Phylogeographic and diversification patterns of the white-nosed coati (Nasua narica): Evidence for south-to-north colonization of North America

Sergio F. Nigenda-Morales⁎, Matthew E. Gompperb, David Valenzuela-Galván⁵, Anna R. Layd, Karen M. Kapheime, Christine Hassf, Susan D. Booth-Binczikg, Gerald A. Binczikh, Ben T. Hirschi, Maureen McColginj, John L. Koprowskik, Katherine McFaddendl, Robert K. Wayn, Klaus-Peter Koepfl

⁎ Corresponding authors at: National Laboratory of Genomics for Biodiversity, Center for Research and Advance Studies, Irapuato, Guanajuato 36821, Mexico (S.F. Nigenda-Morales). Smithsonian Conservation Biology Institute, National Zoological Park, Washington, D.C. 20008, USA (K.-P. Koepfl).

E-mail addresses: snigenda@ucla.edu (S.F. Nigenda-Morales), koepfl@si.edu (K.-P. Koepfl).

1 Deceased.

ARTICLE INFO

Keywords:
Procyonidae
Mitochondrial haplotypes
Microsatellites
Great American Biotic Interchange
Biogeographical models
Evolutionary history

ABSTRACT

White-nosed coatis (Nasua narica) are widely distributed throughout North, Central, and South America, but the patterns of temporal and spatial diversification that have contributed to this distribution are unknown. In addition, the biogeographic history of procyonid species in the Americas remains contentious. Using sequences from three mitochondrial loci (Cytochrome b, NAHD5 and 16S rRNA; 2201 bp) and genotypes from 11 microsatellite loci, we analyzed genetic diversity to determine phylogeographic patterns, genetic structure, divergence times, and gene flow among Nasua narica populations throughout the majority of the species’ range. We also estimated the ancestral geographic range of N. narica and other procyonid species. We found a high degree of genetic structure and divergence among populations that conform to five evolutionarily significant units. The most southerly distributed population (Panama) branched off much earlier (∼3.8 million years ago) than the northern populations (< 1.2 million years ago). Estimated gene flow among populations was low and mostly northwards and westwards. The phylogeographic patterns within N. narica are associated with geographic barriers and habitat shifts likely caused by Pliocene-Pleistocene climate oscillations. Significantly, our findings suggest the dispersal of N. narica was south-to-north beginning in the Pliocene, not in the opposite direction during the Pleistocene as suggested by the fossil record, and that the most recent common ancestor for coati species was most likely distributed in South or Central America six million years ago. Our study implies the possibility that the diversification of Nasua species, and other extant procyonid lineages, may have occurred in South America.

https://doi.org/10.1016/j.ympev.2018.11.011
Received 23 November 2017; Received in revised form 12 November 2018; Accepted 16 November 2018
Available online 20 November 2018
1055-7903/ © 2018 Elsevier Inc. All rights reserved.
1. Introduction

Phylogeographic studies illuminate the historical geography of genetic lineages and are essential to understand the evolutionary process of lineage diversification (Riddle et al., 2008; Hickerson et al., 2010). One of the regions with a paucity of phylogeographic studies is Middle America, the area lying between the United States and South America (Winker, 2011). Because of its biological diversity, variety of environments, dynamic climate, complex geological and biogeographic history, Middle America is considered an exceptional region where biological processes including lineage diversification, dispersion, speciation and extinction occur at high rates (Marshall, 2007; Hardy et al., 2013; Gutiérrez-García and Vázquez-Domínguez, 2013). Within Middle America, Central America has been important in shaping the present biodiversity in both North and South America (Woodburne et al., 2006; Webb, 2006; Morgan, 2008) due to its pivotal role in the Great American Biotic Interchange (GABI). This major intercontinental migration event was facilitated by the emergence of the Isthmus of Panama, which allowed previously isolated taxonomic lineages to migrate from North America to South America and vice versa across the land bridge, thereby forever altering the evolutionary histories of both continents (Marshall et al., 1982; Stehli and Webb, 1985; Webb, 2006). Phylogeographic research in Central America has allowed the juxtaposition of the biogeographic history of a variety of taxa with the complex geological events and paleoenvironmental changes that have occurred over the last several million years, thereby increasing our understanding of the origins of Neotropical biodiversity (Daza et al., 2016; Gutiérrez-García and Vázquez-Domínguez, 2013; Bagley and Johnson, 2014).

Procyonids (Mammalia: Carnivora: Procyonidae) are a group of mammals that include the olingos (Bassaricyon spp.), ringtails (Bassariscus spp.), raccoons (Procyon spp.), coatis (Nasua spp. and Nasuella spp.) and kinkajous (Potos flavus) (Nowak, 2005). The diversification and biogeographic history of procyonids in the New World is controversial and the group has figured prominently in understanding the history of the GABI (Keoepfl et al., 2007, Eizirk, 2012; Soibelzon and Prevosti, 2013; Forasiepi et al., 2014). The fossil record indicates that they dispersed from North America into South America two separate times and were among the very first groups of North American mammals to colonize South America. The first dispersion event occurred in the Late Miocene (5–7.3 Mya) with the appearance of the fossil genus Cyonasa in South America, long before the closure of the Isthmus of Panama and the major migration events of the GABI approximately 2.4 to 2.8 Mya according to the standard GABI model (Marshall et al., 1979; Webb, 1985; Webb, 2006; Woodburne, 2010). All descendants from that first colonization apparently went extinct by the end of the Middle Pleistocene (Marshall, 1985; Soibelzon and Prevosti, 2013). The second dispersion of procyonids into South America is thought to be the one made by the ancestors of the extant genera during the last major pulse of the GABI in the late Pleistocene after 0.125 Mya (sensu Woodburne, 2010). Because of the gap in the fossil record of over 500 ky (thousand years) between the extinct procyonids and the appearance of the extant species (Soibelzon and Prevosti, 2013; Forasiepi et al., 2014), the living species are not considered to be descendants of the procyonids that originally invaded South America (Marshall, 1985; Webb, 1985; Baskin, 2004; Soibelzon, 2011). However, studies based on molecular data found that the diversification within the extant genera Nasua and Procyon occurred in the Middle to Late Miocene, temporally coincident with the diversification of the extant genera in South America (Koopfl et al., 2007; Eizirk et al., 2010).

Among coatis, four extant species are currently recognized: the white-nosed coati (Nasua narica), the South American or brown-nosed coati (Nasua nasua), the Western Mountain Coati (Nasuaa olivacea) and the Eastern Mountain Coati (Nasuaa meridensis) (Helgen et al., 2009; Kays, 2009). Nasua narica is the only coati species distributed in North, Central and South America, from Arizona and New Mexico in the United States (U.S.) to Northern Colombia (Gompper, 1995). Across its range, four subspecies are recognized based on body size, cranial features and coat coloration differences, although these descriptions are based on limited quantitative analysis (Hall, 1981; Decker, 1991; Gompper, 1995). The distribution of the subspecies is roughly delimited by geographic barriers and associated with ecological differences: N. n. molaris is distributed north of the Transmexican volcanic belt (TMVB), N. n. narica is distributed south of TMVB, N. n. yucatanica is restricted to the Yucatan Peninsula and N. n. nelsoni is confined to Cozumel Island in Mexico (Hall, 1981; Decker, 1991; Gompper, 1995). The validity of these subspecies (particularly N. n. nelsoni, which is sometimes designated as a distinct species; McFadden et al., 2008), and the genetic structure among N. narica populations, have only been analyzed and tested in a limited part of their distribution (Silva-Caballero et al., 2017). Given the morphological, ecological and geographical differences observed across the wide distribution range of N. narica, phylogeographic analysis could help illuminate the patterns of temporal and spatial diversification of this species.

Due to the prominent role of procyonids at the beginning of the GABI, and the distribution of N. narica in North, Central and South America, a phylogeographic study of this species may provide insight into the contrasting models proposed for the evolution of the Isthmus of Panama and the GABI. The standard model of the GABI assumes the near or complete emergence of the Isthmus of Panama circa 3.0–3.5 million years ago (Mya) in the Middle Pliocene, resulting in the closing of the Central American Seaway (CAS) (Coates and Obando, 1996; Coates and Stallard, 2013; O’Dea et al., 2016). After the emergence of the Isthmus, the fossil record indicates mammalian lineages predominantly migrated south, to colonize South America 2.4–2.8 Mya (Simpson, 1980; Webb, 2006; Woodburne, 2010). An alternative model proposes the appearance of a land bridge and the closure of the CAS 13–15 Mya during the Middle Miocene (Farris et al., 2011; Montes et al., 2012a, 2012b; 2015; Carrillo et al., 2015). Concordant with this model, a recent study proposed that the most significant periods of migration of terrestrial taxa, including mammals, occurred at ca. 20 and 6 Mya, with similar migration rates between North and South America, and that asymmetric migration emerged after 6 Mya, with higher migration from South to North America (Bonac et al., 2015; but see Lessios, 2015; Marko et al., 2015). This model also suggests that environmental processes, not geological features, might be responsible for preventing faunal dispersal (Montes et al., 2015; Bacon et al., 2016). Concordantly, analyses of first appearances of faunal assemblages in South America suggest that faunal migrations associated with GABI began 10 Mya (Carrillo et al., 2015).

Here, we present the first phylogeographic study of N. narica throughout most of its range using sequences from three mitochondrial loci and 11 nuclear microsatellite loci. Our main objective was to test hypotheses concerning the directionality and timing of dispersal for N. narica and the processes driving the divergence of populations, specifically whether a north-to-south dispersal after 3.0 Mya or a south-to-north dispersal timed earlier than 3.0 Mya better explains the biogeographic history of the species. Additionally, we assessed the relative influence of climatic and geological processes on genetic structure. To test these hypotheses we: (1) determined the genetic structure and phylogenetic relationships of N. narica populations throughout most of its range; (2) assessed the geographical and temporal patterns of diversification of populations and related them to geological or climatic processes as well as in the context of comparative phylogeographic studies of other co-distributed taxa; (3) calculated the gene flow between populations to evaluate possible directions of historical population movements; and (4) inferred ancestral distribution areas for populations and tested different dispersal models between geographic regions. Our results have important and novel implications for the diversification and evolutionary history of procyonids in the Americas.
16S rRNA. The tree log likelihood score = -7735.34 (ML) and -7913.75 (BI). The five main clades of *N. narica* are represented by colored rectangles. Each rectangle represents a haplotype and the height of the rectangles denotes the frequency of that particular haplotype. The general sampling locations and number of individuals sampled (in parentheses) are indicated. Numbers shown at nodes are bootstrap support values based on 1000 replicates (ML) and posterior clade probabilities (BI), respectively. Asterisks indicate 100% bootstrap and 1.0 posterior clade probability support. The tree was rooted using *Bassaricyon medius* and *B. alleni*. The scale bar indicates the number of substitutions per site.

2. Materials and methods

2.1. Tissue collection and DNA extraction

We assembled a collection of 85 white-nosed coati samples from throughout most of the range of the species, except for localities in northwestern Colombia (Fig. 1A). Our samples include representatives of the four subspecies of *N. narica* that were collected from different tissues (i.e., whole blood, ear punches, muscle) obtained from animals caught in the field, roadkills, zoo and museum specimens (Table A.1). In addition, we obtained samples from nine South American coatis (*N. nasua*), two Western Lowland olingos (*Bassaricyon medius*) and one Eastern lowland olingo (*B. alleni*) (Table A.1). Genomic DNA was extracted using a QIAamp DNA Mini Kit (Qiagen, Valencia, CA) following the manufacturer's protocol.

2.2. Mitochondrial sequences analyses

2.2.1. Mitochondrial amplification and sequencing

We amplified and sequenced the complete cytochrome-b (*CYTB*) gene and partial regions of the NADH5 and 16S rRNA genes. The *CYTB* gene was amplified in two overlapping segments in two separate polymerase chain reactions (PCRs) using primers L14724 and H15513 (reaction 1) and L15612 and H15915 (reaction 2) (Irwin et al., 1991). The NADH5 and 16S segments were amplified using primers ND5-DF1 and ND5-DR1 (Trigo et al., 2008) and L3259 and H3652 (Sorensen et al., 1999), respectively. PCRs were set up in a volume of 50 µL containing 1 µL of DNA (~0.5 µg), 5 µL 10x PCR buffer, 5 µL of 25 mM MgCl2, 1 µL 10 mM dNTP mix, 1 µL of 25 µM/µL forward and reverse primers, 0.3 µL Taq polymerase (Sigma-Aldrich, St. Louis, MO), and 35.7 µL sterile double-distilled water. Amplifications were conducted in an MWG-Biotech Primus 96 Plus Thermal Cycler (Eurofins Genomics, Huntsville, AL) under the following cycling conditions: 30 cycles of 94 °C for 30 s, 50 or 52 °C for 30 s, 72 °C for 45 s; one cycle of 72 °C for 5 min; and a hold at 4 °C. All PCRs were run with a negative control. Electrophoresis was used to visualize the PCR products in a 1% agarose gel stained with ethidium bromide that included a 100 bp DNA ladder (Promega, Madison, WI). Amplification products were purified using either an Ultra Clean Kit (Mo Bio Laboratories, Carlsbad, CA) or with Exonuclease I and Shrimp Alkaline Phosphatase (Exo-SAP, A Investigator). Cycle sequencing products were purified and then sequenced on a 96-capillary 3130xl DNA Analyzer (Life Technologies, Grand Island, NY) at the UCLA DNA and Genotyping Core Facility. We inspected, edited and assembled forward and reverse sequence reads using Sequencher 3.1 (Gene Codes Corporation, Ann Arbor, MI) or Geneious Pro v7.1.4 (Biomatters Ltd., Auckland, New Zealand; Kearse et al. 2012). CYTB and NADH5 sequences were...
translated into amino acid sequences to verify orthology and exclude the potential presence of NUMTs (nuclear-mitochondrial paralogues), for which none were detected. Sequences of CYTB and NADHS used in a previous study (Koepfli et al., 2007) were downloaded from Genbank and included in the analyses (Table A.1; Bassaricyon alleni (DQ660299 and DQ660230), B. medius (DQ660300 and DQ660231; originally classified as B. gabbii but see Helgen et al., 2013), Nasua narica (DQ660302 and DQ660234, from Barro Colorado Island, Panama), and N. nasua (DQ660303 and DQ660235, from Santa Cruz, Bolivia).

2.2.2. Alignment and phylogenetic analyses

Sequences from each mitochondrial locus were aligned using MAFFT v7.017 (Katoh et al., 2002) in the Geneious v7.1.4 package (Kearse et al., 2012) with the following parameter settings: scoring matrix = 200, PAM/k = 2, gap open penalty = 1.53, offset value = 0.123, and then adjusted by eye. The three alignments were concatenated to generate a data matrix totaling 2201 bp. This alignment contained a total of 97 sequences (N. narica = 85; N. nasua = 9; Bassaricyon alleni = 1; B. medius = 2). We used jModelTest v2.1.4 (Darriba et al., 2012) to select the best-fitting model of DNA substitution using the Bayesian information criterion (BIC) and a starting tree estimated with BIJON (Gascuel, 1997). Phylogenetic trees were estimated using maximum likelihood (ML) and Bayesian inference (BI). The ML tree was reconstructed using raxmlGUI 1.3.1 (Silvestro and Paganelli, 2018) and the auto optimize setting enabled in the Operators window. For the mitochondrial-2 file, all parameter settings were the same as for Mitochondrial-1 except that we used a fossil-based prior set to a log normal distribution with mean = 7.2 Mya and standard deviation = 1.7 Mya, based on the estimated time of the split between the Bassaricyon and Nasua lineages (Eizirik et al., 2010; also see Helgen et al., 2013; and the auto optimize setting enabled in the Operators window. For the mitochondrial-2 file, all parameter settings were the same as for Mitochondrial-1 except that we used a fossil-based prior set to a log normal distribution with mean = 1.75 Mya and standard deviation = 0.15 Mya, to calibrate the split between N. narica and N. nasua, based on the earliest appearance of this genus in Hemphillian/Early Blancan North American Land Mammal Ages (Dalquest, 1978; Baskin, 1982; Emmert and Short, 2018). For both XML files, two independent MCMC analyses were run for 20,000,000 generations, trees and parameters sampled every 2000 generations and the first 10% of these discarded as burn-in. An additional XML file was run for each data set-calibration prior combination without sequence data so that only the priors were sampled. Tracer (Rambaut et al., 2014) was used to inspect the posterior distributions of tree likelihoods, substitution and clock parameters and showed ESS values > 200 for each run of the Mitochondrial-1, Mitochondrial-2 files. The post-burn-in samples of the posterior distribution from the two independent runs of each file were merged using LogCombiner (Bouckaert et al., 2014) and FigTree (Rambaut, 2014) was used to visualize the maximum clade credibility topology and mean node heights.

2.2.2.2. Divergence time estimation

We used BEAST v2.3.1 (Bouckaert et al., 2014) to estimate the divergence times among the phylogroups within N. narica, employing the data matrix containing unique haplotypes from N. narica, N. nasua and the two Bassaricyon species (the Nasuella olivacea sequences were excluded). We set up and generated two different XML files within the BEAUti application, labeled Mitochondrial-1 and Mitochondrial-2, that differed in the calibration priors that were applied. For the Mitochondrial-1 file, the following parameters and settings were employed: HKY + I model of DNA substitution (estimated with jModelTest as above), empirical base frequencies and proportion of invariable sites = Estimate; strict clock model; tree prior = coalescent constant population; gamma distribution (0.001–1000) set for clock rate prior and a 1/X distribution for the population size prior; calibration prior with a truncated normal mean = 7.2 Mya and standard deviation = 1.7 Mya, based on the estimated time of the split between the Bassaricyon and Nasua lineages (Eizirik et al., 2010; also see Helgen et al., 2013; and the auto optimize setting enabled in the Operators window. For the Mitochondrial-2 file, all parameter settings were the same as for Mitochondrial-1 except that we used a fossil-based prior set to a log normal distribution with mean = 1.75 Mya and standard deviation = 0.15 Mya, to calibrate the split between N. narica and N. nasua, based on the earliest appearance of this genus in Hemphillian/Early Blancan North American Land Mammal Ages (Dalquest, 1978; Baskin, 1982; Emmert and Short, 2018). For both XML files, two independent MCMC analyses were run for 20,000,000 generations, trees and parameters sampled every 2000 generations and the first 10% of these discarded as burn-in. An additional XML file was run for each data set-calibration prior combination without sequence data so that only the priors were sampled. Tracer (Rambaut et al., 2014) was used to inspect the posterior distributions of tree likelihoods, substitution and clock parameters and showed ESS values > 200 for each run of the Mitochondrial-1, Mitochondrial-2 files. The post-burn-in samples of the posterior distribution from the two independent runs of each file were merged using LogCombiner (Bouckaert et al., 2014) and FigTree (Rambaut, 2014) was used to visualize the maximum clade credibility topology and mean node heights.

2.2.4. Haplotype genetic diversity and structure

We estimated the genetic diversity among concatenated mtDNA haplotypes by calculating nucleotide and haplotype diversity with ARLEQUIN v3.5.2.2 (Excoffier and Lischer, 2010). Genetic structure was tested using an analysis of molecular variance (AMOVA) with the Tamura-Nei model (which is the closest to HKY + I) to estimate the amount of genetic variation partitioned among and within populations. To further investigate the relationships between mitochondrial haplotypes and haplogroups identified with the AMOVA, we constructed an unrooted median-joining haplotype network with NETWORK v4.6.1.1 (Bandelt et al., 1999) using default parameters.

2.2.5. Estimation of number of migrants

We estimated potential past gene flow and determined the number of migrants per generation with the coalescent-based program MIGRATE-N v3.6.11 (Beerli and Felsenstein, 2001; Beerli, 2006; Beerli and Palczewski, 2010) using the concatenated mtDNA data. We first ran 10 short chains and 4 long chains using a maximum likelihood approach to obtain a better approximation for the theta and migration parameters. The estimates of these parameters were then used as priors for the final Bayesian inference analysis, for which we used the Brownian mutation model, mutation rate estimated from the data, uniform data distribution, a long chain of 5,000,000 steps with four replications, a
burn-in length of 10,000 steps and a static heating scheme with four chains (1, 1.5, 3 and 1 × 10^5).

2.2.6. Biogeographical range reconstruction and dispersal hypothesis testing

To test the biogeographic history of *N. narica* and other procyonid species, we determined the most probable ancestral distribution range at each node of the unique-haplotype phylogenetic tree using two event-based methods with different assumptions, the statistical dispersal-vicariance analysis (S-DIVA; *Yu et al., 2010*) and the Bayesian binary MCMC approach (BBM; *Ronquist & Huelsenbeck, 2003*) implemented in RASP v.3.2 (*Yu et al., 2015*). For the S-DIVA analysis, we used 1000 random trees from the 10,000 BEAST output trees and the BEAST maximum clade probability tree. The BBM analysis was done over the maximum clade probability tree and run for 50,000 cycles and 10 chains, sampling every 100 cycles, discarding 100 trees and setting the evolutionary model to F81 + G (which is the closest to HKY + I) with default parameters. For both analyses, we defined seven geographic regions based on geographic barriers found in previous studies, which limit the distribution and dispersal of vertebrate species and populations (*Helgen et al., 2013; Gutiérrez-García and Vázquez-Domínguez, 2013; Bagley and Johnson, 2014*): North America (NA); North of the Sierra Madre Occidental, western Mexico (WM; western part of TMVB and northwest of Sierra Madre del Sur), central Mexico (CM; from TMVB to the Isthmus of Tehuantepec), northern Central America (NCA; from the Isthmus of Tehuantepec to the Nicaragua depression), middle Central America (MCA; from the Nicaragua depression to the Talamanca mountain range), Panama (PAN; between Talamanca range and the Darién region) and South America (SA; from the Darién region southwards). Both analyses were done allowing a combination of two adjacent areas (i.e. NA + WM, WM + CM, CM + NCA, NCA + MCA, MCA + PAN, PAN + SA).

We implemented a maximum-likelihood dispersal-extinction cladogenesis (DEC) analysis with Lagrange (*Roe and Smith, 2006*) to test three different models for the dispersal of *N. narica* across the Isthmus of Panama. We used the maximum clade credibility tree obtained from the phylogenetic analysis in BEAST, the seven geographic regions previously defined and the maximum size of ancestral ranges restricted to two adjacent areas. The first model (M1) was unconstrained, in which there is no dispersal constraint between geographic regions and assumes the Isthmus of Panama was already present 9.5 Mya, which is the estimated age of the oldest node of our tree (see Results). The second model (M2) also assumes a 9.5 mya presence of the Panamanian Isthmus but defines higher dispersal constraints between regions that are further apart than between those that are more proximate. The third model (M3) reflects the rise of the Isthmus of Panama 3.0 Mya, (i.e., no dispersion between South America and the rest of the regions, and limited dispersal between Panama and all other Central and North America regions before 3.0 Mya) and dispersion restrictions set for after this event. The global maximum likelihood for each model was calculated and compared to obtain the most likely scenario given our data.

2.3. Microsatellites analyses

2.3.1. Microsatellite amplification and genotyping

We successfully amplified 11 microsatellite loci (Table A.2) in 85 *N. narica* individuals that were previously described for *N. nasua* (*Almany et al., 2009*). Loci were amplified in 10 µL reaction volumes using approximately 10−100 ng of genomic DNA on a Peltier Thermal-Cycler (MJ Research PTC-200). The PCR conditions consisted of 1.0 µL of primer mix (0.01 µM forward primer, 0.01 µM dye-labeled M13 primer, 0.2 µM reverse primer.), 0.4 µL 10 ng/ml BSA, 5.0 µL of QIAGEN Mastermix (Qiagen, Valencia, USA) and 2.1 µL of ddH2O. We used multiplex thermocycling profiles for dye-labeled primers and M13 hybrid primers as follows: 95°C for 15 min, 25 cycles at 94°C for 30 s, 55°C for 90 s and 72°C 60 s, followed by 20 cycles at 94°C for 30 s, 50°C for 90 s and 72°C for 60 s, plus a final extension of 60°C for 30 min. All PCR products were electrophoresed on an Applied Biosystems 3730XL DNA Analyzer. Allele sizes were scored automatically using Genemapper v3.7 (*Applied Biosystems, Foster City, USA*) and checked manually with reference to a size standard (LIZ 500). We tested for the presence of null alleles, large allelic dropout and stuttering in our data using MICROCHECKER (*Van Oosterhout et al., 2004*).

2.3.2. Genetic diversity, Hardy-Weinberg equilibrium and linkage disequilibrium

Genetic diversity was measured as allelic diversity in GENALEX v6.5 (*Peckall and Smouse, 2012*) and as the observed (Ho) and expected (He) heterozygosity using ARLEQUIN v3.5.2.2 (*Excoffier and Lischer, 2010*). Significant departure from Hardy-Weinberg equilibrium was tested using exact tests (for heterozygote excess or deficiency) for each locus with GENEPOP v4.5.1 (*Raymond and Rousset, 1995; Rousset, 2008*), which was also used to test for linkage disequilibrium (LD) among loci, applying a log-likelihood ratio test, with an adjusted p-value corresponding to alpha = 0.0009 after Bonferroni correction.

2.3.3. Genetic structure

We implemented a Bayesian clustering method using STRUCTURE v2.3.4 (*Pritchard et al., 2000; Falush et al., 2003*) to infer the number of genetic clusters (K) and assign individuals to clusters based on their multilocus genotypes, without a priori assumptions about sample location and assuming correlated allele frequencies and admixture ancestry between clusters. The K values from 1 through 10 were assessed using a burn-in period of 50,000 iterations and 500,000 sampling iterations for each K value. The stability of the clusters was evaluated using 10 independent runs per K value. The log likelihood values and ΔK parameter of *Evanno et al. (2005)*, calculated with STRUCTURE HARVESTER v0.6.94 (*Earl and vonHoldt, 2012*), were used to determine the most probable number of clusters. We used CLUMPP v1.2.1 (*Jakobsson and Rosenberg, 2007*) with the Greedy algorithm to account for the variation in admixture among individuals over the 10 replicate runs per K in STRUCTURE. We implemented an analysis of molecular variance (AMOVA) based on allele frequencies to examine the genetic variation between and within populations and to calculate the fixation index FST (*Weir and Cockerham, 1984*) between populations using ARLEQUIN (*Excoffier and Lischer, 2010*). AMOVA and FST estimations were calculated using permutation tests of 10,000 randomized runs. Finally, we generated a neighbor-joining tree of all individuals using the microsatellite loci genotypes and the DA genetic distance (*Nei et al., 1983; Takezaki and Nei, 1996*) implemented in POPULATIONS v1.2.30 (*Langella, 1999*).

2.3.4. Estimation of migration rates

Recent gene flow among populations identified with the population structure analysis described above were estimated using a Bayesian MCMC analysis of microsatellite genotypes in BAYESASS v3.0.4 (*Wilson and Rannala, 2003*). For this analysis, we used 20,000,000 iterations, a sampling frequency of 100, a burn-in length of 1,000,000 iterations, and delta values of 0.25, 0.65 and 0.80 for migration rate, allele frequency and level of inbreeding, respectively. To assess convergence of the results, four runs under the same parameters were done using different initial seed numbers.

3. Results

3.1. Mitochondrial sequences

3.1.1. Phylogenetic and divergence time analyses

Among the 85 *N. narica* samples sequenced for 2201 bp across three mitochondrial gene regions, we identified 21 haplotypes that differed by 1 to 237 substitutions (0.046–10.78% uncorrected p-distance; Table A.3). Four haplotypes differing by 2 to 122 substitutions (0.092–5.55% uncorrected p-distance) were observed among the 9 samples of South American regions before 3.0 Mya, (i.e. no dispersion between South America and the rest of the regions, and limited dispersal between Panama and all other Central and North America regions before 3.0 Mya) and dispersion restrictions set for after this event. The global maximum likelihood for each model was calculated and compared to obtain the most likely scenario given our data.
American coatis. *N. narica* differed from South American coatis by 274–314 substitutions (12.45–14.27% uncorrected p-distance; Table A.3). All new sequences were deposited in Genbank (accession numbers: MK135525 – MK135777; MK144297 - MK144326; Table A.1). The 85 *N. narica* sequences were assorted into five clades based on ML and BI phylogenetic analyses (Fig. 1B). Three haplotypes from 13 samples derived from several locations in Panama constituted the earliest branching lineage within the mitochondrial gene tree; these haplotypes were highly divergent from the remaining haplotypes (9.92–10.78% uncorrected p-distance; Table A.3). The middle clade, which we de-noted the Yucatan Peninsula-Guatemala clade, comprised 11 haplotypes from 37 samples collected in Belize, the Yucatan Peninsula region of Mexico, Cozumel Island, Guatemala and Costa Rica. Lastly, seven haplotypes defined three clades containing 35 samples from central Mexico (i.e. state of Morelos), western Mexico (i.e. state of Jalisco) and the southwestern U. S. (i.e. states of Arizona and New Mexico). Node support for clades and subclades across the mitochondrial gene tree was generally high based on bootstrap and posterior clade probability values, indicating a robust phylogenetic signal (retention index = 0.9709, as calculated in PAUP* v.4.0a152, Swofford, 2002).

ML and BI analyses that incorporated CYTB sequences of two *Nasua olivacea* individuals resulted in trees showing these were placed inside the *Nasua narica* clade, where they are joined together with the haplotypes from Panama with high node support (Fig. A.1). This arrangement makes *N. narica* paraphyletic, in addition to that of the genus *Nasua* (Helgen et al., 2009). For the CYTB sequences only, the *N. olivacea* sequences differed by 83–89 substitutions (7.6–7.8% uncorrected p-distances) and 112–130 substitutions (10.24–11.40% uncorrected p-distances) from the Panama haplotypes and remaining haplogroups of *N. narica* to the north, respectively.

Divergence times estimated from the two analytical schemes that differed in the calibration priors employed (Mitochondrial-1 and Mitochondrial-2) were generally congruent. The clade consisting of the Panama haplotypes first diverged ca. 4 Mya (95% highest posterior density [HPD] = 2.0–6.7 Mya and 2.6–5.1 Mya for the Mitochondrial-1 and Mitochondrial-2 analyses, respectively), whereas the other four clades split ca. 1.3 Mya (95% HPD = 0.59–2.1 Mya and 0.78–1.6 Mya for the Mitochondrial-1 and Mitochondrial-2 analyses, respectively; Fig. 2). The split separating the clade of the central Mexico haplotypes (Morelos) from the clade containing the southwestern U. S. + western Mexico (Jalisco) haplotypes occurred ca. 0.89 to 1.0 Mya (95% HPD = 0.462–1.61 Mya). The latter two clades diverged ca. 0.2 Mya (95% HPD = 0.109–0.439 Mya).

### 3.1.3. Biogeographic reconstruction

Although the S-DIVA and BBM biogeographic analyses gave somewhat different results, both analyses identified South America as an area of distribution for the most recent common ancestor (MRCA) of *Nasua* and *Bassaricyon*, while the S-DIVA analysis also identified Panama (i.e. South America + Panama) as a feasible distribution area. Similarly, these were the areas most likely inhabited by the MRCA of *N. nasua* and *N. narica*, and where the split between these species occurred 6 Mya (Fig. 5, Fig. A.2). The split between the Panama population and the rest of the populations within *N. narica* probably occurred in the middle part of Central America and Panama around 4 Mya (Fig. 5, Fig. A.2). The reconstructions of ancestral area using DEG agreed with these results; none of the three DEC models found evidence for areas in North America or Mexico as being part of the ancestral range of *Nasua* species (Fig. A.3). Moreover, among the three DEC models, model M2 had the highest log-likelihood value (lnL = –29.32) (Table A.6), which suggests that dispersion across the Isthmus of Panama likely occurred before 9.5 Mya and dispersion events have been limited between geographic areas further apart (e.g. South America and Mexico and the U.S.) (Fig. A.3). This model was 2.25 likelihood units higher than the M3 model, which represents the rise of the Panamanian isthmus 3.0 Mya and had the lowest log-likelihood value (lnL = –31.57; Table A.6, Fig. A.3).

### 3.2. Microsatellite analyses

#### 3.2.1. Genetic diversity

All 11 microsatellite loci were variable in *N. narica*, with the number of alleles per locus ranging from 3 to 17 (Table A.7). The loci were in Hardy-Weinberg equilibrium; all populations had one to three loci with heterozygote deficit, but no locus showed deviation from Hardy-Weinberg equilibrium across all populations (Table A.8). Also, no evidence of linkage disequilibrium, null alleles, allelic dropout or stuttering was found in our data.

#### 3.2.2. Population structure

The Bayesian clustering analysis showed that K = 5 was the number of clusters with the highest likelihood (–2601.14; Fig. 6A). The Evanno ΔK parameter showed two peaks, at K = 2 and K = 5 (Fig. 6B). At K = 2 the group including samples from the southwestern U.S. and western Mexico was separated from the group containing individuals from central Mexico southwards (Fig. 6C). Given that the ΔK statistic detects the highest level of genetic structure when various hierarchical levels exist (Evanno et al., 2005; Coulon et al., 2008), and the extensive geographic range of our taxon sampling, we tested for genetic structure at higher K values. K = 5 found the same structure detected in the AMOVA of the mtDNA sequences: southwestern U.S. (SWUS), western Mexico (WMEX), central Mexico (CMEX), Yucatan Peninsula-Guatemala (YUCP-GUAT) and Panama (PAN; Fig. 6C). Although K = 6 showed substructure within the Yucatan Peninsula-Guatemala cluster (Fig. 6C), when we ran this cluster alone in STRUCTURE no evidence of substructure was found (Fig. A.4). Therefore, we determined that K = 5 is the most probable number of genetic clusters in our sample. Furthermore, the neighbor-joining tree of allele-sharing distance clearly defined five clusters, with only three individuals mixed between them (Fig. 6D), two of which may represent second generation migrants according to the STRUCTURE plot (Fig. 6C).

Differences in allele frequency (FST) between the five genetic groups were significantly high in all pairwise comparisons, ranging from 0.101 to 0.328 (Table 2), suggesting low gene flow between populations. The lowest FST value was between the Morelos and Yucatan Peninsula populations, which are more distantly separated (1335 Km) than Morelos...
is from the Jalisco population (640 Km). This pattern might be explained by the admixture events between the Morelos and Yucatan-Guatemala populations we detected with the Bayesian clustering analysis and the allele-sharing tree (Fig. 6C; Fig. 6D). The AMOVA of microsatellite loci following the grouping scheme of five populations showed that most of the genetic variation was within individuals (73.75%), followed by the variation among populations (21.21%) and among individuals within populations (5.04%; Table A.9). The cluster
of samples from Morelos showed the highest level of heterozygosity, whereas the Panama cluster had the lowest (Table 3).

Concordant with the mtDNA results, the estimates of recent gene flow based on microsatellite data are low (Table A.10), ranging from 0.0079 to 0.0671, suggesting that very small numbers of individuals have migrated between these populations in recent generations. In general, the migration has occurred northwards and westwards (Fig. 4B), with the Yucatan-Guatemala and Panama populations migrating at a higher rate into the populations in central and western Mexico than vice versa. Although gene flow from the southwestern U.S. population into central and western Mexico may have occurred, the migration rates into these populations were not as high as the migration from southern populations (Fig. 4B; Table A.10).

4. Discussion

4.1. Genetic diversity and structure

Our results agree with a recent study (Silva-Caballero et al., 2017) that there is remarkable genetic differentiation between N. narica populations (Fig. 1B), characterized by high FST values (Table 2), high population structure (Table A.4; Table A.9; Fig. 6), and low gene flow between populations, with less than one migrant per generation (Table A.5; Table A.10; Fig. 4). These observations reflect the history of high diversification commonly observed in taxa inhabiting Middle America, molded by the complex topographical features combined with environmental changes likely resulting from Pliocene-Pleistocene glacial cycles (Dansgaard et al., 1993; Hewitt, 1996; Daza et al., 2010; Bagley and Johnson 2014).

The genetic characteristics of the Panama population suggest that it has been isolated for a long time [e.g. highest FST values (Table 2), lowest migration rates (Table A.5; Table A.10; Fig. 4)] and is the most divergent lineage (Fig. 1B; Fig. 3), with 9.8–10.7% mtDNA divergence from other populations (Table A.3). This level of divergence is twice that which typically defines recognized species of mammals based on the divergence of the CYTB gene (> 5%; Baker and Bradley, 2006) and is commonly observed between procyonid species: olingos (B. alleni, B. medius, B. gabbii) vs. olinguito (B. neblina) = 9.6–11.3% (Helgen et al., 2013); raccoons, P. lotor vs. P. cancrivorus = 10–11% (Helgen et al., 2013), and coatis, N. narica vs. N. nasua = 12.5–14.3% (this study). Although our data are highly suggestive, additional information from autosomal and sex chromosome sequences as well as morphology will be required to confirm whether white-nosed coatis from Panama (and further south) represent a distinct species.

All other N. narica populations have a level of sequence divergence around 2–4% (except between the southwestern U.S. and western Mexico populations), which is in accordance with the recognition of subspecies (Avise and Walker, 1999; Helgen et al., 2009). However, our findings are only marginally concordant with the four-subspecies scheme currently proposed for N. narica (Decker, 1991; Gompper, 1995; Kays, 2009). This is not surprising, as subspecies designations may not be an accurate description of how variation is partitioned across the species, since these have been largely based on a limited number of traits (Gompper, 1995). The southwestern U.S. and western Mexico populations define a clade in the phylogenetic trees (Fig. 1B, 2) that generally conforms to the distribution ascribed to N. n. molaris, north of the TMVB. Nevertheless, south of the TMVB, the inferred phylogroups do not coincide with the ranges suggested for the other three subspecies (Gompper, 1995). We did not find significant evidence indicating coatis from Cozumel Island represent a distinct lineage and therefore a different subspecies (N. n. nelsoni) from those on the Yucatan peninsula (N. n. yucatanica; Table A.4; Table A.9; Fig. 1B; Fig. 3; Fig. 6). These results are largely consistent with the findings of

Table 1
Mitochondrial genetic diversity of the 85 N. narica samples analyzed. Haplogroups as identified in the AMOVA, sample size (N), number of haplotypes, haplotype diversity (h) and nucleotide diversity (π) per sampling locality and haplogroup.

<table>
<thead>
<tr>
<th>Population groups (haplogroups)</th>
<th>Population/locality</th>
<th>N</th>
<th># haps</th>
<th>h</th>
<th>π</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southwestern U.S. (SWUS)</td>
<td>New Mexico</td>
<td>16</td>
<td>3</td>
<td>0.4250 ± 0.1326</td>
<td>0.0011 ± 0.0007</td>
</tr>
<tr>
<td></td>
<td>Arizona</td>
<td>2</td>
<td>2</td>
<td>1.0000 ± 0.5000</td>
<td>0.000460 ± 0.000650</td>
</tr>
<tr>
<td>Western Mexico (WMEX)</td>
<td>Jalisco, Mexico</td>
<td>14</td>
<td>2</td>
<td>0.3626 ± 0.1302</td>
<td>0.001179 ± 0.000754</td>
</tr>
<tr>
<td>Central Mexico (CMEX)</td>
<td>Morelos, Mexico</td>
<td>11</td>
<td>2</td>
<td>0.5455 ± 0.0722</td>
<td>0.000249 ± 0.000249</td>
</tr>
<tr>
<td>Yucatan Peninsula and Guatemala (YUCP-GUAT)</td>
<td>Yucatan, Mexico</td>
<td>8</td>
<td>2</td>
<td>0.2500 ± 0.1802</td>
<td>0.000229 ± 0.000247</td>
</tr>
<tr>
<td></td>
<td>Guatemala</td>
<td>37</td>
<td>11</td>
<td>0.6817 ± 0.0759</td>
<td>0.005024 ± 0.000596</td>
</tr>
<tr>
<td></td>
<td>Costa Rica</td>
<td>13</td>
<td>3</td>
<td>0.2049 ± 0.1558</td>
<td>0.000353 ± 0.000309</td>
</tr>
</tbody>
</table>

Table 2
Pairwise Fst values for the five N. narica populations identified in the AMOVA and Bayesian clustering analysis. Below the diagonal are the Fst values calculated from the concatenated mtDNA sequences and above the diagonal are the values obtained from the analysis of 11 microsatellite loci. All Fst p-values are significant (< 0.0001). SWUS: southwestern U.S.; WMEX: western Mexico, CMEX: central Mexico, YUCP-GUAT: Yucatan Peninsula - Guatemala, PAN: Panama.

<table>
<thead>
<tr>
<th>SWUS</th>
<th>WMEX</th>
<th>CMEX</th>
<th>YUCP-GUAT</th>
<th>PAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWUS</td>
<td>0.15560</td>
<td>0.20061</td>
<td>0.26192</td>
<td>0.32829</td>
</tr>
<tr>
<td>WMEX</td>
<td>0.91038</td>
<td>0.10573</td>
<td>0.16536</td>
<td>0.28029</td>
</tr>
<tr>
<td>CMEX</td>
<td>0.97841</td>
<td>0.99403</td>
<td>0.10155</td>
<td>0.19840</td>
</tr>
<tr>
<td>YUCP-GUAT</td>
<td>0.92574</td>
<td>0.92276</td>
<td>0.91909</td>
<td>0.23303</td>
</tr>
<tr>
<td>PAN</td>
<td>0.99616</td>
<td>0.99847</td>
<td>0.99847</td>
<td>0.97913</td>
</tr>
</tbody>
</table>
McFadden et al. (2008) based on mtDNA control region sequence data that coatis on Cozumel Island may have colonized the island during the Late Pleistocene or Holocene (possibly through human-mediated dispersal). Nonetheless, the number of samples from Cozumel was low (eight samples) and seven of the samples were from pet or captive raised animals (Table A.1) that may have been derived from the mainland and transported to the island. Therefore, we suggest caution in interpreting our results regarding the genetic status of the coati samples from Cozumel. Given the uncertainty about the taxonomic status of the coati population on Cozumel, a more extensive analysis, including more samples and additional loci, will be required to reach any conclusions that could affect the conservation efforts of this population (McFadden et al., 2010). Also, coatis from the Yucatan Peninsula (including Cozumel) group together with individuals from Belize, Guatemala and Panama, which does not correspond to the Yucatan Peninsula group with parts of the predicted range of the N. narica (Gompper, 1991) supports past introgression as a possible scenario based on biogeographic grounds (Toews and Brelsford, 2012). The finding that the haplogroup from Panama is the earliest diverging lineage within N. narica further supports this hypothesis. Additional samples and data, particularly sequences from the nuclear genome, will be required to test this and other hypotheses, and evaluate the evolutionary relationships among species within Nasua and Nasuella.

4.2. Placement of Nasuella olivacea within N. narica.

The analysis of CTYB sequences of two Western Mountain coatis resulted in phylogenies in which these were nested within N. narica and sister to the Panama haplogroup (Fig. A.1). Helgen et al. (2009) reported that Nasuella was monophyletic with N. narica using CTYB sequences, rendering the genus Nasua paraphyletic. With a broader geographic sampling, our study is the first to show that the paraphyly extends to N. narica. This pattern is consistent with past hybridization between the two lineages followed by introgression of Nasuella mitochondrial DNA into Nasua narica. Although we did not sample N. narica from eastern Panama or northern Colombia, the relative proximity or overlap of these locales (in addition to western Panama, from which samples were included) with parts of the predicted range of the Western Mountain Coati (N. olivacea; see Helgen et al., 2009; Decker, 1991) supports past introgression as a possible scenario based on biogeographic grounds (Toews and Brelsford, 2012). The finding that the haplogroup from Panama is the earliest diverging lineage within N. narica further supports this hypothesis. Additional samples and data, particularly sequences from the nuclear genome, will be required to test this and other hypotheses, and evaluate the evolutionary relationships among species within Nasua and Nasuella.

4.3. Phylogeographic pattern and drivers of divergence.

Analyses of maternally-inherited mtDNA sequences and biparentally-inherited microsatellite loci yielded congruent and robust patterns of phylogeographic differentiation. Clade divergence ages and population structure of N. narica tightly correspond with geographic barriers and habitat changes associated with Pliocene-Pleistocene climate oscillations, whose combined actions may have shaped the phylogeographic pattern of this species (Castoe et al., 2009; Bryson et al., 2011; Gutiérrez-García and Vázquez-Domínguez, 2013). The pattern of diversification and estimated divergence time of the Panama clade (~4 Mya, Fig. 1B; Fig. 2) coincides with the final uplift 3–5 Mya of the Talamanca mountain range (Abratis and Wörner, 2001; MacMillan et al., 2006; Mann et al., 2007) that separates northwestern Panama from southern Costa Rica. Previous studies identified this mountain range as a barrier underlying the diversification of a broad range of taxa including amphibians (Wang et al., 2008; Hauswaldt et al., 2011),
reptiles (Castoe et al., 2009; Daza et al., 2010), birds (Cadena et al., 2007; Arbeláez-Cortés et al., 2010) and mammals (Hardy et al., 2013; Arellano et al., 2005; Bradley et al., 2008; Harding and Dragoo, 2012), indicating it may also have been important in isolating the Panama population. The only individual from Costa Rica included in our analyses is more closely related to the geographically more distant Yucatan Peninsula-Guatemala clade than to the Panama clade (Fig. 1A; Fig. 1B; Fig. 3), although it represents the earliest branching haplotype within the former clade (ca. 0.6 Mya; Fig. 2). This suggests that coatis from Costa Rica may represent a distinct population and that the Nicaragua depression, which had marine transgressions during the Pliocene and Pleistocene (Coates and Obando, 1996; Bagley and Johnson, 2014), may have driven its divergence from northern populations, as it did for other mammal species including mice, ocelots and margays (Eizirik et al., 1998; Gutiérrez-García and Vázquez-Domínguez, 2012). Extensive sampling of coatis in Costa Rica as well as from both sides of the Nicaragua depression is needed to further test this hypothesis. Previous research on co-distributed species of reptiles, birds and mammals has identified clades or phylogroups similar to the N. narica Yucatan Peninsula-Guatemala clade, in which the Motagua-Polochic-Jocotán fault system in the south, and the Isthmus of Tehuantepec in the northeast, are geographic barriers defining phylogeographical breaks (Castoe et al., 2009; Daza et al., 2010; Gutiérrez-García and Vázquez-Domínguez, 2013). However, these barriers were formed well before (2.5–6.0 Mya) the divergence of the Yucatan Peninsula-Guatemala clade 1.3 Mya and cannot be responsible for a vicariance event
isolating this clade (Fig. 2; Barrientos et al., 1998; Ortega-Gutiérrez et al., 2007). However, the split of this clade coincides with the sixth North American glaciation that occurred 1.10–1.30 Ma (Barendregt and Duk-Rodkin, 2011; Rutter et al., 2012). Similarly, the Sierra Madre Occidental and Sierra Madre del Sur in northwestern and western Mexico, respectively, and the TMVB in central Mexico are too old (5–35 My; Ferrari et al., 1999; Ferrari et al., 2000; Ferrusquía-Villafranca et al., 2005; Nieto-Samaniego et al., 2006) for their emergence to have caused diversification events of *N. narica* populations 0.2–1.2 Ma (Fig. 2). Yet, the splitting times of the central Mexico, western Mexico, and southwestern U.S. clades are concordant with the seventh North America glaciation (0.99–1.07 Ma) and the Reid glaciation, respectively (0.13–0.28 Ma) (Barendregt and Duk-Rodkin, 2011; Rutter et al., 2012; Fig. 2). These cooling periods may have triggered habitat shifts and dry episodes in North and Central America (Dansgaard et al., 1993; Hodell et al., 2008; Molnar, 2008), reducing and isolating patches of forest in which *N. narica* thrives, possibly leading to the divergence of these clades (Halfter, 1997; Hooghiemstra and van der Hammen, 1998). Phylogeographic patterns of multiple vertebrate species are congruent with the patterns we found for *N. narica* in Middle America (McCormack et al., 2008; Bryson et al., 2011; Hardy et al., 2013; Castañeda-Rico et al., 2014). Taken together, these observations suggest that orogenetic changes combined with climatic and habitat shifts have likely driven phylogeographic breaks within *N. narica*.

### 4.4. Patterns of diversification of *N. narica* and implications for procyonid evolution in the context of the GABI

The earliest fossils of *Nasua* are found in North America and dated to the late Hemphillian (NALMA; 4.7–6.7 Ma) to early Irvingtonian (1.0–1.6 Ma; Baskin, 1982; Dalquest, 1978; Cassiliano, 1999; Emmert and Short, 2018). Based on these remains and other North American procyonid fossils, and the late appearance (i.e. late Pleistocene) of extant species in the South American fossil record, it was proposed that extant procyonids descended from North American lineages that migrated into South America during the Pleistocene following the emergence of the Panamanian isthmus (Baskin, 1982; Baskin, 1989; Baskin, 2003; Forasiepi et al., 2014; Fig. 7). However, our results clearly indicate that cladogenesis in *N. narica* occurred in a south-to-north direction (Fig. 1B), and gene flow between populations has been mostly northwards and westwards (Fig. 4). Furthermore, the earliest bifurcation within the species was in Panama (Fig. 2) before the first episode of the GABI according to the standard model (i.e. 2.4–2.8 Ma, sensu Woodburne, 2010; Fig. 7), and the distribution area of the MRCA between coati and olingo species, and between *N. nasua* and *N. narica*, was estimated to be South America (or South America and Panama), 9.5 Ma and 6 Ma, respectively (Fig. 5, Fig. A.2, Fig. A.3). Collectively, these findings challenge the hypothesis that ancestors of living procyonids, specifically *Nasua* spp., migrated from North America to South America in the late Pleistocene, 0.125 Ma (Webb, 2006; Woodburne, 2010; Fig. 7).

Our results are consistent with studies indicating that the diversification of the South American extinct species *Cyonuana* spp. and *Chapalmaliana* spp. and of extant *Nasua* and *Procyon* species, 5–7 Ma, may have been part of a temporally concordant diversification event predating the GABI (Fig. 2, Fig. 7; Koepfl et al., 2007; Eizirik et al., 2010; Eizirik, 2012, Helgen et al., 2013; Forasiepi et al., 2014; Carrillo et al., 2015). Moreover, a study on the taxonomic revision of olingos (*Bassaricyon* spp.) identified Central America as the most likely origin of dispersion for all extant procyonid genera (Helgen et al., 2013), which is partially consistent with our result, although in our study the analysis also favored South America as an ancestral area of distribution (Fig. 5, Fig. A.2, Fig. A.3). Neither our study nor that of Helgen et al. (2013) found evidence for North America as the origin for extant procyonid lineages. Furthermore, *Nasua* and *Procyon* fossils from 1.5 to 3 Ma were recently discovered in Venezuela, showing the presence of these genera in South America around the time of the full emergence of the Panamanian isthmus (Ruiz-Ramón et al., 2018). Therefore, it is possible that the previously mentioned North American procyonid remains may in fact represent South American lineages that migrated into North America and went extinct in the late Pliocene and Pleistocene due to climatic changes related to glacial cycles, as has been documented for several mammalian species (Martin, 1984; Stuart, 1991; Guthrie, 2003). This pattern of a species having a North American fossil record but in actuality originating in the tropics has been similarly identified in the long-tailed weasel, *Mustela frenata* (Harding and Dragoo, 2012). We acknowledge that this scenario contradicts some fossil record for *Nasua* (and *Procyon*) in North America, where provisional remains of late Hemphillian age and more complete fossils of Blancan age (1.8–4.7 Ma) of *Nasua* have been described (Baskin, 1998; Emmett and Short, 2018). Our understanding of the origin and movement of faunal lineages involved in the GABI, especially within South America during the Miocene and Pliocene, is limited by the paucity of fossil evidence as a result of the taphonomic bias in tropical habitats (e.g., Carrillo et al., 2015).

The almost exclusive distribution of *N. narica* in Central and North America, inhabiting only the most northern part of South America west of the Andes, combined with our divergence dating and ancestral area reconstruction results, suggests that the most probable location for the initial diversification of *Nasua* species may have been the northern Andes. The rapid uplift of the northern Andes during the last 5–10 Myr (Hoorn et al., 2010; Mora et al., 2010) coincides with our results for the divergence time between *N. narica* and *N. nasua* (Fig. 2) and the biogeographic reconstruction that identifies South America as the ancestral range of the common ancestor of these species. Furthermore, this region has played an important role for the diversification of other procyonids (Helgen et al., 2013), including the divergence of two mountain coati species that only inhabit the northern Andean range, *Nasuella olivacea* and *Nasuella meridensis* (Helgen et al., 2009; Helgen et al., 2013). These data suggest that the northern Andes may have acted as both the origin of diversification of currently recognized coati species and a barrier to dispersal (Helgen et al., 2009; Helgen et al., 2013). The discovery of older *Nasua* fossils in Central and South America and a more comprehensive phylogeographic analysis including samples of *N. narica* from northern Colombia are required to further test our hypothesis for the evolution and diversification of *Nasua* in South America.

Finally, our finding that the biogeographic dispersal model with the highest likelihood reflects the presence of the Isthmus of Panama by 9.5 Ma (Table A.6) is consistent with the recently proposed alternative GABI model, which proposes an earlier connection between North/Central America and South America (13–15 Ma) and therefore earlier faunal dispersal between the two continents (Montes et al., 2015; Bacon et al., 2016). This model proposes that climatic or environmental changes, rather than geological events, facilitated or prevented faunal dispersion (Molnar, 2008; Bacon et al., 2015; Montes et al., 2015; Bacon et al., 2016; Fig. 7). Specifically, moist and warm climate existing
in northern South America and Central America before 3.5 Mya favored tropical environments preventing faunal interchange of open-country species that do not thrive in densely forested environments, even when the land bridge was already present (Molnar, 2008; Montes et al., 2012b; Leigh et al., 2014). However, *N. narica* is considered a tropical woodland species adapted to forested habitats (Gompper, 1995) and likely would have easily dispersed through tropical forests before dry savanna-like habitats evolved in the Middle Pliocene (3.0–3.5 Mya; Molnar, 2008; Molnar, 2008; Molnar, 2008; Bacon et al., 2016). Nonetheless, the dating of the GABI and final uplift of the Panamanian isthmus remain contentious (Coates and Stallard, 2013; Montes et al., 2015; Bacon et al., 2015, 2016; O’Dea et al., 2016). If the standard model is accepted, it would not affect our conclusions about the ancestral area of distribution and the direction of diversification or migration. It would only indicate that *N. narica* dispersed north overwater before the final closure of the isthmus (Fig. 7). There is evidence for overwater dispersion of *Nasua* and *Procyon* species to Caribbean islands (McFadden et al., 2008), which renders this explanation plausible.

### 5. Conclusions

Our study reveals that the genetic diversification observed among *Nasua narica* populations has probably been driven by a combined action of geographic barriers and habitat shifts coincident with glacial periods in the northern hemisphere and that the Panama population represents a highly distinct and early branching lineage, warranting further analysis. We demonstrate, contrary to what was previously thought, that the MRCA of coati species inhabited South America around 6 Mya, not North America, and that migration and diversification of *N. narica* occurred in a south-to-north direction. In general, our findings imply that most of the evolutionary history of extant procyonid species may have occurred in South and Central America, instead of North America. This implies a more complex evolutionary history for procyonids in general, and *Nasua* species in particular, than previously acknowledged.

### Acknowledgements

We are indebted to the following individuals that facilitated tissue loans from their institutions or provided tissues: Robb T. Brumfield and Donna L. Dittman, Louisiana State University Museum of Natural Science; Terry L. Yates and Cheryl A. Parmenter, Museum of Southwestern Biology, University of New Mexico; James L. Patton, Museum of Vertebrate Zoology, University of California, Berkeley; James G. Mead, Michael D. Carleton, and Jeremy F. Jacobs, National Museum of Natural History, Smithsonian Institution; Shyamala Ratnayeke, Department of Biological Sciences, Sunway University; and Roland Kays, Nature Research Center, North Carolina Museum of Natural History. We are also grateful to the following institutions and organizations for providing financial support: Louisiana State University; Missouri State University; the US National Science Foundation (NSF) (Grants DEB-1154469, DEB-1247899); the University Museum of Zoology, Cambridge; the Natural Environment Research Council (NERC), UK; the National Geographic Society; the Andrew W. Mellon Foundation; and the Kuwait Institute for Environmental Research. The authors are very grateful to Dr. Brian龍商 for her invaluable help and support with the collection of tissue samples. The authors also acknowledge the generous support of funds from the US National Science Foundation (Grants DEB-1154469, DEB-1247899) and the University Museum of Zoology, Cambridge. The authors are very grateful to Dr. Brian 龍商 for her invaluable help and support with the collection of tissue samples. The authors also acknowledge the generous support of funds from the US National Science Foundation (Grants DEB-1154469, DEB-1247899) and the University Museum of Zoology, Cambridge.
Natural Sciences, North Carolina State University. This research was supported by grants from the United States National Science Foundation [grant 1457106], CONACyT [Doctoral scholarship and project grant J3490-V], Fundación Ecología de Cuixmala, A.C., CIBCyUAE, and UC-MEXUS-CONACyT, United States-Mexico [Doctoral fellowship 210832]. K-P. K. was supported, in part, by Saint Petersburg State University (Genome Russia grant No. 1.52.1647.2016).

Declarations of interest

None.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2018.11.011.

References
