Influences of landscape characteristics and historical barriers on the population genetic structure in the endangered sand-dune subterranean rodent *Ctenomys australis*

Ailin Austrich1 · Matías S. Mora1 · Fernando J. Mapelli2 · Alberto Fameli2 · Marcelo J. Kittlein1

Received: 26 February 2019 / Accepted: 20 May 2020 / Published online: 25 May 2020
© Springer Nature Switzerland AG 2020

Abstract
Understanding the processes and patterns of local adaptation and migration involves an exhaustive knowledge of how landscape features and population distances shape the genetic variation at the geographical level. *Ctenomys australis* is an endangered subterranean rodent characterized by having a restricted geographic range immerse in a highly fragmented sand dune landscape in the Southeast of Buenos Aires province, Argentina. We use 13 microsatellite loci in a total of 194 individuals from 13 sampling sites to assess the dispersal patterns and population structure in the complete geographic range of this endemic species. Our analyses show that populations are highly structured with low rates of gene flow among them. Genetic differentiation among sampling sites was consistent with an isolation by distance pattern, however, an important fraction of the population differentiation was explained by natural barriers such as rivers and streams. Although the individuals were sampled at locations distanced from each other, we also use some landscape genetics approaches to evaluate the effects of landscape configuration on the genetic connectivity among populations. These analyses showed that the sand dune habitat availability (the most suitable habitat for the occupation of the species), was one of the main factors that explained the differentiation patterns of the different sampling sites located on both sides of the Quequén Salado River. Finally, habitat availability was directly associated with the width of the sand dune landscape in the Southeast of Buenos Aires province, finding the greatest genetic differentiation among the populations of the Northeast, where this landscape is narrower.

Keywords *Ctenomys australis* · Population structure · Connectivity · Sand-dune habitat · Dispersal patterns

Introduction
The management of endangered species requires the identification of units that behave independently in terms of population dynamics (Waples and Gaggiotti 2006). One of the ways to study the population dynamics of species and delimit spatially discrete units in highly fragmented landscapes is to quantify the population connectivity and gene flow based on inferences about migration rates (Hanski and Gaggiotti 2004). The movement of individuals (and their genes) will depend mainly on the characteristics of the environment they inhabit, since a fragmented landscape offers a greater degree of heterogeneity that can restrict or promote dispersion (Hanski and Gaggiotti 2004; Crooks and Sanjayan 2006). Beyond the particularities presented by the species, the history and evolution of the landscapes has deeply conditioned the way in which genetic variation has been historically partitioned geographically (Apodaca et al. 2012). In this sense, the study of habitat fragmentation constitutes a fundamental pathway in order to assess the population structure, which influences the level of long-term population genetic variability (Hanski and Gaggiotti 2004). The loss of genetic variability as a result of habitat fragmentation leads both to a decrease in the population connectivity and...
a reduction in effective population sizes (Hanski and Gaggiotti 2004), increasing the effect of genetic drift and levels of inbreeding within such populations (Crispo et al. 2011).

South American subterranean rodents of the genus *Ctenomys* (tuco-tuco) represent excellent models for the study of dispersal patterns because they commonly occupy fragmented habitats and have limited dispersion in relation to the spatial scale of habitat discontinuities (Steinberg and Patton 2000). They are typically characterized by exhibiting small effective population size with low genetic variability and high inter-population divergence (Busch et al. 2000; Lacey 2000).

The sand-dune tuco-tuco (*Ctenomys australis*) is an endemic and endangered subterranean rodent (category endangered by The IUCN Red List of Threatened Species 2018) with a very narrow distributional range in the coastal sand dunes along the Atlantic coast of the Buenos Aires province, Argentina (Zenuto and Busch 1998; Kittlein et al. 2004; Mora et al. 2006, 2010). The individuals of this species are highly specialized in occupying the dune environment and build large burrow systems on the first strip of coastal dunes, on soft soils with scarce plant cover (Vassallo 2004; Mora et al. 2006, 2010). Along the coast, the habitat of *C. australis* is practically linear and mostly continuous over all of its highly restricted distributional range (less than 280 km), interrupted by some towns and cities, streams and a large river, the Quequén Salado River (Mora et al. 2006). Behind the coastal sandy barrier the environment is severely impacted by crops and pasture fields. At smaller spatial scale (few kilometers), the habitat of this species is recurrently interrupted by low inter-dune grasslands with harder soils (not suitable habitat for the occupation of *C. australis*), generating a more heterogeneous environment with numerous effective potential barriers to gene flow that result in several isolated populations along the coast (Mora et al. 2006, 2010; see also Cutrerera and Mora 2017). Additionally, the nearly one-dimensional pattern of distribution along the coast imposes important restrictions on gene flow and on the dynamic of differentiation within this species (Mora et al. 2006; Mora and Mapelli 2010; Cutrerera and Mora 2017). Other factors such as the progressive advance of forest plantations and urbanizations on coastal dunes during the last decades have produced an important reduction and fragmentation of the habitat of this species (Kittlein et al. 2004).

In order to design measures that aim to preserve the sand dune tuco-tuco populations and their habitat, it is necessary to study how the loss and/or fragmentation of its habitat affect their population structure. Particularly, Mora et al. (2010) using microsatellite loci showed moderate genetic structure among different *C. australis* subpopulations associated to a surprisingly fine geographical scale (less than 4 km) on a fragmented sand dune landscape, suggesting that minor discontinuities might limit the dispersal patterns in this species. In this context, we use multilocus genotype data based on microsatellite loci and several Bayesian approaches to characterize the population structure and migration patterns of this species on a spatial scale that covers its entire distribution range. We also analyze the effect of different environmental variables on the degree of genetic structuring of this endemic species. Thus, this approach provides significant results clarifying the relative importance of natural and anthropic barriers, distances among populations and the effect of some landscape variables on the current patterns of gene flow in this coastal endangered species.

**Materials and methods**

**Study area and sampling design**

Sampling was conducted in a sand dune habitat along the south-eastern Atlantic coast, between the localities of Necochea and Pehuen-Có, Buenos Aires province, Argentina (Fig. 1). Due to the difficulties associated with access to sampling sites during the rainy seasons, the study area is not accessible throughout the year. Because of this, and the large geographical extent in which the study was conducted, sampling was carried out during a four-year period, from April 2013 to April 2017. Previous researches show that this subterranean rodent species can live at least four years, having a longer generation time relative to other non-cavimorph rodents. To reduce the overlap in generation times, a sampling period no longer than the generational time of the species was considered using adult individuals of both sexes.

The sand dune habitat had a width that can vary between 200 m and 8 km (Zenuto and Busch 1998), and present scarce vegetation composed mainly of natural grassland (Zenuto and Busch 1995, 1998). A more detailed description can be found in Celsi and Monserrat (2008).

A total of 194 individuals of *C. australis* were captured in 13 sampling sites throughout the entire distribution area, with an average of 15 individuals per locality. We obtained tissue samples (toe snips for subsequent DNA extraction and genetic analyses) from the captured individuals, which were live trapped using Oneida Victor N °0 snap traps (Oneida Victor, Inc., Ltd., Eastlake, OH, USA), with a rubber cover to avoid injuring animals (experience indicates that this procedure neither affects survival nor digging performance of individuals; see Mora et al. 2006, 2010, 2016). Position of captures was determined using GPS. After collection of tissue samples for genetic analyses, animals were immediately released within the same burrow system where they had been captured. The handling of the individuals was carried out taking into account the guidelines of the “American Society of Mammalogists (Animal Care and Use Committee 2016)”.

© Springer
DNA extraction and microsatellite amplification

Total DNA extractions were performed following Mora et al. (2006) and deposited in freezers at −20 °C. Genetic analyses were accomplished using 13 polymorphic microsatellite loci, developed for *Ctenomys haigi* (HAI2, HAI4, HAI5, HAI9, HAI10; Lacey et al. 1999), *Ctenomys socia­bilis* (SOC1, SOC2, SOC3, SOC5, SOC6, SOC8; Lacey 2001), and *Ctenomys torquatus* (TOR5, TOR9; Roratto et al. 2011). PCR amplifications were performed separately for each microsatellite, using either fluorescently labeled primers (SOC1, SOC2, SOC5, SOC6, SOC8, HAI9 and HAI10), or unlabeled M13 primers (SOC3, HAI2, HAI4, HAI5, TOR5 and TOR9). For the former, amplification was carried out in a final volume of 10 µl containing 1 × Taq buffer (750 mM Tris–HCl, 200 mM (NH₄)₂SO₄, 0.1% (v/v) Tween 20), 2.5 mM MgCl₂, 0.6 mM dNTPs, 3 pmol of each primer, 0.45 units of DNA Taq polymerase (Fermentas) and 1 µl of extracted DNA. For the M13 primers, amplifications were carried out in a reaction volume of 10 µl containing 1 × Taq buffer (750 mM Tris–HCl, 200 mM (NH₄)₂SO₄, 0.1% (v/v) Tween 20), 2.5 mM MgCl₂, 0.6 mM dNTPs, 0.167 pmol of forward primer, 3 pmol of reverse primer, 6 pmol of fluorescent M13 primer, 0.16 µg/µl of BSA, 0.45 units of DNA Taq polymerase (Fermentas) and 1 µl of extracted DNA. PCR programs comprised an initial denaturation at 94 °C for 5 min, followed by 30 cycles of 30 s at 94 °C, 30 s at 54 °C (for SOC1, SOC2, HA19, TOR5 and HA110), 56°C (for SOC8) or 58°C (for HA14, TOR9, HA15, SOC5, SOC6, SOC3 and HA12), and 30 s at 72 °C. The final extension was performed at 72 °C for five min. Negative controls were included in all PCRs. Reactions were carried out for each locus separately. Final PCR products were analyzed with a capillary sequencer ABI3100 (MACROGEN, Inc., Korea). Different combinations of multiplexes, relative sizes of fragments and the number of dye labels are shown in Online Resource 1. The fragments were scored with the program GENEIOUS 6.0.6 (Kearse et al. 2012).

Statistical analyses

Genetic diversity and population structure

Genetic diversity was measured as the number of alleles per locus (At) and per sampling site (Ni), observed (Ho) and expected (He) heterozygosity (Nei 1978), using ARLEQUIN 3.1 (Excoffier et al. 2005). Analysis of linkage disequilibrium between pairs of loci and deviations from Hardy–Weinberg equilibrium were tested using ARLEQUIN 3.1. We used 10,000 dememorization steps and 100,000 iterations for the Markov chain implemented by the method of Guo and Thompson (1992), setting an alpha of 0.01 and correcting it by Bonferroni method.

It was also essential to evaluate the presence of null alleles, which arise when mutations in the binding site of the targeted DNA sequence prevent the efficient annealing of at least one primer, resulting in failure of amplification during the PCR reaction. This can cause that samples in the homozygous state do not produce amplification at all, and in the heterozygous state, appear as a homozygous individual for a particular locus (Rico et al. 2017). We used Cervus 3.0 (Kalinowski et al. 2007) to evaluate the presence of null alleles with all sampling localities pooled together, taking
into account a threshold of 0.1, frequency above which the presence of null allele was considered.

We performed a phylogenetic reconstruction using the genetic distances proposed by Bruvo et al. (2004), which are based on the stepwise mutation model that includes allelic repeat score, allowing us to cross-check these results with those performed in STRUCTURE assuming a model of admixture with correlated allele frequencies in our data. This phylogenetic reconstruction was obtained using the package “poppr” in R (Kamvar et al. 2014).

We employed the Slatkin (1995) estimator of pairwise $R_{ST}$ between sampling sites using ARLEQUIN 3.1. This parameter assumes a stepwise mutation model and considers both the size and frequencies of the alleles in each population, proving that it performs well for small samples (Slatkin 1995). Additionally, we used the hierarchical Bayesian method of Foll and Gaggiotti (2006), as implemented in GESTE 2.0, to estimate population specific $F_{ST}$. This method estimates individual $F_{ST}$ values for each local population (in our case the individual sampling sites) using the approach first proposed by Balding and Nichols (1995). The values, estimated for each population individually, represent the genetic differences that exist between a local population (in our case the individual sampling sites) using the Monte Carlo Markov Chain (MCMC). We also ran an additional burn-in of 2.5 × 10^4 iterations, and estimates of $F_{ST}$ local values using 2 × 10^7 additional iterations with a thinning interval of 50. Two independent runs with identical setting values were performed to check for consistency of the estimates.

In order to infer the partitioning of genetic variance within and among populations, different Analyses of Molecular Variance (AMOVAs) were performed using ARLEQUIN 3.1 considering (1) the 13 sampling sites as independent units, and (2) taking into account the most important natural and anthropogenic barriers to gene flow, forming eight different groupings: (1) Two regions limited by the Quequén Salado River, (2) three regions limited by the Quequén Salado River and the Sauce Grande Stream, (3) three regions limited by the Quequén Salado River and the Cristiano Muerto Stream, (4) three regions limited by the Quequén Salado River and the Claromecó Stream, (5) four regions limited by the Quequén Salado River, Sauce Grande Stream and Cristiano Muerto Stream, (6) five regions limited by the Quequén Salado River, Sauce Grande Stream, Cristiano Muerto Stream and Claromecó Stream, (7) seven regions limited by the Quequén Salado River, Sauce Grande Stream, Cristiano Muerto Stream, Claromecó Stream, Los Gauchos Stream delta and urbanizations (Necochea, Claromecó and Monte Hermoso) (see Fig. 1 for the location of rivers, streams and localities).

We used STRUCTURE 2.3.4 (Pritchard et al. 2000) to estimate the number of genetic clusters and assign individuals to them. The program uses a MCMC procedure to estimate the posterior probability that the data fit to the hypothesis of K clusters. At the same time, the program also calculates the fractional membership of each individual to each cluster ($Q$). We set the number of K between one and 15 using the admixture model with correlated allele frequencies, as suggested by the software developers for closely related populations (Falush et al. 2003). We conducted five independent runs for each K value. Preliminary runs showed that convergence was achieved after $1.5 \times 10^5$ iterations. We thus used this as burn-in and based the estimations on $1.5 \times 10^5$ additional iterations. Selection of the most likely number of K was based on the methods of Pritchard and Wen (2003) and the method proposed by Evanno et al. (2005). We considered a threshold of 0.7 in the proportion of membership of each individual to the different clusters (Q) when assigning individuals to clusters. Individuals that had a Q below this value were subjected to the exclusion test implemented in GENECLASS (Cornuet et al. 1999) to determine if they could have originated from unsampled populations. For this purpose we used the Bayesian method reported by Rannala and Mountain (1997) and the simulation algorithm described by Paetkau et al. (2004), both of them implemented in GENECLASS. As suggested by Paetkau et al. (2004), we used an exclusion threshold of 0.01.

Also, we evaluated the demographic history in each sampling site using 2MOD program (Ciofi et al. 1999). The probability that two alleles are identical by descent is described from the frequency distribution of the $F$ values, which can be interpreted as a relative measure of the effect of genetic drift and dispersion on the individual locations (Ciofi et al. 1999). We ran the MCMC simulation for $1 \times 10^6$ iterations, and obtained the posterior distributions of $F$ discarding the initial 10% of data as burn-in. The consistency of estimates was checked using two independents runs.

Finally, we assessed the relationship between the population structure (local $F_{ST}$) and different environmental variables related to the geographic position and environmental connectivity of sampling sites using a generalized linear model (GLM) implemented in GESTE 2.0 (Foll and Gaggiotti 2006). Posterior probabilities associated with each environmental factor, estimated from the number of times the algorithm visited each model, identify those factors that most influence the genetic structure. We analyzed the potential effects of 10 environmental factors on the genetic structure of C. australis. Two of these factors describe the localization of the sampling sites. The -Geographic Longitude- in UTM coordinates was included among these factors. Given the linear distribution of this species on coastal
dues, we included a factor that controls the effect of the geographical position of sampling sites. Thus, for each sampling site we calculated the -Average distance to other sampling sites-. To characterize the landscape features surrounding sampling sites we used a set of eight environmental variables derived from a potential habitat modeling obtained with the MAXENT 3.4.1 program (Phillips et al. 2006). The details of MAXENT setting and the environmental variables implemented for modeling are described in Online Resource 2. Besides, from an artificial surface area of 20 km long by eight km wide (total width of the barrier dunes), we evaluated the following variables: -Optimal habitat area (ha)-, -Suboptimal habitat area (ha)-, -Average size of optimal habitat patch (ha)-, -Standard deviation of optimal habitat patch size (ha)-, -Number of optimal habitat patches-, -Euclidean distance to the nearest neighbor (m)-, and -Connectivity among optimal habitat patches (%)-. We used the FRAGSTATS 4.2.1 program (McGarigal et al. 2012) to obtain these variables. Connectivity, according to McGarigal et al. (2002), represents the number of functional joins between the patches corresponding to the same type of habitat, where each pair of patches is connected according to a specific distance criterion.

$$\text{CONNEC} = \frac{\sum_{i,j,k} C_{ijk}}{(mn(i-1))}$$

Here $c_{ijk}$ represents the joining between patch $j$ and $k$ (0 = unjoined, 1 = joined) of the corresponding patch type ($i$), based on a user-specified threshold distance, and $ni$ is the number of patches in the landscape of the corresponding patch type ($i$). We used a distance value of 400 m, which corresponds to the average distance proposed for a rodent with similar body size to C. australis (Sutherland et al. 2000). Finally, we included the variable—Width of the Southern Barrier-, which was measured as the euclidean distance between the transverse limits of the dunes barrier corresponding to each sampling site.

GESTE runs were performed using the same settings as those for local $F_{ST}$ estimation. A Pearson correlation was used to evaluate the association between locals $F_{ST}$ and the values of significant environmental factors.

**Isolation by distance pattern**

We performed Mantel test (ManTEL 1967) between pairwise estimates of $R_{ST}$ and linear geographic distances among populations to check for the possibility of isolation by distance (IBD, Slatkin 1993). For this analysis, we have considered the results of Structure and Bayesass to define populations. IBD pattern is expected in this species due to its limited dispersal ability and its linear distribution on the sand dune habitat (Wlasiuk et al. 2003; Mapelli et al. 2012; Mora et al. 2006, 2016). Additionally, we evaluated if our data fits to a model of equilibrium between gene flow and local genetic drift (necessary for the establishment of an IBD pattern) using 2MOD program, which calculates the relative likelihoods of two models of population structure, pure drift vs. migration-drift equilibrium. A model of genetic drift is expected when the allele frequencies in each locality are the result of random changes, with no migration between populations. On the other hand, when a migration drift-equilibrium context is observed, population allele frequencies are the result of a balance between gene flow and local genetic drift. The 2MOD uses an MCMC procedure to compare the likelihoods under these two scenarios and obtain the relative probabilities of the data fitting each model. The MCMC simulation was run for $1 \times 10^6$ iterations, and discarding the initial 10% of data as burn-in. The consistency of estimates was checked using two independent runs.

**Historical and contemporary migration rates**

To estimate the short-term patterns of gene flow (the last three-five generations) we used BAYESASS 3.0 (Wilson and Rannala 2003). This program, based on a Bayesian MCMC approach, uses a genetic assignment method and does not make assumptions of genetic balance to estimate recent migration rates, making it more appropriate to estimate dispersal rates in the short-term (Paetkau et al. 2004). The migration rate $m_{ij}$ represents the fraction of individuals in population $i$ that are migrants derived from population $j$ (per generation); it assumes that a part of an individual’s alleles originates through a single migrant ancestor that reached the current (or past) generation (Wilson and Rannala 2003). This method assumes that migration rates between populations may be asymmetric, but they are constant for short periods of time (few generations), and also that migration rates are small (Wilson and Rannala 2003). Although BAYESASS assumes linkage equilibrium between different loci, it allows deviations in the Hardy–Weinberg proportions when introducing an additional inbreeding parameter ($F$). We ran a total of 20 million MCMC iterations, discarding the first 5 million as burn-in, and sampled the chain every 1000 iterations. Various delta values for migration rates ($m$), allele frequencies ($P$), and inbreeding values ($F$) were used, and the chains were assessed for convergence by performing multiple runs with different initial random seeds. When we obtained the acceptance rates recommended by the authors of the program (between 20 and 40%), we performed five runs (each with different seed values) with the same setting as the one made previously. Then, we evaluated the convergence for each run with TRACER 1.6.0 (Rambaut et al. 2014) and selected the run with the lowest Bayesian deviance [see Spiegelhalter et al. (2002) and Faubet et al. (2007) for more details].
To estimate the long-term patterns of gene flow we used MIGRATE 3.6.11 (Beerli and Felsenstein 2001; Beerli 2006). This program uses coalescent theory and MCMC techniques to estimate two parameters from the microsatellite data, θ and M, where θ represents an estimator of the effective population size (4Neµ, for nuclear DNA) and M represents the mutation-scaled immigration rate (m/µ). This coalescent-based approach is most suitable for estimating migration rates over thousands of years or approximately 4Ne generations in the past (Beerli 2008). Due to the linear distribution of this species in the Southeast of Buenos Aires province and the low dispersion rates expected for subterranean rodents, we analyzed a stepping-stone model of migration, where M were estimated only for those neighboring sampling sites. The data were assumed to follow a Brownian motion mutation model. Following the recommendations of the author of MIGRATE 3.6.11 (Peter Beerli comm. pers.), we did initial runs on our data using FST to find the start parameters, and the results of these runs were used as start parameters for subsequent runs. Distributions were estimated with the Bayesian method, specifying five replicates, with four-chain heating at temperatures of 1, 1.5, 3 and 10,000 to increase the efficiency of the MCMC, a sampling increment of 100, 5,000 recorded steps, and a burn-in of 100,000. Because parameter estimates from the final run were similar to the results from the shorter runs, we assumed that the final run had converged (Chiu et al. 2010). To compare estimates from MIGRATE and BAYESASS, we converted the estimates for M from MIGRATE to proportion of migrants (m) for populations by the formula m = M µ, where µ = 5 × 10⁻⁴ (a common microsatellite substitution rate). This method calculates a percentage of probability of the population where the individual was sampled over the highest probability value among all sampled populations, including the population where the individual was sampled (Paetkau et al. 2004). We determined probability values simulating 10,000 individuals and using an exclusion threshold of 0.01 (Paetkau et al. 2004). Second, we used BAYESASS (Wilson and Rannala 2003) that employs a non-equilibrium Bayesian method, and obtained the posterior probabilities of migrant ancestry for each individual. We ran a total of 2 × 10⁷ MCMC iterations, discarding the first 5 × 10⁶ iterations as burn-in and sampled the chain every 10⁰ generations.

**Results**

**Null alleles, Hardy–Weinberg equilibrium, linkage disequilibrium and genetic diversity**

All loci showed significant frequencies of null alleles at a cut-off level greater than 0.1 when sampling localities were pooled together. As suggested by Rico et al. (2017), the high frequencies of null alleles observed in several studies of different vertebrates and invertebrates do not appear to have a significant effect in the estimates of population genetic parameters commonly assessed from microsatellite loci. However, it should be noted that the values considered in this study were at the limit of this cut-off level. The high probability of having null alleles in the complete sample may be due to the inbred genetic structure present in whole populations and not due to a systematic non-amplification of alleles (Kalinowski et al. 2007; Rico et al. 2017). Mora et al. (2010) working at small spatial scale studying *C. australis* showed lower levels of polymorphism in the same loci than those performed here, but with lower probabilities of presence of null alleles. In this study, level of polymorphism obtained for most loci was similar compared to other studies in different *Ctenomys* species (e.g. Własiuk et al. 2003). Based on this evidence, none of them were discarded in subsequent analyses.

No paired comparison between loci showed linkage disequilibrium, suggesting that the 13 loci present independent inheritance. Only 20 cases of a total of 169, product of the comparisons between 13 loci and 13 sampling sites, showed significant deviations from Hardy–Weinberg equilibrium: SOC2 in NC1, NC2, CLA and DUN, SOC6 in NC1, NC2, SCA, DUN, QSR and LCL, SOC5 in LAng and QSR, HAI4 in NC2, SCA, QSR and PCO, TOR9 in LAng, and HAI5 in NC2, LCL and LGS (Online Resource 3; abbreviations of the sampling sites are shown in Fig. 1). All these deviations were caused by an excess of homozygotes.

Genetic diversity parameters for each sampling site are shown in Online Resource 3. The allelic frequencies of microsatellites are shown in Online Resource 4. The loci used in this study showed a total number of alleles per locus between four (TOR 9) and 14 (HAI4) (Online Resource 3), while the average per locus was 8.77 alleles. In general, our results showed moderate levels of polymorphism for all loci analyzed. All sampling sites had at least 11 polymorphic loci. Allele richness increased, in general, towards the Southwest of the species’s distribution; NC1 and CLA presented the lowest levels (1.77 and 2.08, respectively), while LGS and MHO presented the highest levels (3.44 and 3.42,
respectively). The allelic richness by locality for all loci and the mean allelic richness by locality are shown in Fig. S1 (Supplementary Material).

**Population structure**

The results obtained with STRUCTURE (Fig. 2b), where the eleven genetic units were the most likely level of clustering from 13 original sampling sites, are in accord with those observed from phylogenetic reconstruction using the Bruvo genetic distances (Bruvo et al. 2004), that clearly showed also eleven genetic clusters from the total sampling sites, with NC1-NC2 and QSR-LCL as unique clusters (Fig. S2).

Most RST comparisons showed significant differences, except for the neighbors pair QSR/LCL with a 15 km distance between them (Table 1). RST values ranged from 0.04 to 0.64 (0.519–0.763).

**Fig. 2** Assignment probabilities (Q) to different genetic clusters identified by STRUCTURE considering both hypothesis of K, K = 2 (a), and K = 11 (b). Each individual is represented by a vertical bar and each sampling site is labeled as in Fig. 1 separated by black lines.

**Table 1** Pairwise RST estimates (Slatkin 1995) from microsatellite loci showing the genetic differentiation in *C. australis* amongst sampling sites.

|       | NC1  | NC2  | LAng | SCA  | CLA  | DUN  | RET  | QSR  | LCL  | LGS  | SG   | MHO  | PCO  |
|-------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| NC1   | –    | 0.16*|     |      |      |      |      |      |      |      |      |      |      |
| NC2   | 0.36*| –    | 0.47*| 0.36*| 0.47*| 0.47*| 0.47*| 0.47*| 0.47*| 0.47*| 0.47*| 0.47*| 0.47*|
| LAng  | 0.36*| 0.29*| –    | 0.15*| 0.15*| 0.15*| 0.15*| 0.15*| 0.15*| 0.15*| 0.15*| 0.15*| 0.15*|
| SCA   | 0.27*| 0.31*| 0.31*| –    | 0.15*| 0.15*| 0.15*| 0.15*| 0.15*| 0.15*| 0.15*| 0.15*| 0.15*|
| CLA   | 0.79*| 0.76*| 0.76*| 0.76*| –    | 0.62*| 0.62*| 0.62*| 0.62*| 0.62*| 0.62*| 0.62*| 0.62*|
| DUN   | 0.29*| 0.21*| 0.27*| 0.23*| 0.23*| –    | 0.13*| 0.13*| 0.13*| 0.13*| 0.13*| 0.13*| 0.13*|
| RET   | 0.39*| 0.27*| 0.33*| 0.28*| 0.68*| 0.13*| –    | 0.13*| 0.13*| 0.13*| 0.13*| 0.13*| 0.13*|
| QSR   | 0.39*| 0.36*| 0.37*| 0.22*| 0.50*| 0.17*| 0.15*| –    | 0.15*| 0.15*| 0.15*| 0.15*| 0.15*|
| LCL   | 0.41*| 0.37*| 0.40*| 0.26*| 0.48*| 0.18*| 0.16*| 0.16*| 0.16*| 0.16*| 0.16*| 0.16*| 0.16*|
| LGS   | 0.49*| 0.50*| 0.46*| 0.33*| 0.50*| 0.33*| 0.33*| 0.33*| 0.33*| 0.33*| 0.33*| 0.33*| 0.33*|
| SG    | 0.52*| 0.56*| 0.51*| 0.45*| 0.52*| 0.41*| 0.47*| 0.47*| 0.47*| 0.47*| 0.47*| 0.47*| 0.47*|
| MHO   | 0.50*| 0.54*| 0.45*| 0.36*| 0.46*| 0.35*| 0.39*| 0.28*| 0.28*| 0.28*| 0.28*| 0.28*| 0.28*|
| PCO   | 0.50*| 0.54*| 0.41*| 0.34*| 0.43*| 0.41*| 0.44*| 0.31*| 0.34*| 0.34*| 0.34*| 0.34*| 0.34*|

Local FST estimates (right column) between each sampling site are also given.
Values in parenthesis correspond to the 95% highest posterior probability interval (HPDI). Abbreviations for sampling sites are defined in Fig. 1. *P < 0.001
(between QSR and LCL) to 0.79 (between NC1 and CLA). NC1/CLA and NC2/CLA were the most differentiated populations, while QSR/LCL and DUN/RET pairs presented the lowest population differentiation.

Local FST values estimated with GESTE are shown in Fig. 3. Northeast populations in the distribution range of the species (from NC1 to RET) showed the highest local FST values, suggesting a greater degree of isolation relative to the southern localities (from QSR to PCO), which presented lower values and similar among them (Table 1).

When all sampling sites were considered without any level of population subdivision, the AMOVA showed high differentiation among different sampling sites (ΦST = 0.38, p < 0.001; Table 2). AMOVAs performed for localities assembled into groups separated by different types of natural and anthropic barriers also show different significant apportionments of the genetic variance among regional group. When the sampling sites were assembled into eight groups (considering the most important rivers, streams and the larger urbanizations), the population differentiation was almost as high as the particular case in which the barriers were not considered (ΦCT = 0.29, P < 0.001, Table 2). This situation suggests that 29% of the population divergence (those differences that do not take into account populations into groups) is explained by barriers such as rivers and urbanizations. In fact, the differences between populations in different regions are responsible for practically

![Fig. 3 Local FST values (black line) and their corresponding 95% confidence interval calculated with GESTE (Foll and Gaggiotti 2006) are shown. The dotted line in the figure indicates the presence of the Quequén Salado River](image)

**Table 2** Hierarchical analysis of molecular variance (AMOVA) using the sum of squares differences of RST inferences for microsatellite data (Slatkin 1995)

| Resource of variation | Level of subdivision | ΦCT | P   | ΦST | P   |
|-----------------------|----------------------|-----|-----|-----|-----|
| Two regions limited by the Quequén Salado River | [NC1-NC2-LAng-SCA-CLA-DUN-RET] [QSR-LCL-LGS-SG-MHO-PCO] | 0.15 | < 0.01 | 0.42 | < 0.001 |
| Three regions limited by the Quequén Salado River and the Sauce Grande Stream | [NC1-NC2-LAng-SCA-CLA-DUN-RET] [QSR-LCL-LGS] [SG-MHO-PCO] | 0.20 | < 0.001 | 0.42 | < 0.001 |
| Three regions limited by the Quequén Salado River and the Cristiano Muerto Stream | [NC1-NC2-LAng-SCA] [CLA-DUN-RET] [QSR-LCL-LGS-SG-MHO-PCO] | 0.15 | < 0.01 | 0.41 | < 0.001 |
| Three regions limited by the Quequén Salado River and the Claromecó Stream | [NC1-NC2-LAng-SCA-CLA-DUN-RET] [QSR-LCL-LGS-SG-MHO-PCO] | 0.11 | < 0.05 | 0.40 | < 0.001 |
| Four regions limited by the Quequén Salado River, Sauce Grande Stream and Cristiano Muerto Stream | [NC1-NC2-LAng-SCA] [CLA-DUN-RET] [QSR-LCL-LGS-SG-MHO-PCO] | 0.20 | < 0.001 | 0.40 | < 0.001 |
| Five regions limited by the Quequén Salado River, Sauce Grande Stream, Cristiano Muerto Stream and Claromecó Stream | [NC1-NC2-LAng-SCA] [CLA-DUN-RET] [QSR-LCL-LGS] [SG-MHO-PCO] | 0.25 | < 0.001 | 0.40 | < 0.001 |
| Seven regions limited by the Quequén Salado River, Sauce Grande Stream, Cristiano Muerto Stream, Claromecó Stream and urbanizations (Necochea, Claromecó and Monte Hermoso) | [NC1] [NC2-LAng-SCA] [CLA] [DUN-RET] [QSR-LCL-LGS] [SG-MHO-PCO] | 0.27 | < 0.001 | 0.40 | < 0.001 |
| Eight regions limited by the Quequén Salado River, Sauce Grande Stream, Cristiano Muerto Stream, Claromecó Stream, Los Gauchos Stream delta and urbanizations (Necochea, Claromecó and Monte Hermoso) | [NC1] [NC2-LAng-SCA] [CLA] [DUN-RET] [QSR-LCL-LGS] [SG-MHO-PCO] | 0.29 | < 0.001 | 0.39 | < 0.001 |
| No clustering of sampling sites into groups | [NC1] [NC2] [LAng] [SCA] [CLA] [DUN] [RET] [QSR] [LCL] [LGS] [SG] [MHO] [PCO] | 0.38 | < 0.001 |
all the differentiation between populations. In addition, the Quequén Salado River seems to explain at least the 15% of total inter population variation. In conjunction with the $R_{ST}$ comparisons, differentiation among sampling sites was related in some extent to a regional pattern, whereby the historical effect of different barriers (natural and anthropic) cannot be ruled out.

The Bayesian analysis using STRUCTURE identified a strong substructure among all the sampling sites; the logarithm of the data probability [LnP(D)] as a function of K reached a peak for K = 11 (Fig. S3a), although the highest Δk value using the Evanno method was obtained for K = 2 (Fig. S3b). For K = 2 and K = 11, all runs produced identical clustering solutions with similar values of cluster membership (Q) for all individuals. STRUCTURE results revealed for both clustering solutions a strong population structure: 100% and 93.3% of individuals for K = 2 and K = 11, respectively, were assigned to the locality in which were sampled. Thus, we considered these two possible levels of clustering as the most likely scenarios of population subdivision. Under the hypothesis of K = 2, two clusters were defined by the populations of the Northeast and Southwest, which could indicate an important differentiation on both sides of Quequén Salado River (Fig. 2a). For K = 11, almost all sampling sites constituted independent genetic clusters, except for the pairs of neighboring localities NC1-NC2 and LCL-QSR (Fig. 2b). This result is also strongly supported by obtained from analyses such as pairwise $R_{ST}$ estimates and migration rates. It must be highlighted that NC1 and NC2 (distanced at 12.5 km) at the Northeast end of the distribution range of the species, possibly maintain important levels of historical gene flow. It seems that LCL and QSR (distanced by 13 km) also maintained high levels of admixed constituting a unique cluster (Fig. 2b).

There were 13 individuals that could not be assigned to any genetic cluster using the threshold value of Q > 0.7; two of these were from LCL, two from LGS and nine from MHO. These 13 individuals did not belong to any of the 11 clusters according to GENECLASS exclusion test, and consequently were considered as immigrants that came from unsampled populations.

Frequency distribution of F values at the Northeast of the barrier (from NC1 to RET) presented higher mean values than Southwestern localities (from QSR to PCO) (Fig. S4). These results indicate that northeastern populations presented the highest degree of population isolation (the highest $F_{ST}$ values).

### Isolation by distance pattern

According to the STRUCTURE results, eleven genetic units from 13 original sampling sites were identified (Fig. 2b); BAYESASS results also support the hypothesis of eleven genetic units, showing a greater degree of gene flow between
the two pairs of sampling sites that STRUCTURE recognizes as populations (see below the BAYESASS results). Thus, the Mantel test was performed considering eleven populations (genetic units).

Mantel test showed a significant association between pairwise estimates of $R_{ST}$ and linear geographic distances among populations ($R = 0.47, P < 0.01$; Fig. 5). This results shows that genetic differentiation in *C. australis* was consistent with a simple IBD pattern, suggesting an equilibrium between genetic drift and gene flow. Results from 2MOD agreed with those of Mantel tests; the model that best fitted to the data was that who considered an immigration-drift equilibrium. These results supported a demographic framework in which populations have historically evolved in a drift-migration context, which discards the genetic drift model as the only factor involved.

**Historical and contemporary migration rates**

In general, recent estimates of gene flow obtained from BAYESASS were very low, showing a symmetric pattern between pairs of sampling sites. Gene flow estimates ranged from 0.009 to 0.213. Only one case registered a migration rate greater than 0.2 (suggested as a significant migration value by Bergek and Björklund 2007; see also Björklund et al. 2010) with a noticeable pattern of asymmetric gene flow. Such is the case of NC1 and NC2, with a higher proportion of migrants per generation directed from NC1 to NC2 ($m = 0.213$). Two other cases showed lower values but relatively close to the threshold value 0.2; between LCL and QSR, where LCL was the sampling site that received the highest proportion of migrants from QSR ($m = 0.175$), and between MHO and LGS, where the MHO was the locality who received the largest number of migrants from LGS ($m = 0.093$). The other migration rate pairwise comparisons between sampling sites were negligible (the values ranged from 0.009 to 0.02, Table 3).

Estimates of $M$ obtained from MIGRATE also suggested little migration between the most sampling sites, and moderate to high migration among fewer paired sites. Gene flow estimates ranged from 1.59 to 26.2 (Table 3). We calculate the number of migrants per generation ($N_m$) using the formula $N_m = (M\theta)/4$. Many of these values were less than one indicating little historical migration between sampling sites; others showed higher values than one, indicating that hundreds of years ago there was greater connectivity between them. One sampling site pair did exhibit the same pattern of asymmetric migration, historically and contemporary: NC1 showed higher rate of migration into NC2 than vice versa. Another interesting case of asymmetric migration is CLA, a population that presented very low genetic diversity; this population showed a very low recruitment of migrants from neighboring populations (SCA and DUN), but it represents a source population for these same populations (Table 3). Results from Pearson correlation test were significant ($P < 0.01, r = 0.58$) indicating that the two matrices of contemporary and historical migration values are significantly correlated with each other.

Considering 13 individuals that were not assigned to any genetic cluster with STRUCTURE, the First Generation Migrants analysis in GENECLASS (Table 4) only identified one individual of the 13 as a first generation migrant, while BAYESASS identified 11 of them. Even though the information on first generation migrants does not necessarily translate into dispersion distances traveled by these migrants in the same generation, the results support a certain degree of genetic movement in a very limited time span.

**Discussion**

*Ctenomys australis* showed a strong genetic structure in its entire distributional range, with low historical gene flow among different sampling sites (approximately 10–40 km of distance between them). A conjunction of some elements like geographic distances among populations and the most important barriers to gene flow seem to explain in some extend the genetic differentiation in this species. As was originally expected, rivers and streams seem to have been significant barriers on the landscape, limiting the movement of individuals among different populations. Also, we found that, except for the case of NC1 and NC2, the rest of the sampling sites in the Northeast of the species geographical range presented greater isolation among them relative to the populations of the Southeast area. In this context, an isolation by distance pattern was observed in this subterranean rodent, suggesting an equilibrium between gene flow and genetic drift. There were
also some environmental and cartographic variables that have influenced the population structure; geographic location of the sampling sites (described by the Geographic Longitude) and the habitat availability (described by the Width of the Southern Barrier) have affected the pattern of population differentiation, documented in a clear trend SW/NE. In particular, habitat availability in this species is directly related to the thickness of the sand dune barrier in the Southeast of Buenos Aires province (Mora and Mapelli 2010), which determines the connectivity among sampling sites and effective population sizes. It should be noted, the species distributional range on the Northeast of the Quequén Salado River is currently narrower than in the Southwest.

Subterranean rodents tend to have a strong genetic structure as a consequence of their low mobility and their fragmented distributions (Wlasiuk et al. 2003; Mora et al. 2006, 2007; Mapelli et al. 2012; Mora et al. 2016). In C. australis, connectivity among populations is severely limited by its linear distribution on the coast, some important natural barriers and habitat fragmentation (Mora et al. 2006, 2010; Mora and Mapelli 2010).

### Table 3

| Parameter       | M   | N_m | MIGRATE m | 95% confidence interval | BAYESASS m | 95% confidence interval |
|-----------------|-----|-----|-----------|-------------------------|------------|-------------------------|
| NC2 → NC1       | 11.65 | 1.91 | 0.0058    | 0.0031–0.0084           | 0.0134     | 0.2663–0.1601           |
| NC1 → NC2       | 26.20 | 4.30 | 0.0131    | 0.0087–0.0193           | 0.2132     | 0.0385–0.0117           |
| LAng → NC2      | 11.31 | 1.86 | 0.0057    | 0.0029–0.0085           | 0.0101     | 0.0293–0.0091           |
| NC2 → LAng      | 6.15  | 1.01 | 0.0031    | 0.0006–0.0055           | 0.0133     | 0.0386–0.0120           |
| SCA → LAng      | 2.43  | 0.40 | 0.0012    | 0.0000–0.0032           | 0.0134     | 0.0387–0.0119           |
| LAng → SCA      | 1.59  | 0.26 | 0.0008    | 0.0000–0.0025           | 0.0128     | 0.0365–0.0109           |
| CLA → SCA       | 11.24 | 1.85 | 0.0056    | 0.0029–0.0083           | 0.0132     | 0.0381–0.0117           |
| SCA → CLA       | 2.44  | 0.40 | 0.0012    | 0.0000–0.0029           | 0.0111     | 0.0321–0.0099           |
| DUN → CLA       | 2.37  | 0.39 | 0.0012    | 0.0000–0.0029           | 0.0114     | 0.0330–0.0102           |
| CLA → DUN       | 13.76 | 2.26 | 0.0069    | 0.0039–0.0098           | 0.0120     | 0.0345–0.0105           |
| RET → DUN       | 12.86 | 2.11 | 0.0064    | 0.0029–0.0098           | 0.0215     | 0.0525–0.0095           |
| DUN → RET       | 12.53 | 2.06 | 0.0063    | 0.0035–0.0090           | 0.0183     | 0.0502–0.0136           |
| QSR → RET       | 1.99  | 0.33 | 0.0010    | 0.0000–0.0027           | 0.0119     | 0.0346–0.0108           |
| RET → QSR       | 4.98  | 0.82 | 0.0025    | 0.0005–0.0045           | 0.0207     | 0.0538–0.0124           |
| LCL → QSR       | 2.23  | 0.37 | 0.0011    | 0.0000–0.0029           | 0.0110     | 0.0320–0.0100           |
| QSR → LCL       | 9.27  | 1.52 | 0.0046    | 0.0008–0.0092           | 0.1755     | 0.2392–0.1118           |
| LGS → LCL       | 4.99  | 0.82 | 0.0025    | 0.0003–0.0045           | 0.0124     | 0.0357–0.0109           |
| LCL → LGS       | 4.86  | 0.80 | 0.0024    | 0.0000–0.0050           | 0.0118     | 0.0339–0.0103           |
| SG → LGS        | 3.34  | 0.55 | 0.0017    | 0.0000–0.0034           | 0.0123     | 0.0354–0.0108           |
| LGS → SG        | 9.61  | 1.58 | 0.0048    | 0.0015–0.0080           | 0.0185     | 0.0532–0.0162           |
| MH → SG         | 3.08  | 0.51 | 0.0015    | 0.0000–0.0033           | 0.0171     | 0.0502–0.0160           |
| SG → MH         | 15.05 | 2.47 | 0.0075    | 0.0039–0.0115           | 0.0183     | 0.0520–0.0154           |
| PCO → MH        | 8.10  | 1.33 | 0.0040    | 0.0006–0.0085           | 0.0183     | 0.0532–0.0166           |
| MH → PCO        | 5.20  | 0.85 | 0.0026    | 0.0000–0.0053           | 0.0129     | 0.0376–0.0118           |

*M* historical mutation-scaled migration rate (*m*/µ); *N_m* historical number of migrants per generation (*θM*/4); MIGRATE *m* (*M × 0.0005); BAYESASS *m* contemporary migration rate

### Table 4

| Individual | GENECLASS | Source population | BAYESASS | Source population |
|------------|-----------|-------------------|----------|-------------------|
| LCL n° 128 | P = 0     | PCO               | P = 0.825 | PCO               |
| LCL n° 136 | –         | –                 | P = 0.995 | QSR               |
| MHO n° 166 | –         | –                 | P = 0.998 | LGS               |
| MHO n° 167 | –         | –                 | P = 1     | LGS               |
| MHO n° 168 | –         | –                 | P = 1     | LGS               |
| MHO n° 169 | –         | –                 | P = 1     | LGS               |
| MHO n° 170 | –         | –                 | P = 1     | LGS               |
| MHO n° 171 | –         | –                 | P = 0.717 | LGS               |
| MHO n° 172 | –         | –                 | P = 0.967 | LGS               |
| MHO n° 173 | –         | –                 | P = 0.982 | LGS               |
| MHO n° 174 | –         | –                 | P = 0.967 | LGS               |

For GENECLASS, only the P value that did not exceed the 0.01 threshold suggested by Paetkau et al. (2004) are shown; for BAYESASS, only the highest P values are shown, indicating the highest probability of being a first generation migrant from a given source population.
Using allozyme loci, Apfelbaum et al. (1991) found high levels of homozygosity within populations of *C. australis* and an important genetic differentiation between sampling sites, suggesting that this species has shown high levels of population isolation and low effective population sizes. Additionally, Mora et al. (2010) using microsatellite loci observed significant genetic differentiation at lower spatial scales (<4 km) on a small coastal portion of the *C. australis* distribution. Despite the reduced scale in which this latter study was conducted, a pattern of genetic structuring was detected. These authors concluded that the landscape configuration (e.g. habitat availability) and its specific ecological requirements of the habitat would have molded the genetic variation at this reduced spatial scale. In the current study we found that the genetic variability in *C. australis*, such as the mean number of alleles per locus, the number of alleles per population and levels of heterozygosity, was similar or slightly higher that reported in other species of tuco-tucos such as *Ctenomys flamarioni* (Fernandez-Stolz et al. 2007), *Ctenomys rionegrensis* (Wlasiuk et al. 2003), *Ctenomys torquatus* (Gonçalves and de Freitas 2009) and *Ctenomys talarum* (Cutrera et al. 2010). Considering these observations, the *C. australis* populations showed, in general, acceptable levels of genetic diversity. Also, we found that the high levels of population subdivision coincide with those reported in other studies of *Ctenomys*.

Roratto et al. (2014) found high pairwise \( F_{ST} \) estimates between populations in *C. torquatus*, a species with a similar geographic extension that *C. australis* and restricted to lowlands of southern Brazil and northern Uruguay. Mapelli et al. (2012) assessed the genetic structure in *Ctenomys porteousi* on its narrow and non-linear distribution range in central Argentina (around 90 km), and found a strong population structure with pairwise \( F_{ST} \) estimates similar to those observed in *C. australis*. Fernández-Stolz et al. (2007) reported moderate to high population differentiation in a lower spatial scale (around 30 km) in *C. flamarioni*, a species restricted to coastal sand dunes in southern Brazil, and like *C. australis*, exhibits a linear and continuous distribution over a fragmented landscape. Many of these examples in *Ctenomys*, with restricted, fragmented and relatively linear distributions, support the idea of a strong population structure even at very small spatial scales.

Additionally, we found in *C. australis* a different pattern of population structure in both sectors of the dune barrier. On one hand, the Northeast sampling sites showed a lower genetic diversity, with relatively lower mean values of allelic diversity and heterozygosity in comparison to populations from the southwestern side. Our results agree with those published by Apfelbaum et al. (1991) where the authors have reported low levels of genetic diversity in the Necochea locality, and a decreasing pattern of allelic diversity from West to East. Also, the frequency distributions of \( F \) values performed with 2MOD agree with the genetic pattern mentioned above, showing the lowest mean values on Southwest. These results denote greater connectivity between populations of this area and therefore, a lower degree of population differentiation. The Northeast, however, showed the frequency distribution of \( F \) values mainly displaced to the right of the distributions, denoting poor historical connectivity among populations and important local genetic drift, explaining the strong population differentiation observed in this area.

Genetic differentiation among *C. australis* populations was consistent with an IBD pattern, possibly evidencing equilibrium between gene flow and local genetic drift. Mantel test and 2MOD results supported this idea. Furthermore, major differentiation was observed without clustering subpopulations into regions or major hierarchical units. Natural and anthropic barriers (e.g. rivers, streams and urbanizations) explained a high percentage of variation among populations, which shows that low dispersion rates and high habitat specialization play a preponderant role in such population differentiation. The establishment of an IBD pattern would be favored by the linear and one-dimensional distribution of *C. australis* on the coast, since these populations will tend to reach the drift-migration equilibrium more quickly than in species with two-dimensional distributions (Slatkin 1993). Such as *C. australis*, several species of *Ctenomys* with relatively linear distributions have reported an IBD pattern using microsatellite data (C. talarum, by Mora et al. 2007, 2013; C. flamarioni, by Fernández-Stolz et al. 2007; C. pearsoni, by Tomasco and Lessa 2007; C. minutus, by Lopes 2011; C. "chasiquensis", by Mora et al. 2016). A narrow distribution, such as the one presented by these species, restricts their dispersal potential and gene flow in few potential directions, prioritizing the movement of individuals between neighboring populations.

The IBD pattern observed in this study differs from the previous results reported by Mora et al. (2006) using mitochondrial sequences. These latter authors proposed that the species would not have reached at equilibrium between migration and genetic drift (see also Wlasiuk et al. 2003), suggesting that *C. australis* has suffered a recent process of demographic expansion possibly associated with the cycles of the coastal dune formation during the Holocene. Differences in the molecular markers used in these studies, which have distinct mutation rates, may probably explain these contrasting results relative to the accumulation of polymorphisms in natural populations. Mitochondrial DNA is more used to infer historical gene flow, while microsatellite loci primarily estimate contemporary gene flow (Dionne et al. 2008). In this context, different mutation rates for mitochondrial and microsatellite loci (associated with their accumulation of new source of variation) can inflate or reduce gene flow estimates, having an impact on the IBD adjustment.
Most of results of current study agree with eleven genetic units, where the most of the individuals were strongly assigned to their originally were sampled. AMOVAs partially support the hypothesis of K = 2, since Quequén Salado River seems to explain at least the 15% of total inter population variation. The hypothesis of K = 11 was supports by pairwise $R_{ST}$ and genetic distance phylogenetic reconstruction, showing important genetic differentiation among sampling sites. Moreover, BAYESASS supported low migration rates among sampling sites except for the pairs NC1-NC2 and QSR-LCL suggesting the presence of several independent genetic clusters.

Several biological studies have discussed the difficulty of finding the most likely number of genetic clusters (K) from an arrangement of sampling units (see Meirmans 2015 for a detailed discussion of this topic). This author pointed out that the current genetic structures of species are complex due to the underlying demographic, environmental and historical processes. Beyond the inferences of the most likely number of K genetic clusters using the Evanno method (Evanno et al. 2005), Meirmans (2015) observed that, due to the large uncertainty in the estimation of K, there are often few biological reasons to assume only a single value of K. This author emphasizes the idea that different solutions of K clusters might basically reflect different demographic processes and consequently justify our final interpretation of the data. Using human genetic data Kalinowski (2011) also suggest that STRUCTURE does not reliably identify the main genetic clusters within species. This author identified some difficulties attributed to forcing STRUCTURE to place individuals into too few clusters.

To sum up, Evanno et al. (2005) shows the uppermost clustering level, not necessarily the current number of sub-populations. ΔK gives us an optimum of two genetic clusters for our data set, which means that most of the genetic difference is explained by the presence of two clusters, with the Quequén Salado River positioned as the most important historical barrier to gene flow. We observed that value of K = 11 was also very informative about the actual population structure. Bayesian analyses using STRUCTURE clearly showed that most of the sampling sites behaved as independent genetic units.

Another interesting question here is to what extend the population differentiation in C. australis could be explained by historical landscape features, such as natural barriers to dispersion, or due to more contemporary environmental factors. The local $F_{ST}$ values were significantly correlated with two landscape factors: Geographic longitude and Width of the Southern Barrier. We found the lowest local $F_{ST}$ values in the wider sand-dune area at the Southwest, while the highest ones were found to the Northeast, where the barrier width is narrower. Thus, factors related to the geographical position of sampling sites and the habitat availability denotes important levels of isolation of populations. These variables are clearly associated with the availability of natural habitat, which due to natural and anthropic factors is less common and more discontinuous to the Northeast of the barrier (Turno Orellano et al. 2003; Isla et al. 2009). In the Southwest portion on the coast, we found a greater availability of natural habitats, which are distributed more continuously and have greater natural stabilization since the last generation of coastal dune formation (1600–500 years; see Isla et al. 2001; Marcomini and López 2016). These particularities in the landscape have characterized the Southwest area of the barrier and possibly have allowed a more fluid historical dispersal than in the Northeast. Since the presence of C. australis depends strongly on the sand dune habitat availability (Mora et al. 2006; Mora and Mapelli 2010; Cutrera and Mora 2017), it is expected that demographic effect will be stronger on populations located in the Northeast. In fact, in several subterranean rodent species the population size is strongly correlated with the habitat availability (see Mapelli and Kittlein 2009, and Mapelli et al. 2012). It should be noted that other species of subterranean rodents distributed at similar spatial scales to C. australis, or even smaller, have also revealed the importance of landscape characteristics such as the habitat availability (Sato et al. 2014; Kierepka et al. 2016; Biello et al. 2018; Visser et al. 2018). Our results also agreed with those published by Galiano et al. (2014) in C. minutus (a species with a linear coastal geographic distribution of 300 km on the southeastern of Brazil) and Mora et al. (2017) in C. “chasiquensis” (with a distribution of 100 km in central Argentina) using microsatellite loci. These authors observed that those tuco-tucos distributed in areas with higher habitat availability were associated to greater genetic diversity. In summary, although the availability of suitable habitat, the geographic distances among sampling sites, and natural and anthropic barriers seem, in conjunction, to explain the population genetic structure in this species, it is difficult to evaluate independently the effective historical contribution of each of these processes.

Although most of the results indicated relatively low migration rates between sampling sites, some localities have maintained certain levels of gene flow among them. Phylogenetic reconstruction, pairwise $R_{ST}$ values and the Bayesian clustering approach from STRUCTURE are in according to show that some pairs of sampling sites (e.g. NC1-NC2 and QSR-LCL) behave as single populations. Both BAYESASS and MIGRATE also showed high proportion of migrants per generation for these pairs of sampling sites, being NC1-NC2 the only case that presented the highest historic and contemporary migration rate. In both cases (NC1-NC2 and QSR-LCL) an asymmetric gene flow was detected, with East–West directionality; this would indicate that both NC1 and QSR would potentially function as a source population of migrants, increasing their
conclusion values relative to other populations. This is interesting, because NC1 has the highest local F_{ST} value as well as a low genetic diversity, further supporting the idea of an asymmetric gene flow with NC2, being NC1 the location that gives migrants but does not receive. It should be noted that m from MIGRATE showed a moderate value for QSR-LCL (lower than the sampling pair NC1-NC2), but it is interesting that they are currently connected by moderate gene flow. The particular case of CLA also deserves to be highlighted, since this population showed moderate values of gene flow to its neighboring populations but with very little recruitment of migrants. This result is supported by high pairwise R_{ST} and local F_{ST} values, and by their lowest values of genetic diversity. This population is circumscribed by a river, a city (Claromecó) and some forestations that have probably impacted their levels of current gene flow, which has induced a high degree of population isolation. The result of a significant correlation between both estimates of m (BAYESASS and MIGRATE), could suggest that the high levels of structure currently observed in *C. australis* are not a direct consequence of the anthropic fragmentation of the dune barrier, but that the dispersion limited as a biological characteristic of this taxon, together with its linear distribution, could be the main responsible for such structuring (see also an exhaustive discussion in Mora et al. 2006).

As we saw in Table 4, the presence of some first generation migrants (belonging to MHO and LCL) arrived from non-neighboring populations; that is, the individuals with a Q value lower than 0.7, showed the highest probability of being first generation migrants from the mentioned populations. For example, individual n°166 sampled in MH, showed the greatest probability of being a first generation migrant from LGS, and not from their neighboring locations such as SG or PCO. These results have an important implication for the conservation of this threatened species; the studies related to the directionality of gene flow, considering that migration in a generation has not always occurred between neighboring populations, provide valuable information with respect to “genetic source-sink” dynamics in a set of populations, which is highly relevant to conservation efforts (Manier and Arnold 2005).

### Conclusions

*Ctenomys australis* showed strong genetic structure over their entire distributional range. Our outcomes showed that most of sampling sites represent independent genetic units. Beyond the typical low dispersal capacity of subterranean rodents, natural and anthropic barriers, habitat availability and the linear distribution of the sand dune tuco-tuco have had, in conjunction, an important influence on its genetic structure. Also, the evolution of the sandy landscape on this barrier during the last 1600 years appears to have had a great impact on how populations are currently connected by gene flow. The association between population genetic structure and habitat availability was evidenced: northeastern populations of the barrier, where the habitat availability is lower, were most affected by the historical landscape features. However, the advance of urbanization and forestation on this coastal region in the last decades (Turno Orellano and Isla 2004) has strongly impact the habitat availability for this species in this region. Others analysis considering different landscape configurations and different spatial scales allow us to elucidate how these natural and anthropic factors has influenced the population differentiation in this species.

### Acknowledgements

We are grateful to all the members of the Museo Argentino de Ciencias Naturales “Bernardino Rivadavia” for their invaluable support and advice. Financial support was provided by Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, PIP-11220150100066CO), UNMdP (Project EXA903/18) and FONCYT (PICT-201-0427). To these persons and institutions, we express our deep gratitude.

### References

Apfelbaum LJ, Massarini AI, Daleffe LE, Reig OA (1991) Genetic variability in the subterranean rodents *Ctenomys australis* and *Ctenomys portesi* (Rodentia: Octodontidae). Biochem Syst Ecol 19:467–476

Apodaca JJ, Rissler LJ, Godwin JC (2012) Population structure and gene flow in a heavily disturbed habitat: implications for the management of the imperilled Red Hills salamander (*Phaeognathus hubrichti*). Conserv Genet 13:913–923

Balancing DJ, Nichols RA (1995) A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity. Genetica 96:3–12

Beerli P (2006) Comparison of Bayesian and maximum likelihood inference of population genetic parameters. Bioinformatics 22:341–345

Beerli P (2008) MIGRATE documentation (version 3.0). Technical Report

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. Proc Natl Acad Sci USA 98:4563–4568

Bergek S, Björklund M (2007) Cryptic barriers to dispersal within a lake allow genetic differentiation within a population of Eurasian perch (*Perca fluviatilis*). Evolution 61:2035–2041

Biello R, Brunelli A, Sozio G, Havenstein K, Mortelliti A, Ketmaier V, Bertorello G (2018) Genetic structure in the wood mouse and the bank vole: contrasting patterns in a human-modified and highly fragmented landscape. bioRxiv. https://doi.org/10.1101/464057

Björklund M, Ruiz I, Senar JC (2010) Genetic differentiation in the urban habitat: the great tits (*Parus major*) of the parks of Barcelona city. Biol J Linn Soc 99:9–19

Bruvo R, Michielis NK, D’Souza TG, Schulenburg H (2004) A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. Mol Ecol 13:2101–2106

Busch C, Antinuchi CD, del Valle JC, Kittlein MJ, Malizia AI, Valsallo AI, Zenuto RR (2000) Population ecology of subterranean rodents. In: Lacey EA, Patton JL, Cameron GN (eds) Life
underground: the biology of subterranean rodents. University of Chicago Press, Chicago, pp 183–226
Celsi CE, Monserrat AL (2008) La vegetación dunícola en el frente costero de la Pampa Austral (Partido de Coronel Dorrego, Buenos Aires). Multiequina 17:73–92
Chiucchi JE, Gibbs HL (2010) Similarity of contemporary and historical gene flow among highly fragmented populations of an endangered rattlesnake. Mol Ecol 19:5345–5358
Ciofi C, Beaumont MA, Swingland IR, Bruford MW (1999) Genetic divergence and units for conservation in the Komodo dragon Varanus komodoensis. Proc R Soc Lond B 266:2269–2274
Cornuet J, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics 153:1989–2000
Crispo E, Moore JS, Lee-Yaw JA, Gray SM, Haller BC (2011) Broken barriers: human-induced changes to gene flow and introgression in animals. BioEssays 33:508–518
Crooks KR, Sanjayan M (2006) Connectivity conservation. Cambridge University Press, Cambridge
Cutrer AP, Mora MS (2017) Selection on MHC in a Context of Historical Demographic Change in 2 Closely Distributed Species of Tuco-tucos (Ctenomys australis and C. talarum). J Hered 108:628–639
Cutrer AP, Mora MS, Antenucci CD, Vassallo AI (2010) Intra and interspecific variation in home-range size in sympatric tuco-tucos, Ctenomys australis and C. talarum. J Mammal 91(6):1425–1434
Dionne M, Caron F, Dodson JJ, Bernatchez L (2008) Landscape genetics and hierarchical genetic structure in Atlantic salmon: the interaction of gene flow and local adaptation. Mol Ecol 17:2382–2396
Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals structure: a simulation study using the software. Mol Ecol 14:2611–2620
Excoffier L, Laval G, Schneider S (2005) ARLEQUIN ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform 1:47–50
Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587
Faubet P, Waples RS, Gaggiotti OE (2006) Identifying the environmental factors that determine the genetic structure of populations. Genetics 164:1567–1587
Foll M, Gaggiotti OE (2006) Identifying the environmental factors that determine the genetic structure of populations. Genetics 164:1567–1587
Fernández-Stolz GP, Stolz JFB, de Freitas TRO (2007) Bottlenecks and introgressions as origins of individuals structure: a simulation study using the software. Mol Ecol 16:1099–1100
Kalinowski ST (2011) The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. Heredity 106:625–632
Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16:1099–1100
KamvarZN, TabimaJF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281. https://doi.org/10.7717/peerj.281
Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thirier T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649
Kierepka EM, Anderson SJ, Swihart RK, Rhodes OE Jr (2016) Evaluating the influence of life-history characteristics on genetic structure: a comparison of small mammals inhabiting complex agricultural landscapes. Ecol Evol 6:6376–6396
KittleinMJ, VassalloAI, MoraMS, de DuranaF, RicciardiulliMG, TizónFR (2004) Dunas del Sureste Bonaerense. In: Bilenca D, Miñarro F (eds) Identificación de Áreas Variosas de Pastizal en las Pampas y Campos de Argentina. Uruguay y Sur de Brasil, Fundación Vida Silvestre Argentina, Buenos Aires, pp 76–77
Lacey EA (2000) Spatial and social systems of subterranean rodents. In: Lacey EA, Patton JL, Cameron GN (eds) Life Underground: the biology of subterranean rodents. University of Chicago Press, Chicago and London. pp 257–293
Lacey EA (2001) Microsatellite variation in solitary and social tuco-tucos: molecular properties and population dynamics. Heredity 86:628–637
Lacey EA, Maldonado JE, Clabough JP, Matteo MD (1999) Interspecific variation in microsatellites isolated from tuco-tucos (Rodentia: Ctenomyidae). Mol Ecol 8:1753–1768
Lopes CM (2011) Historia evolutiva de Ctenomys minutus e Ctenomys lami (Rodentia, Ctenomyidae) na planicie costeira do Sul do Brasil. [PhD thesis]. [Porto Alegre (Brazil)]: Universidade Federal do Rio Grande do Sul.
Manier MK, Arnold SJ (2005) Population genetic analysis identifies source-sink dynamics for two sympatric garter snake species (Thamnophis elegans and Thamnophis sirtalis). Mol Ecol 14:3965–3976
Mantel N (1967) The detection of disease clustering and a generalized regression approaches. Cancer Res 27:209–220
Mapelli FJ, Kittlein MJ (2009) Influence of patch and landscape characteristics on the distribution of the subterranean rodent Ctenomys porteouisi. Landsc Ecol 24(6):726–733
Mapelli FJ, Mora MS, Mirol PM, Kittlein MJ (2012) Population structure and landscape genetics in the endangered subterranean rodent Ctenomys porteouisi. Conserv Genet 13:165–181
Marcos SC, López RA (2016) Geología de la Costa Marina Bonaerense. In: Atheta J, Celsi CE (eds) La Costa Atlántica de Buenos Aires: Naturalidade e Patrimonio Cultural. Fundación de Historia Natural Félix de Azara, Buenos Aires, pp 20–41
McGarigal K, Cushman SA, Ene E (2002) FRAGSTATS: Spatial pattern analysis program for categorical maps. University of Massachusetts, Amherst
McGarigal K, Cushman SA, Ene E (2012) FRAGSTATS v4: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649
Meirmans PG (2015) Seven common mistakes in population genetics and how to avoid them. Mol Ecol 24:3223–3231
MoraMS, MapelliFJ (2010) Conservación en médanos: Fragmentación del hábitat y dinámica poblacional del tuco–tuco de las dunas. In: Isla FL, Lasta CA (eds) Manual de manejo de barreras
medanosas de la Provincia de Buenos Aires. Universidad de Mar del Plata, Mar del Plata, pp 161–181

Mora MS, Lessa EP, Kittlein MJ, Vassallo AI (2006) Phylogeography of the subterranean rodent *Ctenomys australis* in sand-dune habitats: evidence of population expansion. J Mammal 87:1192–1203

Mora MS, Lessa EP, Cutrera AP, Kittlein MJ, Vassallo AI (2007) Phylogeographical structure in the subterranean tuco-tuco *Ctenomys talarum* (Rodentia: Ctenomyidae): contrasting the demographic consequences of regional and habitat-specific histories. Mol Ecol 16:3453–3465

Mora MS, Mapelli FJ, Gaggiotti OE, Kittlein MJ, Lessa EP (2010) Dispersal and population structure at different spatial scales in the subterranean rodent *Ctenomys australis*. BMC Genet 11:1–14

Mora MS, Cutrera AP, Lessa EP, Vassallo AI, D’Anatro A, Mapelli FJ (2013) Phylogeography and population genetic structure of the Tuco tuco-tuco (*Ctenomys talarum*): integrating demographic and habitat histories. J Mammal 94(2):459–476

Mora MS, Mapelli FJ, López A, Gómez Fernández MJ, Mirol PM, Kittlein MJ (2016) Population genetic structure and historical dispersal patterns in the subterranean rodent *Ctenomys chasiquensis* from the southeastern Pampas region. Argentina Mammal Biol 81(3):314–325

Mora MS, Mapelli FJ, López A, Gómez Fernández MG, Mirol PM, Kittlein MJ (2017) Landscape genetics in the subterranean rodent *Ctenomys chasiquensis* associated with highly disturbed habitats from the southeastern Pampas region, Argentina. Genetica 145:575–591

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure of Canadian polar bears. Mol Ecol 4:347–354

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Mol Ecol 13:55–65

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Model 190:231–259

Pritchard JK, Wen W (2003) Documentation for STRUCTURE software. Version 2. https://www.pritch.bsd.uchicago.edu

Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rambaut A, Suchard MA, Xie W, Drummond AJ (2014) Version v1.6.0. https://beast.bio.ed.ac.uk/Tracer

Rannala B, Mountain JL (1997) Detecting immigration by using Multilocus genotypes. Proc Natl Acad Sci USA 94:1997–9201

Rico C, Cuesta JA, Drake P, Macpherson E, Bernatchez L, Marie AD (2017) Null alleles are ubiquitous at microsatellite loci in the Wedge Clam (*Donax trunculus*). PeerJ 5:e3188. https://doi.org/10.7717/peerj.3188

Roratto PA, Bartholomei-Santos ML, de Freitas TRO (2011) Tetranu-cleotide microsatellite markers in *Ctenomys torquatus* (Rodentia). Conserv Genet Resour 3:725–727

Roratto PA, Fernandes FA, de Freitas TRO (2014) Phylogeography of the subterranean rodent *Ctenomys torquatus*: an evaluation of the riverine barrier hypothesis. J Biogeogr 42:694–705

Sato JJ, Kawakami T, Tasaka Y, Tamenishi M, Yamaguchi Y (2014) A few decades of habitat fragmentation has reduced population genetic diversity: a case study of landscape genetics of the large Japanese field mouse, *Apodemus speciosus*. Mammal Study 39:1–10

Sikes RS, Animal Care and Use Committee of the American Society of Mammalogists (2016) Guidelines of the American Society of Mammalogists for the use of wild mammals in research and education. J Mammal 97:663–688

Slatkin M (1993) Isolation by distance in equilibrium and nonequilibrium populations. Evolution 47:264–279

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462

Spiegelhalter DJ, Best NG, Carlin BR, van der Linde A (2002) Bayesian measures of model complexity and fit. J R Stat Soc B 64:583–616

Steinberg EK, Patton JL (2000) Genetic structure and the geographic of the speciation in subterranean rodents: opportunities and constraints for evolutionary diversification. In: Lacey EA, Patton JL, Cameron GN (eds) Life underground: the biology of subterranean rodents. University of Chicago Press, Chicago, pp 183–226

Sutherland GD, Harestad AS, Price K, Lertzman KP (2000) Scaling of natal dispersal distances in terrestrial birds and mammals. Conserv Ecol 4:16

Tomasco IT, Lessa EL (2007) Phylogeography of the Tuco-tuco *Cteno-mys pearsoni*: mtDNA variation and its Implication for Chromoso-mal Differentiation. In: Kelt DA, Lessa EP, Salazar-Brao J, Patton JL (eds) The quintessential naturalist: honoring the life and legacy of Oliver P. University of California Press, Pearson, Berkeley, Los Angeles and London, pp 859–882

Turno Orellano HA, Isla FI, Juárez VI (2003) Implementación de un SIG en la evaluación de la aptitud para prácticas forestales en el litoral bonaerense. Boletim Paranaense de Geociências 53:27–34

Turno Orellano HA, Isla FI (2004) Developing sinks for CO2 through forestation of temperate coastal barriers: an environmental business. Reg Environ Change 4:70–76

Vassallo AI (1998) Functional morphology, comparative behaviour, and adaptation in two sympatric subterranean rodents of genus *Ctenomys* (Caviomorpha: Octodontidae). J Zool 244:415–427

Visser JH, Bennett NC, van Vuuren BJ (2018) Spatial genetic diversity in the Cape mole-rat, *Georychus capensis*: Extreme isolation of populations in a subterranean environment. PLoS ONE 13:e0194165

Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evol Appl 3:244–262

Waples R, Gaggiotti OE (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol Ecol 15:1419–1439

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163:1177–1191

Wlasiuk G, Garza JC, Lessa EP (2003) Genetic and geographic differentiation in the Rio Negro tuco-tuco (*Ctenomys rionegrensis*): inferring the roles of migration and drift from multiple genetic markers. Evolution 57:913–926

Zenuto RR, Busch C (1995) Influence of the subterranean rodent *Ctenomys pearsonii*: mtDNA variation and its Implication for Chromosomal Differentiation. In: Kelt DA, Lessa EP, Salazar-Brao J, Patton JL (eds) The quintessential naturalist: honoring the life and legacy of Oliver P. University of California Press, Pearson, Berkeley, Los Angeles and London, pp 859–882

Zenuto RR, Busch C (1998) Population biology of the subterranean rodent *Ctenomys australis* (tuco-tuco) in a coastal dune-field in Argentina. Z Säugetierkunde 63:357–367

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.