et al.* (in review). We used the demographic reconstruction analysis for a Bayesian Skyride in Tracer to plot the effective population size as a function of time.

Range expansion

We ran continuous-diffusion phylogeographic analyses in BEAST 1.7.4 with the cyt *b* data set using both a time-homogeneous Standard Random Walk (SRW) model and a time-heterogeneous Relaxed Random Walk (RRW) model with log-normally distributed rate variation. Input xml files were prepared following examples in the tutorials of the BEAST website. After preliminary runs without convergence and low ESS, we subsampled the full cyt *b* data set to select only one individual representing each haplotype found in a given locality, which reduced the data set to 301 individuals. We performed for each model two independent runs of 500 million generations sampled every 25 000 generations, to obtain a total of 20 000 trees from the posterior distribution. Parameter traces were inspected with Tracer to check for stationarity, high ESS and convergence between runs. We compared the fit of SRW and RRW models with marginal likelihood estimations based on path and stepping-stone sampling consisting of a chain of 500 000 steps, sampled every 2000 steps and 100 path steps (Baele *et al.* 2012). Tree files were annotated with TreeAnnotator v1.7.4 (Drummond *et al.* 2012) to estimate a maximum clade credibility (MCC) tree after removing the burn-in. We annotated the MCC tree with the estimated ancestral locations through time to

reconstruct the spatial diffusion pattern and produce an animation for visualization in Google Earth, using SPREAD 1.0.6 (Bielejec *et al.* 2011). We also used TimeSlicer (kindly provided by P. Lemey) to summarize the variation in diffusion rate over time based on the posterior sample of trees at multiple time slices.

All BEAST analyses were run with the Beagle high-performance library (Ayres *et al.* 2012) in the m6 cluster of the Fulton Supercomputing Lab at Brigham Young University (<https://marylou.byu.edu/>). All input xml and output log files from these analyses are available in the Dryad digital repository (see Data Accessibility below).

Palaeo-distribution modelling

We used the maximum entropy machine-learning algorithm implemented by Maxent 3.3.3k (Phillips *et al.* 2006) to infer the potential geographic range of *L. darwini* at the present and at two time periods during the Quaternary (120 and 21 ka) based on palaeo-climatic layers downloaded from the WorldClim database (Hijmans *et al.* 2005). Palaeo-climatic models for the Last Interglacial (120 ka, LIG) followed Otto-Bliesner *et al.* (2006), while we used two models currently available for download for the Last Glacial Maximum (21 ka, LGM): CCSM and MIROC, both developed originally from the Paleoclimate Modelling Intercomparison Project Phase II (PMIP2).

After downloading, all bioclimatic layers (current and past) were cropped to span from 30°41.28' S to 59°18.78' S and from 76°24.39' W to 55°31.89' W, a larger spatial range than the current Monte Desert and Patagonia distributions. We then generated correlation matrices between the climatic variables with ArcGIS version 10 (ESRI) to remove highly correlated variables from the models ($r > 0.85$). We ultimately selected nine of the 19 initial climatic variables (bio2 – Mean Diurnal Range, bio3 – Isothermality, bio4 – Temperature Seasonality, bio6 – Min Temperature of Coldest Month, bio8 – Mean Temperature of Wettest Quarter, bio9 – Mean Temperature of Driest Quarter, bio15 – Precipitation Seasonality, bio16 – Precipitation of Wettest Quarter and bio17 – Precipitation of Driest Quarter). To statistically evaluate model performance, we used the area under the curve (AUC) of the receiver operating characteristic (ROC) plot, a threshold-independent measure of model performance as compared to null expectations.

Because ENMs can produce over-fitted models (often represented by highly complex and irregular variable response curves) and affect model generality and transferability, we did species-specific tuning under the current climatic conditions to increase model performance (Anderson & Gonzales 2011). In the tuning phase, we

explored the strength of protection against over fitting (regularization), which applies a penalty for each term that is included in the model (Phillips *et al.* 2006). We investigated 12 values for the regularization multiplier (from 0.5 to 6.0, in increments of 0.5) under the autofeatures which allows all the set of feature classes (linear, quadratic, hinge, product and threshold) to depend on the number of presence records. We then retained for the past climate projections the best parameter set that produced smoother distributions of the variable response curves, making them more regular and achieving the highest model performance (lowest values of average differences between training and test AUCs) (Anderson & Gonzales 2011). From the investigated values, the intermediate regularization multipliers led to best models, with reduced performances at lower and higher regularization (regularization multiplier selected = 4; results shown in Table S3 and Fig. S1, Supporting information). As reported in other studies (Elith *et al.* 2010), the species-specific smoothing increased model quality and provided more confidence for our projections to the past climate scenarios.

We also explored the Explain tool from Maxent to investigate which variables are driving the reduction in suitability in the palaeo-distribution estimates (see results). Because the Explain tool can only be implemented for additive models, we performed a 10-fold run without the product features for the interactive exploration of predictions. All analyses (including the species-specific tuning) were run with 10 replicates, and in the pictures, we display the pointwise means of the 10 output grids. The occurrence data set included a total of 105 trusted geographic records, and we used a cross-validation testing procedure which divides the occurrence points into training and test points based on the number of replicates (e.g. for the 10-fold cross-validation a test percentage of 10% is used).

Results

The reduced *cyt b* data set (357 bp) for 404 individuals contained 65 haplotypes with a highly variable frequency: 73% of the individuals (295) share the same haplotype (haplotype H1 in cytochrome *b* network of Fig. S2), while 22 haplotypes have a frequency below 2%, and 42 haplotypes are singletons. We successfully amplified and sequenced between 21 and 28 individuals for each ANL locus (see Table S1, Supporting information). We found no evidence of fixed heterozygote positions or recombination in any ANL, and the percentage of heterozygous individuals ranged between 18% and 96%. The gametic phase of heterozygous individuals was resolved for ~90% of variable positions at the 90%-confidence level. The 10% of positions below this

threshold had an average phase-call probability of ~65%. Thus, we used the best allele pairs inferred for each individual in subsequent analyses, which on average had a probability of ~90%, and were consistently inferred between independent runs. We decided to use these best genotype estimates as simulations have shown that omitting unresolved genotypes could underestimate and/or bias parameter estimates (Garrick *et al.* 2010). The number of ANL haplotypes ranged between 6 and 16 and the number of polymorphic sites between 9 and 26 (Table 1). All haplotype networks show a similar topology: one or two high-frequency haplotypes and multiple, rare haplotypes (Fig. S2, Supporting information).

Population structure

The five independent runs in Geneland showed convergence of posterior probability based in burn-in plots and also in the posterior estimates of the number of populations. The mode in the number of populations was five in all runs with an average frequency among MCMC iterations after burn-in of 31% (Fig. S3, Supporting information). The probability maps for the distribution of these five population clusters partition the range of *L. darwinii* into central (clusters 1 and 3), southern (cluster 2) and northern regions (clusters 4 and 5) (Fig. 2). Input and output files of Geneland are deposited in Dryad (see Data Accessibility below).

Demographic expansion

The EBSF analysis shows an initial period of stability between 95 and 55 ka followed by exponential growth in population size since ~55 ka until the present (Fig. 3). The final population size is about 50 times higher than the initial population size before the demographic expansion (Fig. 3). The Skyride reconstruction based on *cyt b* only resulted in a similar pattern of population growth that started ~55 ka and also a very similar final population size (Fig. 3). Five ANL plus *cyt b* had significant *D* and *F_s* values, except for a nonsignificant *F_s* value for A8E (Table 1). Two additional ANL also have negative, but nonsignificant, *D* and *F_s* values (A1D and A12F). Based on mismatch distributions and the raggedness index, the model of demographic expansion was rejected only for three ANL (A6D, A9C and A12C) while the model of spatial expansion was not rejected for any loci (Table 1).

Range expansion

The comparison of the SRW and RRW models based on marginal likelihood estimates indicates that the variable diffusion rate model fits the data better than the

Table 1 Locus information, samples sizes, summary statistics and GenBank Accession numbers of sequences used in this study

Locus	N	bp	H	s	k	π	Hd	D	Fs	DE	SE
cyt b	404	357	65	59	0.596	0.00167 \pm 0.00014	0.466 \pm 0.032	-2.66019	-28.16556	0.69	0.82
A1D	56	707	12	23	3.589	0.00508 \pm 0.00092	0.597 \pm 0.074	-0.9025	-0.62422	0.08	0.71
A6D	56	641	10	10	3.244	0.00506 \pm 0.00021	0.655 \pm 0.041	1.37301	0.2098	0	0.25
A8E	56	625	5	14	1.464	0.00236 \pm 0.00054	0.525 \pm 0.053	-1.54731	1.73821	0.93	0.3
A8F	48	604	8	19	5.447	0.00911 \pm 0.00077	0.650 \pm 0.059	0.86834	7.3238	0.12	0.5
A9C	52	481	7	12	4.483	0.00932 \pm 0.00079	0.679 \pm 0.049	2.02204	4.05001	0.02	0.53
A9E	44	622	8	9	0.914	0.00148 \pm 0.00035	0.499 \pm 0.090	-1.60038	-8.41053	0.89	0.97
A12C	48	682	8	18	4.837	0.00716 \pm 0.00066	0.749 \pm 0.036	0.6099	2.68505	0.03	0.19
A12D	58	687	6	9	2.404	0.00350 \pm 0.00038	0.649 \pm 0.043	0.64451	2.07738	0.06	0.49
A12F	44	715	10	19	3.186	0.00448 \pm 0.00084	0.695 \pm 0.060	-0.8753	-0.80804	0.15	0.59
B5B	52	646	16	26	7.126	0.01110 \pm 0.00100	0.868 \pm 0.033	0.77841	1.22323	0.1	0.28
B6B	54	415	10	20	2.106	0.00508 \pm 0.00108	0.542 \pm 0.077	-1.6391	-1.5793	0.99	0.67
B6D	50	550	16	18	1.92	0.00349 \pm 0.00058	0.765 \pm 0.057	-1.64329	-8.85054	1	0.89
B7D	52	532	12	22	6.81	0.01297 \pm 0.00073	0.817 \pm 0.036	1.27839	-2.55038	0.66	0.91
B7E	54	685	12	13	1.361	0.00199 \pm 0.00028	0.709 \pm 0.062	-1.54735	-6.17009	0.92	0.86
Locus	GenBank										
cyt b	KC150146–KC150517, AY367797, AY389190, AY389193, AY389198–AY389200, AY389205–AY389207, AY389212–AY389214, AY389220–AY389224, AY389227–AY389230, AY389232–AY389235, AY389238, AY389239, JF272771, JN683141–JN683143										
A1D	KC150518–KC150539, JN682528, JQ659062–JQ659066										
A6D	KC150540–KC150566, JN682616										
A8E	KC150567–KC150594										
A8F	KC150595–KC150617, JN682659										
A9C	KC150618–KC150633, JN682697, JQ659106, JQ659107, JQ659109–JQ659115										
A9E	KC150634–KC150654, JN682732										
A12C	KC150655–KC150678										
A12D	KC150679–KC150706, JN682770										
A12F	KC150707–KC150728										
B5B	KC150729–KC150753, JN682938										
B6B	KC150754–KC150775, JQ659155, JQ659157–JQ659160										
B6D	KC150776–KC150800										
B7D	KC150801–KC150826										
B7E	KC150827–KC150853										

N, sample size; bp, number of base pairs; H, number of haplotypes; s, number of segregating sites; k, average number of differences between sequences; π , nucleotide diversity \pm standard deviation; Hd, haplotype diversity \pm standard deviation; D, Tajima's D; Fs, Fu's Fs; DE, Arlequin's demographic expansion statistic; SE, Arlequin's spatial expansion statistic. Values in bold correspond to significant values at the 0.05 probability level.

constant rate model. The \log_e marginal likelihood values for the RRW runs were -3855.8 and -3855.2 , and for the SRW runs were -4275.4 and -4274.8 , for path and stepping-stone sampling, respectively. The resulting \log_e Bayes Factor ($=\log_e$ RRW $-\log_e$ SRW) ~ 420 represents very strong evidence in favour of the RRW model (Kass & Raftery 1995). The selected model was used to reconstruct phylogeographic diffusion patterns based on the MCC tree, which shows the origin of the expansion located in northern Río Negro Province (37.716° S, 67.957° W) and dated at ~ 95 ka (see supplemental KML file in Dryad; see Data Accessibility below). However, the uncertainty associated with this ancestral location of *L. darwinii* is large because the 95%-HPD area ranges

between 36.794 and 39.980° S latitude and from 66.594 to 69.285° W longitude.

The reconstruction suggests three phases in the expansion with initial long-distance colonizations starting at ~ 85 ka: (i) northwestward, to the four-corner region among Neuquén, Mendoza, La Pampa, and Río Negro Provinces; and (ii) southeastward to central Río Negro (Fig. 4a). During the second phase, which was already completed around 55 ka, there were three additional long-distance colonizations towards: (iii) southwestern Neuquén; (iv) eastern Río Negro; and (v) northeastern Chubut Province (Fig. 4b). In the last phase, contiguous range expansions took place in all regions colonized in the previous phases up to the

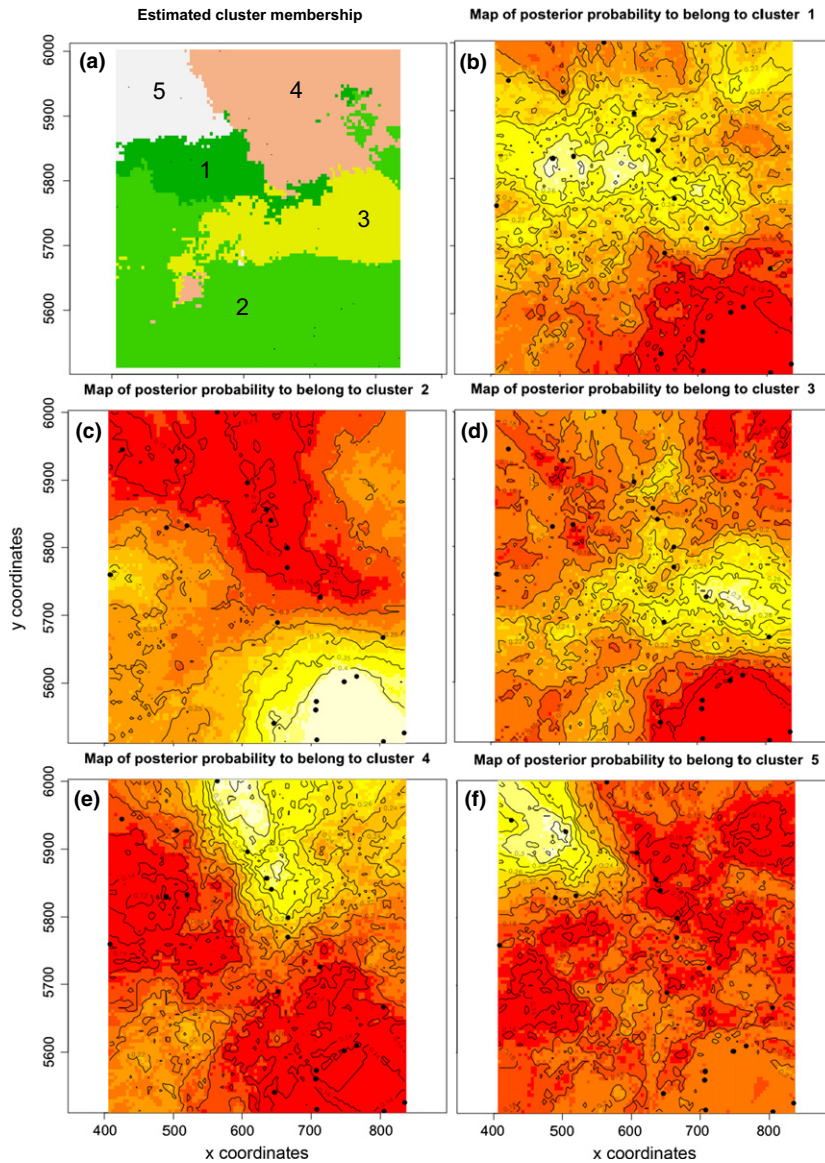


Fig. 2 Maps of cluster membership and posterior probability for each cluster based on Geneland analyses. The estimated cluster membership represents the modal cluster assignment of each pixel (a) and coloured contours show regions of high (light yellow) to low (red) posterior probability of membership to cluster 1 (b), cluster 2 (c), cluster 3 (d), cluster 4 (e) and cluster 5 (f).

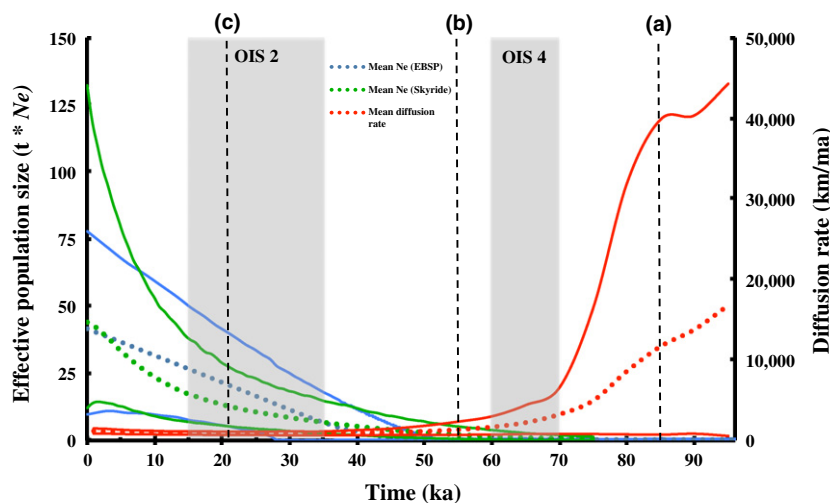


Fig. 3 Variation through time in effective population size based on EBSF analyses (blue), Skyride analyses (green) and in diffusion rate based on RRW analyses (red). Continuous lines above and below the mean parameter values represent 95%-HPD. The vertical dotted lines A, B and C are time slices that correspond with those shown in phylogeographic reconstructions (Fig. 4) and ENM suitability maps (Fig. 5). The two most recent glacial episodes, the Oxygen Isotope Stage 2 (OIS2) and OIS4, are shown as grey bars.

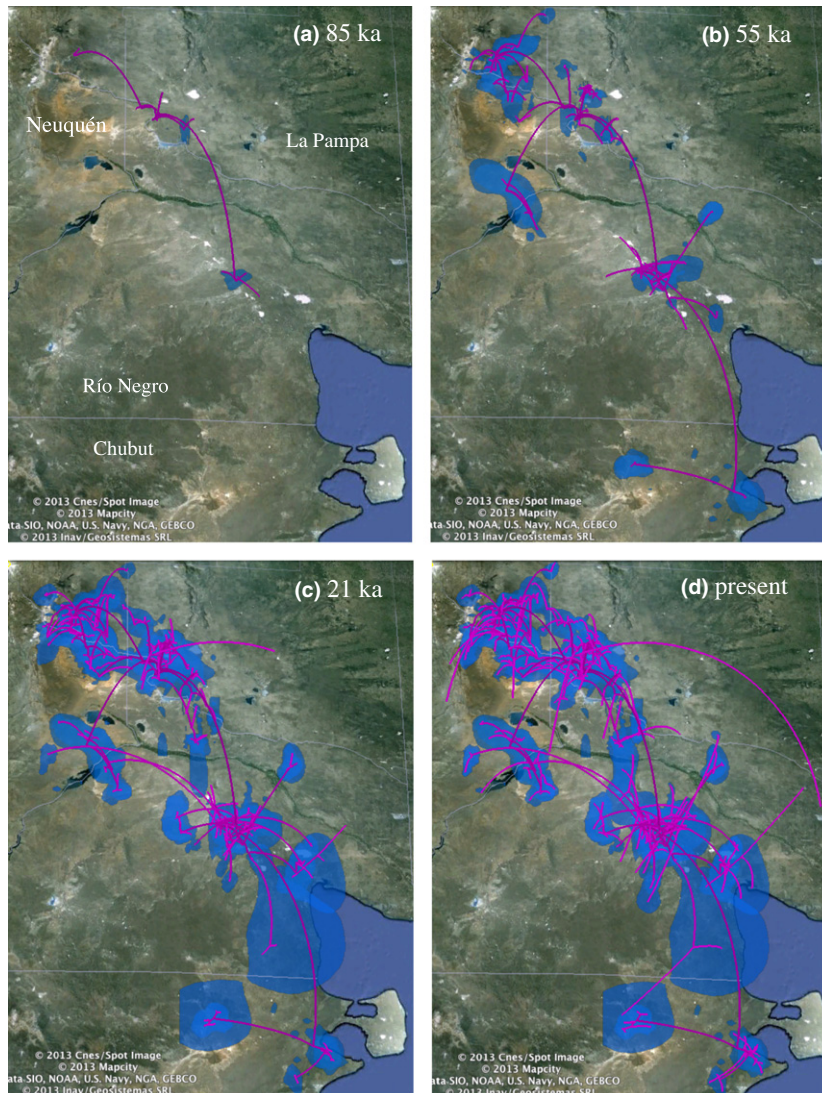


Fig. 4 Spatial projection of the diffusion pattern through time, based on the maximum clade credibility (MCC) tree, estimated with a Bayesian phylogeographic analysis in BEAST (RRW model) at four time slices: (a) 85 ka, (b) 55 ka, (c) 21 ka and (d) present time. The red lines represent the branches of the MCC tree with a gradient from dark to light colour indicating older vs. younger branches. The blue regions represent the 80%-HPD uncertainty in the location of ancestral branches with a gradient between clear and dark representing older vs. younger diffusion events.

LGM (Fig. 4c) and with little subsequent change in the distribution until today (Fig. 4d).

The mean diffusion rate for the analysis was ~ 1108 km per million years ($=1.1$ m/year; 95%-HPD: 0.9–1.4), but there was considerable variation over time. Diffusion rates were as high as ~ 17 m/year at the start of rapid range expansion, but declined to <1 m/year during the phase of demographic expansion since 55 ka (Fig. 3). An animation of the geographic diffusion over time is provided as a KML file in the Dryad digital repository for visualization in Google Earth (see Data Accessibility below).

Palaeo-distribution modelling

The average test AUC for the 10 replicate runs conducted under each regularization multiplier value in the tuning phase varied very little (from 0.954 to 0.961; Table S3, Supporting information). The average test

AUC was 0.961 (standard deviation of 0.013) for the 10 replicate runs that applied the selected regularization multiplier of 4 on the past climate projections, indicating overall high and consistent performance for the models. The clamping grids produced by Maxent have shown that variables were constrained to remain within the range of values in the training data only in the case of 21 ka based on the CCSM model. However, regions of uncertain predictions because of the method of extrapolation are located mostly to the northwest of the suitable areas in this time period and mostly do not affect the core of the prediction range (Fig. S4, Supporting information).

Palaeo-distribution modelling of *L. darwini* at 120 ka suggests a very restricted distribution in the northwestern portion of the present distribution albeit with moderate predictive power (Fig. 5a). Predictions at the LGM based on the MIROC climatic model suggest an even

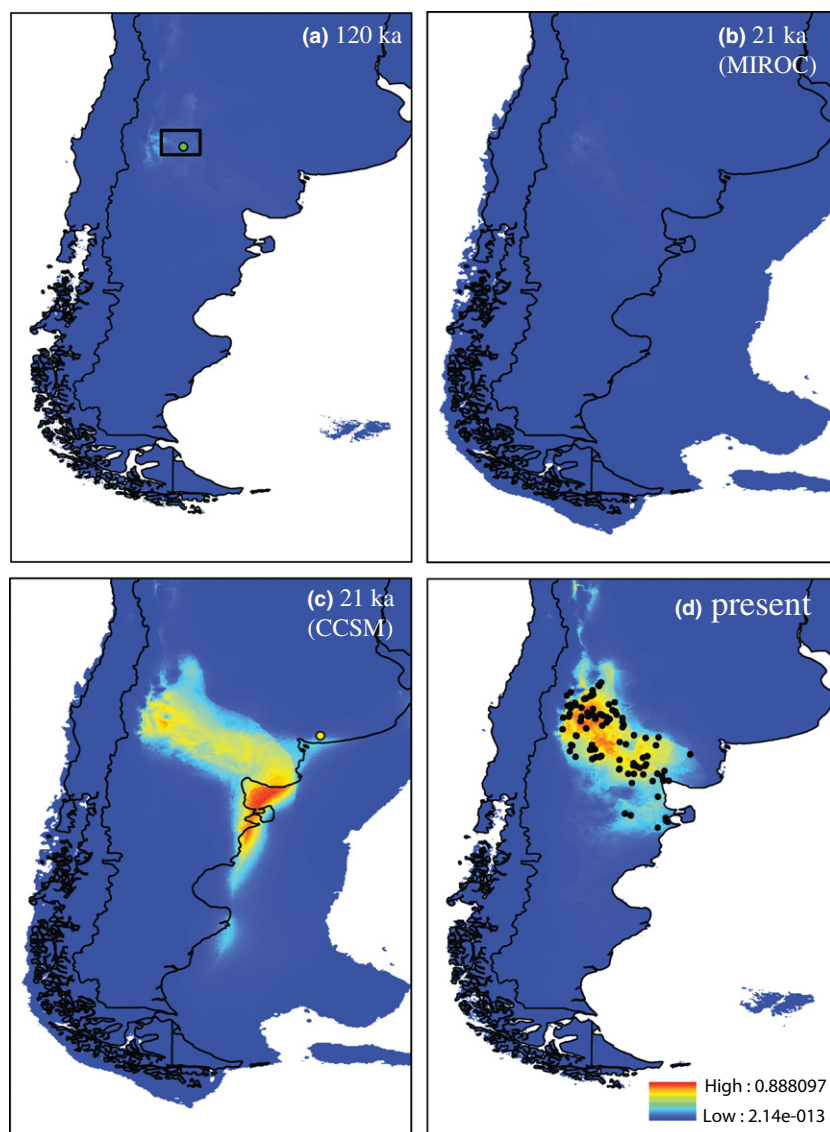


Fig. 5 Modelled ranges of *Liolaemus darwini* across Quaternary climatic fluctuations and present climatic scenarios. Warmer colours of the logistic output format correspond to regions with higher probability of occurrence. Black dots in the logistic output for the current climatic scenario (lower right) represent the occurrence data set implemented in the distribution modelling analyses. Notice the lower probabilities of potential occurrence on the logistic output for 120 ka and 21 ka (MIROC) when compared with the other models. The rectangle over the LIG prediction (120 ka) shows the estimated 95%-HPD area for the origin of the expansion, with the green dot inside the rectangle representing the mean estimate in northern Río Negro Province dated to ~95 ka. The yellow dot in the map of the CCSM model (21 ka) shows the location of the *L. darwini* fossil found in southern Buenos Aires Province and dated to the Pleistocene–Holocene boundary. Note that the dark blue shading represents the extent of land area above sea level, which included the Atlantic continental shelf at the LGM.

smaller range, but those based on the CCSM suggest a southeastern range expansion with high-suitability values (Fig. 5b, c). The ENM at the present time occupies a larger high-suitability area than at the LGM suggesting range expansions towards the north and south, reaching coastal Chubut Province (Fig. 5d).

Results from the analysis of variable contributions (Table S4, Supporting information) indicate that the variables that have higher contribution for the model are bio16 (precipitation of wettest quarter), bio4 (temperature seasonality), bio2 (mean diurnal range) and bio3 (isothermality). The interactive exploration of predictions (Explain tool) indicates that bio3 and bio4 had the most negative contribution to logit of the palaeo-distribution estimates with reduced suitability, which means they are the most limiting variables for the presence of *L. darwini* (Fig. S5, Supporting information).

These results indicate that contrasting between environmental conditions (seasonality and temperature evenness) probably limited *L. darwini* distributions during the different past time periods more than other variables that represent mean environmental conditions.

Discussion

Our study takes advantage of novel developments in Bayesian phylogeographic inference that have been applied successfully to many viral epidemics (Lemey *et al.* 2009b; Raghwani *et al.* 2011; Lam *et al.* 2012; Pybus *et al.* 2012), and a few animal and plant species (Chiari *et al.* 2012; Escobar-Garcia *et al.* 2012) to infer the diffusion dynamics of *L. darwini* over continuous space and time. In addition to reconstructing the range expansion pattern spatially and temporally, this

approach also allowed modelling the demographic expansion using BSP, effectively decoupling the two components of the expansion process, which were confounded in previous approaches. Instead of a past 'averaged' estimate of range distribution, this Bayesian approach provides a dynamic reconstruction that can be directly compared with similar reconstructions based on species-distribution modelling approaches (e.g. ENM). For example, Bayesian diffusion and palaeo-distribution models have been used before to identify potential locations of glacial refugia and range expansions in invertebrates and plants (Smith *et al.* 2011; Marske *et al.* 2012).

The combination of these approaches in our study supports the hypothesis that the demographic history of *L. darwini* was influenced by climatic changes during the late Pleistocene and Holocene. First, both phylogeographic reconstruction and ENM suggest a restricted range of *L. darwini* in northern Río Negro Province in the mid LIG at about 95 ka. Despite substantial uncertainty in the estimated ancestral location of *L. darwini*, which can be enclosed within a square of $\sim 2^\circ$ of latitude and longitude, it occurs within an area of predicted relative suitability for *L. darwini* at 120 ka (see Fig. 5a). The high uncertainty is not unexpected because deep phylogeographic histories (i.e. thousands of years) and the lack of serially sampled data close to the root node hamper the ability of this Bayesian method to infer the geographic location of the root with high precision (Lemey *et al.* 2009a). Second, multiple long-distance colonizations via high diffusion rates increased substantially the distribution of *L. darwini* to the west and southeast between 95 and 55 ka (Fig. 4b) but without population growth (Fig. 3, see time slices A and B). Subsequently, diffusion rates decreased, but population size began its expansion in a region of high suitability until the LGM (Figs 4c and 5b, c; see time slices B and C in Fig. 3).

Both Bayesian phylogeography and ENM models suggest only modest increases in distribution range after the LGM and up to the present time, but a large increase in population size (Figs 3, 4d, 5d). The direction and timing of range expansion coincided between both approaches as they suggest a major southeastern expansion during the termination of the LIG, but basically no further major increases in range size after the LGM. Independent evidence further corroborates the palaeo-distribution model at the LGM because a fossil of *L. darwini* dated to the 'late Pleistocene to early Holocene' has been described from southern Buenos Aires Province as marginally predicted by the CCSM model (Albino 2005)(Fig. 5c). Consequently, fossil evidence supports the CCSM model (instead of the geographically more restricted MIROC prediction), which as discussed above, is consistent with

the range distribution inferred with Bayesian phylogeography at the LGM.

After these phases of range and demographic expansions in the recent past, the current population structure of *L. darwini* seems to be partitioned into five clusters arranged into a north-to-south axis (Fig. 2). This geographic partitioning probably has an historical component because it appears associated with the biogeographic subdivision of the Monte Desert into districts and subdistricts based on vegetation distribution (Roig *et al.* 2009). In addition, the altitude and the current climatic zonation of the Monte Desert, in particular, total annual rainfall (Labraga & Villalba 2009), also show gradients in a north-to-south direction that could partially limit gene flow and account for some of the differentiation between the inferred population clusters within *L. darwini*.

Demographic history during the Late Pleistocene

A number of previous studies of lizards and other vertebrates, terrestrial and aquatic, have revealed demographic expansions that took place during the Late Pleistocene but before the LGM. Based on NCPA and test summary statistics, Yoke *et al.* (2006) found a strong signal of a recent demographic expansion in *Cnemidophorus longicauda*, a lizard that co-occurs with *L. darwini* in the southernmost half of the Monte Desert. More recently, a time-calibrated BSP analysis showed that the population size of *L. gracilis*, another lizard codistributed with *L. darwini*, remained relatively stable until ~ 100 ka when a demographic expansion started without subsequent changes until the Holocene (Olave *et al.* 2011). The onset of this expansion occurred in the middle of the LIG period known as the Oxygen Isotope Stage 5 (OIS 5, ~ 70 – 140 ka) and continued unaffected through the late glaciations of the Pleistocene (OIS 4, 60 – 70 ka; OIS 2, 15 – 35 ka), which were less severe than the penultimate glacial episode (OIS 6, ~ 140 – 180 ka) (Rabassa & Clapperton 1990; Rabassa *et al.* 2005, 2011; Rabassa 2008). In contrast, a clade of the lizard *L. petrophilus*, from northern Patagonia and distributed south of the *L. darwini*'s range, showed demographic stability during the Pleistocene (Fontanella *et al.* 2012).

Very similar patterns of demographic expansions during the LIG with little or no subsequent effect of late glaciations, especially the LGM, have also been detected with BSP analyses in fish and rodents from the Monte Desert and the adjacent Patagonian steppe. Populations of the fish *Percichthys trucha* in this region started a demographic expansion ~ 80 ka, near the end of OIS 5, with a slow down during OIS 2 (Ruzzante *et al.* 2008). In addition, three rodents (*Phyllotis xanthopygus*, *Abrothrix olivaceus* and *Eligmodontia typus*) began

demographic expansions around ~53–63 ka, during the interglacial OIS 3, and continued growing over the last glaciation OIS 2 (Lessa *et al.* 2010; Abud 2011; Da Silva 2011; Riverón 2011). Therefore, the evidence from multiple codistributed taxa in the Monte Desert region suggests congruent demographic stability in the penultimate glaciation, population expansions starting in the LIG and continuing through the LGM, before stabilizing again during the Holocene.

Demographic stability in multiple codistributed species in the Monte Desert is consistent with palaeo-climatic modelling, geological dating and pollen records suggesting colder, drier and windier climates during the cooling phases of the Pleistocene (Clapperton 1993; Rabassa *et al.* 2005; Labraga & Villalba 2009). Pollen records indicate a dominance of typical Andean and Patagonian steppe, herbaceous taxa including Grammineae and Asteracea, in the Monte Desert during the LGM (Labraga & Villalba 2009), suggesting that during cooling phases, ranges shifted towards glacial refugia at lower latitudes. The smaller and more northerly restricted range of *L. darwini* at the onset of the LIG (Figs 4a and 5a) also supports the prediction that these climatic and biogeographic shifts occurred multiple times during the Pleistocene, especially during the longer and more intense penultimate glaciation (i.e. OIS 6). Further, the dominance of typical Monte Desert taxa, such as Poaceae, creosote (*Larrea*) bush and mesquite (*Prosopis*) trees, in warmer and more humid climates (Labraga & Villalba 2009), is consistent with demographic expansions of Monte Desert species during the relatively long and benign LIG. However, the continuous population growth over OIS 4 and 2 seems an unexpected pattern even though the last glaciations were less severe and shorter than the OIS 6 (Rabassa 2008). This apparent conflict between phylogeographic patterns in multiple species and palaeo-climatic predictions for the LGM could be derived from the assumption of simultaneous spatial and demographic expansions when using BSP and test statistics.

While previous studies have not been able to distinguish between spatial and demographic expansions, we could model and estimate relevant parameters of these processes separately with a Bayesian phylogeography method. In fact, the temporal reconstruction of both processes suggests that they did not occur simultaneously in *L. darwini*, which can account for unexpected results seen in previous studies. We found that the demographic expansion in *L. darwini*, even during the last glaciations, took place with modest increases in range distribution (Figs 3 and 4b,c). In contrast, *L. darwini* expanded its range dramatically during the LIG in spite of a relatively stable population size (Figs 3 and 4a,b). Therefore, the counter-intuitive inference

of demographic expansion during unfavourable glacial climates for several Monte Desert taxa could be accounted for a model of population growth without range expansion during the last glaciation. The Bayesian 'phylodynamic' approach applied in our study has given novel insights about the independence in terms of timing between range and demographic expansions and their association with Pleistocene climatic events, which has not been possible with other methods in recent studies.

Diffusion vs. dispersal rates

The historical estimates of diffusion rates from Bayesian phylogeography provide an opportunity for comparison with known field estimates of dispersal distance in other lizards. Even though the dispersal dynamics of stable vs. expanding populations can be very different, we were interested in exploring whether estimates of diffusion rates were realistic in comparison with the observed dispersal abilities of lizards in the field. In *Liolaemus*, most capture-mark-recapture studies have actually quantified the home range of adults (Rocha 1999; Frutos *et al.* 2007; Kacoliris *et al.* 2009), but at least one study in *L. koslowskyi* (Frutos & Belver 2007), a member of the *L. darwini* group, also measured the shift of the home range between two consecutive breeding attempts (i.e. breeding dispersal; Clobert *et al.* 1994; Olsson & Shine 2003). Most individuals of *L. koslowskyi* maintained their home ranges from one year to the next, and males dispersed further than females, an average of ~12 m/year (Frutos & Belver 2007). However, long-term studies of home-range shifts including hatchling/juvenile stages (i.e. natal dispersal), which represent a significant portion of the total dispersal distance during a lifetime, have not been carried out in any species of *Liolaemus*.

Very few studies have quantified natal and/or breeding dispersal in other lizards, including *Sceloporus occidentalis* (Massot *et al.* 2003), *Lacerta agilis* (Olsson *et al.* 1997; Ryberg *et al.* 2004), *Lacerta vivipara* (Le Galliard *et al.* 2003), *Niveoscincus microlepidotus* (Olsson & Shine 2003) and *Xantusia vigilis* (Davis *et al.* 2011). Based on a multi-year study, Davis *et al.* (2011) found larger dispersal distances in juveniles (~50 m) in comparison with adults (~25 m) of *X. vigilis*, a relatively sedentary, desert-dwelling lizard. In *Sceloporus undulatus*, an oviparous semi-arboreal species, the mean dispersal distance of hatchlings was also ~50 m (Massot *et al.* 2003). In *Lacerta agilis*, a viviparous species, mean dispersal distance of adult females measured as the shift in home-range location was ~10 m (Ryberg *et al.* 2004). In general, dispersal distances vary among species co-occurring in the same habitats (Hokit *et al.* 1999), but also within species, due to body size, age, sex

and social interactions (Le Galliard *et al.* 2003; Massot *et al.* 2003; Cote & Clobert 2007). Based on the first phase of range expansion, between 95 and 55 ka, mean diffusion rates were as high as 20 m/year, but during the demographic expansion phase (from 55 ka to the present), rates were <1 m/year. Despite differences in habitat and dispersal abilities, our range of diffusion rates for *L. darwinii* is somewhat lower than the dispersal rates for *Xantusia* and *Sceloporus* (albeit of the same order of magnitude), similar to estimates for *Lacerta*, and especially close to the estimates for the congeneric and closely related *L. koslowskyi*. A good fit between dispersal rates estimated from current field data and phylogeographic estimates is not expected as dispersal dynamics generally differ between historical range expansions and ongoing range shifts (Phillips *et al.* 2008). Therefore, based on the comparisons made above, our Bayesian diffusion rates are probably realistic estimates of the historical dispersal dynamics and provide further support to the temporal scale of the phylogeographic reconstruction.

Relaxed Bayesian diffusion models

Our model comparisons based on Bayes factors suggest that a time-heterogeneous rate of dispersal fits the data better than a standard, time-homogeneous model. While the geographic area occupied by *L. darwinii* consists of a relatively homogenous landscape of Monte Desert habitat, the climatic changes during the late Quaternary suggest that favourable conditions for dispersal varied over time. The log-normal distribution used to model the diffusion rate variation is an extremely relaxed model, but it is possible that more restrictive models of over-dispersion, such as gamma-distributed or autocorrelated rates, can also fit better than the standard Brownian model (Lemey *et al.* 2010). This possibility becomes more plausible given the observed pattern of range expansion, which indicates high diffusion rates at deep branches presumably associated with sudden changes in habitat suitability during demographic history. In particular, autocorrelated models might applied well to the observed diffusion dynamics where some branches in specific parts of the distribution consistently have higher diffusion rates leading to a net south-eastward range expansion. Furthermore, simulations and real data from invasive species suggest that dispersal rates in the expanding front are higher than near the origin of expansion due to strong spatial selection on dispersal ability (Phillips *et al.* 2008), which also makes appealing the implementation of autocorrelated models of diffusion rate in phylogeography. Finally, future developments of Bayesian phylogeographic models are directed towards incorporating spatial heterogeneity via

dispersal constraints derived from ENM (Lemey *et al.* 2010). Indeed, our comparison of phylogeographic reconstructions and ENM further encourages a full integration of these approaches for achieving more biologically realistic models of range expansion and for disentangling this process from demographic expansion.

Acknowledgements

We want to dedicate this research work to our close colleague Dr. Natalia Feltrin, who recently passed away on March 7th, in a car accident. She developed her undergraduate honour's thesis and her PhD project in our research group for the last 8 years; and during that time, we shared many moments, including travels to the US and all around Patagonia, as well as discussions about academia and life in general. Natalia was a happy, optimistic and hard-working person, and a good example of how much one can accomplish with perseverance and team work. Her passing is a profound loss for our group, as we all loved her very much and she will always be with us.

We thank two anonymous reviewers whose constructive comments and suggestions greatly improved an earlier version of the manuscript. We thank B. Noonan for his help in ANL development. A.C. acknowledges a postdoctoral fellowship from the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, Argentina), financial support from the BYU Department of Biology and Graduate Studies Office, as well as support from the Society for Systematic Biologists and the Society for the Study of Amphibians and Reptiles. F.P.W. was supported by fellowships from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/Fulbright no. 15073722-2697/06-8) and Atracción de Jovens Talentos from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq: 374307/2012-1). L.J.A. and M.M. acknowledge several grants under programmes PIP & IBoI (CONICET) and FONCYT PICT (ANPCYT) (Argentina).

Specimens were collected following local regulations and permits issued by Administración de Parques Nacionales, Direcciones Provinciales de Fauna de Río Negro, Chubut, Neuquén, La Pampa, Mendoza and Administración de Areas Protegidas de Neuquén and Mendoza (Argentina). JWS, Jr acknowledges laboratory support from the BYU Jessie Knight Fellowship. Funding for this study was provided by the NSF 'Partnership for International Research and Education' award (OISE 0530267) for collaborative research on Patagonian biodiversity, granted to the following institutions (listed alphabetically): BYU, Centro Nacional Patagónico, Dalhousie University, Darwinion Botanical Institute, George Washington University, Universidad Nacional de Córdoba, Universidad Austral de Chile, Universidad Nacional del Comahue and the Universidad de Concepción.

References

- Abdala CS (2007) Phylogeny of the *boulengeri* group (Iguania: Liolaemidae, *Liolaemus*) based on morphological and molecular characters. *Zootaxa*, **1538**, 21–33.
- Abud C (2011) *Variación Genética y Estructura Filogeográfica del Abrothrix Olivaceus en la Patagonia Argentina y sur Chileno Master's Thesis*. Universidad de la República, PEDECIBA.

- Albino AM (2005) A Late Quaternary lizard assemblage from the southern Pampean region of Argentina. *Journal of Vertebrate Paleontology*, **25**, 185–191.
- Allicock OM, Lemey P, Tatem AJ *et al.* (2012) Phylogeography and population dynamics of dengue viruses in the Americas. *Molecular Biology and Evolution*, **29**, 1533–1543.
- Anderson RP, Gonzales I Jr (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: an implementation with Maxent. *Ecological Modelling*, **222**, 2796–2811.
- Araujo MB, Peterson AT (2012) Uses and misuses of bioclimatic envelope modeling. *Ecology*, **93**, 1527–1539.
- Ayres DL, Darling A, Zwickl DJ *et al.* (2012) BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Systematic Biology*, **61**, 170–173.
- Baele G, Lemey P, Bedford T *et al.* (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*, **29**, 2157–2167.
- Beaumont MA, Panchal M (2008) On the validity of nested clade phylogeographical analysis. *Molecular Ecology*, **17**, 2563–2565.
- Bielejec F, Rambaut A, Suchard MA, Lemey P (2011) SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics*, **27**, 2910–2912.
- Bloomquist EW, Lemey P, Suchard MA (2010) Three roads diverged? Routes to phylogeographic inference. *Trends in Ecology and Evolution*, **25**, 626–632.
- Bouckaert R, Lemey P, Dunn M *et al.* (2012) Mapping the origins and expansion of the Indo-European language family. *Science*, **337**, 957–960.
- Breitman MF, Avila LJ, Sites JW Jr, Morando M (2012) How lizards survived blizzards: phylogeography of the *Liolaemus lineomaculatus* group (Liolaemidae) reveals multiple breaks and refugia in southern Patagonia, and their concordance with other co-distributed taxa. *Molecular Ecology*, **21**, 6068–6085.
- Camargo A, Morando M, Avila LJ, Sites JW Jr (2012) Species delimitation with abc and other coalescent-based methods: a test of accuracy with simulations and an empirical example with lizards of the *Liolaemus darwini* complex (Squamata: Liolaemidae). *Evolution*, **66**, 2834–2849.
- Carnaval AC, Moritz C (2008) Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography*, **35**, 1187–1201.
- Carstens BC, Richards CL (2007) Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution*, **61**, 1439–1454.
- Chan LM, Brown JL, Yoder AD (2011) Integrating statistical genetic and geospatial methods brings new power to phylogeography. *Molecular Phylogenetics and Evolution*, **59**, 523–537.
- Chiari Y, van der Meijden A, Mucedda M *et al.* (2012) Phylogeography of Sardinian cave salamanders (genus *Hydromantes*) is mainly determined by geomorphology. *PLoS ONE*, **7**, e32332.
- Clapperton CM (1993) Nature of environmental changes in South America at the Last Glacial Maximum. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **101**, 189–208.
- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.
- Clobert J, Massot M, Lecomte J *et al.* (1994) Determinants of dispersal behavior: the common lizard as a case study. In: *Lizard Ecology: Historical and Experimental Perspectives* (eds Vitt L & Pianka R), pp. 183–206. Princeton University Press, Princeton.
- Cosacov A, Johnson LA, Paiaro V *et al.* (2012) Precipitation rather than temperature influenced the phylogeography of the endemic shrub *Anarthrophyllum desideratum* in the Patagonian steppe. *Journal of Biogeography*, **40**, 168–182.
- Cote J, Clobert J (2007) Social personalities influence natal dispersal in a lizard. *Proceedings of the Royal Society B*, **274**, 383–390.
- Da Silva CC (2011) *Filogeografía del Género Eligmodontia (Rodentia: Cricetidae) en la Patagonia Argentina* Master's Thesis. Universidad de la República, PEDECIBA.
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.
- Davis AR, Corl A, Surget-Groba Y, Sinervo B (2011) Convergent evolution of kin-based sociality in a lizard. *Proceedings of the Royal Society B*, **278**, 1507–1514.
- Drummond AJ, Ho SY, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *Plos Biology*, **4**, e88.
- Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.
- Elith J, Graham CH (2009) Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, **32**, 66–77.
- Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677–697.
- Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330–342.
- Escobar-Garcia P, Winkler M, Flatscher R *et al.* (2012) Extensive range persistence in peripheral and interior refugia characterizes Pleistocene range dynamics in a widespread Alpine plant species (*Senecio carniolicus*, Asteraceae). *Molecular Ecology*, **21**, 1255–1270.
- Etheridge R (1993) Lizards of the *Liolaemus darwini* complex (Squamata: Iguania: Tropiduridae) in northern Argentina. *Bollettino del Museo Regionale di Scienze Naturali*, **11**, 137–199.
- Etheridge R (2001) A new species of *Liolaemus* (Reptilia: Squamata: Tropiduridae) from Mendoza Province, Argentina. *Cuadernos de Herpetología*, **15**, 3–15.
- 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 HE (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.
- Flot JF (2010) SeqPHASE: a web tool for interconverting phase input/output files and fasta sequence alignments. *Molecular Ecology Resources*, **10**, 162–166.
- Fontanella FM, Feltrin N, Avila LJ, Sites JW Jr, Morando M (2012) Early stages of divergence: phylogeography, climate modeling, and morphological differentiation in the South

- American lizard *Liolaemus petrophilus* (Squamata: Liolaemidae). *Ecology and Evolution*, **2**, 792–808.
- Fraser CI, Nikula R, Ruzzante DE, Waters JM (2012) Poleward bound: biological impacts of Southern Hemisphere glaciation. *Trends in Ecology and Evolution*, **27**, 462–471.
- Frutos N, Belver LC (2007) Dominio vital de *Liolaemus koslowskyi* Etheridge, 1993 (Iguania: Liolaemini) en el noroeste de la provincia de La Rioja, Argentina. *Cuadernos de Herpetología*, **21**, 83–92.
- Frutos N, Camporro LA, Avila LJ (2007) Ámbito de hogar de *Liolaemus melanops* Burmeister, 1888 (Squamata: Liolaemini) en el centro de Chubut, Argentina. *Gayana*, **71**, 142–149.
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitch-hiking, and background selection. *Genetics*, **147**, 915–925.
- Garrick RC, Dyer RJ, Beheregaray LB, Sunnucks P (2008) Babies and bathwater: a comment on the premature obituary for nested clade phylogeographical analysis. *Molecular Ecology*, **17**, 1401–1403.
- Garrick RC, Sunnucks P, Dyer RJ (2010) Nuclear gene phylogeography using PHASE: dealing with unresolved genotypes, lost alleles, and systematic bias in parameter estimation. *BMC Evolutionary Biology*, **10**, 118.
- Guillot G, Estoup A, Mortier F, Cosson JF (2005a) A spatial statistical model for landscape genetics. *Genetics*, **170**, 1261–1280.
- Guillot G, Mortier F, Estoup A (2005b) Geneland: a program for landscape genetics. *Molecular Ecology Notes*, **5**, 712–715.
- Harpending RC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.
- Heled J, Drummond AJ (2008) Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, **8**, 289.
- Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.
- Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B*, **359**, 183–195.
- Hewitt GM, Ibrahim KM (2001) Inferring glacial refugia and historical migrations with molecular phylogenies. In: *Integrating Ecology and Evolution in a Spatial Context, BES Symposium Volume* (eds Silvertown J & Antonovics J), pp. 271–294. Blackwells, Oxford.
- Hickerson MJ, Carstens BC, Cavender-Bares J *et al.* (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, **54**, 291–301.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Ho SY, Shapiro B (2011) Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources*, **11**, 423–434.
- Hokit DG, Stith BM, Branch LC (1999) Effects of landscape structure in Florida scrub: a population perspective. *Ecological Applications*, **9**, 124–134.
- Kacolicis FP, Williams JD, De Arcaute CR, Cassino C (2009) Home range size and overlap in *Liolaemus multimaculatus* (Squamata: Liolaemidae) in Pampean coastal dunes of Argentina. *South American Journal of Herpetology*, **4**, 229–234.
- Kass RE, Raftery AE (1995) Bayes Factors. *Journal of the American Statistical Association*, **90**, 773–795.
- Knowles LL (2008) Why does a method that fails continue to be used? *Evolution*, **62**, 2713–2717.
- Knowles LL (2009) Statistical phylogeography. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 593–612.
- Labraga JC, Villalba R (2009) Climate in the Monte Desert: past trends, present conditions, and future projections. *Journal of Arid Environments*, **73**, 154–163.
- Lam TT, Hon CC, Lemey P *et al.* (2012) Phylodynamics of H5N1 avian influenza virus in Indonesia. *Molecular Ecology*, **21**, 3062–3077.
- Le Galliard J-F, Ferrière R, Clobert J (2003) Mother-offspring interactions affect natal dispersal in a lizard. *Proceedings of the Royal Society B*, **270**, 1163–1169.
- Lemey P, Rambaut A, Drummond AJ, Suchard MA (2009a) Bayesian phylogeography finds its roots. *PLoS Computational Biology*, **5**, e1000520.
- Lemey P, Suchard M, Rambaut A (2009b) Reconstructing the initial global spread of a human influenza pandemic: a Bayesian spatial-temporal model for the global spread of H1N1pdm. *PLoS Currents*, **1**, RRN1031.
- Lemey P, Rambaut A, Welch JJ, Suchard MA (2010) Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology and Evolution*, **27**, 1877–1885.
- Lemmon AR, Lemmon EM (2008) A likelihood framework for estimating phylogeographic history on a continuous landscape. *Systematic Biology*, **57**, 544–561.
- Lessa EP, D'Elia G, Pardiñas UF (2010) Genetic footprints of late Quaternary climate change in the diversity of Patagonian-Fuegian rodents. *Molecular Ecology*, **19**, 3031–3037.
- Marske KA, Leschen RA, Buckley TR (2012) Concerted versus independent evolution and the search for multiple refugia: comparative phylogeography of four forest beetles. *Evolution*, **66**, 1862–1877.
- Martin DP, Lemey P, Lott M *et al.* (2010) RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics*, **26**, 2462–2463.
- Massot M, Huey RB, Tsuji J, van Berkum FH (2003) Genetic, prenatal, and postnatal correlates of dispersal in hatchling lizards (*Sceloporus occidentalis*). *Behavioral Ecology*, **14**, 650–655.
- Minin VN, Bloomquist EW, Suchard MA (2008) Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Molecular Biology and Evolution*, **25**, 1459–1471.
- Morando M, Avila LJ, Baker J, Sites JW Jr (2004) Phylogeny and phylogeography of the *Liolaemus darwini* complex (Squamata: Liolaemidae): Evidence for introgression and incomplete lineage sorting. *Evolution*, **58**, 842–861.
- Noonan BP, Yoder AD (2009) Anonymous nuclear markers for Malagasy plated lizards (*Zonosaurus*). *Molecular Ecology Resources*, **9**, 402–404.
- Olave M, Martinez LE, Avila LJ, Sites JW Jr, Morando M (2011) Evidence of hybridization in the Argentinean lizards *Liolaemus gracilis* and *Liolaemus bibronii* (Iguania: Liolaemini): an integrative approach based on genes and morphology. *Molecular Phylogenetics and Evolution*, **61**, 381–391.
- Olsson M, Shine R (2003) Female-biased natal and breeding dispersal in an alpine lizard, *Niveoscincus microlepidotus*. *Biological Journal of the Linnean Society*, **79**, 277–283.

- Olsson M, Gullberg A, Tegelström H (1997) Determinants of breeding dispersal in the sand lizard, *Lacerta agilis* (Reptilia, Squamata). *Biological Journal of the Linnean Society*, **60**, 243–256.
- Otto-Bliesner BL, Marshall SJ, Overpeck JT *et al.* (2006) Simulating Arctic climate warmth and icefield retreat in the Last Interglaciation. *Science*, **311**, 1751–1753.
- Pearson RG, Thuiller W, Araújo MB *et al.* (2006) Model-based uncertainty in species range prediction. *Journal of Biogeography*, **33**, 1704–1711.
- Petit RJ (2008) The coup de grace for the nested clade phylogeographic analysis? *Molecular Ecology*, **17**, 516–518.
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Phillips BL, Brown GP, Travis JM, Shine R (2008) Reid's paradox revisited: the evolution of dispersal kernels during range expansion. *American Naturalist*, **172**(Suppl 1), S34–S48.
- Pybus OG, Suchard MA, Lemey P *et al.* (2012) Unifying the spatial epidemiology and molecular evolution of emerging epidemics. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 15066–15071.
- Rabassa J (2008) Late Cenozoic glaciations in Patagonia and Tierra del Fuego. In: *Late Cenozoic of Patagonia and Tierra del Fuego* (ed. Rabassa J), pp. 151–204. Elsevier, Amsterdam.
- Rabassa J, Clapperton CM (1990) Quaternary glaciations of the Southern Andes. *Quaternary Science Reviews*, **9**, 153–174.
- Rabassa J, Coronato AM, Salemme M (2005) Chronology of the Late Cenozoic Patagonian glaciations and their correlations with biostratigraphic units of the Pampan region (Argentina). *Journal of South American Earth Sciences*, **20**, 81–103.
- 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.
- Raghwanji J, Rambaut A, Holmes EC *et al.* (2011) Endemic dengue associated with the co-circulation of multiple viral lineages and localized density-dependent transmission. *Plos Pathogens*, **7**, e1002064.
- Rambaut A, Drummond AJ (2009) *Tracer v1.5.0, MCMC Trace Analysis Tool*. Available from <http://beast.bio.ed.ac.uk/Tracer>.
- Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.
- Raxworthy CJ, Ingram CM, Rabibisoa N, Pearson RG (2007) Applications of ecological niche modeling for species delimitation: a review and empirical evaluation using day geckos (*Phelsuma*) from Madagascar. *Systematic Biology*, **56**, 907–923.
- Richards CL, Carstens BC, Knowles LL (2007) Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. *Journal of Biogeography*, **34**, 1833–1845.
- Rissler LJ, Apodaca JJ (2007) Adding more ecology into species delimitation: ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). *Systematic Biology*, **56**, 924–942.
- Riverón S (2011) *Estructura Poblacional e Historia Demográfica del "Pericote Patagónico", Phyllotis Xanthopygus (Rodentia: Sigmodontinae) en Patagonia Argentina Master's Thesis*. Universidad de la República, PEDECIBA.
- Rocha CFD (1999) Home range of the tropidurid lizard *Liolaemus lutzae*: sexual and body size differences. *Revista Brasileira de Biologia*, **59**, 125–130.
- Roig FA, Roig-Juñent S, Corbalán V (2009) Biogeography of the Monte Desert. *Journal of Arid Environments*, **73**, 164–172.
- Ruzzante DE, Walde SJ, Gosse JC *et al.* (2008) Climate control on ancestral population dynamics: insight from Patagonian fish phylogeography. *Molecular Ecology*, **17**, 2234–2244.
- Ryberg K, Olsson M, Wapstra E *et al.* (2004) Offspring-driven local dispersal in female sand lizards (*Lacerta agilis*). *Journal of Evolutionary Biology*, **17**, 1215–1220.
- Sede SM, Nicola MV, Pozner R, Johnson LA (2012) Phylogeography and paleodistribution modelling in the Patagonian steppe: the case of *Mulinum spinosum* (Apiaceae). *Journal of Biogeography*, **39**, 1041–1057.
- Sérsic AN, Cosacov A, Cocucci AA *et al.* (2011) Emerging phylogeographical patterns in plants and terrestrial vertebrates from Patagonia. *Biological Journal of the Linnean Society*, **103**, 475–494.
- Smith CI, Tank S, Godsoe W *et al.* (2011) Comparative phylogeography of a coevolved community: concerted population expansions in Joshua trees and four yucca moths. *PLoS ONE*, **6**, e25628.
- Stephens M, Smith N, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.
- Swofford DL (2002) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4*. Sinauer Associates, Sunderland, Massachusetts.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.
- Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **13**, 789–809.
- Templeton AR (2008) Nested clade analysis: an extensively validated method for strong phylogeographic inference. *Molecular Ecology*, **17**, 1877–1880.
- Thomé MT, Zamudio KR, Giovanelli JG *et al.* (2010) Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. *Molecular Phylogenetics and Evolution*, **55**, 1018–1031.
- Thomson RC, Wang IJ, Johnson JR (2010) Genome-enabled development of DNA markers for ecology, evolution and conservation. *Molecular Ecology*, **19**, 2184–2195.
- Walker RS, Ribeiro LA (2011) Bayesian phylogeography of the Arawak expansion in lowland South America. *Proceedings of the Royal Society B*, **278**, 2562–2567.
- Waltari E, Hijmans RJ, Peterson AT *et al.* (2007) Locating pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS ONE*, **2**, e563.
- Warren DL, Glor RE, Turelli M (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868–2883.
- Werneck FP, Costa GC, Colli GR, Prado D, Sites JW Jr (2011) Revisiting the historical distribution of Seasonally Dry Tropical Forests: new insights based on palaeodistribution modelling and palynological evidence. *Global Ecology and Biogeography*, **20**, 272–288.
- Werneck FP, Gamble T, Colli GR, Rodrigues MT, Sites JW Jr (2012a) Deep diversification and long-term persistence in the South American 'dry diagonal': integrating continent-wide phylogeography and distribution modeling of geckos. *Evolution*, **66**, 3014–3034.

- Werneck FP, Nogueira C, Colli GR, Sites JW Jr, Costa GC (2012b) Climatic stability in the Brazilian Cerrado: implications for biogeographical connection of South American savannas, species richness, and conservation in a biodiversity hotspot. *Journal of Biogeography*, **39**, 1695–1706.
- Yoke MM, Morando M, Avila LJ, Sites JW Jr (2006) Phylogeography and genetic structure in the *Cnemidophorus longicauda* complex (Squamata, Teiidae). *Herpetologica*, **62**, 420–434.
- Yu Y, Harris AJ, He X-J (2011) RASP (Reconstruct Ancestral State in Phylogenies) 2.0 beta. Available at: <http://mnh.scu.edu.cn/soft/blog/RASP>.

A.C. designed the study, performed phylogeographic analyses, prepared figures and wrote the manuscript. F.P.W. performed palaeo-distribution modelling, prepared figures and wrote the manuscript. M.M. collected field samples and sequenced DNA data. J.W.S., Jr provided laboratory space, supplies, equipment and facilitated BYU computing resources. L.J.A. collected field samples, and provided laboratory space and equipment.

Data accessibility

DNA sequences: GenBank Accession numbers (see Table 1). Voucher and locality information: uploaded as online supporting information.

Alignments of DNA sequences and input/outfiles files from analyses PHASE, TCS, GENELAND, BEAST and MAXENT analyses: Dryad entry doi: 10.5061/dryad.m81b5.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Individual variable response curves in Maxent analyses.

Fig. S2 Haplotype network for cyt *b*, maximum credibility tree (MCC) of cyt *b*, and haplotype networks of nuclear ANL.

Fig. S3 Burn-in plot of log-posterior probability and the relative frequency of the number of clusters in Geneland analyses.

Fig. S4 Clamping grids of Maxent analyses.

Fig. S5 Interactive exploration of predictions using the 'Explore' tool in Maxent analyses.

Table S1 Voucher information of specimens used in this study, including locality data, population clusters assignments estimated with Structure using cyt *b*, and haplotype codes for cyt *b* and ANL used in Figures 1 and S2.

Table S2 Primer sequences of ANL used in this study.

Table S3 Summary of the results from the species-specific tuning in Maxent analyses.

Table S4 Analysis of variable contributions in Maxent analyses.