Phylogeographical spatial diffusion analysis reveals the journey of Geoffroy’s cat through the Quaternary glaciations of South America

MARÍA JIMENA GÓMEZ FERNÁNDEZ, ALBERTO FAMELI, JULIO ROJO GÓMEZ, JAVIER A. PEREIRA and PATRICIA MIROL*

Grupo de Genética y Ecología en Conservación y Biodiversidad, División Mastozoología, Museo Argentino de Ciencias Naturales ‘Bernardino Rivadavia’, CONICET. Ángel Gallardo 470, 1425, Buenos Aires, Argentina

Received 10 October 2019; revised 18 December 2019; accepted for publication 19 December 2019

Leopardus geoffroyi is a small feline with a widespread distribution in a broad array of habitats. Here we investigate its evolutionary history to characterize the phylogeographical patterns that led to its present distribution using mitochondrial DNA from 72 individuals collected throughout its entire range. All haplotypes conformed to a monophyletic group, including two clades with a central/marginal disposition that is incongruent to the proposed subspecies. Spatial diffusion analysis showed the origin of the species within the oldest and more diverse central clade. A Bayesian Skyline Plot combined with a dispersal through time plot revealed two population increases at 190 000–170 000 and 45 000–35 000 years ago, the latter period accompanied by an increase in the dispersal rate. Species distribution models showed similar patterns between the present and Last Interglacial Period, and a reduction of high-probability areas during the Last Glacial Maximum (LGM). Molecular evidence confirms L. geoffroyi as a monotypic species whose origin is located in Central Argentina. The last glaciation had little effect on the pattern of distribution of the species: the population and range expansion that started before the LGM, although probably being halted, continued after the glaciation and resulted in the presence of this felid in the far south of Patagonia.

ADDITIONAL KEYWORDS: Leopardus geoffroyi, mitochondrial DNA, Quaternary glaciations, South America, spatial diffusion analysis, species distribution modelling

INTRODUCTION

South America harbours 11 species of wild felids, representing 27% of the global diversity of this group (Kitchener et al., 2017; but see Nascimento & Feijó, 2017 and Ruiz-García et al., 2017 for two other possible species recently described). Eight of these species are included in the ‘Ocelot’ lineage and are currently placed in a single genus, Leopardus, based on the recent radiation, natural hybridization and the close similarity in skull morphology between them (Johnson et al., 2006; Kitchener et al., 2017). The most common felid species of this genus in the southern cone of South America is Geoffroy’s cat Leopardus geoffroyi (D’Orbigny & Gervais, 1844), a small feline (~4–5 kg) with a widespread and continuous distribution from Bolivia and Brazil to southern Patagonia in Argentina and Chile (Pereira et al., 2015). This species lives in a broad array of natural habitats, including scrublands, dry forests, savannas, grasslands, marshlands and steppes of the subtropical and temperate Neotropics, from 0 to 3800 m a.s.l. (Cuyckens et al., 2016). Geoffroy’s cat exhibits behavioural plasticity, exploiting human-dominated habitats, and it is currently categorized as of Least Concern by the IUCN (Pereira et al., 2015).

Based on the morphological studies of Pocock (1940), Cabrera (1958) and Ximenez (1973), four subspecies of Geoffroy’s cat were recognized by Ximenez (1975): L. g. geoffroyi (D’Orbigny & Gervais, 1844) in Patagonia and central Argentina; L. g. salinarum (Thomas, 1903) in north-western Argentina; L. g. paraguai (Pocock, 1940) in southern Brazil, Uruguay, and eastern Argentina and Paraguay; and L. g. euxanthus (Pocock, 1940) in Bolivia and northern Argentina. Johnson et al. (1999)
subsequently revisited the taxonomy of Geoffroy’s cat by assessing patterns of DNA sequence variation using three mitochondrial genes (16S rRNA, ATP8 and NADH-5) and 20 microsatellite loci of 38 captive specimens throughout its distribution. These authors found a lack of geographical structure in the species and estimated that extant lineages of L. geoffroyi diverged 2.0 Mya. Furthermore, Nascimento (2014) reassessed Geoffroy’s cat taxonomic units by using external and craniodental morphology in a sample of 200 specimens housed in museums, detecting a high degree of morphological variation but no evidence of any subspecific division. Kitchener et al. (2017) proposed that three lines of correlated evidence are required to support a given taxonomy: morphological, genetic and biogeographical. Given the knowledge available on L. geoffroyi at that time, they suggested that ‘until further genetic data are available, it is probably best to treat L. geoffroyi as a monotypic species’ (p. 57, Kitchener et al., 2017).

A well-supported taxonomy is critical to determining the conservation status of any group of organisms. The phylogenetic approach is now commonly used to establish species/subspecies limits based on reciprocal monophyly. However, there are numerous examples of incongruence between these inferences and those derived from other sources such as morphology (Knowles & Carstens, 2007). Indeed, gene trees cannot always be equated to species trees. Instead, they should be combined with historical information and the evolutionary context and dynamics of the lineage in order to establish its taxonomy.

One of the major breakpoints in the evolutionary history of many species of plants and animals were the glaciations of the Quaternary. The effects of these glaciations on the evolution of biodiversity have been extensively studied in the Northern Hemisphere (Hewitt, 2000, 2004) but remain neglected for the extensively studied in the Northern Hemisphere glaciations on the evolution of biodiversity have been the glaciations of the Quaternary. The effects of these history of many species of plants and animals were in order to establish its taxonomy.

The phylogenetic approach is now commonly used to characterize the phylogeographical patterns that led to its current distribution. We use a large dataset covering the entire geographical range of the species to perform explicit phylogeographical analyses using spatial diffusion models and project these to understand the incidence of past major climate changes. We expect to clarify Geoffroy’s cat taxonomy not only through a molecular phylogeny but also by including its demographic history.

In this study we investigate the evolutionary history of L. geoffroyi using mitochondrial DNA (mtDNA) to perform explicit phylogeographical analyses using spatial diffusion models and project these to understand the incidence of past major climate changes. We expect to clarify Geoffroy’s cat taxonomy not only through a molecular phylogeny but also by including its demographic history.

MATERIAL AND METHODS

SAMPLING AND GENETIC DATA

We assessed sequence variation in mtDNA for 72 individuals collected throughout the range of L. geoffroyi in Bolivia, Brazil, Uruguay and Argentina (Fig. 1). An area of extensive hybridization between L. geoffroyi and L. guttulus has been identified at their geographical contact zone in southern Brazil (Trigo et al., 2008). To avoid including samples of potential hybrids that may influence our analysis, samples of ‘pure’ Geoffroy’s cats from Brazil were provided by T. Trigo.
Muscle samples were collected from road-killed animals and preserved in 96% ethanol, while blood samples were obtained from wild animals captured for ecological studies and kept in Tris-EDTA. Museum samples consisted of ~0.5 cm² of footpads or skin. We extracted DNA from blood using DNAzol (ThermoFisher Scientific) following the manufacturer instructions, and from other tissues by following a modified SDS–proteinase K–ClNa protocol (Miller et al., 1988), with a previous wash and hydration step with NaCl 100 mM, Tris-HCl 10 mM, EDTA 1 mM (TNE) solution for museum samples.

We amplified 1306 bp comprising a fragment of the NADH gene (NAD, primers ND5-DF1 and ND5-DR1, Trigo et al., 2008) and two fragments of the Control Region (CR, primers Thr-L 5′-GAA

Figure 1. Samples collected for the study. In grey, the distribution range of Leopardus geoffroyi according to Cuyckens et al. (2016). Black squares, Peripheral Clade; white circles, Central Clade. The abbreviations identify the geographical areas, and are included in Supporting Information Table S1.
and aligned with BIOEDIT 7.0.5 (Hall, 1999). We sequenced the CR in two separate fragments to avoid amplification of a repetitive region that produces heteroplasmy. PCR amplifications for all fragments were performed in a final volume of 20 µL, containing DNA template (1–2 µL), 0.4 mM each primer, 0.2 mM dNTP, 1× KAPA2G Buffer A, 1.5 mM MgCl₂, 0.2 mg/mL BSA and 0.03 units of KAPA2G Fast DNA Polymerase (KAPA Biosystems). The cycling for CR consisted of 94 °C for 5 min, followed by 10 cycles of denaturing at 94 °C for 40 s and annealing at 65 °C for 20 s, with 1 °C decrements at every cycle. This was followed by 10 cycles of denaturing at 94 °C for 40 s and annealing at 56 °C for 20 s, and finally 20 cycles of denaturing at 94 °C for 20 s and annealing at 54 °C for 10 s For NAD the cycling protocol consisted of 94 °C for 5 min, followed by 12 cycles of denaturing at 94 °C for 40 s, annealing at 58 °C for 20 s with 1 °C decrements from 58 to 46 °C at every cycle, and 30 cycles of denaturing at 94 °C for 40 s and annealing at 48 °C for 30 s Negative controls were included in all PCR runs to check for contamination.

We visualized PCR products on 1.2% agarose gels and purified the amplicons with the enzymes Exonuclease I and Shrimp Alkaline Phosphatase (ThermoFisher Scientific). Both strands of each product were sequenced using an ABI3100 capillary sequencer (MACROGEN Inc.), with the primers used for PCR. Finally, the chromatograms were visualized and aligned with BIOEDIT 7.0.5 (Hall, 1999).

GENETIC VARIABILITY AND PHYLOGENETIC INFERENCE

We tested for incongruence of substitution rates of the combined dataset with the partition homogeneity test (Farris et al., 1995), as implemented in PAUP* (Swofford, 1998). The result of this test showed that the sequence for the two loci were congruent (P = 0.30), and thus a concatenated fragment of mtDNA adding 1306 bp was used. We calculated several descriptors of genetic variability, including the number of variable sites (S), number of haplotypes (H), nucleotide diversity (π) and haplotype diversity (Hd) using ARLEQUIN 3.5 (Excoffier & Lischer, 2010).

We performed a Genetic Landscape Shape Interpolation analysis using the software ALLELES IN SPACE (Miller, 2005) to assess the existence of barriers to gene flow and areas of high genetic richness occurring in our dataset. We created two different connectivity networks: the pairwise location-based network and the Delaunay triangulation-based network. We chose a grid of 60 cells (east–west) × 120 cells (north–south) to represent the study area. The output was overlapped on a map of the area using DIVA-GIS v.7.5.0 (http://www.diva-gis.org/).

A Bayesian phylogenetic analysis was performed using BEAST 1.8.2 (Drummond et al., 2012) for the concatenated dataset (CR + NAD). We ran the analysis using a relaxed uncorrelated molecular clock and HKY+G model of nucleotide substitution, chosen after running jModeltest (Posada, 2008). We used three calibration points obtained from Johnson et al. (2006): lineage ocellot of 2.91 Mya (95% highest posterior density (HPD): 2.02–4.25 Mya); Leopardus pardalis–Leopardus weidii, 1.58 Mya (95% HPD: 1.01–1.24 Mya) and Leopardus geoffroyi–Leopardus guigna, 0.74 Mya (95% HPD: 0.41–1.21 Mya), with normal prior distribution and a Yule tree prior. The analysis was performed for 10⁶ iterations and parameters were sampled every 10 000 iterations. Acceptable convergence to stationarity was checked using the program TRACER 1.7.0 (Rambaut et al., 2018). A maximum credibility tree was summarized using TreeAnnotator 1.8.2 (Drummond et al., 2012). The sequences of L. pardalis (Accession No. NC_028315.1) and L. guigna (Accession No. NC_028321.1) were included as outgroups. We also constructed maximum likelihood and parsimony trees with PAUP v.4.0 (Swofford, 1998). The best-fitting model of nucleotide substitution was selected with jModeltest as before. Maximum parsimony and maximum likelihood heuristic searches were conducted with 1000 random sequence addition replicates. Relationships among haplotypes were studied with a network constructed with TCS (Clement et al., 2000) implemented in the software PopArt (http://popart.otago.ac.nz).

Finally, we calculated the fixation index ΦST using analysis of molecular variance (AMOVA) (Excoffier et al., 1992) implemented in ARLEQUIN to evaluate the differentiation among groups detected with the previous analyses. The correlation between genetic and geographical distances was assessed with a Mantel test (Mantel, 1967) with 1000 permutations using the program ALLELES IN SPACE (Miller, 2005).

DEMографИЧЕСКИЙ ИСТОРИЯ

To analyse the demographic history of L. geoffroyi, we used two complementary approaches. First, and in order to estimate changes in population size over time, we constructed Bayesian skyline plots (BSPs) as implemented in BEAST 1.8.2 (Drummond et al., 2012). The runs were performed for 5 × 10⁷ iterations and parameters were sampled every 5000 iterations. We selected the most adequate nucleotide substitution model with jModeltest (Posada, 2008) and applied a relaxed molecular clock model with a mean rate of 0.0354 substitution/sites/Myr (SD = 0.000396) following the estimates of Trigo et al. (2013). We
ran two independent chains with identical settings, further combined into a single string, and discarding the 10% burn-in using the program LOG COMBINER 1.6.2 (Drummond & Rambaut, 2007). The results were visually inspected with the program TRACER 1.7.0 (Rambaut et al., 2018). The effective sample size was used to assess the convergence of the strings with a threshold value of 200. In addition, we performed model comparisons (Constant size vs. BSP) within the Bayesian phylogenetic and phylogeographical analysis framework using the stepping-stone (SS) and path sampling (PS) marginal likelihood estimators (Xie et al., 2011; Baele et al., 2012) in BEAST.

Second, we used ARLEQUIN 3.0 (Excoffier et al., 2005) to test for sudden demographic expansion using mismatch distribution analyses (Rogers & Harpending, 1992). We employed the sum of squared deviations (SSD) statistic and the raggedness index (Rg) defined by Harpending (1994) to test the goodness of fit of the observed mismatch distribution to that expected under the sudden expansion model. Additionally, to test for deviations from neutrality (as would be expected under demographic expansion) we used Fu’s F (Fu, 1997) and R2 (Ramos-Onsins & Rozas, 2002). The significance of these statistics was tested with DNAsp v.5 (Rozas & Rozas, 1995).

**Ancestral distribution estimates**

We used the continuous diffusion model implemented in BEAST (Lemey et al., 2010) to estimate the patterns of ancestral distribution of the group. We ran the analysis with a smaller dataset, selecting only one individual representing each haplotype found in a given locality. We performed runs for each dispersal process (Homogeneous Brownian, Gamma, Cauchy and Lognormal) and compared them through a Marginal Likelihood Estimation analysis. The most appropriate model (Gamma RRW model) for the dataset was finally used for the phylogeographical analysis, with the same calibration points as before. The maximum clade credibility tree was obtained with TreeAnnotator 1.8.2 (Drummond et al., 2012) and used as input in SPREAD3 (Bielejec et al., 2016) to generate time-calibrated reconstruction of the diffusion process. Finally, we 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.

**Species distribution models**

To study the probable changes in the distribution of *L. geoffroyi* through time, we created SDMs using the maximum entropy method in MaxENT v.3.4.0 (Phillips et al., 2017). Climatic variables for current and past conditions were downloaded from WorldClim (http://www.worldclim.org). We ran the distribution model for current conditions, and then projected onto palaeoclimatic models for the Last Interglacial period (LIG; 120 000–140 000 years BP) and for the LGM (22 000 years BP). Current and LIG models were run using the finer spatial resolution of 30 arc seconds, while the projections for the LGM were run at 2.5 arc minutes resolution. Past climatic models were taken from the Model for Interdisciplinary Research On Climate (MIROC) for the LGM and Otto-Bliesner et al. (2006) for the LGI. We also included three topographic variables (altitude, slope and aspect), derived from a digital elevation model produced by the NASA—Shuttle Radar Topographic Mission.

In addition to the 72 samples used in this work, we included 266 presence records reported by Cuyckens et al. (2016). We tested for correlation between environmental variables using Pearson’s correlation coefficient (*r*) implemented in ENM Tools (Warren et al., 2010; Warren & Seifert, 2011). For model calibration, we varied the regularization multiplier values following Anderson & Gonzales (2011). This parameter penalizes complex models that include many features and compel MaxENT to concentrate on only those with the highest explanatory capability (Phillips et al., 2006). We tested 10 regularization values from 0.5 to 5. For each run we split the data as 75% (training data) and 25% (test data), for 10 replicates (run type = bootstrap) and 2500 maximum iterations. We evaluated the performance of the different values using Akaike’s information criterion (AIC) using ENM Tools (Warren et al., 2010; Warren & Seifert, 2011).

Based on the results of the calibration, we ran 10 different models including the one with all the variables and nine more combining non-correlated (*r* < 0.8) variables with >1% contribution to the model. We then evaluated model performance with the Model Selection tool in ENM Tools. Finally, we carried out model projections using the selected environmental variables and regularization value through the MESS analysis implemented in MaxENT.

**RESULTS**

**Genetic variability**

We sequenced 1306 bp in 72 individuals and identified 56 haplotypes (Supporting Information, Table S1). Only 10 haplotypes were shared between individuals. There were 85 variable sites, 47 of which were phylogenetically informative. Nucleotide diversity (*π*) was 0.09, haplontypic diversity 0.982 and mean pairwise difference 8.920.
Figure 2 shows the results of the Genetic Landscape Shape Interpolation according to the Delaunay triangulation (the pairwise location-based network was essentially the same). The map reveals higher genetic diversity at the centre of the distribution, with some small areas of low diversity embedded within it. The peripheral part of the distribution, especially the region in Brazil and Uruguay, shows lower diversity than the centre.

Phylogenetic Analyses
The phylogenetic trees constructed using Bayesian inference, maximum likelihood and parsimony were consistent in their main topology. Therefore, we only show the Bayesian tree constructed with BEAST (Fig. 3). All haplotypes corresponding to *L. geoffroyi* were grouped into a monophyletic cluster with maximum posterior probability. As suggested by other studies (Johnson et al., 1999, 2006), *L. guigna* is confirmed as the sister species of Geoffroy’s cat. The estimated time to the most recent common ancestor (TMRCA) for the species is 0.78 Mya (95% HPD 0.47–1.14 Mya).

The topology of the tree indicates two moderately supported clades (0.5 and 0.7 posterior probability).

The estimated TMRCAs for the two clades were 0.64 Mya (95% HPD = 0.35–0.98 Mya) and 0.59 Mya (95% HPD = 0.32–0.94 Mya) respectively. A close examination to the geographical origin of samples within each clade (Fig. 1) shows no evident geographical pattern or relationship with the four subspecies described for *L. geoffroyi*. However, again there appears to be a central/marginal pattern, as indicated by the variable levels of genetic diversity analysed earlier: the central clade (C) is slightly older than the peripheral one (P). This younger clade includes all samples from Bolivia, Brazil and Uruguay, while the older clade comprises only Argentinian samples.

The haplotypic network (Fig. 4) showed similar geographical structure to the one observed in the phylogenetic trees. In this case, each clade (Central and Peripheral) appears subdivided, with no evident geographical correlation. When comparing both C and P clades using AMOVA, there were significant differences ($\Phi_{ST} = 0.397$, $P < 0.001$). Mantel tests within each clade did not show any correlation between geographical and genetic distances.

Demographic History
The mismatch distribution analysis for the entire dataset resulted in a unimodal distribution with mode around 13 pairwise differences. Fu’s $F$ and $R^2$ were also significant, indicating expansion [$F = -0.364$, $P < 0.005$; $R^2 = 0.097$ (95% confidence interval 0.047–0.159), $P < 0.05$]. The sum of squared deviation (SSD) and raggedness index (Rg) statistics showed demographic expansion (SSD = 0.00047, $P = 0.99$; Rg = 0.0023, $P = 0.97$) as well as spatial expansion (SSD = 0.0017, $P = 0.61$; Rg = 0.0023, $P = 0.97$). Each individual clade showed also a demographic history of expansion. The TMRCA was 0.44 Mya for the Central clade and 0.37 Mya for the Peripheral clade.

The BSP combined with the plot for dispersal rate through time (Fig. 5) revealed two periods of population growth in effective size. The first one would have begun ~190–170 kya, producing a 10-fold increase in effective size. After a period of relative stability, a second growth accompanied by an increase of dispersal rate, i.e. population and geographical expansions, occurred at ~45–35 kya, also resulting in a 10-fold increase in effective size. Thus, the final mean population size was 100 times higher than before the two expansions.

Diffusion Analysis
The spatial diffusion rate for the group was 2058 km/Myr (95% HPD = 1524–2634 km/Myr), roughly 2 m/year. The model inferred that the expansion of the species would have begun at ~500 kya, near the time of the previously suggested origin of the two clades found...
in the phylogenetic analysis. The model also located its geographical origin in Central Argentina (Fig. 6A), from where the lineage diversification started towards the north and the south (Fig. 6B). The spatio-temporal reconstruction of the diversification indicates further expansions in the same directions, along with expansions to the east and west between 300 and 150 kya, colonizing Paraguay, Bolivia, Uruguay, Brazil and the Andes (Fig. 6C, D). In particular, the LIG expansions (between 150 and 120 kya) showed the advance towards the south of Brazil and Bolivia (Fig. 6E). The last and highest lineage diversification, during the LGM (~20 kya), arrived in the previously unoccupied Patagonia in Argentina and Chile (Fig. 6F).

Species distribution model
Based on Pearson’s correlations and contributions to the full model, we selected eight of the 19 bioclimatic variables producing three models with different combinations of non-correlated variables (Supporting Information, Table S2). Area under the curve (AUC) values between models were similar and >0.8. The most accurate model according to the AIC is shown in Figure 7 and Table S2. The bioclimatic variables that contributed the most to the selected model were Temperature Seasonality (Bio04), followed by Precipitation of Wettest Month (Bio13), Min. Temperature of Coldest Month (Bio06), Precipitation of Driest Quarter (Bio17) and Precipitation Seasonality (Bio15).

The palaeo-distribution obtained during the LIG (120 kya) shows a very sharp limit among good/very good and poor niche conditions for the presence of the species, following an especially dry subtropical region of the Andes known as the ‘dry diagonal’. South of this diagonal, the probability of occurrence of the species is very low. Conversely, a high probability of occurrence is predicted north of the diagonal, where the diffusion analysis inferred the origin and initial expansion of the lineage.

The model for the LGM (21 kya) showed a reduction of high-probability areas, as well as total presence areas, which are restricted to the north of 37°S. The probability of presence to the west, where the elevation is increasingly higher, is similar to that shown by the LIG model.
The distribution model for the present climatic conditions showed a similar distribution to the LIG model, with some differences. First, and importantly, the area south to the dry diagonal now has higher probabilities of occurrence, especially the coastal zone of Patagonia. Second, the areas with the highest probabilities of presence are located at the eastern and northern edges of the distribution range, including coastal areas of eastern Argentina, Uruguay, South Brazil and Paraguay. Third, favourable areas diminished in the Gran Chaco and Cerrado in Bolivia.

**DISCUSSION**

**HOW MANY, IF ANY, SUBSPECIES?**

The Ocelot lineage is one of the two lineages of felids living in South America. It diversified between 8 and 2.9 Mya, after the faunal exchange between North and South America facilitated by the emergence of the Panamanian land bridge (Johnson et al., 2006). The lineage includes eight *Leopardus* species: *L. pardalis*, *L. weidii*, *L. jacobita*, *L. colocola*, *L. tigrinus*, *L. guigna*, *L. guttulus* and *L. geoffroyi*. The taxonomy and genetics of most of these species have been studied in depth; however, that is not the case for *L. geoffroyi*. Therefore, we present here the first exhaustive (regarding sampling and geography) study of the genetic variability, genetic structure and intraspecific taxonomic status of *L. geoffroyi*. We also show results that help to elucidate the demographic history of the species.

Based on samples collected in the wild from all four putative subspecies of *L. geoffroyi*, we found no phylogenetic evidence supporting any subdivision of the species. Furthermore, we detected two clades congruent with the spatial diffusion model we obtained, but completely dissimilar to the geographical range of the four described subspecies. Given that the morphological data (Nascimento, 2014) also do not appear to support the subspecies distinction, we are proposing to accept *L. geoffroyi* as a monotypic

![Figure 4. Haplotypic network based on mtDNA in Leopardus geoffroyi. Haplotypes are represented by circles, the size of which is proportional to their frequency. Lines on branches indicate the number of mutational steps separating haplotypes; colours indicate previously proposed subspecies. Lines encircle Central and Peripheral Clades. Names of the geographic areas, as in Figure 1, are given in parentheses.](https://academic.oup.com/biolinnean/article-abstract/129/3/603/5731807)
species. The only previous phylogeographical study on the species, performed based on samples from captive individuals (whose geographical origin certainty depends on the zoo’s records) also supported this taxonomic arrangement (Johnson et al., 1999).

**GENETIC VARIABILITY AND DEMOGRAPHIC HISTORY**

The criteria used by the IUCN to categorize species in the Red List include small and fragmented populations, suffering or having suffered population decline. Translating these criteria to genetic terms means considering the effects of genetic drift and therefore low genetic diversity. The Ocelot lineage comprises species in all four IUCN Red List categories, with haplotype diversity between 0.46 and 0.98, and nucleotide diversity varying by more than one order of magnitude, from 0.0030 to 0.098 (see references in the next paragraph).

The most threatened species in the lineage, *L. jacobita*, has the lowest level of variation ($\pi = 0.0030$, $H = 0.46$, Ruiz García et al., 2013). We found that *L. geoffroyi*, categorized as LC according to the IUCN, has a nucleotide diversity of 0.090 and a haplotype diversity of 0.98. These values are similar to those found in the margay, *L. weidii* (Eizirik et al., 1998), species NT according to the IUCN Red List, and half of that of the kodkod (*L. guigna*: Johnson et al., 1999; Napolitano et al., 2012), its sister species and categorized as Vulnerable. The remaining species of the group have intermediate levels of variability, regardless of their conservation status. Johnson et al. (1999) reported for Geoffroy’s cat a nucleotide diversity of 0.027, three times lower than the one found here. We believe that this difference lies in the fact that all samples used in Johnson’s study came from captive individuals, where inbreeding is quite common (Frankham et al., 2010).

The low magnitude of the nucleotide diversity associated with high haplotype diversity indicates a pattern of recent and rapid expansion, consisting of many haplotypes differentiated by only a few mutations. This pattern is also found in other species of the group such as *L. colacola* (Cossios et al., 2009; Santos et al., 2018) and *L. guigna* (Napolitano et al., 2014). The geographical distribution of this genetic variability is not clinal, as suggested by morphology (Nascimento, 2014), but rather is arranged in a central–peripheral pattern where the highest variability is located at the centre (Fig. 2). Theory predicts that the geographical origin of a lineage will show the highest genetic diversity, which would be reduced as its expansion occurs, being lowest at the edges of the distributional range (Eckert et al., 2008). Therefore, the geographical distribution of genetic variability found in our work suggests an origin of the species in Central Argentina, a result also shown by the phylogenetic reconstruction (Fig. 3) and the spatial diffusion model (Fig. 6).
The haplotype network does not show any evident expansion: there are no widespread haplotypes (the most frequent one was found seven times, most of the haplotypes being represented only once), and most of the branches connecting them have between two and five nucleotide differences. Despite the network resembling a pattern of population stability, the mismatch analysis, Fu’s $F$ and $R^2$ showed signals of demographic expansion. However, these analyses are shown to be effective only in cases of recent expansions and they fail to distinguish between demographic (It is now possible to discriminate between the two situations using Bayesian phylogeographical inference and SDMs (Lemey et al., 2009, 2010; Camargo et al., 2013; Baranzelli et al., 2017). We are aware of the limitations of using only mtDNA to reconstruct demographic history, but different lines of evidence produced concordant results, as detailed below.

We found that the separation between $L. Geoffroyi$ and its sister species, $L. guigna$, occurred ~783 000 years ago (Fig. 3), similar to the divergence time estimated with mitogenomics for both species (0.68–0.52 Mya: Ruiz-García et al., 2017). The origin of the species is located in Central Argentina (Fig. 6A). From this geographical origin, the Central Clade originated ~640 000 years ago and diversified 200 000 years later (Fig. 6B). The Peripheral Clade arose from the central part of the distribution 590 000 years ago (Fig. 6B), starting its diversification 370 000 years ago (Fig. 6C). In Figure 6 we found a possible explanation for the existence of these two non-clinal clades, since the diffusion model shows long branches leading to the north and the south – the Peripheral Clade – as well as more nodes and shorter branches around the origin representing the Central Clade. The diversification of both clades seems to continue to the present.
LATE PLEISTOCENE DEMOGRAPHIC HISTORY AND PALAEO-DISTRIBUTION MODELLING

Climate change produced shifts of the different biomes. Sea level fell during glacial periods, and in the southern part of South America, the submarine platform became partially exposed, from 10 to 140 m during full glacial periods. There was an increase in extreme temperatures as well as reduction in precipitation, mainly in the Pampas and Patagonia, from approximately 34°S to the south. These glacial episodes led to the formation of sand dunes in these regions. Regarding vegetation cover, displacement of the forest led to shrubby steppe environments. In addition, the grassland was reduced, replaced by the monte and steppe ecosystems (Rabassa et al., 2005).

Our results show the origin of *L. geoffroyi* in central Argentina, the current Espinal biome (thorny deciduous shrubland forests), during the Pleistocene around 758 000 years ago. The species shows very high ecological plasticity, living in a broad array of natural habitat types, including scrublands, dry forests, savannas, grasslands, marshlands and steppes of the subtropical and temperate Neotropics (Cuyckens et al., 2016). This ecological plasticity constitutes an obvious advantage during climate change. Although the existence of stable areas, i.e. refugia, during the glaciations in South America has been described only for Patagonia (Sersic et al., 2011), it is safe to assume that in more temperate zones, like the Espinal, the Pampas and the Monte biomes, they must have existed as well. Moreover, for species associated with open vegetation, such as Geoffroy’s cat, previous study indicates that during glacial periods they could have shrunk, maintained or even expanded their geographical ranges (Turchetto-Zolet et al., 2013).

According to the Skyline plot analyses (Fig. 5), a high rate of increase in effective number without range expansion occurred for *L. geoffroyi* before the LIG, starting ~170 000 years ago. Bayesian diffusion analysis (Fig. 6) shows that the lineage then extended its range northwards, reaching Paraguay, and southwards to the Monte-Desert and Patagonian Steppe ecoregions. This range expansion is shown as a slightly but continuous increase in the curve of the Skyline plot. After the LGM, there is a period of stability in effective population size ($N_e$).

Strong evidence of population growth combined with range expansion through time was obtained during the LIG approaching the LGM. Spatio-temporal reconstruction indicated that such an expansion would have started ~40 000 years ago, during the LIG, and lasted until the beginning of the LGM. The resulting $N_e$ was ~100 times higher after the two expansions. Diffusion analysis shows that around this time, the species extended its range further north to Bolivia and south to coastal Patagonia. At this time, the species crosses the Dry Diagonal, a transition zone between two circulation systems found in South America that influences the climate. The Dry Diagonal has been...
shown to be a clear divisor between ‘good’ and ‘poor’ niches for many species (Bruniard, 1982). However, it does not seem to have acted as a barrier either for Geoffroy’s cat or for L. colocola (Santos et al., 2018), which show central/peripheral and west/east differentiation, respectively.

Leopardus colocola also suffered two episodes of population expansion at approximately the same geological time as Geoffroy’s cat (Santos et al., 2018), 200 and 60–50 kya, the latter period coinciding with the expansion of the savanna in Brazil. Leopardus geoffroyi populations expanded 180 000 and 40 000 years ago, being the late Pleistocene expansion coincident with the advance of the Monte vegetation to the north (Baranzelli et al., 2017). However, a very important difference exists between these two felid species: while L. colocola has a genetic structure indicating four subspecies with distributions historically linked to glacial episodes, we found that L. geoffroyi constitutes a monotypic species distributed in many different biomes. Moreover, we found that late glaciations had little or no effect on the demographic expansions that occurred during the LIG, as shown in other vertebrates such as lizards (Olave et al., 2011; Camargo et al., 2013), fish (Ruzzante et al., 2008) and rodents (Lessa et al., 2010),

The SDM based on climatic variables shows that during the LIG, a large portion of South America was suitable for the species. After the LGM the distribution shifted southwards, and the whole region of Patagonia became habitat with different degrees of suitability for the species. The range expansion to the south that started before the LGM (Figs 5, 6E) probably continued after the glaciation, without significant retreat, and resulted in the presence of this felid in the far south of Patagonia. Thus, our comparisons using independent approaches such as SDM and Bayesian reconstruction of the diffusion process agreed in the spatial direction of the range expansion.

This study is the first to include samples of Geoffroy’s cat collected in the field from its entire distribution. We show that it is a monotypic species with one of the highest levels of genetic variability within the Ocelot lineage and demonstrate through different lines of evidence its evolutionary history of demographic and spatial expansions. These results, in agreement with other phylogeographical studies of co-distributed species, will fill an important gap in the knowledge of the effect of glacial cycles in the Southern Hemisphere. They will also contribute to the conservation of this species in highlighting the existence of a unique evolutionary significant unit.

ACKNOWLEDGEMENTS

We thank T. C. Trigo, A. Abba, G. Aprile, M. Carrera, S. de Bustos, A. Di Giacomo, N. Fracassi, J. Liotta, M. Monteverde, M. Morales, R. Palacios, H. Pastore, A. Paviolo, N. Pelegrin, V. Raimondi, M. Superina and A. Ximenez for providing samples. We also thank Museo Argentino de Ciencias Naturales ‘Bernardino Rivadavia’ (D. Flores), Colección Boliviana de Fauna (I. Moya), Museo de La Plata (M. Merino), Colección de Mamíferos de la Fundación Azara (Y. Davies) and Grupo de Ecología Comportamental de Mamíferos (E. Casanave and M. Lucherini) for allowing us access to samples in their collections. We also thank four anonymous reviewers for their insightful comments. Funding was provided by FONCyT (PICT 2016-4087), CONICET (PICT 114-201101-00050), Amersfoort Zoo, Feline Conservation Federation and Le Parc des Felines.

REFERENCES


© 2020 The Linnean Society of London, Biological Journal of the Linnean Society, 2020, 129, 603–617
SUPPORTING INFORMATION

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

Table S1: Geographical location and haplotype of the samples included in this study.
Table S2: Percentage contribution of climatic variables to the past and current potential distributions in each one of the three models with the highest AUC values.

SHARED DATA

Sequences are available at GenBank (https://www.ncbi.nlm.nih.gov/genbank/), under Accession Numbers MN793073- MN793128 (NADH) and MN849194 - MN849300 (Control Region).