Conservation Genetics, Demographic History, and Climatic Distribution of the Nine-Banded Armadillo (*Dasypus novemcinctus*): An Analysis of Its Mitochondrial Lineages

Maria Clara Arteaga, Jaime Gasca-Pineda, Rafael Bello-Bedoy, Luis E. Eguiarte and Rodrigo A. Medellin

Abstract For species widely distributed on the American continent, we can expect to find divergent lineages because historical events, such as the Andean uplift, the emergence of the Panama isthmus, and the Milankovitch climate cycles, affected genetic connections among populations, as well as population sizes. There are only one mammalian species on the American continent, which shows a transcontinental distribution (from South to North America) and had a South American origin, the nine-banded armadillo (*Dasypus novemcinctus*). By sampling 242 *D. novemcinctus* from 16 countries, we explored the number of mitochondrial lineages across its complete range and assessed the climatic distribution and demographic history. We found four well-supported and highly divergent monophyletic lineages, which show a parapatric distribution. The geographic distribution of three of those lineages is wide, and the climatic conditions that they occupy are very different. In the middle of the Pleistocene, strong demographic expansions occurred in populations from three lineages, while in the Last Glacial Maximum (LGM), population sizes of the

M. C. Arteaga (✉) · J. Gasca-Pineda
Departamento de Biología de La Conservación, Centro de Investigación Científica Y de Educación Superior de Ensenada (CICESE), Ensenada, Baja California, Mexico
e-mail: arteaga@cicese.mx

J. Gasca-Pineda
e-mail: jaimegasca@yahoo.com

R. Bello-Bedoy
Facultad de Ciencias, Universidad Autónoma de Baja California, Ensenada, Baja California, Mexico
e-mail: rbello@uabc.edu.mx

L. E. Eguiarte
Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Apartado Postal 70-275, Ciudad de México 04510, Mexico
e-mail: fruns@unam.mx

R. A. Medellin
Departamento de Ecología de La Biodiversidad, Instituto de Ecología, Universidad Nacional Autónoma de México, Apartado Postal 70-275, México Ciudad de México 04510, Mexico
e-mail: medellin@iecologia.unam.mx

J. Ortega and J. E. Maldonado (eds.), *Conservation Genetics in Mammals*,
https://doi.org/10.1007/978-3-030-33334-8_7

four lineages apparently were stable, despite the dramatic changes detected in available areas with a suitable habitat. The geographic distribution of the mitochondrial lineages does not correspond with the six subspecies defined by morphological characteristics. Our results provide tools for planning future studies utilizing ecological, morphological, and genomic analyses to elucidate the number of taxonomic units, which should be recognized. A clear delineation of the distribution and status of these units will be invaluable for the conservation management of the nine-banded armadillo.

Keywords Cingulata · Climatic distribution · Control region · Genetic diversity · Genetic structure

1 Introduction

After the breakup of Gondwana in the mid-Cretaceous, at around 130 mya, South America remained isolated for several millions of years (Pascual and Ortíz-Jaureguizar 2007). An astonishing diversity of mammalian fauna evolved during that time (Simpson 1980), as a result of its complex geological and climatic history. Important events took place in South America during the Miocene, such as the Andean uplift (Poulsen and Jeffery 2011), the Amazonian drainage changes (Hoorn et al. 2010), and marine incursions (Lovejoy et al. 2006). The Milankovitch climatic cycles (Gibbard and Van Kolfschoten 2004) also affected, although indirectly, the South America subcontinent. In the late Miocene, South and North America were connected by the emergence of the Isthmus of Panama (Coates et al. 2004). Several species participated in dispersion waves in both directions and increased their distributional ranges in a process which began about 10 mya and is known as the Great American Biotic Interchange (GABI; Marshall et al. 1982; Webb 1976).

Because geologic and climatic factors affected habitat distribution in South America, genetic connections among populations, as well as population sizes, were also affected and influenced the speciation processes in many taxa. These events had a deep impact on biotic diversity and distribution as has been recorded for mammals (Costa 2003; Ditchfield 2000; Lessa et al. 2003; Patton et al. 2000), birds (Brumfield and Capparella 1996; Weir and Price 2011), and plants (Antonelli et al. 2009). For widely distributed species on the American continent, we can expect to find divergent phylogroups because wide geographic distribution can lead to genetic differentiation due to historical events and physical barriers, as mentioned previously, generating isolation by distance (e.g., Eizirik et al. 2001; Tchaicka et al. 2007), and we can also expect differentiation by local adaptation.

There is only one mammal in America which shows a continental distribution and has a South American origin: the nine-banded armadillo, *Dasypus novemcinctus* (Fig. 7.1). Nine-banded armadillo colonized North America after the emergence of the Isthmus of Panama (Simpson 1980), and currently its distribution ranges from 35° south to 43° north (Feijó et al. 2018; Taulman and Robbins 2014), living across

Fig. 7.1 Nine-banded armadillo, *Dasypus novemcinctus*. Photograph by Guillermo Ferraris, used with permission

a broad diversity of habitats, such as savannas, dry forests, and rain forests (Feijó et al. 2018). Six *D. novemcinctus* morphological subspecies have been described across its distribution (McBee and Baker 1982; Wetzel et al. 2007). Arteaga et al. (2012) reported two divergent mitochondrial lineages of nine-banded armadillos in Mexico and the southeastern USA. Whereas the mtDNA data showed high genetic divergence between the lineages (4.7%), nuclear data indicated that there is evidence of contemporary gene flow. Genetic data of the North American populations as well as individuals from French Guiana do not fit completely with the morphological classification (Huchon et al. 1999; Arteaga et al. 2012; Gibb et al. 2016; Feijó et al. 2019).

Because the nine-banded armadillo has a wide geographical and environmental distribution and originated in South America (Simpson 1980), we consider it an appropriate model species to explore the effect which historical and geologic events have had on its genetic diversity pattern so we can better understand how lineages are geographically distributed across a continental-wide scale. By correlating the distribution of lineages across the geographical range of the species with different climates, we may predict which lineage can be more tolerant or sensitive to rapid environmental changes. This information provides us with the tools to make inferences about the effects which current climatic changes have on the genetic diversity

of this species. This, in turn, can be used for the appropriate management of the species as a whole by helping to define long-term policies for its conservation.

In this study, we aimed to specifically address the following questions: (i) How many mitochondrial lineages are there across the entire range of *D. novemcinctus*?; (ii) was the Pleistocene an important period of expansion for this species?; and (iii) are there differences in the climate distribution of mitochondrial lineages? Finally, we compare mitochondrial lineages distribution with the described morphological subspecies, and we discuss recommendations for its conservation.

2 Methods

2.1 Sample Collection, PCR Amplification, and Sequencing

We collected tissue samples of 222 *D. novemcinctus* from 153 localities in 16 countries from South, Central, and North America, in accordance with the guidelines of the American Society of Mammalogy. Seventy-four samples were collected in the field from wild populations, which consisted of a small piece of ear (3 mm^2) cut from each individual. One hundred and forty-eight samples were collected from museum specimens. In the collections, from each armadillo specimen, we cut a small piece (5 mm^2) of dry ventral tissue using sterile surgical scissors and stored it in a sterile plastic bag. Next, in the laboratory, in order to clean and hydrate the tissues, we immersed each piece in a tube with PBS solution at 1% (8 g of NaCl, 0.2 g of KCl, 1.44 g of NaHPO, and 0.24 g of KH2PO4 with a pH of 7.4), and we placed the tubes in a thermomixer for 24–72 h at 56°C. We changed the PBS solution of each tube at least once during that time. This procedure was done in a work table not used for the manipulation of polymerase chain reaction (PCR) products. Total genomic DNA was extracted from each sample using a DNeasy Kit (Qiagen, Inc., Valencia, CA, USA) and then stored at 4°C.

Via PCR, we amplified a segment of the mitochondrial control region: A fragment size of 377 bp was amplified using D2 (5´-ATTTYGGCGCTATGTAATTCG-3´; F. Delsuc, Université Montpellier II) and R1 (5´- GGTTGCTGGTTTCTCGGGAGTTGG-3´; designed in our laboratory), and another fragment of 228 bp was amplified using Fw1 (5´-CTACGCACAGACATAACGAGTACC-3´; designed in our laboratory) and N4 (5'-GGCATAAGTCCATCGAGATGTC-3´; designed in our laboratory). A region of 93 bp was overlapped between both fragments. For each reaction, 2.5 U of *Taq* polymerase was added per 25 μl of reaction volume. The final concentration was 1X buffer, 0.4 μM of each primer, 0.15 mM of dNTP, and 2 mM of $MgCl_2$. The thermal profile for amplification consisted of an initial denaturation cycle at 95°C for 3 min, followed by 30 cycles at 94°C for 30 s, 62°C for 45 s, and 72°C for 120 s, and a final extension at 72°C for 10 min. Reactions using both sets of primers were run at the same previously described conditions. Amplifications were performed in

a Perkin-Elmer GeneAmp PCR system 9600 (Applied Biosystems, Foster City, CA, USA). The sequences were aligned using the muscle algorithm in Geneious v6.1.6 (Drummond et al. 2010), and the final alignment was edited visually.

2.2 *Diversity, Genetic Structure, and Estimation of Demographic Expansion*

We used mitochondrial data to explore the phylogeographic structure by means of a phylogenetic tree and haplotype network. In order to increase the geographical sampling, haplotypes obtained by Huchon et al. (1999) from French Guyana ($n = 12$) and the USA ($n = 8$) were added to the analysis. Based on the Akaike information criteria (AIC) in JModeltest 0.1.1 (Posada 2008), we determined the evolution model for the control region.

Phylogenetic relationships among haplotypes were inferred using the likelihood method in RaxML 7.0.3 (Stamatakis 2006). Furthermore, we estimated the genealogical relationships among haplotypes using the median-joining algorithm of NETWORK 4.516 (http://www.fluxus-engineering.com). The median-joining method uses a maximum parsimony approach to search for all the shortest, least complex phylogenetic trees from a given dataset (Bandelt et al. 1999). When internal node haplotypes are not sampled, the median-joining method provides the best estimate of the true genealogy (Cassens et al. 2005).

Estimates of nucleotide diversity (π), haplotype diversity (h), and number of haplotypes were obtained in DnaSP 5.0 (Librado and Rozas 2009). Also, these estimators were obtained for each of the major mitochondrial lineages identified in this study (see results). We employed an analysis of molecular variance (AMOVA) (Excoffier et al. 1992) to assess the distribution of mtDNA variation. AMOVA was performed with ARLEQUIN 3.11 (Excoffier et al. 2005) by partitioning the total sum of the squares into components representing variations among the lineages and within lineages. In addition, we estimated pairwise F_{ST}, and their significance was evaluated by comparing the observed value with the distribution of the values obtained from 10,000 random permutations.

We estimated the historical demographic trends of the major mitochondrial lineages identified in this study. We applied a coalescent-based method known as the Bayesian skyline plot (BSP) (Drummond et al. 2005), as implemented in Beast 1.8.4 (Drumond et al. 2012; Drummond and Rambaut 2007). From a sample of gene sequences, given a specified nucleotide-substitution model (Drummond et al. 2005), the BSP model uses standard Markov chain Monte Carlo (MCMC) sampling procedures to estimate the posterior distribution of effective population size through time. The BSP constructs a model of demographic history based on how the number of coalescent events over a given interval differs from that expected under a neutral model for a panmictic population (Crandall et al. 2008). The BSP was performed for each lineage using the complete dataset of haplotypes with the HKY + I + G model

(see Results). We implemented the uncorrelated lognormal relaxed clock model of rate variation (Drummond et al. 2007), and the mutation rate considered was 1.9×10^8 (Arteaga et al. 2012). The defaults for other priors were used, and the program was repeatedly run optimizing the scale factors of the a priori function. MCMC tests were run for 5×10^7 steps and sampled every 5000 steps. Convergence of the chains and effective sample size of each parameter were evaluated with the Tracer 1.7 program (Rambaut et al. 2018).

Tajima's D (Tajima 1989), Fu and Li's (Fu and Li 1993), and Fu's (Fu 1997) neutrality tests were computed for each lineage with ARLEQUIN 3.11 (Excoffier et al. 2005). Negative values are expected in populations which have undergone demographic expansions, while positive values are expected in those which have recently experienced bottlenecks. Values are expected to be near zero in stable population sizes.

2.3 *Lineage Distribution Relative to Climate and Elevation Data*

To explore the distribution of the mitochondrial lineages across different climates, we collected climate data from 156 unique *D. novemcinctus* localities using the geographical information system (GIS). Locality information came from GPS devices for wild-caught individuals, from georeferenced museum specimens (153 localities recorded by the authors), and from a previous publication (three localities recorded by Huchon et al. (1999)). The GIS data included bioclimatic variables (http://worldclim.org/bioclim), which have a resolution of 1 km^2 and describe aspects of temperature, precipitation, seasonality, as well as the potentially biologically limiting extremes of these variables (e.g., Bio6, minimum temperature of the coldest month).

Using ArcMap, we first extracted GIS data from 19 bioclimatic variables layers (*Bio1–Bio19*) at armadillo occurrence points. Using JMP 5.01 software (SAS, Cary, NJ, USA), we conducted a correlation test to remove highly correlated variables, which could bias subsequent analyses (Graham et al. 2004). Specifically, if two variables showed a correlation coefficient higher than 0.75, we considered them highly correlated, and for each pair of correlated variables, we selected the variable which was more temporally inclusive (e.g., preferring mean temperature over mean temperature of the driest quarter) or those likely to be most relevant to armadillos (e.g., temperature variables, given that armadillos are poor thermoregulators (McNab 1980)) and their distribution might be partially limited by environmental temperatures). The remaining six climate variables included mean diurnal range of temperature, minimum temperature of the warmest month, minimum temperature of the coldest month, precipitation of the driest month, precipitation seasonality, and precipitation of the warmest quarter. Finally, from an altitude layer with a resolution of 1 km^2, we also obtained altitude values from the 156 unique localities.

To determine if the lineages are occurring in areas with different climatic conditions, we used a MANOVA. To preserve their interpretability and test their individual importance, we separately analyzed the environmental variables, rather than processing them through their principal components (e.g., Hof et al. 2010). We characterized the climatic distribution of each lineage by calculating the mean and the maximum and minimum values of the six environmental variables (e.g., Olalla-Tarragán et al. 2011). We also performed an ANOVA to explore if lineages occur across different altitudes. Data from lineage I were not included in these analyses, nor were they used for the lineage distribution model, because we only had one georeferenced locality for this lineage (haplotypes were obtained from GenBank; Huchon et al. 1999).

Geographically explicit predictions of climatic niches can often provide a good starting point for exploring regions of suitable climatic conditions, so we generated distribution models for the lineages (lineage distribution models: LDMs). LDMs estimate the relationship between lineage records at sites and the environmental and/or spatial characteristics of those sites (Peterson 2001). The projection of these models to the past climatic conditions allows explore changes in the distribution of the areas with suitable climatic features for each lineage, considering the assumption that the climatic niche which is being modeled has been conserved through time (Peterson 2003). Data of the Last Glacial Maximum (LGM), at ca. 21 kya, were derived from a dataset representing simulations run using one general circulation model, the Community Climate System Model (CCSM3; Collins et al. 2006). Data layers of current and past condition were collected from WorldClim (www.worldclim.org).

We constructed LDMs for lineages II, III, and IV using MAXENT 3.2.1 (Phillips et al. 2006). The LDMs and projections to the past conditions were created using the six bioclimatic variables layers selected. Individual environmental variables, rather than the axes from the principal component analysis (PCA), were used to construct the models in order to know the relative contribution of each variable to the LDMs. These models generate geographic predictions of lineage distributions based on known sampling points (Peterson 2001). MAXENT 3.2.1 implements a maximum entropy algorithm to model species distributions using information on lineage occurrences and environmental variables. It generates LDMs using only present records. We used the default convergence threshold (10^{-5}) and maximum number of iteration (500) values (Pearson et al. 2007), using 25% of the localities for model training and 75% for model testing. To assess model performance, we used the area under the receiver operating characteristic curve (AUC) (Mertz 1978) as a measure of overall classification accuracy (capacity to discriminate between occupied and unoccupied records). The AUC can vary from 0.5, indicating no discrimination capacity, to 1, indicating perfect discrimination capacity. MAXENT produces probability predictions of occurrence (Phillips et al. 2006).

3 Results

3.1 Diversity, Genetic Structure, and Demographic Expansion

Control region sequences data (387 bp) analyzed consists of 242 *D. novemcinctus* (222 individuals collected by the authors and 20 sequences recorded by Huchon et al. (1999)). Ninety-five variable nucleotide sites were found, defining a total of 127 haplotypes (Appendix 1). Based on AIC, the mitochondrial DNA (mtDNA) control region best fits an HKY + I + G model of molecular evolution.

The phylogenetic tree and median-joining network separated the haplotypes into four lineages, with disjunct geographical distribution (Fig. 7.2). Lineage I is restricted

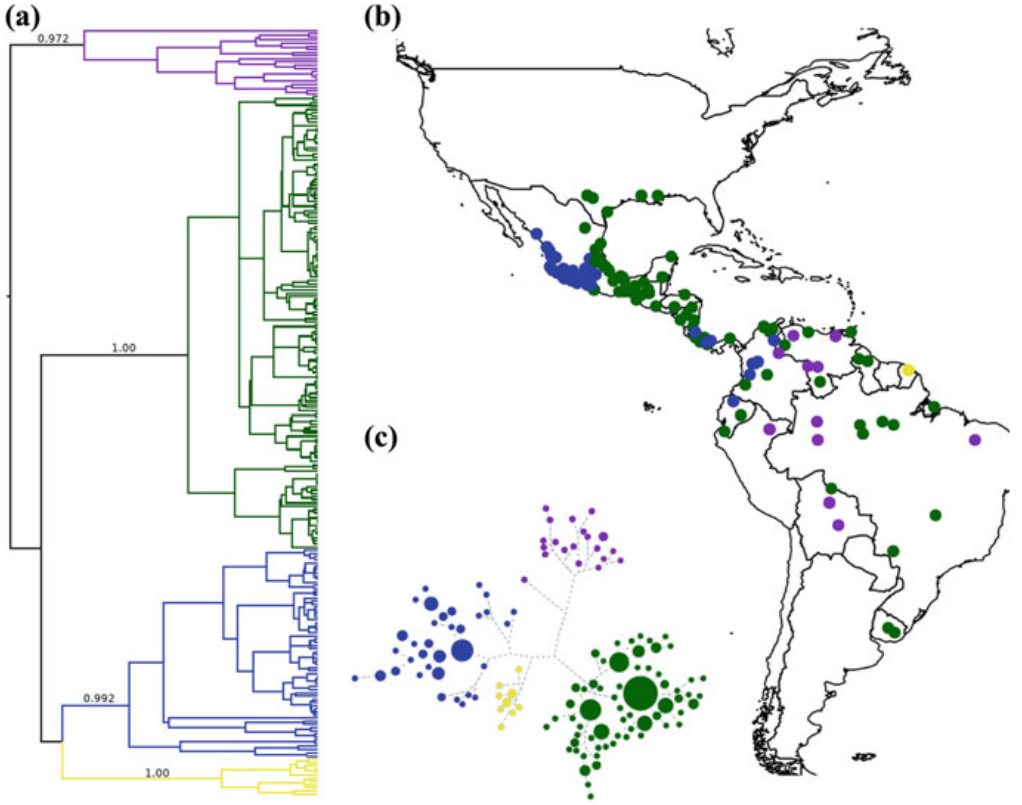

Fig. 7.2 **a** Maximum likelihood phylogenetic tree of *D. novemcinctus* haplotypes based on 387 bp of the mitochondrial control region showing the four major lineages. Numbers above the branches indicate maximum likelihood bootstrap higher than 0.8. **b** Spatial location of the complete dataset of samples analyzed. **c** Median-joining networks based on 387 bp of the mitochondrial control region showing the four major lineages. Circle size is proportional to haplotype frequencies; dots represent the number of mutational steps between haplotypes. In the tree, the map, and the network, the colors indicate the samples from each lineage detected: lineage I: yellow; lineage II: purple; lineage III: blue; and lineage IV: green

Table 7.1 Analyzed sample sizes, number of haplotypes, haplotype diversity (h), nucleotide diversity (π), and neutrality tests of the four armadillo lineages, based on 387 bp of the mitochondrial control region

	Sample size	N. haplotypes	h	π	Tajima's D test	Fu and Li test	Fu´s test
Lineage I	12	10	0.969	0.009	−1.127	−1.427	−4.364*
Lineage II	21	20	0.980	0.025	−0.82	−0.889*	−4.842*
Lineage III	68	37	0.960	0.021	−1.147	−2.878	−14.314*
Lineage IV	141	62	0.918	0.014	−2.762*	−0.758	−39.229*
Total	242	127	0.966	0.030	−2.102*	0.423	−67.144*

*$P < 0.005$

to French Guiana, lineage II occurs only in South America, and lineages III and IV occupy South, Central, and North America, but lineage III occurs on the west side of the continent, while lineage IV has a wider distribution. Internal branches of the phylogenetic tree which join haplotypes as lineages have strong support, but external branches show weak support (Fig. 7.2).

High haplotype and nucleotide diversities were recorded for all datasets and for each lineage (Table 7.1). The AMOVA showed that most of the genetic variance was found among lineages (64.46%; $\varphi_{ST} = 0.644$, $p < 0.001$), whereas within-lineage variance was relatively low (35.54%). Pairwise F_{ST} values among lineages were high (ranging from 0.23 to 0.39), except between lineages I and III (Appendix 2).

According to the BSPs, demographic expansions of lineages II, III, and IV began approximately 300 kya, and also lineage IV had a higher and more recent population growth (Fig. 7.3). In contrast, lineage I, restricted to French Guiana, showed a slight expansion which began 150 kya (Fig. 7.3). Tajima's D, Fu and Li's, and Fu's neutrality tests supported demographic expansion in all lineages, as the values were negative in all cases, and in many cases, they were significant (Table 7.1).

3.2 Lineage Distribution Relative to Climate and Elevation Data

From the 156 unique *D. novemcinctus* localities used to collect GIS data on climate, 11 localities are from lineage II, 51 are from lineage III, and 94 are from lineage IV. Lineage I was not included in the analyses because we had only one georeferenced locality (haplotypes were obtained from GenBank, Huchon et al. 1999).

The climatic distribution of the three lineages is wide (Fig. 7.4), and the MANOVA showed that climatic conditions of the areas occupied by them are significantly different (Wilks' lambda value = 0.488, approx. $F = 17.173$; $p < 0.0001$). We also

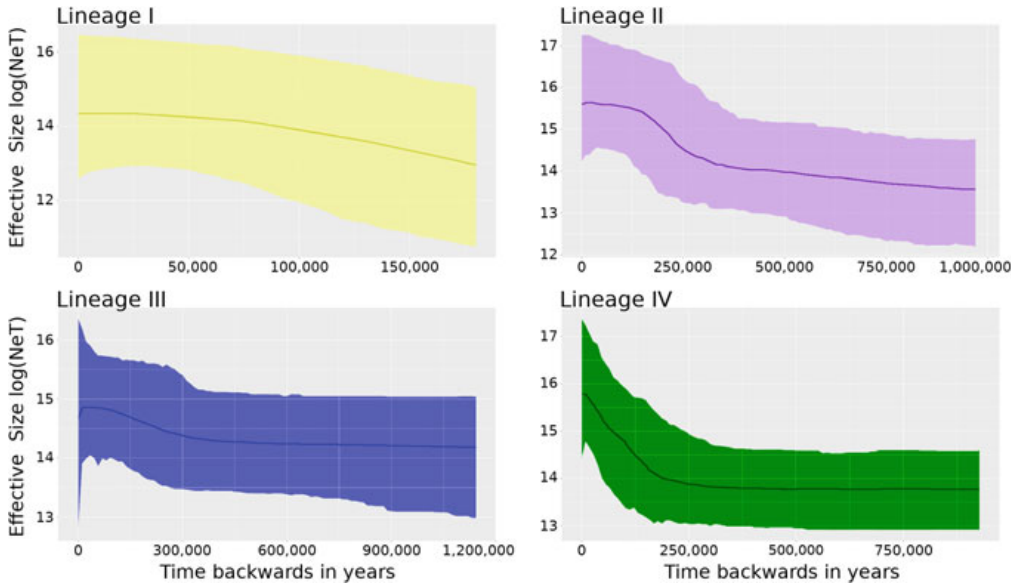

Fig. 7.3 Bayesian skyline plots showing the historical demographic trends for the four mitochondrial lineages detected. Along the *y*-axis, the estimated population sizes are expressed in units of $Ne\tau$, the product of the effective population size per generation length. Solid lines are median estimates, whereas shaded areas represent confidence intervals

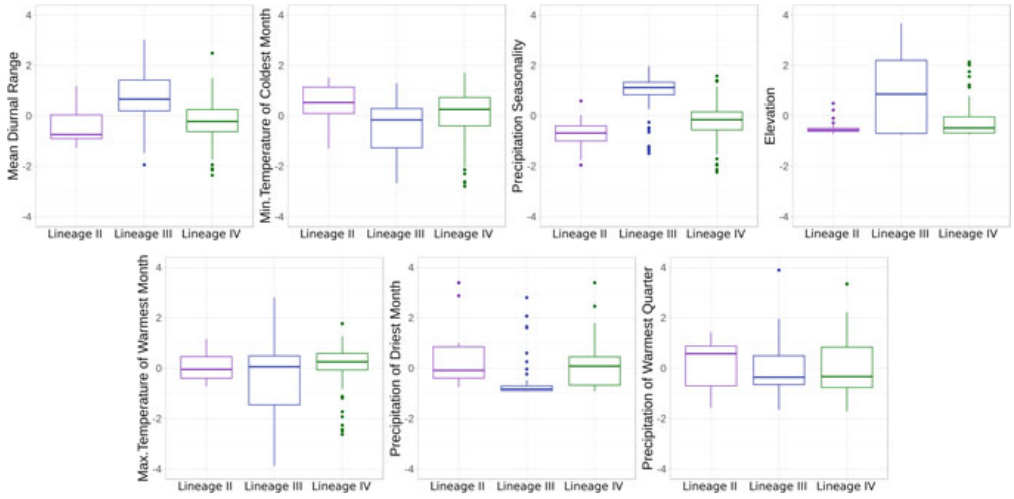

Fig. 7.4 Boxplot chart of the variance (*y*-axis) of each climatic variable and the altitude. The mean of each variable was standardized to zero, and the values are presented in the variance unit. The bigger boxplot represents high tolerance

found a significant difference in the altitudinal distribution of lineages ($F = 39.852$; $p < 0.0001$). While lineage II, which is restricted to South America, is distributed under 1000 m above sea level (masl), the other two lineages reach higher altitudes, and, in particular, lineage III occurs at the highest altitudes across the species range, reaching up to 2853 masl.

The LDMs show a parapatric distribution pattern (Fig. 7.5). The minimum temperature of the coldest month was the most relevant environmental variable for predicting the climatic distribution of lineages II (89.1%) and IV (48.7%), and precipitation seasonality (61.3%) was the most important variable for predicting the distribution of lineage III. The predictive power of occurrence was high for the three models (the AUC of the test data were 0.870, 0.898, and 0.906 for lineages II, III, IV, respectively). The projections to the LGM show changes in the distribution of suitable conditions between these two times (Fig. 7.5). Comparing both times, in the present, there is a reduction in areas with suitable conditions for lineage III, particularly in Colombia and Ecuador. On the other hand, for lineages II and IV, the models suggest an increase in areas, specifically in South America and the southeastern part of USA for lineage IV, and in eastern Brazil for lineage II (Fig. 7.5).

4 Discussion

This is the first genetic study of *D. novemcinctus* across its complete geographical range. We found four highly divergent and diverse mitochondrial lineages. Two lines of evidence support the existence of these lineages: the high support and long length of the internal branches of the phylogenetic tree, and the distribution of the haplotypes in the network. The lineages show, in general, a parapatric distributions, with two lineages occurring in South, Central, and North America (lineages III and IV), while the other two are restricted to South America (lineages I and II). Currently, the highest genetic richness of the nine-banded armadillo is the north of South America, with the four lineages distributed there, and we find three of these lineages specifically in the northwestern region (Fig. 7.2). Our results show a wider geographic distribution of lineages III and IV, which were recorded previously in North America (Arteaga et al. 2012). The finding of several lineages in the geographic distribution of *D. novemcinctus* fits with our expectation for widely distributed species. Future studies using higher genomic covering and a more even sampling could better determine the relationship among the armadillo's lineages and estimate the divergence time between them.

4.1 Demographic Expansion and Climatic Distribution

In the middle of the Pleistocene, strong demographic expansions occurred in three lineages, while the French Guiana lineage showed a very slight demographic expansion in the last 150 kya. This is consistent with its very localized distribution. Although several warm/cool oscillations characterized the middle of the Pleistocene (Hewitt 2011), it seems that they did not negatively affect armadillo populations, and the recent demographic expansions could be related with the geographical range expansion which occurred during the colonization processes of new areas. Specifically,

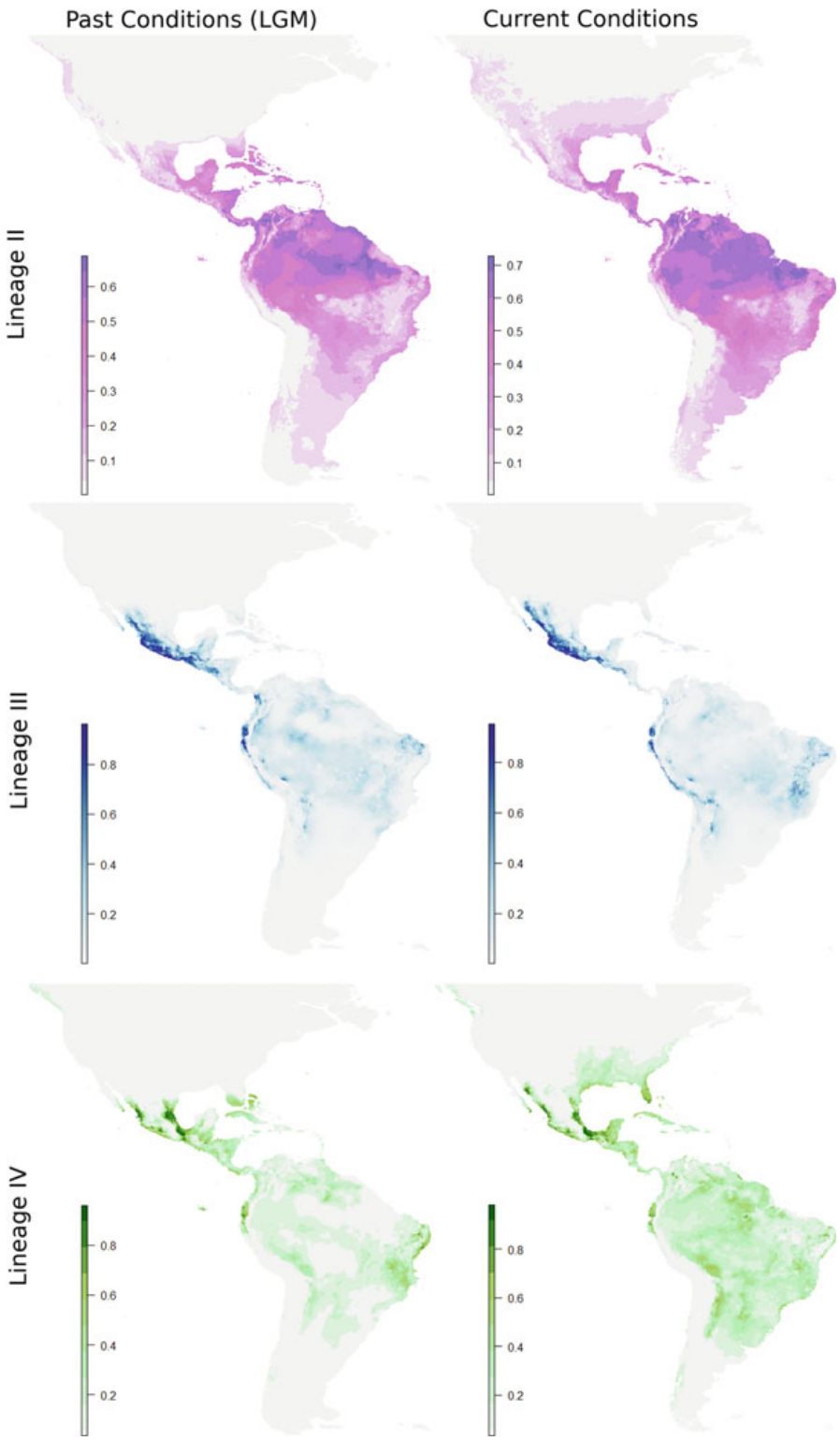

Fig. 7.5 LDMs considering the current climatic conditions and the Last Glacial Maximum (LGM) conditions. High levels of predicted habitat suitability are indicated by darker colors

range expansions of lineages III and IV could have happened during the GABI. Lineage IV colonized more northern latitudes and more recently founded the populations in the USA. The same lineage reached the highest latitudes to the south, being currently the most widely distributed lineage inside the geographical range of the nine-banded armadillo. This fact is congruent with the largest population size detected in this lineage, compared to the others. Since the LGM, ca. 21 kya, population sizes of three lineages have been stable (Fig. 7.3), despite that the lineages have experienced dramatic changes in available suitable areas up to the present time (Fig. 7.5).

Perhaps, the capacity of the species to colonize a diversity of habitats, such as forests or more open areas such as savannas, also allowed populations to be resilient to climatic changes. For instance, although habitat distribution changed as a consequence of climatic oscillation, armadillo populations could occupy the same areas with the new environmental conditions. This hypothesis is partially supported by the wide climatic distribution observed in the three analyzed lineages. The resiliency which these lineages apparently showed to past climatic changes can occur in the face of current climatic changes. At the same time, these current changes can benefit populations inhabiting specific areas inside the geographic range. For example, the northern limit of the range of the species, corresponding to lineage IV, is still expanding (Feng and Papeş 2015; Freeman and Genoways 1998; Hofman 2009; Taulman and Robbins 2014). As the environmental temperature increases, individuals will colonize areas which were not climatically available in the past.

The most important variable predicting the distribution of lineages II and IV was minimum temperature of the coldest month. This makes sense when we consider that the nine-banded armadillo has a low capacity for thermoregulation (McNab 1980), and its distribution is significantly restricted by low temperatures. In accordance with this, Taulman and Robbins (1996, 2014) indicated that the increase of the geographic range of this species in the northern area of its distribution is limited by temperature and rain. On the other hand, precipitation seasonality was the most important variable predicting the distribution of lineage III, which is found primarily in tropical seasonal forests. The seasonality of precipitation is highly related with the productivity of this biome and is characterized by soils rich in microinvertebrates. Perhaps, the distribution of lineage III in tropical seasonal forests was mainly driven by the availability of food to satisfy their insectivorous diet (Redford 1985).

The broad geographic distribution of the nine-banded armadillo across the American continent covers an extensive range of latitudes and has a wide range of climatic conditions. We explored the climatic distribution of the lineages, but we did not test if the lineages show a difference in climatic niche because it has already been reported that at least two of the lineages show niche conservatism (Arteaga et al. 2011). The idea behind testing differences among climatic conditions is to explore how the lineages are distributed across the complete range of the species, and a cause–effect pattern is not established with our approach.

The parapatric distribution of the lineages is probably more a consequence of the historical events than a cause of the difference in climatic affinities. Physical barriers can maintain the allopatric pattern of some regions. It is interesting that we did not

observe lineage II individuals to the west of the Andes. Because this lineage was recorded up to 800 masl, some high mountains in the Andes probably work as a barrier to dispersal, while lineage IV individuals apparently are able to cross them. Similarly, lineage III also shows a wide altitudinal range, and we expect that the Andes does not act as a barrier; however, we did not observe lineage III individuals to the east side of the Andes. This is possibly more related to the lack of available areas with suitable climatic conditions, as was observed in the distribution model for this lineage. The presence of individuals from two or three lineages in the same region, such as the northwest of South America and the north of Costa Rica, supports the fact that climatic niche affinities do not maintain the parapatric distribution of the mitochondrial lineages.

4.2 Taxonomical Approach

The geographic distribution of genetic mitochondrial lineages does not conform to the subspecies distribution of the nine-banded armadillo described based on morphological characteristics (McBee and Baker 1982; Wetzel et al. 2008). The four subspecies described in Central and South America are included inside the geographical range of the four genetic lineages, but boundaries among subspecies and lineages are not the same. For instance, in the geographical range of the subspecies *D. novemcinctus novemcinctus*, the haplotypes of lineages I, II, and III are included. Similarly, the haplotypes of lineages II, III, and IV are present in the geographical range of the subspecies *D. n. fenestratus*. In contrast, in North America, there are two described subspecies, *D. n. mexicanus* found in eastern Mexico and the USA and *D. n. davisi* which is restricted to western Mexico, occurring from the Balsas Basin north to Morelos (McBee and Baker 1982; Wetzel et al. 2008). As mentioned by Arteaga et al. (2012), the allopatric distribution of the lineages in Mexico corresponds to the geographic distribution of these two subspecies; nevertheless, the presence of both lineages in Central America and Colombia suggests that the morphological subspecies may occur in a wider geographical area.

Congruence among information from nuclear and mitochondrial markers should be used to evaluate extant subspecies classifications and to confirm that the mitochondrial lineages may potentially represent different species. However, strong mitochondrial patterns, such as those detected in this study, can be very useful to propose a specific area of research which is needed in order to solve taxonomical problems. The results from our genetic analysis provide tools for planning future studies to determine if each mitochondrial lineage is also recovered as monophyletic groups using nuclear data. Widely distributed species, such as the nine-banded armadillo, which occur across very different ecological conditions are good study models to answering ecological, evolutionary, and systematic questions. Exploring the distribution and level of genetic diversity shows the high potential of this model to test different historical and ecological hypotheses. Our extensive sampling allowed us to examine, in detail, the distribution of mtDNA variation in the nine-banded armadillo,

one of the most widely distributed mammals in America, providing information about structure within its geographical range.

4.3 Conservation Recommendations and Management

The population sizes of the lineages seem to be stable. Although some populations can be small locally because of hunting and increased habitat fragmentation (Escamilla et al. 2000), others are actively expanding, particularly the ones in the southern USA. The recent expansion of the armadillo's geographical range in the USA is a result of the life history of the species, its human-mediated introduction, and the areas with suitable conditions for it to thrive (Feng and Papeş 2015; Taulman and Robbins 2014).

The unique restricted lineage, lineage I, which is distributed in French Guiana, is clearly the only one which should be a concern for conservation at this moment. More sampling efforts should be concentrated in localities in and near French Guiana to better define the distribution of this differentiated haplotype group.

Detailed ecological and population genetics analysis using genomic methods will also be needed to analyze the conservation genetics of the species (e.g., Castellanos-Morales et al. 2016). These data, coupled with more detailed morphological (e.g., Billet et al. 2017; Feijó et al. 2018; Hautier et al. 2017) and climatic analysis, will help to better define the taxonomic status of the found lineages, to determine if they correspond to different subspecies, or even to different species, and to analyze them for local climatic adaptations.

Acknowledgements We wish to thank all the people and institutions that provided biological samples for this study (Appendix 3). We thank the Dirección General de Vida Silvestre for the permits to collect samples of armadillos. We are grateful to Juan Pablo Jaramillo and Daniel Piñero for valuable comments which improved the manuscript. MCA is grateful to Sigma Xi and The American Society of Mammalogists for research grants. This chapter was written during a sabbatical leave of LEE at the University of Minnesota, Department of Plant and Microbial Biology, in Peter Tiffin's laboratory, with support from a scholarship from PASPA, DGAPA, and UNAM.

Appendix 1

Control region haplotypes identified in 222 *D. novemcinctus* samples from South, Central, and North American countries.

Haplotype ID	GenBank accession number	Number of samples	Countries from samples origin
H1	JN602096	1	México
H2	JN602097	2	México
H3	JN602098	1	México
H4	JN602099	1	México
H5	JN602100	2	México
H6	JN602101	1	México
H7	JN602102	12	México
H8	JN602169	1	México
H9	JN602103	1	México
H10	JN602104	2	México
H11	JN602105	4	México
H12	JN602106	6	México
H13	JN602107	1	México
H14	JN602108	1	México
H15	JN602109	11	México
H16	JN602110	3	México
H17	JN602111	6	México
H18	JN602112	1	México
H19	JN602113	1	México
H20	JN602114	2	México
H21	JN602115	2	México
H22	JN602116	2	México
H23	JN602117	2	México
H24	JN602118	2	México
H25	JN602119	1	México
H26	JN602120	1	México
H27	JN602121	1	México
H28	JN602122	1	México
H29	JN602123	1	México
H30	JN602124	2	México
H31	JN602125	1	México
H32	JN602170	3	México
H33	JN602126	1	México
H34	JN602127	1	México
H35	JN602128	1	México
H36	JN602129	2	México
H37	JN602130	5	México

(continued)

(continued)

Haplotype ID	GenBank accession number	Number of samples	Countries from samples origin
H38	JN602131	3	México
H39	JN602132	1	México
H40	JN602133	3	México
H41	JN602134	1	México
H42	JN602135	1	México
H43	JN602136	1	México
H44	JN602137	1	México
H45	JN602138	1	México
H46	JN602139	2	México
H47	JN602140	1	México
H48	JN602172	4	México
H49	JN602141	2	México
H50	JN602142	1	México
H51	JN602143	2	México
H52	JN602144	1	México
H53	JN602145	1	México
H54	JN602146	26	México, Colombia, Brazil, Honduras, Costa Rica, Nicaragua, Panamá, Venezuela, Guyana.
H55	JN602147	1	México
H56	JN602148	4	México
H57	JN602149	1	México
H58	JN602150	1	México
H59	JN602151	1	México
H60	JN602152	1	México
H61	JN602153	1	México
H62	JN602154	1	México
H63	JN602155	1	México
H64	JN602156	1	México
H65	JN602157	1	México
H66	JN602158	1	México
H67	JN602159	1	México
H68	JN602160	1	México
H69	JN602161	3	México, USA
H70	JN602162	1	México
H71	JN602163	1	Costa Rica

(continued)

(continued)

Haplotype ID	GenBank accession number	Number of samples	Countries from samples origin
H72	JN602164	1	Costa Rica
H73	JN602165	1	Nicaragua
H74	JN602171	1	Colombia
H75	JN602166	1	Colombia
H76	JN602167	1	Colombia
H77	JN602168	2	Colombia
H78	JN602173	2	Colombia, Panamá
H79	JN602174	1	Colombia
H80	JN602175	1	Nicaragua
H81	JN602176	1	Guatemala
Dn1	MH760470	1	Ecuador
Dn2	MH760471	1	Ecuador
Dn3	MH760472	1	Venezuela
Dn4	MH760473	1	Panamá
Dn5	MH760474	1	Guyana
Dn6	MH760475	1	Venezuela
Dn7	MH760476	1	Venezuela
Dn8	MH760477	1	Brazil
Dn9	MH760478	1	Venezuela
Dn10	MH760479	1	Venezuela
Dn11	MH760480	1	Panamá
Dn12	MH760481	1	Panamá
Dn13	MH760482	1	Bolivia
Dn14	MH760483	2	Bolivia
Dn15	MH760484	1	Bolivia
Dn16	MH760485	1	Bolivia
Dn17	MH760486	1	Perú
Dn18	MH760487	1	Perú
Dn19	MH760488	1	Perú
Dn20	MH760489	1	Perú
Dn21	MH760490	1	Venezuela
Dn22	MH760491	1	Venezuela
Dn23	MH760492	1	Bolivia
Dn24	MH760493	1	Bolivia
Dn25	MH760494	1	Brazil
Dn26	MH760495	2	Brazil, Trinidad and Tobago

(continued)

(continued)

Haplotype ID	GenBank accession number	Number of samples	Countries from samples origin
Dn27	MH760496	1	Venezuela
Dn28	MH760497	1	Perú
Dn29	MH760498	1	Venezuela
Dn30	MH760499	6	Brazil, Ecuador, Colombia, Nicaragua
Dn31	MH760500	1	Uruguay
Dn32	MH760501	1	Guyana
Dn33	MH760502	1	Ecuador
Dn34	MH760503	1	Brazil
Dn35	MH760503	4	Brazil, Uruguay
Dn36	MH760504	2	Perú, Brazil

Haplotypes 1-81 are recorded in Arteaga et al. 2012. There are 10 haplotypes in the 12 samples from French Guiana and two haplotypes in the eight samples from USA (Huchon et al. 1999).

Appendix 2

Pair-Fst between mitochondrial lineages of *D. novemcinctus.*

	Lineage I	Lineage II	Lineage III
lineage I			
lineage II	0.232*		
lineage III	−0.032	0.390*	
lineage IV	0.349*	0.283*	0.311*

*$P < 0.001$

Appendix 3

Dasypus novemcinctus samples obtained from different museums.

Collection names	Collection numbers of specimens	Total samples
Mammal collection of American Museum of Natural History, USA (AMNH)	26,007, 171,919, 171,921, 207,420, 24,053, 24,054, 176,676, 176,675, 7278, 182,075, 14,663, 33,148, 15,463, 139,318, 98,462, 99,173, 99,174, 247,662, 255,867, 255,866, 255,868,255,863, 211,674, 211,668, 98,463, 99,172, 28,422, 15,461, 136,252, 127,562, 128,131, 144,828, 78,517, 33,151, 24,230, 126,136, 18,890, 46,555, 67,710, 7519, 205,727, 205,726, 42,914, 98,465, 133,267, 91,709, 93,117, 143,013, 91,707, 133,328, 95,113, 209,944, 133,274, 133,330	54
Colección Zoológica Regional, Instituto de Historia Natural y Ecología, México (CZRMA)	1581, 2255, 43, 50	4
Colección de la Reserva de la Biósfera Los Tuxtlas, México	a, b	2
Colección Mastozoológica del Sureste de México, México (ECOSUR-SC)	52, 575, 980, 1270, 1570, 1572	6
Colección de Mamíferos del Museo de Zoología “Alfonso L. Herrera”, México (MZFC)	4254, 5078, 5079, 5075, 4900, 5077	6
Colección Mastozoológica HMAN Instituto Tecnológico Agropecuario de Hidalgo, México (HMAM)	540, 544, 536, 543, 547, 534, 541, 535, 545, 531, 539, 537, 546, 532, 548	15
Colección Nacional de Mamíferos, Instituto de Biología, UNAM, México (IBUNAM)	1153, 3592, 10,069, 11,535, 17,037, 16,559, 31,600, 43,121, 16,520, 15,583, 27,275, 16,558, 14,528, 43,122, 14,527, 37,069, 16,551, 16,496	18
Colección de mamíferos, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, México (ENCB)	40,868, 16,019, 21,096, 39,110, 35,629, 15,099, 26,577, 21,669, 26,067, 36,004, 34,214	11
Colección de Mamíferos de la Universidad Michoacana San Nicolás de Hidalgo, México (UMSNH)	1323, 394, 887, 1983, c	4

(continued)

(continued)

Collection names	Collection numbers of specimens	Total samples
Mammal collection of Smithsonian National Museum of Natural History, USA (USMN)	553,928, 281,290, 281,285, 281,291, 281,288, 337,563, 442,787, 406,709, 406,706, 442,799, 296,615, 259,485, 406,703, 406,704, 374,827, 338,783, 337,718, 361,234, 11,383, 250,321, 305,595, 339,669, 339,668, 337,564, 442,790, 578,437, 324,266, 324,265	28
TOTAL		148

a, b, and c are specimens without assigned numbers

References

Antonelli A, Nylander JAA, Persson C et al (2009) Tracing the impact of the Andean uplift on Neotropical plant evolution. Proc Natl Acad Sci USA 106:9749–9754

Arteaga MC, McCormack JE, Eguiarte LE et al (2011) Genetic admixture in multidimensional environmental space: asymmetrical niche similarity promotes gene flow in armadillos (*Dasypus novemcinctus)*. Evolution 65:2470–2480

Arteaga MC, Piñero D, Eguiarte LE et al (2012) Genetic structure and diversity of the nine-banded armadillo in Mexico. J Mammal 93:547–559

Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Billet G, Hautier L, De Thoisy B et al (2017) The hidden anatomy of paranasal sinuses reveals biogeographically distinct morphotypes in the nine-banded armadillo (*Dasypus novemcinctus*). PeerJ 5:e3593

Brumfield RT, Capparella AP (1996) Historical diversification of birds in northwestern South America: a molecular perspective on the role of vicariant events. Evolution 50:1607–1624

Cassens I, Mardulyn P, Milinkovitch MC (2005) Evaluating intraspecific "network" construction methods using simulated sequence data: do existing algorithms outperform the global maximum parsimony approach? Syst Biol 54:363–372

Castellanos-Morales G, Gutiérrez-Guerrero YT, Gámez N et al (2016) Use of molecular and environmental analyses for integrated in situ and ex situ conservation: The case of the Mexican prairie dog. Biol Conserv 204:284–295

Coates AG, Collins LS, Aubry MP et al (2004) The geology of the Darien, Panama, and the late Miocene-Pliocene collision of the Panama arc with northwestern South America. Geol Soc Am Bull 116:1327–1344

Collins WD, Bitz CM, Blackmon ML et al (2006) The community climate system model version 3 (CCSM3). J Clim 19:2122–2143

Costa L (2003) The historical bridge between the Amazon and the Atlantic Forest of Brazil: a study of molecular phylogeography with small mammals. J Biogeogr 30:71–86

Crandall ED, Frey MA, Grosberg RK et al (2008) Contrasting demographic history and phylogeographical patterns in two Indo-Pacific gastropods. Mol Ecol 17:611–626

Ditchfield AD (2000) The comparative phylogeography of Neotropical mammals: patterns of intraspecific mitochondrial DNA variation among bats contrasted to nonvolant small mammals. Mol Ecol 9:1307–1318

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214

Drummond AJ, Rambaut A, Shapiro B et al (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol 22:1185–1192
Drummond AJ, Ashton B, Buxton S et al (2010) Geneious v6.1.6 http://www.geneious.com/. Accessed 20 May 2017
Drummond AJ, Suchard MA, Xie D et al (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29:1969–1973
Escamilla A, Sanvincente M, Sosa M et al (2000) Habitat mosaic, wildlife availability and hunting in the tropical forest of Calakmul, Mexico. Conserv Biol 14:1592–1601
Eizirik E, Kim JH, Menotti-Raymond M et al (2001) Phylogeography, population history and conservation genetics of jaguars (*Panthera onca*, Mammalia, Felidae). Mol Ecol 10:65–79
Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distance among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479–491
Excoffier L, Laval G, Schneider S (2005) Arlequin v3.0: An integrated software package for population genetics data analysis. Evol Bioinform 1:47–50
Feijó A, Patterson BD, Cordeiro-Estrela P (2018) Taxonomic revision of the long-nosed armadillos, Genus Dasypus Linnaeus, 1758 (Mammalia, Cingulata). PLoSOne 13(4):e0195084
Feijó A, Vilela JF, Cheng J et al (2019) Phylogeny and molecular species delimitation of long-nosed armadillos (Dasypus: Cingulata) supports morphology-based taxonomy. Zool J Linnean Soc 186:813–825
Feng X, Papeş M (2015) Ecological niche modelling confirms potential north-east range expansion of the nine-banded armadillo (Dasypus novemcinctus) in the USA. J Biogeogr 42:803–807
Freeman PW, Genoways HH (1998) Recent northern records of the nine-banded armadillo (Dasypodidae) in Nebraska. Southwest Nat 43:491–495
Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. Genet 133:693-709
Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915–925
Gibbard P, Van Kolfschoten T (2004) The Pleistocene and Holocene epochs. In: Gradstein F, Ogg J, Smith A (eds) A geologic time scale 2004. Cambridge University Press, Cambridge, UK, pp 441–452
Gibb GC, Condamine FL, Kuch M et al (2016) Shotgun mitogenomics provides a reference phylogenetic framework and timescale for living Xenarthrans. Mol Biol Evol 33:621–642
Graham CH, Ron SR, Santos JC et al (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in Dendrobatid frogs. Evolution 58:1781–1793
Hautier L, Billet G, De Thoisy B et al (2017) Beyond the carapace: skull shape variation and morphological systematics of long-nosed armadillos (genus Dasypus). PeerJ 5:e3650
Hewitt GM (2011) Quaternary phylogeography: the roots of the hybrid zones. Genetica 139:617–638
Hof C, Rahbek C, Araújo MB (2010) Phylogenetic signals in the climatic niches of the world's amphibians. Ecography 33:242–250
Hofmann JE (2009) Records of nine-banded armadillos, *Dasypus novemcinctus*, in Illinois. Trans Ill State Acad Sci 102:95–106
Hoorn C, Wesselingh FP, Ter Steege H et al (2010) Amazonia through time: andean uplift, climate change, landscape evolution, and biodiversity. Sci 330:927–931
Huchon D, Delsuc F, Catzeflis FM et al (1999) Armadillos exhibit less genetic polymorphism in North America than in South America: nuclear and mitochondrial data confirm a founder effect in *Dasypus novemcinctus* (Xenarthra). Mol Ecol 8:1743–1748
Lessa EP, Cook JA, Patton JL (2003) Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. Proc Natl Acad Sci USA 100:10331–10334
Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452
Lovejoy NR, Albert JS, Crampton WGR (2006) Miocene marine incursions and marine/freshwater transitions: evidence from Neotropical fishes. J South Am Earth Sci 21:5–13

Marshall LG, Webb SD, Sepkoski JJ et al (1982) Mammalian evolution and the Great American Interchange. Science 215:1351–1357
McBee K, Baker RJ (1982) Dasypus novemcinctus. Mamm Species 162:1–9
McNab BK (1980) Energetics and the limits to a temperate distribution in armadillos J Mammal 61:606–627
Mertz CE (1978) Basic principles in ROC analysis. Semin Nucl Med 8:283–298
Olalla-Tárraga MA, Mclnnes L, Bini LM et al (2011) Climatic niche conservatism and the evolutionary dynamics in species range boundaries: global congruence across mammals and amphibians. J Biogeogr 38:2237–2247
Pascual R, Ortiz-Jaureguizar E (2007) The Gondwanan and South American episodes: two major and unrelated moments in the history of the South American mammals. J Mamm Evol 14:75–137
Patton JL, Da Silva MNF, Malcolm JR (2000) Mammals of the rio Juruá and the evolutionary and ecological diversification of Amazonia. Bull. Am. Mus. Nat. Hist. 244:1–306
Pearson RG, Raxworthy CJ, Nakamura M et al (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J Biogeogr 34:102–117
Peterson AT (2001) Predicting species geographic distributions based on ecological niche modelling. Condor 103:599–605
Peterson AT (2003) Predicting the geography of species' invasions via ecological niche modelling. Q Rev Biol 78:419–433
Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modelling of species geographic distributions. Ecol Modell 190:231–259
Posada D (2008) JModelTest: Phylogenetic model averaging. Mol Biol Evol 25:1253–1256
Poulsen CJ, Jeffery ML (2011) Climate change imprinting on stable isotopic compositions of high-elevation meteoric water cloaks past surface elevations of major orogens. Geology 39:595–598
Rambaut A, Drummond AJ, Xie D et al (2018) Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Syst Biol 67:901–904
Redford KH (1985) Food habits of armadillos (Xenarthra: Dasypodidae). In: Montgomery GG (ed) The evolution and ecology of armadillos, sloths, and vermilinguas. Smithsonian Institute Press, Washington, pp 429–437
Simpson GG (1980) Splendid Isolation. The Curious History of South American Mammals. Yale University Press, New Haven, Connecticut
Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690
Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595
Taulman JF, Robbins LW (1996) Recent range expansion and distributional limits of the nine-banded armadillo *(Dasypus novemcinctus)* in the United States. J Biogeogr 23:635–648
Taulman JF, Robbins LW (2014) Range expansion and distributional limits of the nine-banded armadillo in the United States: an update of Taulman & Robbins (1996). J Biogeogr 41:1626–1630
Tchaicka L, Eizirik E, De Oliveira TG et al (2007) Phylogeography and population history of the crab-eating fox (*Cerdocyon thous*). Mol Ecol 16:819–838
Webb SD (1976) Mammalian faunal dynamics of the Great American Interchange. Paleobiology 2:220–234
Weir JT, Price M (2011) Andean uplift promotes lowland speciation through vicariance and dispersal in Dendrocincla woodcreepers. Mol Ecol 20:4550–4563
Wetzel RM, Gardner AL, Redford KH et al (2008) Order Cingulata. In: Gardner AL, Creighton G (eds) Mammals of South America. Marsupials, Xenarthrans, Shrews, and Bats, vol 1. The University of Chicago Press, Chicago, pp 128–156