

UCLA

UCLA Electronic Theses and Dissertations

Title

Phenotypic and gene expression variation in the Virginia opossum (*Didelphis virginiana*) and phylogeography of the white-nosed coati (*Nasua narica*)

Permalink

<https://escholarship.org/uc/item/5t38z5hx>

Author

Nigenda Morales, Sergio Fabian

Publication Date

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Phenotypic and gene expression variation in the Virginia
opossum (*Didelphis virginiana*) and phylogeography of
the white-nosed coati (*Nasua narica*)

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctor of Philosophy
in Biology

by

Sergio Fabian Nigenda Morales

2016

© Copyright by

Sergio Fabian Nigenda Morales

2016

ABSTRACT OF THE DISSERTATION

Phenotypic and gene expression variation in the Virginia
opossum (*Didelphis virginiana*) and phylogeography of
the white-nosed coati (*Nasua narica*)

by

Sergio Fabian Nigenda Morales

Doctor of Philosophy in Biology

University of California, Los Angeles, 2016

Professor Robert Wayne, Chair

Phenotypic variation may arise due to the action of natural selection or the plastic response of individuals to environmental changes. In addition, stochastic processes also play an important role in shaping the evolutionary history of species and populations. Both high phenotypic and genetic variation occur in areas with high climatic diversity or with abundant geographic barriers, which usually are biologically diverse. The Neotropical and Nearctic biogeographical realms in Central and North America are among the regions with the highest biodiversity in the world. However, very limited genetic studies have been performed in wild populations of mammals to study the evolutionary process shaping their diversity. In this dissertation we explore the phenotypic diversity and gene expression pattern in the Virginia opossum (*Didelphis virginiana*) and the phylogeographic structure in the white-nosed coati (*Nasua narica*), two

mammal species inhabiting Central and North America. In chapter 1 we determined the association of phenotypic variables for body size, extremity size, skin and coat coloration, and environmental variables using a random forest algorithm. This analysis show the phenotypic geographic variation in opossums follows three ecogeographic rules and may be driven by natural selection in response to selective pressures to adapt to seasonal and colder environments in North America. In chapter 2 we focus on the variation in skin pigmentation. We use a gene coexpression approach to analyze the skin gene expression differences between tropical and temperate opossum populations, which indicates pigmentation differences arise from expression changes in melanocytic and immune related genes. These findings suggest two possible explanations: 1) an adaptive explanation that there is a trade-off between high production of eumelanin in tropical populations for protection against pathogens and allocation of energy towards growth and energy storage, which are probably more immediate and greater selective pressures to survive in temperate environments than pathogen resistance; 2) a plastic explanation that depigmentation is the result of an autoimmune response against melanocytes triggered by low temperatures in northern latitudes. Finally, in chapter 3 we analyzed several nuclear and mitochondrial loci in the white-nosed coati under a phylogeographic framework. We found strong genetic structure across the distribution range of the species. The phylogenetic, population structure and migration rate analyses indicate the Panama population is the deepest split, which diverged 4 million years ago, and the migration of the species occurred mainly northwards into North America. These results challenge previous hypotheses about the time of emergence of the Isthmus of Panama and the origin and diversification of extant procyonid species. This dissertation contributes to better understand the evolutionary processes that have shaped the phenotypic and genetic diversity of mammal species in Central and North America.

The dissertation of Sergio Fabian Nigenda Morales is approved.

Janet S. Sinsheimer

Kirk Edward Lohmueller

Thomas Bates Smith

Robert Wayne, Committee Chair

University of California, Los Angeles

2016

To my family (Má, Pá, Jordan, Daniel and Adriana) who have
always supported me in any possible way. I love you all.

TABLE OF CONTENT

LIST OF TABLES	ix
LIST OF FIGURES	xi
ACKNOWLEDGMENTS	xiii
VITA/BIOGRAPHIC SKETCH	xvi
CHAPTER 1: Phenotypic adaptation to temperate environments in a North American marsupial	1
Abstract	1
Introduction	2
Methods	5
Museum specimens	5
Body dimension measurements	6
Skin pigmentation measurements.....	7
Coat and facial coloration measurements.....	7
Environmental data.....	8
General data analysis	9
Correlation with latitude.....	10
Association with environmental variables (random forest analysis).....	10
Results	11
Body and extremities size variation with latitude	12
Skin pigmentation and coat coloration variation with latitude	12
Correlation between phenotypes and environment	13
Body and extremities size	14
Skin and coat pigmentation.....	14
Discussion	15
Bergmann’s rule	17
Allen’s rule	18
Gloger’s rule.....	19
Sex coloration differences	21
Phenotypic plasticity	21
Tables and Figures	24
Supplementary Material	30

Appendix 1.1	39
References	46
CHAPTER 2: Transcriptomics of geographic skin pigmentation variation in the Virginia opossum	54
Abstract	54
Introduction	55
Methods	59
Sampling and phenotypic data collection	59
RNA extraction, library preparation and sequencing	61
Quality check, mapping, filtering and normalization	61
Data analysis	62
Sample-based clustering and principal component analysis	62
Differential expression analysis and weighted gene co-expression network analysis (WGCNA)	63
Results	65
Sequencing, mapping and filtering	65
Sample-based clustering and PCA	66
Differential expression and WGCNA	67
Discussion	70
Pigmentation and Immune genes drive gene expression variation	71
Skin pigmentation variation hypotheses	73
Pathogen resistance hypothesis	74
Environmental stress hypothesis	76
Wnt signaling pathway regulates opossum depigmentation	77
Summary and conclusions	79
Tables and Figures	81
Supplementary Material	90
References	92
CHAPTER 3: Phylogeography and diversification pattern of the white-nosed coati (<i>Nasua narica</i>)	106
Abstract	106
Introduction	107
Methods	112

Tissue collection and DNA extraction	112
Mitochondrial sequences	112
Mitochondrial amplification and sequencing	112
Alignment and phylogenetic analyses.....	114
Divergence time estimation	115
Haplotype genetic diversity and haplogroups structure	116
Number of migrants estimation.....	117
Microsatellites	117
Microsatellite amplification and genotyping	117
Genetic diversity, Hardy-Weinberg equilibrium and linkage disequilibrium.....	118
Genetic structure	119
Migration rates estimates	120
Results	120
Mitochondrial sequences	120
Phylogenetic and divergence time analyses	120
Population genetic analyses	122
Microsatellite analyses	123
Genetic diversity	123
Population structure	123
Discussion	125
Genetic diversity and structure	125
Phylogeographic pattern and drivers of divergence.	128
Patterns of diversification of <i>Nasua narica</i> and implications for the evolution of procyonids in the context of the Great American Biotic Interchange	131
Tables and Figures	137
Supplementary Material	146
Appendix 3.1	156
References	159

LIST OF TABLES

Table 1.1	Pearson correlation values between dimension traits and latitude, and their p-values	24
Table 1.2	Pearson and Spearman correlation between skin pigmentation, face and torso coat coloration traits and latitude	25
Table 1.3	Environmental models derived from random forest analysis for each phenotypic trait and the percentage of variance they explain	26
Table 1.S1	Total number of male and female specimens analyzed for all dimension and pigmentation traits	30
Table 1.S2	List of the bioclimatic variables measured in our study	31
Table 1.S3	ANOVA and post hoc Tuckey’s test results comparing body size and extremities size traits between the different latitudinal groups	32
Table 1.S4	ANOVA and Kruskal-Wallis test results for skin pigmentation, face and torso coat coloration traits	33
Table 2.1	Geographic and phenotypic information on the sampled opossums used for the transcriptome analysis	81
Table 2.2	Gene ontology terms and KEEG pathways enriched in the gene modules of differentially expressed genes	82
Table 2.3	Top 25 hub genes in the pink, yellow, turquoise, blue and green modules. Hub genes were identified with WGCNA as the genes with higher connectivity and module membership (kME) in that particular module	84
Table 2.S1	Analysis of variance for the skin pigmentation traits between the three populations sampled	90
Table 2.S2	Distances in Kilometers separating the three populations analyzed.	91
Table 3.1	Mitochondrial genetic diversity	137
Table 3.2	Pairwise F_{ST} values for the five white-nosed coati populations identified in the AMOVA and Bayesian clustering analysis	138
Table 3.3	Average genetic diversity for all five white-nosed coati populations based on analysis of 11 microsatellite loci	139
Table 3.S1	Primer sequences and repeat motifs of the 11 microsatellite loci used in this study	146
Table 3.S2	AMOVA results of the concatenated mtDNA sequence data based on different grouping schemes	147
Table 3.S3	Number of migrants per generation estimated with MIGRATE-N based on the concatenated mitochondrial sequences	148

Table 3.S4	Summary statistics for the 11 microsatellite loci used in this study	149
Table 3.S5	Observed and expected heterozygosity, and deviations from Hardy-Weinberg equilibrium for 11 microsatellite loci in five white-nosed coati populations	150
Table 3.S6	GENECLASS2 results	151
Table 3.S7	AMOVA results for the microsatellite data based on different grouping schemes	152
Table 3.S8	Inferred migration rates among white-nosed coati populations based on the microsatellite data	153

LIST OF FIGURES

Figure 1.1	Geographic range of the Virginia opossum	27
Figure 1.2	Importance scores for each environmental variable used as input to random forest models for a) body length, b) tail length, c) proportion of tail pigmentation and d) temporal region lightness	28
Figure 1.3	Geographic variation of the most important environmental variables associated with variation of the phenotypic traits	29
Figure 1.S1	Phenotypic variation between Virginia opossum specimens from Southern (shown in the superior part in all images) and Northern populations (inferior part in all images)	34
Figure 1.S2	Picture showing the sites on the opossum skin from which reflectance measurements were taken	35
Figure 1.S3	Scatter plots of the relationship between all phenotypic traits and latitude, showing the non-linear loess function line, which indicates the trend of the relationship	36
Figure 1.S4	Importance scores for each environmental variable used as input to random forest models	37
Figure 1.S5	Scatter plot of body length residuals (from the regression with latitude) and latitude	38
Figure 2.1	Clustering of the 13 individuals based on their gene expression profiles from 14,289 genes with at least three mapped reads in all individuals.	85
Figure 2.2	Correlation between skin gene modules and the pigmentation traits.	86
Figure 2.3	Gene coexpression network of the 1,186 genes differentially expressed between pigmented and depigmented opossums and module detection.	87
Figure 2.4	Schematic overview of the signaling pathways involved in melanocyte development and melanogenesis.	88
Figure 2.5	Depigmented skin phenotypes in humans and Virginia opossums.	89
Figure 3.1	Sampling localities and phylogenetic analysis of <i>Nasua narica</i>	140
Figure 3.2	Divergence times among <i>Nasua narica</i> lineages estimated from the concatenated sequences of three mitochondrial loci.	141
Figure 3.3	Median-joining network showing the phylogeographic structure of the 21 haplotypes of white-nosed coati.	142
Figure 3.4	Schematic summary of gene flow among white-nosed coati populations.	143
Figure 3.5	Genetic subdivision structure among white-nosed coati populations based on the analysis of 11 microsatellites in 85 individuals.	144

Figure 3.6	Temporal range for extinct and extant procyonid genera, procyonid dispersal events, GABI hypotheses and Isthmus of Panama uplift hypotheses.	145
Figure S3.1	Haplotype Neighbor-Net network based on the patristic distance between 21 haplotypes of the concatenated mitochondrial sequences	154
Figure S3.2	Structure analysis results for the YUCP-GUAT cluster.	155

ACKNOWLEDGMENTS

First and foremost, I would like to thank to the members of my doctoral committee Janet Sinsheimer, Thomas Smith, Kirk Lohmueller and Robert Wayne for their guidance, support and important comments during the preparation of this dissertation. I would like to acknowledge my adviser Bob Wayne for the opportunity he gave me to be part of his amazing lab all these years, he was a great mentor that encouraged me to look for and build my own path in academia.

My dissertation work would not have been possible without the support of my collaborators. David Valenzuela Galvan, James Beasley and Hugo Ruiz Piña, who provided field and logistical support, collected samples and helped me obtaining collecting permits. I hope to continue collaborating with you all in the future. I would like to thank Ryan Harrigan for his help with the environmental and random forest analyses used in chapter one of my dissertation, he provided very important and useful discussion on data analysis methods and data organization. Klaus Peter Koepfli was the most important collaborator for the third chapter, and a great friend through the years, many times he played a role as an unofficial academic and life advisor.

Field assistants were very instrumental and important for the data collection for my dissertation. Nicole Corpuz was the most important person helping me as a field assistant helping me to go through the vaults of old museum specimens or setting traps in the profound of the Mexican jungle, our work relationship ended up in an amazing friendship that have taken us to several adventures in different places in the world, I am looking forward to our adventures to come! I want to acknowledge Danelly Solalinde and Juan Carlos Martínez for their help during fieldwork to obtain opossum samples in Morelos, Mexico. In particular Danelly was very instrumental for additional laboratory work too.

Past and present members of the Wayne lab were crucial for the completion of my dissertation, I would like to thank them all for their support through different stages of my dissertation, our discussions about science and life encouraged me to keep going. I was lucky to have been working side by side with such highly qualified group of people. Also, the EEB graduate students were exceptional, their passion for science and nature really made an impression on me.

The Mexican Consejo Nacional de Ciencia y Tecnología (CONACYT) and the University of California Institute for Mexico (UCMEXUS) provided me with the Doctoral Fellowship that allowed me to study at UCLA. UCMEXUS funded my research by granting me a Small Grant and a Dissertation Research Grant. Also, a UCMEXUS-CONACYT Collaborative Grant to Dr. Robert Wayne and Dr. David Valenzuela Galvan funded my dissertation research. The Ecology and Evolutionary Biology Department at UCLA provided support through several research and travel grants.

Many people that made my life in Los Angeles a lot easier and fun. Julia Barske, Nicole Corpuz, Shauna Price, Graham Slater, Daniel Greenfield, Rena Schweizer, Clare Marsden, Jacqueline Robinson and Rachel Johnston, thank you for all the fun and your friendship.

I would like to thank my family that always supported and encouraged me through the different stages of my dissertation: Sergio Nigenda Capito and María del Carmen Morales, thank you for your love and your encouragement to do the things I am passionate about to the best of my capabilities; My brothers, Jordan and Daniel, thank you for your support, the fun moments and for always been there for me, I am there for you too. Also, I want to thank the Uscanga-Morales and Morales-Angulo families for allowing me to do fieldwork in their properties, and

for one or two roadkill opossum Christmas gifts. My family in-law provided us a place to work and stay during the final months of the writing process, Don Arturo, Olga, Arturo and Erika, you rock, thank you very much for all your help and support.

Finally, I would like to thank my wife Adriana Garmendia. We met when I was doing fieldwork for my thesis work and life was different after that moment. Thank you for your support, the long-distance relationship for several years and for standing by me through the difficult times that a dissertation work implies. We have been through a lot already, and we will face more challenges together, but as we have done until now, we will come on top. Thank you for deciding to share your life with me. We did it!

VITA/BIOGRAPHICAL SKETCH

SERGIO FABIÁN NIGENDA MORALES

EDUCATION

- **Ph.D.** Candidate. University of California, Los Angeles. Department of Ecology and Evolutionary Biology. October 2011
- **B.S.** Marine Biology, Autonomous University of South Baja California, Mexico. Thesis: Major Histocompatibility complex (DQB-1) polymorphism in the Gulf of California fin whale population. May 2005

AWARDS, FELLOWSHIPS, GRANTS

- 2012 - UCMEXUS Dissertation Research Grant
- 2011 - UCMEXUS Small Student Grant
- 2008 - UCMEXUS-CONACYT Doctoral Fellowship
- 2002 - Award for best student in the Marine Biology cohort 1997-2001 at Universidad Autónoma de Baja California Sur

PUBLICATIONS

- Chabot C and Nigenda S. 2011. Characterization of 13 microsatellite loci for the tope shark, *Galeorhinus galeus*, discovered with next-generation sequencing and their utility for eastern Pacific smooth-hound sharks (*Mustelus*). *Conservation Genetic Resources* 3(3): 553-555.
- Alter SE, Flores-Ramirez S, Nigenda S, Urban-R. J, Rojas-Bracho L and Palumbi SR.

2009. Mitochondrial and nuclear genetic variation across calving lagoons in Eastern North Pacific gray whales. *The Journal of Heredity* 100: 34-46.

- Nigenda-Morales S, Flores-Ramirez S, Urban-R. J and Vazquez-Juarez R. 2008. MHC DQB-1 polymorphism in the Gulf of California fin whale (*Balaenoptera physalus*) population. *The Journal of Heredity* 99: 14-21.

PRESENTATIONS AT PROFESSIONAL MEETINGS

- Oral presentation – *Phenotypic variation and differential gene expression between tropical and temperate opossum (Didelphis virginiana) populations. 2015 Genome 10K Conference, Santa Cruz, California, U.S.A. March 2015*
- Oral presentation – Phenotypic variation and differential genes expression in opossum (*Didelphis virginiana*) populations inhabiting different environments. **Evolution 2014**, Raleigh, North Carolina, U.S.A. June 2014.
- Oral presentation – Asociación entre variables ambientales y morfológicas en el tlacuache (*Didelphis virginiana*). **XI Congreso Mexicano de Mastozoología**. Xalapa, Veracruz, Mexico. October 2012
- Poster presentation – Understanding the demographic history and adaptive evolution to different environments of the North American marsupial, the Virginia opossum (*Didelphis virginiana*). **American Genetic Association Annual Symposium 2011 (Genomics and Biodiversity)**. Guanajuato, Mexico. July 2011
- Oral presentation – Major histocompatibility complex DQB-1 polymorphism in the Gulf of California fin whale population. **16th Biennial Conference on the Biology of Marine Mammals**, San Diego, California, U.S.A. December 2005

CHAPTER 1

Phenotypic adaptation to temperate environments in a North American marsupial

Abstract

Geographic phenotypic variation along environmental gradients can provide evidence of natural selection and local adaptation. Body and extremity size, and skin and coat pigmentation are traits that in combination may affect many of the physiological attributes and overall fitness of individuals. The Virginia opossum (*Didelphis virginiana*) is a marsupial that shows strong phenotypic variation in many traits, has a tropical origin and recently expanded its geographic range into temperate North America. We performed non-parametric regression analysis of environmental data from remote-sensing databases and phenotypic measurements of museum specimens to identify environmental predictors and possible adaptive significance of phenotypic variation. We found that temperature seasonality is an important predictor of body size variation, with larger sizes in more seasonal environments. Low temperatures predicted 8-38% of variation in extremity size, with smaller extremities in Northern populations. In contrast, precipitation, temperature seasonality and low temperatures were strong predictors of skin and coat pigmentation variation. Specifically, we found darker individuals distributed in warmer environments with more precipitation seasonality. These results show that the Virginia opossum follows three ecogeographic patterns codified under Bergmann's, Allen's and Gloger's rule and that phenotypes of American marsupials may be greatly affected by environmental variation. Our findings suggest that larger bodies, smaller extremities and lighter coloration of opossums in Northern populations may have facilitated the adaptation and expansion of this marsupial to the more seasonal and colder environments of temperate North America.

Introduction

Clinal geographic variation is a gradual change in one or more traits along an environmental gradient (Endler 1977; Brown and Lomolino 1998). These changes could arise as local adaptation within environmental gradients across the geographic range of a species (Mayr 1956) and contribute to phenotypic divergence among populations. In general, individuals within populations are expected to be continually or periodically under selection such that local adaptation to environmental conditions occurs (Gibbs and Grant 1987; Wigginton and Dobson 1999).

Some of the most conspicuous traits that respond to changes in the environment are body size, coloration and body extremity dimensions, especially in species with large geographic ranges (Millien *et al.* 2006). These traits are arguably very important as they can affect numerous physiological and ecological processes in animals (Burt 1981; Calder 1984; Caro 2005; Lomolino and Perault 2007; Nudds and Oswald 2007; Tattersall *et al.* 2012). The geographic variation on these traits may follow general ecogeographic patterns such as Bergmann's rule where individuals inhabiting higher latitudes (colder conditions) tend to have larger body size than those in lower and warmer latitudes (Mayr 1956; James 1970). Similarly, the presence of individuals with longer body extremities in warmer climates and shorter ones in colder climates defines Allen's rule (Allen 1877; Mayr 1956; Ray 1960). Finally, according to Gloger's rule, populations in warm and humid regions are more pigmented than populations in cooler and drier regions (Gloger 1833; Caro 2005). These ecogeographic rules are generally thought to be the result of adaptation resulting from selective pressures imposed by environmental gradients in temperature, precipitation, humidity, altitude, seasonality, primary productivity or ultraviolet

radiation (Mayr 1956; Burt 1981; Ashton *et al.* 2000; Ashton *et al.* 2002; Caro, 2005; Millien *et al.* 2006; Jablonski and Chaplin 2010; Tattersall *et al.* 2012; Koski and Ashman 2015).

Depending on how environmental variables correlate with the variation in these traits, several hypothesis have been proposed to explain ecogeographical patterns in body size, including thermoregulation, primary productivity, food seasonality-fasting endurance and food quality (Rosenzweig 1968; James 1970; Boyce 1979; Lindstedt and Boyce 1985; McNab 2010).

Similarly, heat conservation has been advanced as an explanation for size change in extremities (Allen 1877; Mayr 1956); whereas concealment, thermoregulation, pathogen resistance, prevention of cold injury and sexual selection have been proposed as factors driving skin and coat pigmentation variation (Burt 1981; Mackintosh 2001; Caro 2005). Therefore, finding environmental variables associated with the geographic variation in these traits is important to elucidate the selective pressures underlying the phenotypic change and adaptation.

The ecogeographic patterns mentioned above are widely accepted to occur in mammals (Ashton 2000; Meiri and Dayan 2003), however, very limited studies have been made in marsupial mammals (Yom-Tov and Nix 1986; Lindenmayer *et al.* 1995; Quin *et al.* 1996; Cooper 1998). Marsupials represent the ancestor group of Eutherian mammals, they have lower metabolism and body temperatures than Eutherians, probably making them more susceptible to selective pressures related to environmental fluctuation (McNab 1978; Tyndale-Biscoe 2005). To our knowledge, no study have explored the effect of environmental variables on phenotypic patterns in American marsupials. The Virginia opossum is a nocturnal marsupial species widely distributed from northwestern Costa Rica to southern Ontario and British Columbia in Canada. It is the only American marsupial not confined to tropical and temperate habitats of the Neotropical Realm (Gardner and Sunquist 2003; Patton and Costa 2003). The species originated in tropical

Central America, and the fossil record and paleoclimate data suggest that this species has expanded its range northwards into the seasonal temperate habitats of North America during the last 15,000 – 11,000 years (Guilday 1958; Graham *et al.* 1996; Bartlein *et al.* 1998; Morgan 2008; Graham and Lundelius 2010). Although widely distributed, this species is mostly absent in deserts, xeric environments and habitats with extreme low temperatures (McManus 1974; Gardner and Sunquist 2003). The Virginia opossum has relatively poor thermoregulatory capabilities at low temperatures due to its high thermo-neutral temperature (Lustick and Lustick 1972). Because of their sensitivity to cold conditions opossums do not forage when temperatures drop below -4°C , which causes high rates of mortality in northern populations due to starvation during cold weather (Brocke 1970; McNab 2002; Kanda 2005). Also, the fur density of individuals in northern populations increase during the colder months (Gardner 1973), which might be an adaptation to low temperatures in seasonal habitats. These observations indicate the species is sensitive to extreme temperatures, and that accessibility to resources is important for its survival in cold and seasonal environments.

Great geographic variation has been reported for several body dimension and coloration traits in the Virginia opossum across its geographic range. Southern populations have lower body weights and longer tails, (although apparently they do not necessarily have shorter bodies) than those in northern and more temperate climes (Gardner 1973; Gardner and Sunquist 2003). The naked ears, feet and tail have greater proportion of the skin depigmented in northern individuals. The pelage coloration of the face and the coat coloration of the dorsal part of the torso, which is due to the underfur, are lighter in northern populations as well. Also, two main color phases of guard hair, gray and black, have been described independent of the overall torso coloration (Allen 1901; Gardner 1973; Gardner and Sunquist 2003), which are found throughout the

distribution range of the species. Previous research has only generally described patterns of phenotypic variation across the distribution range of the Virginia opossum and without regard to environmental variables. Consequently, the selective pressures driving the extensive geographic variation in this species and other marsupials are poorly understood.

Due to its large distribution, sensitivity to environmental variables and substantial geographic phenotypic variation, the Virginia opossum is a good model species to study how the environment shapes phenotypes in marsupial mammals. In this study we aim to: 1) assess specific patterns of geographic variation in several body and extremity dimensions, skin pigmentation and coat coloration traits across the distribution range of the Virginia opossum; and 2) test if the geographical variation is associated with environmental variables, in order to identify selective pressures underlying mechanisms of local adaptation. To this end, we used several environmental predictors corresponding to temperature, precipitation, elevation, moisture, and vegetation coverage among others that have been suggested to affect coloration, body and extremity size variation.

Methods

Museum specimens.

We examined 352 (163 females and 189 males) study skins of Virginia opossum museum specimens to perform analysis of the body and extremity dimensions, and 348 (163 females and 185 males) for skin pigmentation and coat coloration analysis. The specimens belong to five different natural history collections (*i.e.* Smithsonian Museum of Natural History, American Museum of Natural History, Colección Nacional de Mamíferos at Universidad Nacional

Autónoma de México, Museum of Vertebrate Zoology at University of California-Berkeley and Museo de Zoología “Alfonso L. Herrera” at Universidad Nacional Autónoma de México). These specimens were collected over 145 years (1865-2010) by different collectors along the geographic range of the species (Figure 1.1), from southern Nicaragua to the northern east coast of the United States (U.S.), and include some specimens from the Western US in California. We obtained the collection coordinates for each specimens using two different approaches: 1) the coordinates provided by the original collectors; or 2) when this information was not available, we located the collecting locality provided by the original collector using the Google Earth program and extracted the geographic coordinates. Appendix 1.1 provides information about the locality, state, country and coordinates in which the specimens were collected. In order to avoid the confounding effects of age, all the specimens measured for this study were adults. We determined the age following Gardner’s (1982) age estimation method based on tooth eruption. Specifically, if the upper and lower mandibles of the specimen’s skull had all the molars erupted they were considered adults older than 10 months and included in our analysis. For each specimen we took measurements to identify variation in body dimension, proportion of pigmented skin and, body and facial coat coloration traits. Twelve phenotypic traits were measured in total, four body dimension traits, four skin pigmentation traits, three facial and one torso coat coloration trait (Table 1.S1). The variation in some of these traits is shown in Figure 1.S1. Because some specimens had damaged body parts, not all specimens were measured for all traits. In Table 1.S1 we show the total number of male and female specimens that were used in the analysis for each trait.

Body dimension measurements

To quantify variation in body size, we measured body length (head and body; Figure 1.S1a) and for the appendages variation we measured the tail length (base to tip; Figure 1.S1b), posterior aspect of the ear length (from base to most superior part of the pinna) and hindfoot length (heel to tip of middle digit; Table 1.S1). We measured the posterior aspect of the ear because most of the specimens had their ears folded and it was impossible to measure the anterior part of the pinnae, which is the more common measurement. For simplicity, hereafter, we will refer to the posterior aspect of ear length as ear length. All measurements were taken in centimeters to the nearest 0.1 decimal.

Skin pigmentation measurements

We measured four skin pigmentation traits by recording the proportion of the tail, ear, ventral and dorsal aspect of the hindfoot's middle digit that was visible pigmented (Table 1.S1; Figure 1.S1b-c). We measured the proportion of pigmentation on both the ventral and dorsal part of the middle digit because it has been shown that populations in the U.S. have hindfeet with light pigmentation on the sole and ventral part of the toes, and in the most northern populations, the light pigmentation reaches the dorsal part of the toes (Allen 1901; Gardner 1973).

Coat and facial coloration measurements

Body and facial coat coloration measurements from the museum skins were taken using a Minolta ChromaMeter CR-200 (Minolta, Osaka, Japan), which is a tristimulus colorimeter with a xenon light source. The colorimeter measures the reflectance of the xenon flash light and shows the color in a three-dimensional color space, the CIELab color space. The lightness axis (L^*) expresses the color brightness, ranging from a value of 100 for a white surface and 0 for a black

surface. The a^* and b^* axes are the two color coordinates, with the a^* axis ranging from green (negative values) to red (positive values) and the b^* axis ranging from yellow (positive values) to blue (negative values) (Fullerton and Keiding 1997; Clarys *et al.* 2000). Since the phenotypes of coat coloration in the Virginia opossum range from dark to light (Figure 1.S1a) we only used the information of the lightness axis (L^*) in our analysis. The probe with the light source was held at 90° angle to the surface parallel to the body axis, and the reflectance readings were recoded manually. All the reflectance measurements in each site were measured in triplicate and averaged to calculate the total lightness value per site. We recorded reflectance values of lightness from three facial traits (rostrum, cheek and temporal regions) and from the torso of the body (Table 1.S1; Figure 1.S2). The reflectance measurements from the torso were taken from nine sites (*i.e.* three on the right flank, three on the dorsal stripe, and three on the left flank; Figure 1.S2) and the mean of these measurements was the average torso lightness value.

Environmental data

Since different environmental variables related to temperature, precipitation, moisture and vegetation coverage have been reported to affect the size and coloration in different mammal species (Ashton *et al.* 2000; Ashton 2002; Chaplin 2004; Caro 2005), we obtained information on 24 environmental variables from the locations in which the specimens were collected. Nineteen bioclimatic variables that are related to variation in temperature, precipitation and seasonality were mined from the WorldClim database (Hijmans *et al.* 2005; Table 1.S2). In addition, five variables were extracted from remote-sensing databases: 1 and 2) the normalized difference vegetation index maximum value (NDVIMAX) and standard deviation (NDVISTD; Table 1.S2) which are related to vegetation density, measure of greenness (Buermann *et al.*

2008) and productivity (Tucker and Sellers 1986) were taken from the MODIS database; 3) the vegetation continuous field product from the MODIS database as an estimate of the percentage of tree coverage (TREECOV; Hansen *et al.* 2002; Table 1.S2); 4) the monthly backscatter measurements that capture attributes related to surface moisture and roughness (ROUGH) from the QuickScat database (Long *et al.* 2001; Table 1.S2); and 5) elevation information from the SRTM (Shuttle Radar Topography Mission; Table 1.S2) database.

General data analysis

We carried out Shapiro-Wilk normality and Bartlett homoscedasticity tests for all the traits and analyzed them accordingly. All measurements of body dimension were normally distributed for both sexes and because there is significant sexual dimorphism in body and appendages size in opossums (Gardner and **Sunquist** 2003, Astúa 2010), we performed all the analyses of dimension measurements on residuals controlled for sex. To determine if there were differences in skin pigmentation and coat coloration traits between males and females we used analysis of variance (ANOVA) or Wilcoxon rank-sum test if the data was normally or non-normally distributed, respectively. We found that all traits had homogeneity of variance, but only cheek and torso lightness were normally distributed and were also the only coloration traits statistical different between males and females (cheek lightness: $F= 29.85$, $p\text{-value} < 0.001$; coat lightness: $F = 8.85$, $p\text{-value} = 0.003$). Since our study is the first quantitative study of pigmentation in Virginia opossum we decided to analyze the coloration traits with significant differences between males and females separately, and the rest of the pigmentation traits were analyzed together for both sexes. All these tests were performed in the R statistical framework (R core team 2015).

Correlation with latitude

To determine the relationship between the phenotypic traits and latitude we performed Pearson or Spearman correlation analyses, if the data was normally distributed, and plotted the trait values against latitude using the non-linear regression loess function to graphically show the trend of the correlations. In addition, to detect differences in mean values between groups of opossums along the latitudinal gradient, we divided the data into eight latitudinal ranges, including four latitudes degrees each one (*e.g.* 11 to 14°N, 15 to 21°N, 22 to 25°N, etc.), and performed analysis of variance (ANOVA) or Kruskal-Wallis test between the individuals in these ranges, depending if the trait had normal distribution or not, respectively. These analyses were carried out using R statistical package (R core team 2015).

Association with environmental variables (random forest analysis)

To estimate the proportion of trait variation explained by the environmental variables and to identify which of these are the most important predictors of the phenotypic variation observed, we ran random forest analyses with the randomForest software in R (Liaw and Wiener 2002; Prasad *et al.* 2006; Strobl *et al.* 2009), taking as predictor variables the 24 environmental variables (Table 1.S2) and correlated them with each of the phenotypic traits (response variables). Decision trees (regression or classification) and random forest methods have no *a priori* assumptions about the relationship between predictor and response variables, allowing for the possibility to analyze non-linear relationships with complex interactions and are less susceptible to spatial autocorrelation (Breiman 2001, Strobl *et al.* 2009, Evans *et al.* 2010). Random forests analyses are iterations of large number of decision trees, which recursively partition the data into binary homogeneous groups splitting the response variable by the predictor

variable explaining most of the remaining variance. The amount of variation in the response variable explained by each predictor is incorporated in the model. Applying a randomized bootstrapping (bagging) method, random forest analysis uses a subset of both response and predictor variables randomly permuted to construct each regression tree and assess the robustness of the model based on the remaining data not included in the tree. If the accuracy of the model decreases when a variable is left out of the model, then that variable is an important predictor of the data (Breiman 2001, Prasad *et al.* 2006; Strobl *et al.* 2009).

We ran 10,000 regression trees for each random forest run, removing the least important variables after each run, until we identified the model that explained the largest amount of out of bag variation for each trait, which results in two to four predictors that explain the majority of phenotypic variation attributed to the environment. After identifying the best models, we added latitude and longitude to these models as predictor variables and re-ran random forest to test if they significantly improve the variance explained, which would indicate that distance is more important in explaining the variation than the actual environmental variables.

Taking the random forest results of the best model for each phenotypic variable, we applied partial dependence functions to graphically show the effect of each environmental predictor in the best models on the phenotypic traits after taking into account the average effects of all the other predictors in the model (Elith *et al.* 2008). These graphs are representations of the effects of each predictor and provide useful information for interpretation of the models (Friedman and Meulman 2003).

Results

Body and extremities size variation with latitude

We found significant correlation with latitude for all dimension measurements, especially for tail length (Table 1.1). The correlations with latitude for body size ($r = 0.314$, $P < 0.001$) and hindfoot length ($r = 0.284$, $P < 0.001$) were positive, whereas the correlations for tail length ($r = -0.613$, $P < 0.001$) and ear length ($r = -0.172$, $P < 0.023$) were negative (Table 1.1). In morphological studies, hindfoot length size is often used as an alternative to represent body size because they frequently covary positively (Suttie and Mitchell 1983; Wheatley 2007; Martin *et al.* 2013). Consequently, since we found a positive correlation with latitude for these measurements and they were highly correlated with each other ($r = 0.547$; p -value < 0.001), we decided to use hindfoot length as a proxy for body size in subsequent analysis, instead as a measurement of extremity size. Body length and hindfoot length increased after 24°N and had significantly higher values above 35°N and above 27°N , respectively (Table 1.S3; Figure 1.S3a; Figure 1.S3d). In contrast, tail length was larger below 24°N and significantly decreases above 27°N (Table 1.S3; Figure 1.S3b). For ear length, smaller values occur only in latitudes above 39°N (Table 1.S3; Figure 1.S3c). In general, opossums with smaller bodies and hindfeet but larger tails were distributed in latitudes below 25°N , and larger individuals with larger feet and shorter tails were found in higher latitudes.

Skin pigmentation and coat coloration variation with latitude

There was a significant negative correlation between all skin pigmentation traits and latitude (Table 1.2; Figure 1.S3e-h). The correlation was strong especially for tail pigmentation ($\rho = -0.700$, $P < 0.001$) and inferior toe pigmentation ($\rho = -0.583$, $P < 0.001$; Table 1.2). For facial lightness traits the correlation with latitude followed an inverse pattern (*i.e.* positive correlation),

which was particularly noticeable for the rostrum ($\rho=0.679$, $P < 0.001$; Table 1.2) and temporal region ($\rho=0.718$, $P < 0.001$). Also, weak but significant correlation was found for torso lightness in both sexes (Females: $r= 0.298$, $P < 0.001$; Males: $r= 0.251$, $P < 0.001$; Table 1.2). Skin pigmentation was higher below 26°N for all traits (Figure 1.S3e-h), whereas above 27°N the skin traits were significantly less pigmented, especially the tail, ear and inferior aspect of the toe (Table 1.S4; Figure 1.S3e-g). The superior aspect of the toe had significant less pigmentation above 31°N (Table 1.S4; Figure 1.S3h). Rostrum and temporal regions of the face were darkly or lightly pigmented below $25\text{-}26^{\circ}\text{N}$ or above $31\text{-}33^{\circ}\text{N}$, respectively, with a steep cline and significant difference between these latitudes (Table 1.S4; Figure 1.S3i-j). This is a common pattern we observed across many body dimension and pigmentation traits where opossums below 24°N or above 31°N are phenotypically similar, suggesting that latitudinal variation is mainly driven by individuals in the range between these latitudes (Table 1.S3; Table 1.S4; Figure 1.S3). Finally, males were lighter than females for cheek ($F= 29.85$, $p\text{-value} = < 0.001$) and torso coloration ($F= 8.84$, $p\text{-value} = 0.031$) across the distribution range, especially in lower latitudes ($< 15^{\circ}\text{N}$; Figure 1.S3k-l). For both males and females cheek lightness was lower below 24°N with a significant increase above 35°N (Table 1.S4). Torso lightness increased in both sexes around $27\text{-}29^{\circ}\text{N}$ (Figure 1.S3l), however, only males showed consistent significant differences towards lighter coloration, which occurred above 31°N (Table 1.S4). Overall, individuals with higher proportion of their skin pigmented with darker faces and torso were distributed in lower latitudes whereas less pigmented individuals with lighter faces and torso were found in higher latitudes.

Correlation between phenotypes and environment

Body and extremities size

The best environmental models explained 18% and 19% of the total variation in hindfoot and body length, respectively, and 38% of tail length variation (Table 1.3). The most important predictor of body size and hindfoot length was annual temperature range (Bio7; Table 1.3; Figure 1.2a; Figure 1.S4b), which is a variable that measures seasonality (Wigginton and Dobson 1999). The temperature of the coldest quarter (Bio11) was the most important predictor of tail length (Table 1.3; Figure 1.2b), whereas only eight percent of the ear length variance was explained by the model and its main predictor was annual mean temperature (Bio1; Table 1.3; Figure 1.S4a). Taken together the distribution of the opossums and the most important variables of the best models indicate that larger opossums (above 25°C) were present where there is higher variation in temperature range (Figure 1.3a). For the extremity size variation, tail length was reduced as the temperature of the coldest quarter decreases (Figure 1.3b). In summary, the models indicated that opossums have larger bodies and hindfeet in highly seasonal environments and show smaller tails in locations where the mean temperature during the winter was relatively low. Also, individuals with smaller ears are observed where temperature is low throughout the year.

Skin and coat pigmentation

For skin and coat pigmentation, the environmental models explained 34 to 60% of the variation in skin pigmentation, 19 to 74% in face coloration and 10 to 12 % in torso coloration (Table 1.3). The most important predictors for explaining phenotypic variation in skin

pigmentation traits were consistently the temperature of the coldest quarter and precipitation seasonality (Bio15; Table 1.3; Figure 1.2c, Figure 1.S4c-e). According to the distribution of these predictors and of the opossums, individuals with more depigmented skin (distributed above 25°C) were found in localities with relatively low temperatures during the coldest quarter (i.e. winter; Figure 1.3b) and precipitation seasonality (Figure 1.3c). For all face coloration traits, the important predictors always were temperature seasonality (Bio4) and temperature of the coldest quarter (Table 1.3; Figure 1.2d, Figure 1.S4f-h). Lighter coloration for all face traits (*i.e.* become less pigmented) were found as the seasonality of temperatures increased (Figure 1.3d) and temperatures of the coldest quarter diminished (Figure 1.3b). In contrast, the environmental models weakly explained the variation in torso lightness for males (10%) and females (12%), with temperature seasonality and annual precipitation (Bio12) as the most important predictors (Table 1.3; Figure 1.S4i-j). A pattern of lighter torso coloration was found where temperature seasonality was higher and annual precipitation lower. Taken together these results show that opossums have a higher proportion of their skin pigmented and their torso darker in humid tropical environments (where conditions are warmer, less seasonal and with higher precipitation variability), whereas they are less pigmented towards seasonal, drier and colder environments. In addition, opossum's faces are lighter in colder and more seasonal locations.

Latitude and longitude had a small effect on the environmental models for all the phenotypic variables. The maximum improvement by adding these variables to the models was 3% (Table 1.3), indicating that distance is not significantly more important than environmental variables for explaining the geographic variation observed.

Discussion

Due to the lower body temperature and metabolism that marsupials have compared to Eutherian mammals, they might be more susceptible to the effects of environmental variables and would be under high selective pressure to adaptively respond to environmental changes. For the Virginia opossum, this seems to be true. Our study provides the first quantitative data on the association of body dimensions and pigmentation variation with environmental variables across the distribution range of the Virginia opossum. We show that environmental variables largely explain the phenotypic variation observed, especially for appendages size and skin coloration. As we discuss below, the phenotypic variation in the Virginia opossum conform to the three main ecogeographical patterns: Bergmann's, Allen's and Gloger's rules.

Our finding of a correlation between body size and latitude related to environmental variables was unexpected because previous studies have not found such a relationship (Sunquist and Gardner 2003). The difference between Gardner and Sunquist's (2003) study and ours could be due to the different types of data used. These authors used measurements reported in the literature by the original collectors, whereas we obtained the data directly measuring the study skins. It is well known that a great amount of variation is introduced in the data due to different scientists taking measurements, even when they are well-defined standardized measurements (Harper 1994; Gosler *et al.* 1998; Kuczynski *et al.* 2003). In our study, only one investigator made all the measurements for consistency and to minimize error; however, it is also know that the skin is elastic and spurious variation in body length could be introduced by different curators in preparing the skin (Knox 1980; Bjordal 1983). To test if our data was systematically biased due to differences in skin preparation, with a tendency towards stretching skins of opossums from higher latitudes, we calculated the residuals of the linear regression between body size and latitude, and plotted the residuals against latitude. A positive relationship would be expected if

there was a bias towards higher latitudes. We did not find significant relationship between the residuals and latitude (Figure 1.S5), arguing against preparation bias. Also, using the lower first molar area as body size surrogate, Koch (1986) detected a similar trend of increasing size at higher latitudes as we found. Finally, since many researchers prepared skins from specimens distributed in both, high and low latitudes, a consistent bias towards stretching the skins only in opossums from higher latitudes seems unlikely. Therefore, we maintain our results for body size reflect a biological difference between high and low latitude populations.

Bergmann's rule

Our finding of association of larger hindfoot and body size with environmental predictors of seasonality (Bio7 and Bio4; Table 1.3; Figure 1.2a; Figure 1.3a; Figure 1.S4b), is consistent with the resource seasonality (also known as fasting endurance) hypothesis as an explanation for Bergmann's rule. This hypothesis suggests that natural selection favors larger individuals in regions of greater seasonality where food availability and energy demands are less predictable (Boyce 1979; Lindstedt and Boyce 1985). Larger individuals accumulate more fat and metabolize it at lower rates than smaller individuals, thus having a greater fasting endurance and survival probability during seasonal resource shortages (Boyce 1979; Lindstedt and Boyce 1985; Millar and Hickling 1990). Therefore, individuals able to obtain enough resources and large body size during times of favorable conditions have higher probability of survival during the season of scarce food availability. For the opossum, one of the most important causes of mortality is starvation due to harsh climate conditions in the winter, especially in the northern part of its distribution (Kanda and Fuller 2004; Kanda 2005). The main factors predicting if opossums would survive the winter are body weight and size (Brocke 1970; Kanda 2005). Larger

individuals save more energy than small ones because they have lower metabolism at low temperatures (McNab 1978). Moreover, under fasting conditions opossums change from using carbohydrates as energy source to lipid storage, allowing them to survive for longer during the winter (Weber and O'Connor 2000). Some Australian marsupials have showed to conform with Bergmann's pattern in response to temperature variability and food availability (Yom-Tov and Nix 1986; Quin *et al.* 1996). Therefore, our data suggests that the selective pressure of weather-driven food seasonality and availability may have resulted in an adaptive phenotypic response in the Virginia opossum towards larger bodies in more seasonal and temperate environments above 24°N.

Allen's rule

We observed a geographic pattern in ear and tail length toward smaller sizes in higher latitudes (Figure 1.S3) which was explained by environmental variables related to temperature (Table 1.3; Figure 1.2b; Figure 1.3b; Figure 1.S4a). This pattern is predicted by Allen's rule, the adaptive explanation for this pattern is that in warmer environments natural selection favors individuals with larger body appendages that increase surface area to dissipate heat via conduction, whereas in colder climates shorter appendages would be favored to reduce heat loss (Allen 1877; Millien *et al.* 2006; Tattersall *et al.* 2012). Allen's pattern has been previously found in two Australian marsupial species (*i.e.* kangaroos; Yom-Tov and Nix 1986) while other three species did not show consistency with this pattern [the brush-tail possum (*Trichosurus vulpecula*), the mountain brush-tail possum (*T. caninus*) and the Southern brown bandicoot (*Isodon obesulus*); Yom-Tov and Nix 1986; Lindenmayer *et al.* 1995; Cooper 1998]. Therefore, we show that Allen's rule also applies to American marsupials. It is possible that during the expansion of the opossum into

North America, the new selective pressures imposed by colder environments favored individuals with reduced body extremities that conserved heat better, which may be a more significant challenge for this species given its tropical origin and completely naked appendages (*i.e.* without any fur to insulate them). Physiological studies have suggested that the Virginia opossum is poorly adapted to cold climates because of high conductivity (*i.e.* its skin facilitates heat transfer) and low metabolism, and relies on behavioral mechanisms and heat production instead of heat conservation as primary thermoregulatory strategy (Lustick and Lustick 1972; Hsu *et al.* 1988). However, all physiological studies have used individuals from northern populations above 24°N [*i.e.* Florida (McNab 1978), Ohio (Lustick and Lustick 1972; Hsu *et al.* 1988), Michigan (Brocke 1970), New York (McManus 1969)], which as our results and other studies have revealed are phenotypically similar. Consequently, research is needed to explore the physiological characteristics of tropical populations. As we have shown, northern populations may have phenotypic adaptations to colder climates favoring heat conservation compared to populations in the south. However, these adaptations might not be sufficient for effective thermoregulation in the northern part of the opossum range where heat production is also needed to survive.

Gloger's rule

To our knowledge, this study is the first to explore the association between environmental variables and pigmentation in marsupials. Our findings that skin pigmentation traits, face and torso coloration are associated with temperature and precipitation variables across the geographic range of the Virginia opossum (Table 1.3; Figure 1.2c-d; Figure 1.3b-d; Figure 1.54d-j) are concordant with Gloger's rule. Two non-mutually exclusive hypotheses explaining this pattern are consistent with our results: the pathogen resistant and cold injury hypotheses. The

pathogen resistant hypothesis suggests that the higher pathogenic incidence in humid tropical environments is the selective pressure driving the increased pigmentation observed in tropical populations, because highly pigmented skin, hair or feathers confer better resistance to pathogenic infection (Wasserman 1965; Mackintosh 2001; Burt and Ichida 2004). Accumulating evidence support that melanocytes, melanosomes and melanin function as integral part of the innate immune system response against invading skin pathogens (Mackintosh 2001; Elias 2007). Darkly pigmented melanocytes acidify the epidermis contributing to enhance cutaneous antimicrobial defense (Elias 2007) and eumelanin has antibiotic properties (Mackintosh 2001). There are several examples of highly pigmented individuals that are more resistant to pathogens. For example, darker humans are less prone to bacterial and fungal infections than individuals with light skin (Mackintosh 2001), black feathers are more resistant to bacterial degradation than light ones (Burt and Ichida 2004; Gunderson *et al.* 2008), and darker greenfinches (*Carduelis chloris*) have higher survival to protozoan infections than paler ones (Männiste and Hōrak 2014). In contrast, the cold injury hypothesis suggests that lightly pigmented skin evolved as an adaptation to resist injuries derived from exposure to cold environmental conditions (Post *et al.* 1975). Melanocytes are sensitive to low temperatures, with damage occurring below -4°C (Gage 1979; Page and Shear 1988). Epidemiological studies have revealed that individuals with darker skin are more susceptible to peripheral (*i.e.* in the hands, feet and face) cold injuries and frostbite (Candler and Ivey 1997; DeGroot *et al.* 2003; Burgess and Macfarlane 2009; Maley *et al.* 2014). In addition, experiments with cold injury in guinea pigs resulted in greater degree of damage in pigmented versus non-pigmented skin in the same individuals (Post *et al.* 1975). Therefore, it is possible that the dark pigmentation of the opossum evolved as a response to pathogenic selective pressures in the warm and humid climate of the tropics and as the opossum dispersed into dryer

and colder climates where pathogen selection was relaxed and cold injury pressures increased, a change towards lighter pigmentation was favored.

Sex coloration differences

In general, females were darker than males for torso and cheek coloration. These findings could indicate sexual selection for lighter males, especially in the case of the cheek lightness that has been proposed to be a diagnostic character of the species (Gardner 1973). The fact that only 10 – 12 % of the torso lightness variance was explained by our environmental model suggests that other variables we did not consider in our analysis (*i.e.* predation, sexual selection) may have more important effects than temperate seasonality and annual precipitation on torso coloration.

Phenotypic plasticity

There is the possibility that the patterns we observed might be explained as a plastic phenotypic response to environmental variation (Price *et al.* 2003; Ghalambor *et al.* 2007). In some bird and mammal species (Teplitsky *et al.* 2008; Ozgul *et al.* 2009; Ozgul *et al.* 2010; Husby *et al.* 2011), including one marsupial (Riek and Geiser 2012), differences in body size within a species was found to be a plastic response to changes in temperature. Our results for body size variation did not detect low or high temperatures as important predictors, making this alternative hypothesis unlikely. Experimental studies in mice (Ashoub 1958; Harland 1960; Noel and Wright 1970; Serrat *et al.* 2008), rats (Lee *et al.* 1969), and pigs (Weaver and Ingram 1969; Heath 1984) have shown that body extremity size of genetically similar individuals (*i.e.* siblings) varies depending on the temperature at which they are reared, resulting in larger and shorter

extremities in warm and cold conditions, respectively. However, the only similar experiment performed in a marsupial species (*Sminthopsis crassicaudata*) did not find differences in extremity size (Riek and Geiser 2012). Finally, Himalayan rabbits (Kaufman 1925), Himalayan mouse (Kidson and Fabian 1979), Siamese cats (Iljin and Iljin 1930) among other mammals (Robinson 1973) show acrosematic pigmentation, with darker pigmentation on the ears, feet, tail and face whereas the rest of the body is lighter. This pattern is due to temperature differences in the skin of the appendages and face compared to the core of the body, in which the appendages are the coolest (Stern 1968). Nevertheless, the pattern of pigmentation in the Virginia opossum is the opposite of the acrosematic pattern. The role of plasticity in the Virginia opossum phenotype variation should be further explored through common garden studies may be by rearing individuals with distinct phenotypes in different environmental conditions and obtaining backcrossed generations to assess the heritability of the traits.

Our study is the first to quantitatively measure phenotypic variation in the Virginia opossum or any American marsupial and associate it with environmental factors. We have shown that temperature and precipitation variables have been very important in shaping the geographic variation of body size, extremity size, and skin and coat coloration in this species. Also, we advance the knowledge on the evolution of phenotypic traits in marsupials providing evidence that selective pressures from environmental variables greatly influence their phenotypic variation. In the Virginia opossum, variation conforms to three main ecogeographic patterns: Bergmann's, Allen's and Gloger's rule. This phenotypic divergence in the opossum has occurred relatively recently in last 15,000 years that the species has inhabited temperate environments of North America (Guilday 1958; Graham *et al.* 1996; Morgan 2008). Although we cannot completely rule out the possibility that adaptive phenotypic plasticity has played a role in driving

these phenotypic patterns, the evidence we have presented better supports an adaptive response through the recent action of natural selection. Further research on developmental physiology, population structure, demographic history and gene expression would be needed to further test ideas about phenotypic variation in this marsupial species. In conclusion, we have provided evidence suggesting that the Virginia opossum have evolved phenotypic adaptations to temperate and seasonal environments as the species was expanding its range northwards, and perhaps facilitating the invasion of North America.

Tables and Figures

Table 1.1. Pearson correlation values between dimension traits and latitude, and their p-values.

Trait	Pearson correlation	p-value
Body length	0.3135801	<0.001
Tail length	-0.6133094	<0.001
Ear length	-0.1722173	0.0234
Hindfoot length	0.2842715	<0.001

Table 1.2. Pearson and Spearman correlation between skin pigmentation, face and torso coat coloration traits and latitude. F = Females, M = Males, N/A = Not applicable.

Trait	Pearson correlation	Spearman correlation	p-value
Tail pigmentation	N/A	-0.7005385	<0.0001
Ear pigmentation	N/A	-0.5588478	<0.0001
Inferior toe pigmentation	N/A	-0.5834413	<0.0001
Superior toe pigmentation	N/A	-0.5481903	<0.0001
Rostrum lightness	N/A	0.6792412	<0.0001
Cheek lightness			
F	0.517	N/A	<0.0001
M	0.512	N/A	<0.0001
Temporal lightness	N/A	0.7187652	<0.0001
Torso lightness			
F	0.298	N/A	0.00013
M	0.251	N/A	0.0005

Table 1.3. Environmental models derived from random forest analysis for each phenotypic trait and the percentage of variance they explain. In parenthesis is the variance explained when latitude and longitude were added to the model. F = Females, M = Males.

Traits	Best models	% variance explained
<u>Body Dimensions</u>		
Body length	Bio7, Bio2, NDVISTD	19 (19)
Tail length	Bio11, Bio19, ELEV	38 (40)
Ear length	Bio1, Bio15, ROUGH	8 (7)
Hindfoot length	Bio7, Bio4, Bio12	18 (19)
<u>Skin pigmentation</u>		
Tail pigmentation	Bio11, Bio4, Bio15	56 (58)
Ear pigmentation	Bio4, Bio15, Bio19, ELEV	53 (55)
Inferior toe pigmentation	Bio15, Bio11, Bio4, Bio18	60 (60)
Superior toe pigmentation	Bio11, Bio15, Bio19	34 (36)
<u>Face coloration</u>		
Rostrum region lightness	Bio11, Bio4, Bio1, ELEV	71 (71)
Cheek lightness		
F	Bio4, Bio11, Bio18	19 (22)
M	Bio4, Bio11, NDVIMAX, Bio18	38 (38)
Temporal region lightness	Bio4, Bio11, Bio18	74 (75)
<u>Torso lightness</u>		
F	Bio12, Bio4, TREECOV	12 (12)
M	Bio4, Bio12, Bio15	10 (9)

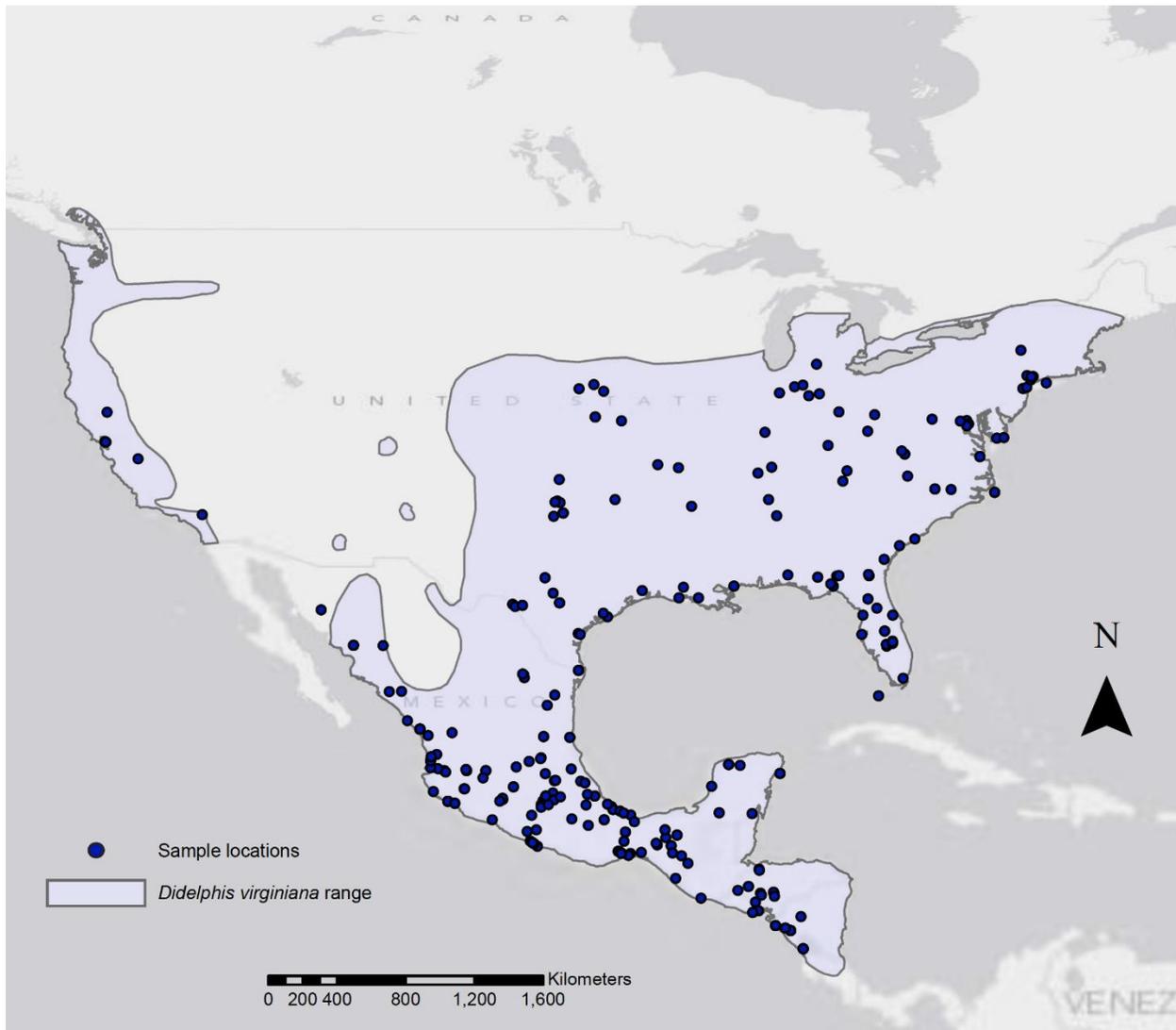


Figure 1.1. Geographic range of the Virginia opossum. The shaded light blue area represents the reported distribution range for the species (Cuarón *et al.* 2008) and the collecting localities of the museum specimens analyzed are marked by blue dots.

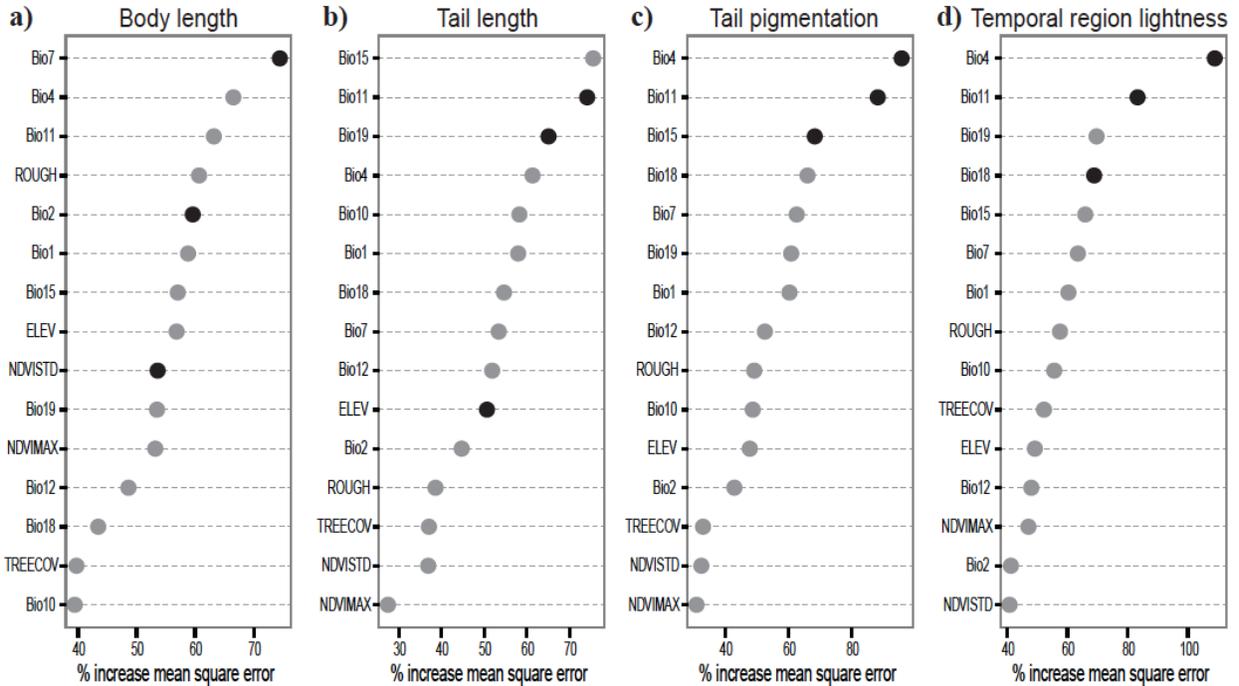


Figure 1.2. Importance scores for each environmental variable used as input to random forest models for a) body length, b) tail length, c) proportion of tail pigmentation and d) temporal region lightness. Variables with higher mean square error (calculated as the average increase in squared residuals when the variable is permuted) are more important. Variables shown with a black circle are those that remained important as the model was refined.

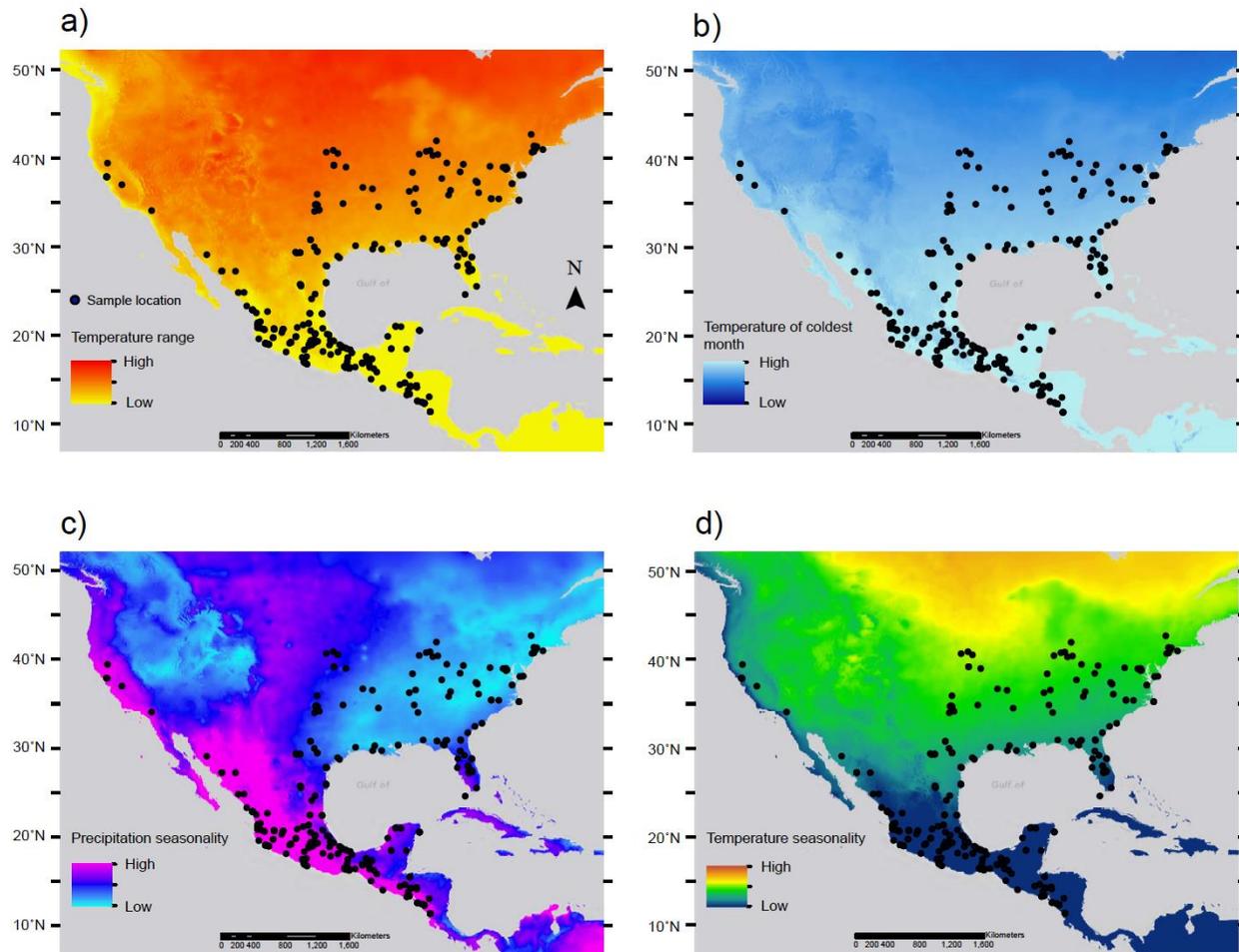


Figure 1.3. Geographic variation of the most important environmental variables associated with variation of the phenotypic traits. a) Temperature range, b) temperature of the coldest month, c) precipitation seasonality and d) temperature seasonality. The figures show the distribution of the Virginia opossum specimens.

Supplementary Material

Table 1.S1. Total number of male and female specimens analyzed for all dimension and pigmentation traits. N = Total number of specimens, F = Females, M = Males.

Trait	N			No. of specimens analyzed		
	F	M	Total	F	M	Total
Body length	163	189	352	163	185	348
Tail length	163	189	352	162	186	348
Ear length	163	189	352	84	101	185
Hindfoot length	163	189	352	160	185	345
Relative tail length	163	189	352	162	182	342
Relative ear length	163	189	352	84	101	185
Relative hindfoot length	163	189	352	160	185	345
Tail pigmentation	159	186	345	157	183	340
Ear pigmentation	159	186	345	134	159	293
Inferior toe pigmentation	159	186	345	151	183	334
Superior toe pigmentation	159	186	345	155	183	338
Rostrum lightness	159	186	345	159	186	345
Cheek lightness	159	186	345	159	186	345
Temporal lightness	159	186	345	159	186	345
Torso lightness	159	186	345	159	186	345

Table 1.S2. List of the bioclimatic variables measured in our study. The temperature and precipitation variables are derived from the monthly temperature and rainfall values. The variables represent annual trends, seasonality and extreme or limiting environmental factors.

Bio1	= Annual mean temperature
Bio2	= Mean diurnal range (mean of monthly (max temp - min temp))
Bio3	= Isothermality (Bio2/Bio7) (* 100)
Bio4	= Temperature seasonality (standard deviation *100)
Bio5	= Max temperature of warmest month
Bio6	= Min temperature of coldest month
Bio7	= Temperature annual range (Bio5-Bio6)
Bio8	= Mean temperature of wettest quarter
Bio9	= Mean temperature of driest quarter
Bio10	= Mean temperature of warmest quarter
Bio11	= Mean temperature of coldest quarter
Bio12	= Annual precipitation
Bio13	= Precipitation of wettest month
Bio14	= Precipitation of driest month
Bio15	= Precipitation seasonality (coefficient of variation)
Bio16	= Precipitation of wettest quarter
Bio17	= Precipitation of driest quarter
Bio18	= Precipitation of warmest quarter
Bio19	= Precipitation of coldest quarter
NDVIMAX	= Normalized difference vegetation index maximum value
NDVISTD	= Normalized difference vegetation index standard deviation
TREECOV	= Vegetation continuous field product (Tree coverage)
ELEV	= Elevation
ROUGH	= Surface moisture and roughness

Table 1.S3. ANOVA and post hoc Tuckey's test results comparing body size and extremities size traits between the different latitudinal groups. P-values for the post hoc tests between latitudinal groups for all dimension traits are shown. F values are shown in parenthesis below each trait. NS = Not significant. BOL= Body length, HFTL= Hindfoot length, TAL= Tail length, Ear length= ERL. The latitudinal groups are shown as capital letters: A= 11-14°N, B= 15-18°N, C= 19-22°N, D= 23-26°N, E=27-30°N, F= 31-34°N, G= 35-38°N, H= 39-42°N.

Latitudinal groups comparison	Body size		Extremities size	
	BOL (F = 7.62)	HFTL (F = 8.40)	TAL (F = 33.79)	ERL (F = 2.91)
A-B	NS	NS	NS	NS
A-C	NS	NS	NS	NS
A-D	NS	NS	NS	NS
A-E	NS	< 0.001	< 0.001	NS
A-F	NS	< 0.01	< 0.001	NS
A-G	< 0.05	< 0.001	< 0.001	NS
A-H	< 0.001	< 0.001	< 0.001	NS
B-C	NS	NS	NS	NS
B-D	NS	NS	NS	NS
B-E	NS	< 0.01	< 0.001	NS
B-F	NS	< 0.05	< 0.001	NS
B-G	< 0.05	< 0.05	< 0.001	NS
B-H	< 0.001	< 0.05	< 0.001	< 0.01
C-D	NS	NS	NS	NS
C-E	NS	< 0.01	< 0.001	NS
C-F	< 0.05	< 0.05	< 0.001	NS
C-G	< 0.01	< 0.05	< 0.001	NS
C-H	< 0.001	< 0.05	< 0.001	< 0.01
D-E	NS	< 0.05	NS	NS
D-F	NS	< 0.05	< 0.001	NS
D-G	NS	< 0.05	< 0.001	NS
D-H	< 0.001	< 0.05	< 0.001	NS
E-F	NS	NS	NS	NS
E-G	NS	NS	< 0.01	NS
E-H	< 0.01	NS	< 0.001	< 0.05
F-G	NS	NS	NS	NS
F-H	NS	NS	NS	< 0.05
G-H	NS	NS	NS	NS

Table 1.S4. ANOVA and Kruskal-Wallis test results for skin pigmentation, face and torso coat coloration traits. P-values for the post hoc tests between latitudinal groups for skin pigmentation and coat coloration traits are shown. F or χ^2 values are shown in parenthesis below each trait. NS = Not significant. TPP = Tail pigmentation, EPP = ear pigmentation, TTP = Toe inferior aspect pigmentation, TSP = Toe superior aspect pigmentation, ROL = Rostrum lightness, TEL = Temporal region lightness, CLF = Cheek lightness in females, CLM = Cheek lightness in males, TLF = Torso lightness in females, TLM = Torso lightness in males. The latitudinal groups are shown as capital letters: A= 11-14°N, B= 15-18°N, C= 19-22°N, D= 23-26°N, E=27-30°N, F= 31-34°N, G= 35-38°N, H= 39-42°N.

Latitudinal groups comparison	Skin pigmentation				Face coloration				Torso coloration	
	TPP ($\chi^2=194.04$)	EPP ($\chi^2=151.95$)	TTP ($\chi^2=187.39$)	TSP ($\chi^2=146.32$)	ROL ($\chi^2=209.48$)	TEL ($\chi^2=223.47$)	CLF (F=8.25)	CLM (F=13.49)	TLF (F=3.47)	TLM (F=5.82)
A-B	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
A-C	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
A-D	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
A-E	<0.001	<0.05	<0.001	<0.001	<0.05	<0.001	<0.05	NS	NS	NS
A-F	<0.001	<0.05	<0.001	<0.001	<0.001	<0.001	NS	NS	NS	<0.05
A-G	<0.001	<0.05	<0.001	<0.001	<0.001	<0.001	<0.001	<0.01	<0.05	NS
A-H	<0.001	<0.001	<0.05	<0.05	<0.001	<0.001	<0.001	<0.05	NS	NS
B-C	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
B-D	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
B-E	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NS	NS	NS	NS
B-F	<0.001	<0.05	<0.001	<0.001	<0.001	<0.001	NS	NS	NS	<0.01
B-G	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NS	NS
B-H	<0.001	<0.001	<0.001	NS	<0.001	<0.001	<0.01	<0.001	NS	<0.05
C-D	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
C-E	<0.001	<0.001	<0.001	<0.05	<0.001	<0.001	<0.05	NS	NS	NS
C-F	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NS	NS	NS	<0.001
C-G	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NS	NS
C-H	<0.001	<0.001	<0.001	NS	<0.001	<0.001	<0.001	<0.001	NS	<0.01
D-E	NS	NS	<0.001	NS	NS	<0.05	NS	NS	NS	NS
D-F	<0.05	<0.05	<0.001	NS	<0.05	<0.001	NS	NS	NS	NS
D-G	<0.001	<0.05	<0.001	<0.05	<0.001	<0.001	NS	NS	NS	NS
D-H	<0.05	<0.001	<0.001	NS	<0.001	<0.001	NS	NS	NS	NS
E-F	NS	NS	NS	NS	NS	NS	NS	NS	NS	<0.001
E-G	NS	NS	NS	<0.05	<0.05	NS	NS	<0.001	NS	NS
E-H	NS	NS	NS	NS	<0.05	NS	NS	<0.01	NS	<0.001
F-G	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
F-H	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
G-H	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

a)



b)



c)



Figure 1.S1. Phenotypic variation between Virginia opossum specimens from Southern (shown in the superior part in all images) and Northern populations (inferior part in all images). a) Differences in body size (*i.e.* head and body) and coat coloration of the face and dorsal part of the torso. b) Variation in tail length, proportion of tail, inferior and superior aspects of the hindfoot toes pigmentation. c) Dissimilarities in the proportion of ear pigmentation.



Figure 1.S2. Picture showing the sites on the opossum skin from which reflectance measurements were taken. Nine sites were measured on the torso and three sites on the face (*i.e.* rostrum, cheek and temporal regions).

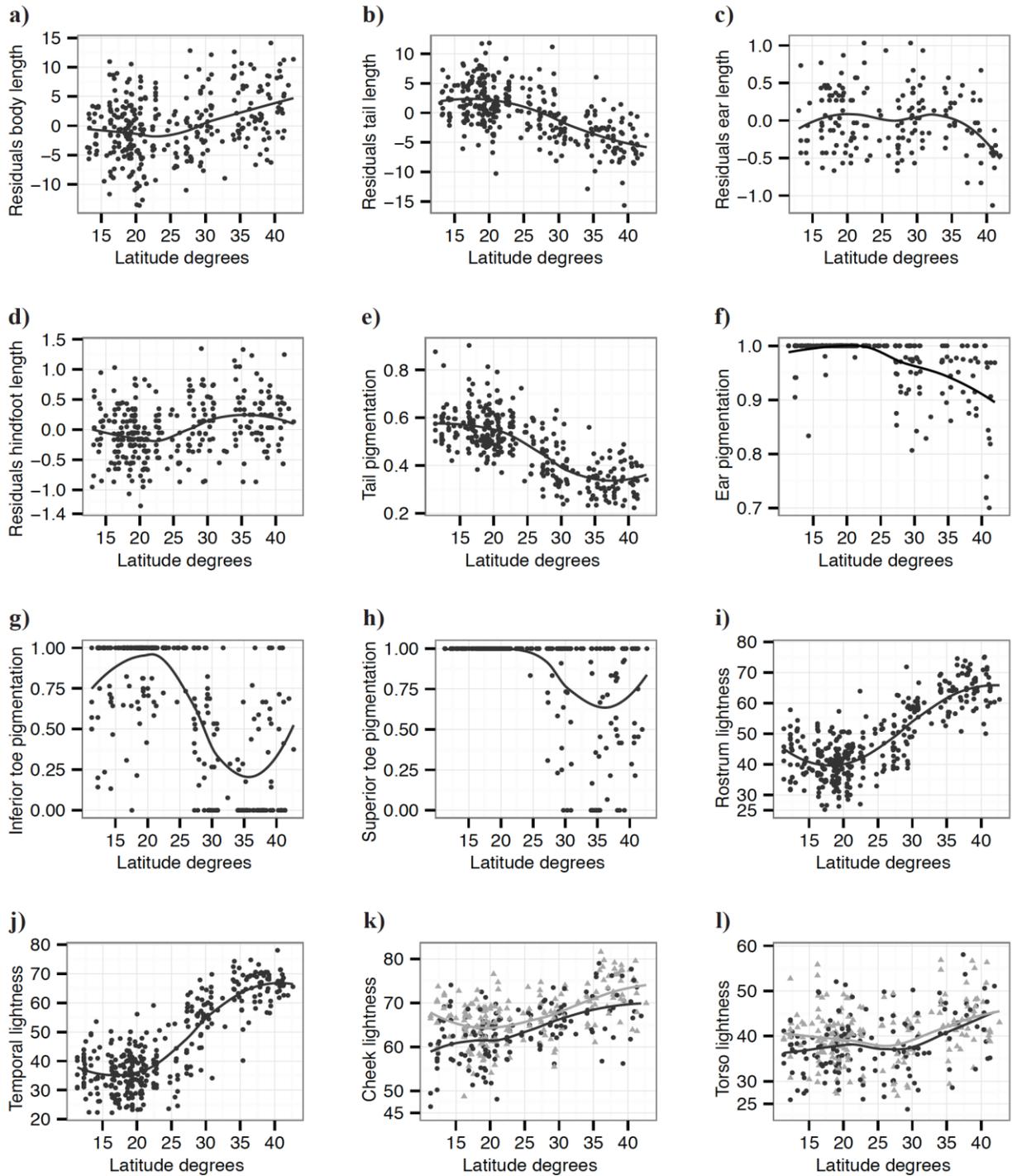


Figure 1.S3. Scatter plots of the relationship between all phenotypic traits and latitude, showing the non-linear loess function line, which indicates the trend of the relationship. a) body length, b) tail length, c) ear length, d) hindfoot length, e) proportion of tail pigmentation, f) proportion of ear pigmentation, g) proportion of toe inferior pigmentation, h) proportion of toe superior pigmentation, i) rostrum lightness, j) temporal region lightness, k) cheek lightness, l) torso lightness. For cheek and torso lightness, the gray triangles and line represent the data for males while the black circles and line represent the data for females.

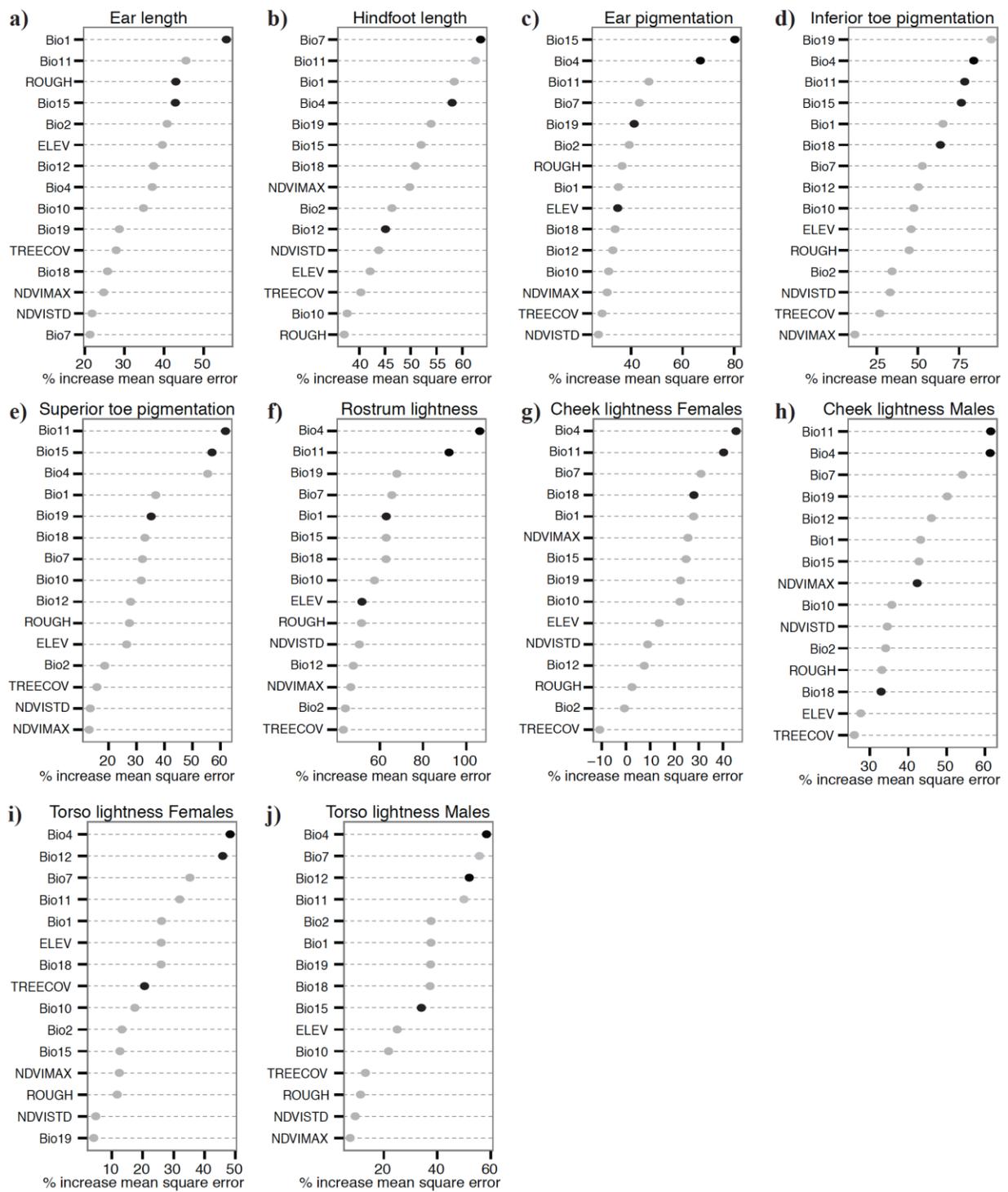


Figure 1.S4. Importance scores for each environmental variable used as input to random forest models for a) ear length, b) hindfoot length, c) proportion of ear pigmentation, d) proportion of toe inferior pigmentation, e) proportion of toe superior pigmentation, f) rostrum lightness, g) female cheek lightness, h) male cheek lightness, i) female torso lightness, j) male torso lightness. Variables shown with a black circle are those that remained important as the model was refined.

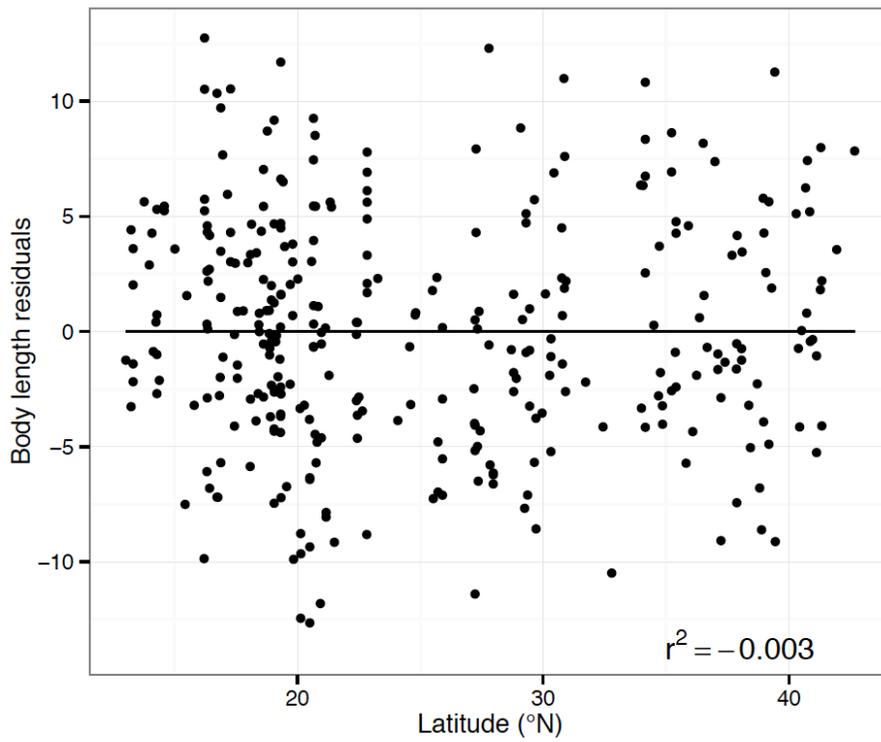


Figure 1.S5. Scatter plot of body length residuals (from the regression with latitude) and latitude. There is not trend showing the residuals increasing towards highest latitudes.

Appendix 1.1. List of all the museum specimens used in our study. The specimen numbers were recorded as shown in the specimen tags. In collection, AMNH = American Museum of Natural History, CNMA = Colección Nacional de Mamíferos (UNAM, Mexico), MALH = Museo de Zoología “Alfonso L. Herrera” (UNAM, Mexico), MVZ = Museum of Vertebrate Zoology (UC-Berkeley), USNM = Smithsonian Museum of Natural History. In country, BEL = Belize, GUA = Guatemala, HON = Honduras, MEX = Mexico, NIC = Nicaragua, SLV = El Salvador, USA = United States of America. In sex, F = Females, M = Males. N/A = Not available.

Collection	Specimen Number	Collecting year	Country	State	Latitude	Longitude	Sex
AMNH	42270	1917	USA	IN	40.380755	-84.864867	F
AMNH	135053	1942	USA	KS	39.179033	-96.597909	M
AMNH	142645	1943	USA	KS	39.179033	-96.597909	M
AMNH	142644	1943	USA	KS	39.179033	-96.597909	F
AMNH	121735	1935	USA	MI	41.937622	-85.011236	F
AMNH	9071/7378	1894	USA	NJ	40.670676	-74.208794	M
AMNH	3596/2791	1891	USA	NY	40.716771	-74.011413	F
AMNH	16664	1901	USA	NY	41.332225	-73.987078	M
AMNH	16608	1900	USA	NY	42.667253	-74.319322	M
AMNH	16666	1901	USA	NY	41.332225	-73.987078	M
AMNH	180142	1959	USA	NY	41.11812	-73.798402	M
AMNH	180141	1959	USA	NY	41.11812	-73.798402	M
AMNH	1233	1887	USA	NY	40.958687	-72.992222	F
AMNH	146632	1950	USA	NY	41.292679	-73.669186	F
AMNH	129388	1938	USA	NY	41.273651	-73.77428	M
AMNH	16594	1899	USA	NC	35.221196	-75.686696	M
AMNH	16595	1899	USA	NC	35.221196	-75.686696	M
AMNH	16596	1899	USA	NC	35.221196	-75.686696	M
AMNH	147741	1951	USA	NC	35.401311	-78.81566	M
AMNH	147743	1951	USA	NC	35.401311	-78.81566	F
AMNH	147744	1951	USA	NC	35.401311	-78.81566	F
AMNH	149377	1952	USA	TN	35.809671	-83.637353	F
AMNH	8653	1895	USA	TX	29.446438	-98.472335	M
AMNH	8651	1895	USA	TX	29.446438	-98.472335	M
AMNH	14827	1897	USA	TX	29.446438	-98.472335	F
AMNH	126141	1936	HON	FCM	14.284046	-87.257673	F
AMNH	126189	1937	HON	FCM	14.284046	-87.257673	F
AMNH	146585	1949	BEL	N/A	18.400705	-88.389935	M
AMNH	128477	1937	HON	FCM	14.284046	-87.257673	M
AMNH	128476	1937	HON	FCM	14.284046	-87.257673	F
AMNH	123286	1935	HON	FCM	14.08223	-87.214262	M
AMNH	126193	1936	HON	LPZ	14.246064	-87.955458	F
AMNH	172153	1956	MEX	CHI	17.155095	-92.900016	M
AMNH	172160	1956	MEX	CHI	16.728409	-92.627128	M
AMNH	126191	1936	HON	LPZ	14.139676	-87.908227	F
AMNH	128974	1937	HON	LEA	14.585282	-88.58484	F
AMNH	128976	1937	HON	LEA	14.585282	-88.58484	F
AMNH	25181	1905	MEX	JAL	20.654136	-102.329135	M
AMNH	25182	1905	MEX	JAL	20.654136	-102.329135	M
AMNH	25183	1905	MEX	JAL	20.654136	-102.329135	M
AMNH	16625	1900	MEX	JAL	20.659814	-103.346833	F
AMNH	16626	1900	MEX	JAL	20.659814	-103.346833	F
AMNH	16629	1900	MEX	JAL	20.659814	-103.346833	F
AMNH	26017	1905	MEX	JAL	19.709972	-103.470899	M
AMNH	25184	1905	MEX	JAL	20.657412	-104.509528	M

AMNH	25188	1905	MEX	JAL	20.657412	-104.509528	M
AMNH	25852	1905	MEX	JAL	20.577607	-104.432894	F
AMNH	145179	1947	MEX	OAX	16.22253	-94.855967	M
AMNH	145951	1948	MEX	OAX	16.318396	-94.753993	M
AMNH	145180	1947	MEX	OAX	16.22253	-94.855967	M
AMNH	145181	1947	MEX	OAX	16.22253	-94.855967	F
AMNH	145629	1948	MEX	OAX	16.22253	-94.855967	M
AMNH	145175	1947	MEX	OAX	16.426591	-95.410184	F
AMNH	145176	1947	MEX	OAX	16.426591	-95.410184	M
AMNH	145956	1949	MEX	OAX	16.332628	-95.22265	M
AMNH	145634	1948	MEX	OAX	16.417467	-95.332708	M
AMNH	145955	1949	MEX	OAX	16.332762	-95.227662	M
AMNH	145958	1949	MEX	OAX	16.332628	-95.22265	F
AMNH	24717	1904	MEX	SIN	22.835248	-105.787437	F
AMNH	24718	1904	MEX	SIN	22.835248	-105.787437	M
AMNH	24720	1904	MEX	SIN	22.835248	-105.787437	F
AMNH	131205	1938	USA	FL	30.778987	-84.95078	F
AMNH	131206	1938	USA	FL	30.778987	-84.95078	F
AMNH	91144	1929	USA	GA	30.86181	-82.258406	F
AMNH	182029	1951	USA	FL	27.791546	-82.641085	F
AMNH	100088	1931	USA	GA	30.840257	-83.96581	F
AMNH	93125	1930	USA	GA	30.873395	-83.84229	M
AMNH	24722	1904	MEX	SIN	22.835248	-105.787437	M
AMNH	24725	1904	MEX	SIN	22.835248	-105.787437	F
AMNH	24726	1904	MEX	SIN	22.835248	-105.787437	M
AMNH	24727	1904	MEX	SIN	22.835248	-105.787437	M
AMNH	146770	1949	MEX	TAM	24.082936	-99.110274	F
AMNH	172164	1956	MEX	VER	18.775154	-95.759372	F
AMNH	172157	1956	MEX	VER	17.438123	-95.023441	M
AMNH	172159	1956	MEX	VER	17.438123	-95.023441	F
AMNH	30524	1905	MEX	YUC	20.974654	-89.613131	M
AMNH	1899/1164	1889	USA	FL	29.63898	-82.32214	F
AMNH	1900/1165	1889	USA	FL	29.63898	-82.32214	F
AMNH	163976	1951	USA	FL	28.795785	-82.589785	F
AMNH	243428	1941	USA	FL	27.181218	-81.35071	F
AMNH	243429	1941	USA	FL	27.267314	-81.364121	F
AMNH	243430	1941	USA	FL	27.267314	-81.364121	M
AMNH	131201	1938	USA	FL	30.315848	-84.134981	F
AMNH	131204	1938	USA	FL	30.434785	-84.271829	M
CNMA	34858	1992	MEX	DF	19.323422	-99.185176	M
CNMA	43374	2006	MEX	DF	19.3125	-99.188056	F
CNMA	43372	2005	MEX	DF	19.3125	-99.188056	M
CNMA	4254	1989	MEX	DF	19.332633	-99.185517	F
CNMA	33417	1990	MEX	DF	19.332633	-99.185517	F
CNMA	23070	1998	MEX	DF	19.317839	-99.184141	M
CNMA	45114	2008	MEX	DF	19.320667	-99.177722	M
CNMA	45113	2009	MEX	DF	19.322528	-99.194722	F
CNMA	45116	2008	MEX	DF	19.325278	-99.189167	F
CNMA	4159	1986	MEX	GUA	20.8421917	-100.7442474	F
CNMA	21866	1985	MEX	GUA	21.140256	-100.066275	M
CNMA	3522	1979	MEX	GUE	16.6900005	-99.6255569	M
CNMA	15637	1974	MEX	GUE	16.8272228	-99.8047256	M
CNMA	45117	2009	MEX	GUE	16.972472	-100.014722	M
CNMA	11687	1970	MEX	GUE	18.3152771	-99.9300003	M
CNMA	42921	2005	MEX	JAL	19.561389	-105.083333	M
CNMA	3788	1982	MEX	MEX	19.481741	-98.823527	M

CNMA	9758	1967	MEX	MOR	19.0705547	-99.3308334	M
CNMA	26454	1985	MEX	MOR	18.8677769	-99.4402771	F
CNMA	26460	1985	MEX	MOR	18.8516674	-99.4122238	M
CNMA	1193	1952	MEX	NVL	25.673344	-100.341523	M
CNMA	45141	2007	MEX	OAX	18.133611	-97.825278	M
CNMA	2475	1956	MEX	OAX	16.3669453	-94.1944427	F
CNMA	3790	1984	MEX	PUE	20.106589	-97.356429	M
CNMA	21991	1985	MEX	SLP	22.4381561	-99.3056946	F
CNMA	21989	1985	MEX	SLP	22.4381561	-99.3056946	F
CNMA	21990	1985	MEX	SLP	22.4381561	-99.3056946	M
CNMA	45120	2008	MEX	SIN	22.821278	-105.787556	F
CNMA	26116	1975	MEX	TAM	24.6095352	-98.7221832	F
CNMA	45124	2008	MEX	VER	18.619164	-95.661222	M
CNMA	38011	1995	MEX	VER	20.0162354	-97.1416092	M
CNMA	45123	2008	MEX	VER	18.619164	-95.661222	M
CNMA	45125	2008	MEX	VER	18.619164	-95.661222	F
CNMA	45126	2008	MEX	VER	18.619164	-95.661222	F
CNMA	45127	2008	MEX	VER	18.619164	-95.661222	M
MALH	3466	1984	MEX	QUE	21.332467	-99.436397	M
MALH	3465	1985	MEX	DF	19.321558	-99.190207	F
MALH	3472	1984	MEX	MOR	18.738847	-99.443999	F
MALH	3468	1982	MEX	DF	19.321558	-99.190207	F
MALH	1397	1983	MEX	QUE	21.284622	-99.470073	F
MALH	3475	1984	MEX	VER	18.901625	-95.955748	M
MALH	3473	1984	MEX	VER	19.330151	-96.628094	M
MALH	45	1975	MEX	VER	18.335131	-94.750667	M
MALH	3476	1984	MEX	VER	19.330151	-96.628094	F
MALH	925	1983	MEX	GUE	17.474755	-100.173364	F
MALH	931	1985	MEX	GUE	17.555115	-99.685619	M
MVZ	100067	1943	MEX	MCH	19.14769116	-101.4421185	F
MVZ	100074	1943	MEX	MCH	19.20953492	-101.45667	M
MVZ	91164	1940	MEX	NVL	25.52559527	-100.30833	F
MVZ	85261	1939	MEX	SON	27.225	-109.25889	M
MVZ	85262	1939	MEX	SON	27.225	-109.25889	F
MVZ	121179	1955	MEX	VER	18.53222222	-95.30722222	F
MVZ	85264	1939	MEX	SON	27.225	-109.25889	M
MVZ	85263	1939	MEX	SON	27.225	-109.25889	F
MVZ	81582	1937	USA	TX	30.089	-94.144	F
MVZ	33460	1923	USA	CA	37.8724816	-122.2453117	F
MVZ	47148	1931	USA	CA	37.8634074	-122.2148874	M
MVZ	51974	1932	USA	CA	37.8761067	-122.2546404	M
MVZ	97345	1941	USA	CA	37.8788042	-122.265829	M
MVZ	98145	1942	USA	CA	37.89053	-122.27374	F
MVZ	126168	1960	USA	CA	37.85754	-122.2316	F
MVZ	220246	2006	USA	CA	37.870098	122.2449	F
MVZ	138994	1968	USA	CA	39.4184056	-122.1628702	F
MVZ	104560	1946	USA	CA	36.9841	-120.5341	M
MVZ	77023	1937	USA	CA	34.056	-117.182	F
MVZ	21990	1892	USA	KS	38.9633	-95.235	M
MVZ	106578	1946	USA	NE	40.645	-97.4503	F
MVZ	114854	1944	USA	NE	40.5105	-96.1637	F
MVZ	114842	1944	USA	NE	40.8619	-96.6701	F
MVZ	129731	1961	USA	NC	36.0816	-80.2506	M
MVZ	80746	1938	USA	OK	35.90302	-98.48426	M
MVZ	81408	1937	USA	OH	39.4397	-83.8367	M
MVZ	31840	1918	USA	VA	38.8822	-77.1714	M

MVZ	14704	1910	USA	VA	37.104081	-76.46406	M
MVZ	14705	1910	USA	VA	37.104081	-76.46406	F
MVZ	130275	1927	SLV	CHO	14.38333	-89.13333	F
MVZ	98152	1942	SLV	MON	13.76667	-88.21667	M
MVZ	98155	1942	SLV	SML	13.31667	-88.06667	M
MVZ	130298	1925	SLV	SML	13.31667	-88.06667	F
MVZ	130299	1925	SLV	SML	13.31667	-88.06667	F
MVZ	130302	1925	SLV	SML	13.31667	-88.06667	M
MVZ	130310	1926	SLV	SML	13.23333	-88.36667	M
MVZ	130311	1926	SLV	SML	13.23333	-88.36667	F
USNM	76717	1895	GUA	HUE	15.801345	-91.75282	F
USNM	275678	1947	GUA	ESC	13.975257	-91.056542	M
USNM	19463	1891	HON	COR	15.501307	-88.028649	F
USNM	148748	1901	HON	COR	15.432392	-88.011173	M
USNM	181261	1913	MEX	CAM	18.449316	-90.106533	M
USNM	181262	1913	MEX	CAM	18.449316	-90.106533	M
USNM	76202	1895	MEX	CHI	16.763288	-93.364336	F
USNM	76203	1895	MEX	CHI	16.859337	-93.412168	M
USNM	76209	1895	MEX	CHI	16.728409	-92.627128	F
USNM	76211	1895	MEX	CHI	17.280126	-92.31423	M
USNM	76213	1895	MEX	CHI	17.280126	-92.31423	F
USNM	76214	1895	MEX	CHI	17.280126	-92.31423	M
USNM	76716	1895	MEX	CHI	16.204405	-92.087984	M
USNM	78001	1896	MEX	CHI	15.015922	-92.393241	M
USNM	133187	1904	MEX	CHI	16.347409	-92.549229	M
USNM	32635	1892	MEX	COL	19.052582	-104.3199	F
USNM	32636	1892	MEX	COL	19.052582	-104.3199	M
USNM	32637	1892	MEX	COL	19.052582	-104.3199	M
USNM	32639	1892	MEX	COL	19.052582	-104.3199	F
USNM	32645	1892	MEX	COL	19.052582	-104.3199	M
USNM	32646	1892	MEX	COL	19.052582	-104.3199	F
USNM	33226	1892	MEX	COL	19.052582	-104.3199	M
USNM	33264	1892	MEX	COL	18.940487	-103.959676	F
USNM	33265	1892	MEX	COL	18.940487	-103.959676	M
USNM	33266	1892	MEX	COL	18.940487	-103.959676	F
USNM	33268	1892	MEX	COL	18.940487	-103.959676	M
USNM	96819	1899	MEX	DUR	24.812327	-106.739015	M
USNM	70617	1895	MEX	GUE	16.883054	-99.887943	M
USNM	70618	1895	MEX	GUE	16.883054	-99.887943	F
USNM	70619	1895	MEX	GUE	16.883054	-99.887943	M
USNM	70620	1895	MEX	GUE	16.883054	-99.887943	F
USNM	126715	1903	MEX	GUE	18.078074	-101.991479	F
USNM	26418	1893	MEX	HID	20.131751	-98.74074	M
USNM	26419	1893	MEX	HID	20.131751	-98.74074	F
USNM	26420	1893	MEX	HID	20.13823	-98.673559	M
USNM	81726	1896	MEX	HID	20.489274	-99.214157	M
USNM	33517	1892	MEX	JAL	19.704722	-103.447259	F
USNM	34338	1892	MEX	JAL	20.718654	-103.364894	F
USNM	34339	1892	MEX	JAL	20.718654	-103.364894	M
USNM	88142	1892	MEX	JAL	20.76192	-104.84995	F
USNM	51506	1893	MEX	DF	19.115792	-98.760162	M
USNM	20443	1892	MEX	MCH	20.27995	-102.473747	F
USNM	35526	1892	MEX	MCH	19.806169	-100.890701	F
USNM	35527	1892	MEX	MCH	19.806169	-100.890701	F
USNM	35528	1892	MEX	MCH	19.806169	-100.890701	M
USNM	126167	1903	MEX	MCH	19.057792	-101.607015	M

USNM	51125	1893	MEX	MOR	18.878628	-99.058899	F
USNM	25558	1891	MEX	NVL	25.71894	-100.404588	F
USNM	25735	1891	MEX	NVL	25.71894	-100.404588	M
USNM	88143	1897	MEX	NAY	21.503094	-104.898364	F
USNM	91169	1897	MEX	NAY	22.501016	-105.363347	F
USNM	512183	1976	MEX	NAY	20.797837	-105.22071	F
USNM	523002	1977	MEX	NAY	21.166813	-105.227379	M
USNM	523003	1977	MEX	NAY	21.166813	-105.227379	M
USNM	553883	1981	MEX	NAY	21.384945	-105.190793	M
USNM	65954	1894	MEX	OAX	18.09047	-96.131845	M
USNM	65955	1894	MEX	OAX	18.09047	-96.131845	F
USNM	69798	1894	MEX	OAX	17.798372	-96.962691	F
USNM	73490	1895	MEX	OAX	16.318736	-95.240935	M
USNM	73491	1895	MEX	OAX	16.961302	-95.096382	F
USNM	73492	1892	MEX	OAX	16.318736	-95.240935	M
USNM	55579	1893	MEX	PUE	19.287397	-98.430622	F
USNM	92978	1898	MEX	PUE	20.736921	-97.85191	M
USNM	96225	1898	MEX	SIN	27.202507	-107.712006	F
USNM	96820	1899	MEX	SIN	24.790704	-107.386673	M
USNM	96821	1899	MEX	SIN	23.269672	-106.433106	F
USNM	98077	1898	MEX	SIN	22.835248	-105.787437	F
USNM	33705	1892	MEX	SON	29.079003	-110.946764	F
USNM	100509	1900	MEX	TAB	17.55885	-92.943884	F
USNM	100511	1900	MEX	TAB	17.55885	-92.943884	M
USNM	92963	1898	MEX	TAM	22.400273	-97.935081	M
USNM	92964	1898	MEX	TAM	22.400273	-97.935081	F
USNM	94092	1898	MEX	TAM	22.400273	-97.935081	M
USNM	7846	N/A	MEX	VER	18.855177	-97.091325	M
USNM	54989	1893	MEX	VER	19.417195	-97.007315	M
USNM	65956	1894	MEX	VER	18.428974	-95.117066	M
USNM	78123	1896	MEX	VER	17.984231	-94.550149	M
USNM	90988	1897	MEX	ZAC	22.642443	-104.102195	M
USNM	332425	1962	NIC	MAN	12.276025	-86.36593	M
USNM	332426	1962	NIC	MAN	12.276025	-86.36593	M
USNM	332429	1962	NIC	MAN	12.276025	-86.36593	M
USNM	332430	1962	NIC	MAN	12.276025	-86.36593	F
USNM	332431	1962	NIC	MAN	12.276025	-86.36593	M
USNM	332432	1962	NIC	MAN	12.276025	-86.36593	F
USNM	332433	1962	NIC	MAN	12.276025	-86.36593	M
USNM	332434	1962	NIC	MAN	12.276025	-86.36593	F
USNM	334580	1962	NIC	MAN	12.276025	-86.36593	M
USNM	334581	1962	NIC	MAN	12.276025	-86.36593	F
USNM	334582	1962	NIC	LEO	12.41721	-86.645137	F
USNM	337522	1963	NIC	CHN	12.550278	-87.166108	F
USNM	337525	1963	NIC	CHN	12.550278	-87.166108	F
USNM	337526	1963	NIC	CHN	12.550278	-87.166108	F
USNM	337531	1963	NIC	MAT	13.018815	-85.835485	F
USNM	337653	1962	NIC	LEO	12.41721	-86.645137	F
USNM	337654	1962	NIC	LEO	12.41721	-86.645137	M
USNM	337841	1963	NIC	RIV	11.324662	-85.70739	F
USNM	337842	1963	NIC	RIV	11.324662	-85.70739	M
USNM	337843	1963	NIC	RIV	11.324662	-85.70739	M
USNM	337846	1963	NIC	RIV	11.324662	-85.70739	F
USNM	337847	1963	NIC	RIV	11.324662	-85.70739	F
USNM	11422	1865	MEX	YUC	20.974654	-89.613131	F
USNM	11425	1865	MEX	YUC	20.974654	-89.613131	F

USNM	108495	1901	MEX	QTR	20.504898	-86.925979	F
USNM	108496	1901	MEX	QTR	20.504898	-86.925979	F
USNM	108497	1901	MEX	QTR	20.504898	-86.925979	F
USNM	108499	1901	MEX	QTR	20.504898	-86.925979	M
USNM	172068	1902	MEX	YUC	20.933487	-89.016588	M
USNM	100531	1900	MEX	CAM	19.843214	-90.50253	M
USNM	18342	1890	USA	TX	29.36306	-100.929094	M
USNM	24359	1890	USA	TX	29.242485	-100.790865	M
USNM	14909	1877	USA	TX	25.900031	-97.472453	F
USNM	31909	1891	USA	TX	27.836771	-97.492465	M
USNM	31415	1891	USA	TX	27.7917	-97.395474	M
USNM	32691	1892	USA	TX	25.900031	-97.472453	F
USNM	33131	1892	USA	TX	25.900031	-97.472453	F
USNM	33132	1892	USA	TX	25.900031	-97.472453	M
USNM	63130	1893	USA	TX	29.301431	-100.41757	F
USNM	63131	1893	USA	TX	29.301431	-100.41757	M
USNM	143137	1893	USA	TX	29.301431	-100.41757	M
USNM	61841	1895	USA	FL	27.35493	-81.043848	M
USNM	61843	1895	USA	FL	27.324717	-81.034807	M
USNM	61844	1895	USA	FL	27.324717	-81.034807	M
USNM	63997	1894	USA	FL	27.433725	-81.034504	M
USNM	63998	1894	USA	FL	27.388499	-81.037855	F
USNM	71753	1895	USA	FL	24.567871	-81.761412	F
USNM	79113	1895	USA	FL	28.793453	-81.027767	M
USNM	79116	1896	USA	FL	28.793453	-81.027767	M
USNM	111239	1902	USA	FL	27.966209	-81.451962	M
USNM	111262	1902	USA	FL	27.966209	-81.451962	F
USNM	263699	1938	USA	FL	29.155318	-81.861353	F
USNM	348115	1970	USA	FL	30.907265	-86.516541	M
USNM	111274	1902	USA	FL	27.966209	-81.451962	M
USNM	231584	1919	USA	FL	25.493869	-80.482784	M
USNM	33060	1892	USA	GA	31.724426	-81.474712	F
USNM	223934	1917	USA	GA	30.927461	-82.295943	M
USNM	34290	1892	USA	LA	30.261134	-91.990976	F
USNM	139776	1905	USA	LA	29.702158	-92.211705	M
USNM	247171	1926	USA	LA	29.706483	-91.195489	M
USNM	33126	1892	USA	MS	30.311638	-89.339036	M
USNM	33849	1892	USA	MS	30.311638	-89.339036	F
USNM	186544	1885	USA	SC	32.433683	-80.674068	F
USNM	32431	1892	USA	TX	28.703537	-95.962361	F
USNM	32433	1892	USA	TX	28.904226	-96.183146	F
USNM	203535	1914	USA	AL	34.851795	-87.527773	F
USNM	207189	1915	USA	AL	33.995097	-87.099042	M
USNM	197061	1900	USA	AR	34.496837	-91.561318	M
USNM	347670	1967	USA	IN	38.362158	-87.710899	M
USNM	347672	1962	USA	IN	40.290583	-85.423302	M
USNM	347673	1962	USA	IN	40.429833	-86.953337	F
USNM	506572	1968	USA	IN	40.746064	-86.170982	M
USNM	506573	1968	USA	IN	40.840152	-85.731262	M
USNM	268016	1938	USA	KY	37.680097	-84.41814	F
USNM	116197	1902	USA	MD	38.977128	-77.157349	M
USNM	397176	1970	USA	MD	38.06781	-75.567968	M
USNM	397187	1970	USA	MD	38.06781	-75.567968	M
USNM	235291	1920	USA	MO	36.507346	-92.244956	M
USNM	34867	1892	USA	MO	36.668084	-93.334739	F
USNM	360907	1963	USA	NC	35.371924	-77.96403	M

USNM	252999	1928	USA	OH	39.290862	-81.976513	F
USNM	36179	1892	USA	OK	34.847296	-95.566814	F
USNM	132448	1904	USA	OK	34.151706	-98.278902	M
USNM	132449	1904	USA	OK	34.151706	-98.278902	M
USNM	132450	1904	USA	OK	34.151706	-98.278902	M
USNM	132452	1904	USA	OK	34.151706	-98.278902	M
USNM	132453	1904	USA	OK	34.151706	-98.278902	F
USNM	132455	1904	USA	OK	34.767786	-98.592081	M
USNM	273946	1942	USA	OK	34.688344	-98.449463	F
USNM	273966	1938	USA	OK	34.722082	-98.700329	F
USNM	173088	1911	USA	SC	32.785077	-79.861392	M
USNM	34864	1892	USA	TN	36.237234	-88.087727	F
USNM	267411	1937	USA	TN	36.356057	-83.415341	M
USNM	267413	1937	USA	TN	36.538467	-87.360104	M
USNM	112025	1900	USA	TX	29.953021	-98.803901	F
USNM	186548	1886	USA	TX	30.752681	-99.234574	M
USNM	186547	1885	USA	TX	30.752681	-99.234574	F
USNM	348326	1972	USA	TX	33.962807	-98.775435	F
USNM	251440	1934	USA	MD	38.091101	-75.20495	M
USNM	19153	1890	USA	VA	38.804835	-77.046921	M
USNM	293195	1944	USA	VA	37.228202	-80.390484	M
USNM	349915	1971	USA	VA	38.707186	-77.156004	M
USNM	567500	1936	USA	VA	37.228202	-80.390484	M
USNM	589209	1972	USA	VA	37.392989	-80.559771	F
USNM	260240	1936	USA	WV	38.429163	-82.336847	F
USNM	396183	N/A	USA	WV	39.056714	-78.960104	M
USNM	600217	2010	USA	VA	38.947802	-77.49075	M

References

- Allen JA. 1877. The influence of physical conditions in the genesis of species. *Radical Review* 1:108–140.
- Allen JA. 1901. A preliminary study of the North American opossums of the genus *Didelphis*. *Bulletin of the American Museum of Natural History* 14:149 - 195.
- Ashoub MA. 1958. Effect of two extreme temperatures on growth and tail-length of mice. *Nature* 181:284.
- Ashton KG. 2002. Patterns of within-species body size variation of birds: strong evidence for Bergmann's rule. *Global Ecology and Biogeography* 11:505 – 523.
- Ashton KG, Tracy MC, Queiroz A. 2000. Is Bergmann's rule valid for mammals? *The American Naturalist* 156:390-415.
- Astúa D. 2010. Cranial sexual dimorphism in New World marsupials and a test of Rensch's rule in Didelphidae. *Journal of Mammalogy* 91:1011-1024.
- Bartlein PJ, Anderson KH, Anderson PM, Edwards ME, Mock CJ, Thompson RS, Webb RS, Webb III T, Whitlock C. 1998. Paleoclimate simulations for North America over the past 21,000 years: features of the simulated climate and comparisons with paleoenvironmental data. *Quaternary Science Reviews* 17:549-585.
- Bjorndal H. 1983. Effects of deep freezing, freeze-drying and skinning on the body dimensions of house sparrows (*Passer domesticus*). *Cinclus* 6:105–108.
- Boyce MS. 1979. Seasonality and patterns of natural selection for life histories. *American Naturalist* 114:569–583.
- Breiman L. 2001. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science* 16:199–231.
- Brocke RH. 1970. The winter ecology and bioenergetics of the opossum, *Didelphis marsupialis*, as distributional factors in Michigan. Ph.D. thesis, Michigan State University.
- Brown JH, Lomolino MV. 1998. *Biogeography*. 2nd ed. Sinauer Associate, Sunderland, Massachusetts. 691 pp.
- Buermann W, Saatchi S, Smith TB, Zutta BR, Chaves JA, Mila B, Graham CH. 2008. Predicting species distributions across the Amazonian and Andean regions using remote sensing data. *Journal of Biogeography* 35:1160–1176.

- Burgess, JE, Macfarlane F. 2009. Retrospective analysis of the ethnic origins of male British army soldiers with peripheral cold weather injury. *Journal of the Royal Army Medical Corps* 155:11–15
- Burt EH. 1981. The adaptiveness of animal colors. *BioScience* 31:723-729.
- Burt EH, Ichida JM. 2004. Gloger's rule, feather-degrading bacteria, and color variation among song sparrows. *The Condor* 106:681-686.
- Calder WA. 1984. Size, function and life history. Harvard University Press, Cambridge, Massachusetts. 431 pp.
- Candler WH, Ivey H. 1997. Cold weather injuries among U.S. soldiers in Alaska: a five-year review. *Military Medicine* 162:788– 791.
- Caro T. 2005. The adaptive significance of coloration in mammals. *BioScience* 55:125-136.
- Chaplin G. 2004. Geographic distribution of environmental factors influencing human skin coloration. *American Journal of Physical Anthropology* 125:292-302.
- Clarys P, Alewaeters K, Lambrecht R, Barel AO. 2000. Skin color measurements: comparison between three instruments: the Chromameter, the DermaSpectrometer and the Mexameter. *Skin Research and Technology* 6:230–238.
- Cooper ML. 1998. Geographic variation in size and shape in the southern brown bandicoot, *Isodon obesulus* (Peramelidae: Marsupialia), in Western Australia. *Australian Journal of Zoology* 46:145-152.
- Cuarón AD, Emmons L, Helgen K, Reid F, Lew D, Patterson B, Delgado C, Solari S. 2008. *Didelphis virginiana*. The IUCN Red List of Threatened Species. Version 2015.1.
- Elias PM. 2007. The skin barrier as an innate immune element. *Seminars in Immunopathology* 29:3-14.
- Elith J, Leathwick JR, Hastie T. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology* 77:802-813.
- Evans SN, Hower V, Pachter L. 2010. Coverage statistics for sequence census methods. *BMC Bioinformatics* 11:430.
- Endler JA. 1977. Geographic variation, speciation, and clines. Princeton University Press, Princeton, NJ. 246 pp.
- DeGroot DW, Castellani JW, Williams JO, Amoroso PJ. 2003. Epidemiology of U.S. army cold weather injuries, 1980 - 1999. *Aviation, Space, and Environmental Medicine* 74:564-570.

- Friedman JH, Meulman JJ. 2003. Multiple additive regression trees with application in epidemiology. *Statistics in Medicine* 22:1365–1381.
- Fullerton A, Keiding J. 1997. A comparison between a tristimulus colorimeter (Minolta ChromaMeter CR-200) and two spectrophotometers (Minolta Spectrophotometer CM-508i and CM-2002). Quantification of UV-B induced erythema in a hairless guinea pig. *Skin Research and Technology* 3:237-241.
- Gage AA. 1979. What temperature is lethal for cells? *Journal of Dermatologic Surgery and Oncology* 5:459-464.
- Gardner AL. 1973. The systematics of the genus *Didelphis* (Marsupialia: Didelphidae) in North and Middle America. Museum, Special Publication of the Museum Texas Tech University 4:1-81.
- Gardner AL. 1982. Virginia opossum, pp. 3-36. In: Chapman JA, Feldhamer GA (eds). *Wild mammals of North America: biology, management, and economics*. Johns Hopkins University Press, Baltimore, MD.
- Gardner AL, Sunquist ME. 2003. Opossum, *Didelphis virginiana*, pp. 3-29. In: Feldhamer GA, Thompson BC, Chapman JA (eds.). *Wild Mammals of North America*, 2nd ed. Johns Hopkins University Press, Baltimore, MD.
- Ghalambor CK, McKay JK, Carroll SP, Reznick DN. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* 21:394-407.
- Gibbs HL, Grant PR. 1987. Oscillating selection on Darwin's finches. *Nature* 327:511–513.
- Gloger CL. 1833. *Das Abändern der vögel durch einfluss des klimas*. August Schulz, Breslau, Poland.
- Gosler AG, Greenwood JJ, Baker JK, Davidson NC. 1998. The field determination of body size and condition in passerines: a report to the British Ringing Committee. *Bird Study* 45:92-103.
- Graham RW, Lundelius Jr EL. 2010. FAUNMAP II: New data for North America with a temporal extension for the Blancan, Irvingtonian and early Rancholabrean. FAUNMAP II Database, version 1.0. www.ucmp.berkeley.edu/faunmap/index.html.
- Graham RW, Lundelius Jr EL, Graham MA, Schroeder EK, Toomey III RS, Anderson E, Barnosky AD, Burns JA, Churcher CS, Grayson DK, Guthrie D, Harington CR, Jefferson JT, Martin LD, McDonald HG, Morlan RE, Semken Jr HA, Webb D, Werdelin L, Wilson M. 1996. Spatial response of mammals to late Quaternary environmental fluctuations. *Science* 272:1601-1606.

- Guilday JE. 1958. The prehistoric distribution of the opossum. *Journal of Mammalogy* 39:39-43.
- Gunderson AR, Frame AM, Swaddle JP, Forsyth MH. 2008. Resistance of melanized feathers to bacterial degradation: Is it really so black and white? *Journal of Avian Biology* 39:539–545.
- Hansen M, DeFries R, Townshend J, Sohlberg R, Dimiceli C, Carroll M. 2002. Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. *Remote Sensing of Environment* 83:303–319.
- Haper DGC. 1994. Some comments on the repeatability of measurements. *Ring and Migration* 15:84-90.
- Harland SC. 1960. Effect of temperature on growth in weight and tail-length of inbred and hybrid mice. *Nature* 186:446.
- Heath ME. 1984. The effects of rearing-temperature on body conformation and organ size in young pigs. *Comparative Biochemistry and Physiology Part B* 77:63-72.
- Hijmans R, Cameron S, Parra J, Jones P, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965–1978.
- Hsu M, Harder JD, Lustick SI. 1988. Seasonal energetics of opossums (*Didelphis virginiana*) in Ohio. *Comparative Biochemistry and Physiology Part A* 90:441- 443.
- Husby A, Hille SM, Visser ME. 2011. Testing mechanisms of Bergmann’s rule: phenotypic decline but no genetic change in body size in three Passerine bird populations. *American Naturalist* 178:202–213.
- Iljin NA, Iljin VN. 1930. Temperature effects on the colour of the Siamese cat. *Journal of Heredity* 21:309-318.
- Jablonski NG, Chaplin G. 2010. Human skin pigmentation as an adaptation to UV radiation. *Proceedings of the National Academy of Sciences* 107:8962-8968.
- James FC. 1970 Geographic size variation in birds and its relationship to climate. *Ecology* 5:365–390.
- Jones M. 1997. Character displacement in Australian dasyurid carnivores: size relationships and prey size patterns. *Ecology* 78, 2569-2587.
- Kanda LL. 2005. Winter energetics of Virginia opossums *Didelphis virginiana* and implications for the species' northern distributional limit. *Ecography* 28:731-744.
- Kanda LL, Fuller TK. 2004. Demographic responses of Virginia opossums to limitation at their

- northern boundary. *Canadian Journal of Zoology* 82:1126-1134.
- Kaufman L. 1925. Experimental study in the partial albinism in Himalayan rabbits. *Biologia Generalis Vienna* 1: 7-21.
- Kidson SH, Fabian BC. 1979. Pigment synthesis in the Himalayan mouse. *Journal of Experimental Zoology* 210: 145-152.
- Knox A. 1980. Post-mortem changes in wing lengths and wing formulae. *Ring and Migration* 3:29-31.
- Koch PL. 1986. Clinal geographic variation in mammals: implications for the study of chronoclines. *Paleobiology* 12:269-281.
- Koski MH, Ashman T. 2015. Floral pigmentation patterns provide an example of Gloger's rule in plants. *Nature Plants* 1: 1-5.
- Kuczynski L, Tryjanowski P, Antczak M, Skoracki M, Hromada M. 2003. Repeatability of measurements and shrinkage after skinning: the case of the great grey shrike *Lanius excubitor*. *Bonner Zoologische Beiträge* 51:127-130.
- Lee MM, Chu PC, Chan HC. 1969. Magnitude and pattern of compensatory growth in rats after cold exposure. *Journal of Embryology and Experimental Morphology* 21:407-416.
- Liaw A, Wiener M. 2002. Classification and Regression by randomForest. *R News* 2:18-22.
- Lindstedt SL, Boyce MS. 1985. Seasonality, fasting endurance, and body size in mammals. *American Naturalist* 125:873-878.
- Lomolino MV, Perault DR. 2007. Body size variation of mammals in a fragmented, temperate rainforest. *Conservation Biology* 21:1059-1069.
- Long DG, Drinkwater MR, Holt B, Saatchi S, Bertoia C. 2001. Global ice and land climate studies using scatterometer image data. *EOS, Transactions, American Geophysical Union* 82:503.
- Lustick S, Lustick DD. 1972. Energetics in the opossum, *Didelphis marsupialis virginiana*. *Comparative Biochemistry and Physiology Part A*. 43:643-647.
- Mackintosh JA. 2001. The antimicrobial properties of melanocytes, melanosomes and melanin and the evolution of black skin. *Journal of Theoretical Biology* 21:101-113.
- Maley MJ, Eglin CM, House JR, Tipton MJ. 2014. The effect of ethnicity on the vascular responses to cold exposure of the extremities. *European Journal of Applied Physiology* 114:2369-2379.

- Männiste M, Hõrak P. 2014. Emerging infectious disease selects for darker plumage coloration in greenfinches. *Frontiers in Ecology and Evolution* 2:4.
- Martin JG, Festa-Bianchet M, Côté SD, Blumstein DT. 2013. Detecting between-individual differences in hind-foot length in populations of wild mammals. *Canadian Journal of Zoology* 91:118-123.
- Mayr E. 1956. Geographical character gradients and climatic adaptation. *Evolution* 10:105–108.
- McManus JJ. 1969. Temperature regulation in the opossum, *Didelphis virginiana marsupialis*. *Journal of Mammalogy* 50:550-558.
- McManus JJ. 1974. *Didelphis virginiana*. *Mammalian Species* 40: 1-6.
- McNab BK. 1978. The comparative energetics of Neotropical marsupials. *Journal of Comparative Physiology* 125:115-128.
- McNab BK. 2002. *The physiological ecology of vertebrates*. Cornell University Press, Ithaca, NY.
- McNab BK. 2010. Geographic and temporal correlations of mammalian size reconsidered: a resource rule. *Oecologia* 164:13–23.
- Meiri S, Dayan T. 2003. On the validity of Bergmann's rule. *Journal of Biogeography* 30: 331-351.
- Millar JS, Hickling GJ. 1990. Fasting endurance and the evolution of mammalian body size. *Functional Ecology* 4:5-12.
- Millien V, Lyons SK, Olson L, Smith FA, Wilson AB, Yom-Tov Y. 2006. Ecotypic variation in the context of global climate change: revisiting the rules. *Ecology Letters* 9: 853–869.
- Morgan GS. 2008. Vertebrate fauna and geochronology of the great american biotic interchange in North America, 93-140 pp. In: Lucas SG, Morgan GS, Spielmann JA, Prothero DR (eds.). *Neogene Mammals*. New Mexico Museum of Natural History and Science Bulletin 44.
- Noel JF, Wright EA. 1970. The effect of environmental temperature on the growth of vertebrae in the tail of the mouse. *Journal of Embryology and Experimental Morphology* 24:405–410.
- Nudds RL, Oswald SA. 2007. An interspecific test of Allen's rule: evolutionary implications for endothermic species. *Evolution* 61:2389-2848.
- Ozgul A, Tuljapurkar S, Benton TG, Pemberton JM, Clutton- Brock TH, Coulson T. 2009. The dynamics of phenotypic change and the shrinking sheep of St. Kilda. *Science* 325:464–467.

- Ozgul A, Childs DZ, Oli MK, Armitage KB, Blumstein DT, Olson LE, Tuljapurkar S, Coulson T. 2010. Coupled dynamics of body mass and population growth in response to environmental change. *Nature* 466: 482–485.
- Page EH, Shear NH. 1988. Temperature-dependent skin disorders. *Journal of the American Academy of Dermatology* 18:1003-1019.
- Patton JL, Costa LP. 2003. Molecular phylogeography and species limits in rainforest didelphid marsupials of South America, pp. 21-29. In: Jones M, Dickman C, Archer M (eds). *Predators with pouches: the biology of carnivorous marsupials*. CSIRO publishing, Collingwood, Australia.
- Post PW, Daniels F, Binford RT. 1975. Cold injury and the evolution of “white” skin. *Human Biology* 47:65–80.
- Prasad AM, Iverson LR, Liaw A. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* 9:181–199.
- Price TD, Qvarnström A, Irwin DE. 2003. The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society of London B: Biological Sciences* 270:1433-1440.
- Quin DG, Smith AP, Norton TW. 1996. Eco-geographic variation in size and sexual dimorphism in sugar gliders and squirrel gliders (Marsupialia: Petauridae). *Australian Journal of Zoology* 44:19–45.
- R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Ray C. 1960. The application of Bergmann’s and Allen’s rules to the poikilotherms. *Journal of Morphology* 106:85–108.
- Riek A, Geiser F. 2012. Developmental phenotypic plasticity in a marsupial mammal. *Journal of Experimental Biology* 215:1552–1558.
- Robinson R. 1973. Acromelanic albinism in mammals. *Genetica* 44: 454-458.
- Rosenzweig ML. 1968. Net primary productivity of terrestrial communities: prediction from climatological data. *American Naturalist*. 102: 67–74.
- Stern C. 1968. *Genetic Mosaics and Other Essays*. Harvard University Press, Cambridge, MA.
- Strobl C, Malley J, Tutz G. 2009. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods* 14:323–348.

- Suttie JM, Mitchell B. 1983. Jaw length and hind foot length as measures of skeletal development of Red deer (*Cervus elaphus*). *Journal of Zoology* 200:431-434.
- Tattersall GJ, Sinclair BJ, Withers PC, Fields PA, Seebacher F, Cooper CE, Maloney SK. 2012. Coping with thermal challenges: physiological adaptations to environmental temperatures. *Comprehensive Physiology* 2:2151–2202.
- Teplitsky C, Mills JA, Alho JS, Yarrall JW, Merila J. 2008. Bergmann's rule and climate change revisited: Disentangling environmental and genetic responses in a wild bird population. *Proceedings of the National Academy of Sciences of the United States of America* 105:13492–13496.
- Tucker CJ, Sellers PJ. 1986. Satellite remote sensing of primary production. *International Journal of Remote Sensing* 7:1395–1416.
- Tyndale-Biscoe H. 2005. *Life of marsupials*. CSIRO Publishing, Australia. 464 pp.
- Wasserman HP. 1965. Human pigmentation and environmental adaptation. *Archives of Environmental Health* 11:691–694.
- Weaver ME, Ingram DL. 1969. Morphological changes in swine associated with environmental temperature. *Ecology* 50:710–713.
- Weber JM, O'Connor T. 2000. Energy metabolism of the Virginia opossum during fasting and exercise. *Journal of Experimental Biology* 203:1365–1371.
- Wheatley M. 2007. Relating red squirrel body size to different conifer cone morphologies within the same geographic location. *Journal of Mammalogy* 88:220-225.
- Wigginton JD, Dobson FS. 1999. Environmental influences on geographic variation in body size of western bobcats. *Canadian Journal of Zoology* 77: 802–813.
- Yom-Tov Y, Nix H. 1986. Climatological correlates for body size of five species of Australian mammals. *Biological Journal of the Linnean Society* 29: 245-262.

CHAPTER 2

Transcriptomics of geographic skin pigmentation variation in the Virginia opossum

Abstract

Skin and coat pigmentation are two of the best-studied examples of traits under natural selection given their readily quantifiable fitness interactions with the environment (*e.g.*, camouflage) and signaling with other organisms (*e.g.* warning coloration). Previous morphological studies have found that skin pigmentation variation in the Virginia opossum (*Didelphis virginiana*) is associated with variation in precipitation and temperatures across its distribution range following Gloger's rule (lighter pigmentation in temperate environments). To investigate the molecular mechanism associated with skin pigmentation variation, we used RNA-Seq and quantified gene expression of wild opossums from tropical and temperate populations. Using differential expression analysis and a coexpression network approach, we found that expression variation in genes with melanocytic and immune functions are significantly associated with the degree of skin pigmentation variation and follow Gloger's rule. We identified several candidate genes underlying pigment variation, including genes in the melanogenesis pathway and the major histocompatibility complex. Further, we found evidence suggesting that the Wnt/ β -catenin signaling pathway might be regulating the depigmentation observed in temperate populations. Given that the opossum only invaded temperate environments in the U.S. recently, within the last 15, 000 years, our findings suggest rapid phenotypic evolution in new environments. Based on our study results we present two hypotheses that may explain Gloger's rule pattern of skin pigmentation variation in opossums skin: 1) changes in pathogen diversity supporting a pathogen resistant hypothesis; and 2) thermal stress associated with temperate environments.

Introduction

A fundamental goal of evolutionary biology is to determine the molecular mechanisms underlying the generation and maintenance of phenotypic diversity in natural populations (Hoekstra 2006). To accomplish this goal, the interplay between the genetic, morphological, and physiological factors that interact with the ecological environment of individuals needs to be understood. Ultimately, these factors direct the patterns of evolutionary change in phenotypic variation in wild populations (Hubbard *et al.* 2010; Hancock *et al.* 2011; Schweizer *et al.* 2016a).

Pigmentation is among the most conspicuous and consequential phenotypic characteristics in vertebrate population. Environmental and biotic factors and interactions with other individuals are the prime factors that affect pigmentation variation within populations (Lai *et al.* 2008). Therefore, pigmentation patterns may be under natural selection in populations (Caro 2005; Larison *et al.* 2015). Since pigmentation patterns frequently exhibit strong variation between populations that can be quantified (Endler 1990), research on pigmentation variation and its adaptive significance have more than a 100-year history (Bennett and Lamoreux 2003; Hoekstra 2006). However, little progress has been made to specifically test adaptive hypotheses explaining pigmentation until recently (Hoekstra 2006). Most notable are studies on mice, which revealed the molecular mechanisms of pigment production and effects on phenotypic variation (Nachman 2005; Hoekstra *et al.* 2006; Mullen and Hoekstra 2008; Linnen *et al.* 2009). Further, field manipulations have been used to directly test adaptive hypotheses explaining pigmentation variation (Vignieri *et al.* 2010; Linnen *et al.* 2013). The advent of next generation sequencing technologies has facilitated progress for investigating the molecular mechanisms and evolutionary forces that shape phenotypic variation (Ungerer *et al.* 2008; Gilad *et al.* 2009;

Eklblom and Galindo 2011). Specifically, when no previous genomic data are available, as is the case for many non-model organisms, sequencing and analysis of the transcriptome and gene expression can be the first step to identify candidate genes affecting the variation in phenotypic traits and develop resources that can be used in downstream analysis (Eklblom and Galindo 2011; Alvarez *et al.* 2015; Todd *et al.* 2016). Gene expression variation in natural populations can play an important role in the evolution of phenotypic differences and is potentially under natural selection, reflecting both adaptive and nonadaptive processes (Oleksiak *et al.* 2002; Manceau *et al.* 2010; Whitehead 2012; Fraser 2013). However, variation in gene expression can also be facilitated by regulatory elements or epigenetic mechanisms that alter expression before genetic variants arise in the populations (West-Eberhard 2005; Espinosa-Soto *et al.* 2011). Consequently, the analysis of gene expression in natural populations may be used to detect mechanisms that underlie local adaptation or epigenetic changes in response to environmental variation (Alvarez *et al.* 2015).

One of the most intriguing and striking observations about pigmentation concerns the change in frequency of dark and light morphs which often follow a specific pattern of geographic distribution with regard to temperature and humidity, called Gloger's rule (Stoner *et al.* 2003; Burt and Ichida 2004; Caro 2005). This ecogeographic rule states that darker individuals are generally found in warmer and humid environments, while lighter morphs are more common in dry and cool regions (Gloger 1833; Gaston *et al.* 2008). Among the proposed selective forces driving Gloger's rule are crypsis, thermoregulation, protection from ultraviolet (UV) radiation, and pathogen resistance (Burt 1981; Burt and Ichida 2004; Caro 2005; Millien *et al.* 2006; Jablonski and Chaplin 2010). However, the molecular mechanisms regulating this ecogeographic pattern in vertebrates are unknown. Critically, most of the studies of geographic variation in

mammals focus on coat coloration (Caro 2005; Hubbard *et al.* 2010) rather than skin pigmentation, which may be an important characteristic impacting fitness in diverse species (mainly primate species; Jablonski and Chaplin 2010; Santana *et al.* 2012). Further, studies have focused on pigmentation in Eutherians (Bellone *et al.* 2008; Linnen *et al.* 2009; Anderson *et al.* 2009; Hubbard *et al.* 2010), but the extent to which such findings can be extended to divergent taxa across mammals is unclear. Consequently, studying skin pigmentation in highly divergent mammals is critical for understanding its molecular basis and evolution.

Marsupials are a large group of mammals representing over 300 species and are evolutionarily important because they form the outgroup to placental mammals (Tyndale-Biscoe 2005). Nevertheless, very few studies on pigmentation variation have been done on this group (Hope and Godfrey 1988; Dawson and Maloney 2004; Dawson *et al.* 2014). The Virginia opossum (*Didelphis virginiana*) is a marsupial of tropical origin that is currently widely distributed throughout Central and North America (Gardner and Sunquist 2003). The fossil record and paleoclimate reconstructions indicate that this species has been expanding its range northwards into more temperate environments during the last 15,000 – 11,000 years (Guilday 1958; Graham *et al.* 1996; Bartlein *et al.* 1998; Morgan 2008; Graham and Lundelius 2010), and shows phenotypic variation in coat and skin pigmentation across its geographic range (Allen 1901; Gardner 1973; McManus 1974; Gardner and Sunquist 2003). Variation in pigmentation in this species has been found to follow Gloger's rule and is associated with precipitation and temperature. Specifically, opossums distributed in tropical latitudes ($< 24^{\circ}\text{N}$), which are exposed to higher precipitation and temperature, have darker coats and higher percentage of pigmented skin on their ears, hindfeet and tails compared to opossums inhabiting drier and cooler temperate environments. Natural selection and phenotypic plasticity have been suggested as possible

processes responsible for this pattern (Nigenda *et al.* unpublished data); however, genetic studies have not been done to test these hypotheses. Interestingly, the Virginia opossum is almost strictly nocturnal and possibly color blind (Jacobs and Williams 2010), which does not support a role for pigmentation variation in camouflage or inter- and intra-specific communication.

The molecular mechanisms of pigmentation in marsupials have not been studied. Further, no previous information at the genomic level is available for the Virginia opossum and the phenotypic variation observed in the Virginia opossum could have arisen relatively recently due to the expansion of the species into temperate environments (Guilday 1958; Graham *et al.* 1996; Gardner and Sunquist 2003). Consequently, as a first step to understand the molecular, genetic and evolutionary mechanisms underlying variation in skin pigmentation in the Virginia opossum, we assessed the transcriptome-wide gene expression of the skin in individuals from wild populations inhabiting tropical and temperate environments, which shows different pigmentation phenotypes. We implemented a differential expression approach and a Weighted Gene Coexpression Analysis (WGCNA; Langfelder and Horvath 2008), which is a method that examines gene expression in terms of coexpression connectivity, identifying modules of coexpressed genes associated to trait variation and the key genes that are likely driving this variation (Oldham *et al.* 2006; Filteau *et al.* 2013; Pardo-Diaz *et al.* 2015). The benefit of a network-based approach is its capability to uncover complex biological mechanisms responsible for the phenotype variation (Weiss *et al.* 2012). This approach has been successfully used to identify modules and genes associated with complex diseases and traits (Ghazalpour *et al.* 2006; Horvath *et al.* 2006; Miller *et al.* 2008; Plaisier *et al.* 2009), but its application in the study of phenotypic variation in wild vertebrate populations has been limited (Filteau *et al.* 2013). Our objective was to link gene expression patterns with skin pigmentation phenotypic variation and

identify potential candidate genes driving this diversity. We found strong positive correlation between skin pigmentation traits and the expression of gene modules associated with pigmentation functions and a negative correlation with the expression of modules related to immune pathways. In addition, we observed that in populations with depigmented skin, genes related to pigmentation processes are down-regulated, while genes involving an immune response are up-regulated. We propose two possible explanations for our findings: 1) depigmentation of the skin in higher latitudes results from a pathogen release mechanism; and 2) the thermal stress imposed by low temperatures in temperate environments triggers an immune response analogous to that known for vitiligo disease, a disorder causing depigmentation in humans.

Methods

Sampling and phenotypic data collection

Based on previous work indicating that Virginia opossums have different skin pigmentation phenotypes above and below 25° N (Nigenda *et al.* unpublished data), we collected phenotypic data and tissue samples from 13 individuals from three different populations during the Spring and Summer of 2013. Two of these populations were located below 25°N, in the states of Morelos (18-19°N; four individuals) and Yucatan (20-21°N; four individuals) in Mexico, and one population was located above this latitude in South Carolina (33°N; five individuals), United States (Table 2.1). To minimize error in calculating expression levels related to age and time of year, we collected samples only from adult individuals older than 10 months old and during months April- July. Age was determined using an approximation technique based on tooth eruption sequence (Gardner 1982).

Opossums were tranquilized using 0.15 ml of Telazol/kg. We collected data on sex and three skin pigmentation traits in all individuals (Table 2.1). We recorded the proportion of pigmented skin of the tail (PTP), left ear (PEP), and ventral aspect of the left hindfoot's middle digit (PTO). These traits have previously been shown to vary between populations along the geographic range of the species (Gardner 1973; Gardner and Sunquist 2003; Nigenda *et al.* unpublished data). Using sterile biopsy punchers, 4 mm skin samples were taken from the most distal part of the naked opossum pinnae, where most individuals from northern populations have their skin depigmented, whereas southern individuals are completely pigmented. The samples were immediately preserved in RNAlater solution for 24 hours at room temperature and then frozen in dry ice or in a -80°C freezer until RNA extraction. Sampling in Mexico was completed under Mexican permit SGPA/DGVS/01396/1. South Carolina sampling was done under the Institutional Animal Care and Use Committee (IACUC) protocol, #A2012 08-004-Y2-A2, and all fieldwork was done in compliance with UCLA's Office of Animal Research Oversight (OARO) protocol 2011-121-01.

To confirm that the opossums sampled exhibited pigmentation differences consistent with previous studies (Garner 1973; Gardner and Sunquist 2003; Nigenda *et al.* unpublished data), we performed ANOVA tests of pigmentation measurements between the three populations. As in previous studies, all pigmentation variables were significantly different between the northern and the southern populations, but no differences were found between the two southern populations (Table 2.S1). Therefore, we grouped the samples into two pigmentation groups, pigmented (Morelos and Yucatan populations) and depigmented (South Carolina population), which were used in the differential gene expression analysis.

RNA extraction, library preparation and sequencing

Total RNA was extracted from each of the 13 skin samples using the TRIzol Plus RNA purification kit (Ambion, ThermoFisher) protocol. The extractions had high quality and quantity as indicated by the Agilent 2100 Bioanalyzer traces (RNA 6000 pico kit) showing RNA integrity numbers (RIN) higher than 7.3. To prepare the cDNA libraries we used the Illumina TruSeq RNA Sample Prep Kit (Illumina, San Diego, CA, USA) according to the manufacturer's protocol. cDNA libraries were quantified on an Agilent 2100 Bioanalyzer High Sensitivity DNA chip. The libraries were randomized and pooled for sequencing as paired-end 100 bp in three lanes of an Illumina HiSeq 2000 platform at the Vincent J. Coates Genomics Sequencing Laboratory (UC-Berkeley). All laboratory work was done under the UCLA Institutional Biosafety Committee protocol IBC# 153.11.0.r.

Quality check, mapping, filtering and normalization

The sequencing reads were filtered to remove reads that did not pass the standard Illumina filter during sequencing. Read quality was checked and visualized with FastQC (Andrews 2010). Low quality reads, short reads and Illumina adapters were eliminated or trimmed using Trim Galore! 0.3.6 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore) so that only reads longer than 50 bp with Phred quality score higher than 20 were kept for further analysis. Using Tophat 2.0.11 (Kim *et al.* 2013), we mapped the reads from each individual to the Gray short-tailed opossum (*Monodelphis domestica*) genome. This species diverged 29-50 million years ago from the Virginia opossum according to the tree of life estimation (TTOL; <http://www.timetree.org>; Hedges *et al.* 2006; Hedges *et al.* 2015), and is the most closely related species to the Virginia

opossum with a good quality genome sequenced and annotated (Mikkelsen *et al.* 2007). To account for the divergence between the Virginia opossum and Gray short-tailed opossum genomes, we used the following Tophat2 parameters to maximize uniquely mapped reads: number of mismatches = 20, gap length = 24, edit distance = 28. Only reads that mapped uniquely were used to construct a read counts table with the tool htseq-count in the program HTSeq 0.6.1 (Anders *et al.* 2014). From these raw read counts, only genes with at least 3 reads mapped in all samples were selected for downstream analysis. However, during a preliminary analysis of the data, we noticed that several reads mapped to genes highly expressed in skeletal muscle. To remove potential effects of the presence of skeletal muscle tissue on skin-specific expression results, 1198 genes highly expressed in skeletal muscle were identified and removed from downstream analyses using a list of genes collated from four databases containing highly expressed genes in human and mouse skeletal muscle: TiGER (Liu *et al.* 2008), TissueDistributionDBS (Kogenaru *et al.* 2010), Tissue Specific Genes Database (TiSGeD; Xiao *et al.* 2010), and Genotype-Tissue Expression (GTEx; The GTEx Consortium 2013). After filtering out these genes, the raw counts of the remaining genes were normalized applying the conditional quantile normalization method as implemented in the cqn software (Hansen *et al.* 2012) to correct for library size and GC content, and gene length. The normalized counts were then transformed to log₂ scale for further analysis.

Data Analysis

Sample-based clustering and principal component analysis

To assess overall covariation of expression levels among samples, we used the normalized counts to perform a principal components analysis (PCA) of the gene expression

variation and a cluster analysis based on the correlation matrix of the expression profiles with the R statistical package (R core team 2015). Since all the samples with depigmented skin came from the South Carolina population, the effects of population and skin pigmentation on gene expression are confounded. Therefore, to determine if a significant fraction of the expression variation observed was mainly due to population structure or to processes related to pigmentation, we applied GO enrichment analysis to the 100 genes with the most positive loadings and 100 genes with the most negative loadings on the first and second principal components (PC1 and PC2). We chose to analyze the most extremely distributed genes at each end of the first two PCs because they are the ones probably driving most of the gene expression variation. In order to compare the result of these analyses to what would be expected by random sampling, we constructed empirical null models based on our data by performing 100 random sampling permutations of the top 100 genes with the most positive and negative loadings on PC1 and PC2, for a total of four null models (*i.e.* PC1 positive, PC1 negative, PC2 positive, PC2 negative); then, GO enrichment analysis was applied to all permutation trials searching for GO terms that have the words “immune”, “immunity” or “pigmentation”. This approach allows us to determine if our result may be significantly driven by biological or random processes. All permutations were done with the function `sample` in the R package (R core team 2015), and the GO analysis was performed using `g:Profiler` web package (Reimand *et al.* 2011) with a false discovery rate (FDR) threshold of 0.01 to control for multiple testing.

Differential expression analysis and weighted gene co-expression network analysis (WGCNA)

To identify genes with that have different expression patterns between depigmented and pigmented individuals we performed differential expression analysis. Differential expression was

calculated with an empirical Bayesian framework based on linear modeling implemented in the program limma-voom (Smyth 2005; Law *et al.* 2014) using an FDR < 0.1. The FDR was calculated with the q-value function in the R package qvalue (Storey and Tibshirani 2003).

Finally, we performed a detailed weighted gene co-expression network analysis (WGCNA; Langfelder and Horvath 2008) to identify groups of differentially expressed (DE) genes associated with pigmentation traits and potentially driving the differences in skin pigmentation. WGCNA is a systems biology approach that uses gene expression levels to identify modules (clusters) of co-expressed genes and their key members associated to phenotypic traits. Because gene modules can correspond to biological pathways and processes, focusing the analysis on modules and their hub genes (*i.e.* highly connected intramodular genes) is a comprehensive approach for biologically meaningful data reduction (Zhang and Horvath 2005; Langfelder and Horvath 2008; Langfelder *et al.* 2013). Briefly, we used the normalized, log₂ transformed counts to analyze the DE genes with the blockwiseModules R function in the WGCNA software for the construction of the weighted gene co-expression network (Langfelder and Horvath 2008). A correlation matrix is computed for the genes and using a power function the correlations are weighted to a power β (Zhang and Horvath 2005). The modules of co-expressed genes are identified as genes with similar patterns of connectivity with other genes applying a dissimilarity measure for hierarchical clustering and a dynamic tree-cutting algorithm (Zhang and Horvath 2005; Langfelder and Horvath 2008). Modules are denoted by colors, genes not assigned to a particular module are allocated to the grey module denoting background genes outside of modules (Langfelder and Horvath 2008; Langfelder *et al.* 2008). For our analysis, the parameters used were: maximum block size = 1186 genes, power (β) = 24, minimum module size

= 25; minimum height for merging modules = 0.15; maximum height for cutting the tree = 1. The remaining parameters were kept at the default settings.

Since our network analysis was on DE genes between skin phenotypes, we expect all modules to be highly associated with the pigmentation traits, which we confirmed using linear correlation between modules and traits. Therefore, in order to identify the modules and genes potentially explaining the variation in skin pigmentation we identified the modules with more intra-modular hub genes, which are the most likely to have an effect on the phenotypes (Han *et al.* 2004; Horvath *et al.* 2006; Langfelder *et al.* 2013). The hub genes of a given module are summarized by their first principal component, the module eigengene, which is employed to estimate module membership measurements to quantify the association of a gene to a given module. For each gene, the module membership value (*i.e.* kME) is the correlation between its expression value and the module eigengene. This measurement allows us to identify phenotype related intramodular hub genes (Langfelder and Horvath 2008; Horvath and Dong 2008; Langfelder *et al.* 2013). We identified the modules with more intramodular hub genes (designated with darker red color in the co-expression network), and calculated the module membership of all the genes to those particular modules to detect the most important hub genes likely driving the association with skin pigmentation. We implemented functional GO enrichment analysis using g:Profiler (FDR < 0.05) to test the biological roles of the modules and to identify intramodular hub genes as candidate driver genes.

Results

Sequencing, mapping and filtering

Sequencing of the transcriptome of 13 Virginia opossums resulted in 54-102 million read fragments per individual after Illumina filtering. After quality check filtering and trimming, 53-97 million reads per individual were suitable for mapping. 71-75 % of reads per individual (38-71 million reads) mapped to the *Monodelphis* genome, with 35-66 million reads uniquely mapped per individual for final use in the read counts table. Of 23,899 annotated genes in the *M. domestica* genome, 15,487 of these genes had more than 3 reads uniquely mapped in all individuals. After removing genes highly expressed in skeletal muscle tissue, 14,289 genes were used in subsequent analysis.

Sample-based clustering and PCA

Clustering analysis of gene expression reveals that the samples mainly cluster by population with only two individuals (YUC3M and MOR3M) from each of the southern populations, clustering separated from their populations of origin (Figure 2.1A). However, at the most basal hierarchical level, the samples are grouped by skin pigmentation phenotype since the pigmented populations (MOR and YUC) are grouped together and are separated from the depigmented samples (SCA; Figure 2.1A). The first two principal components of overall expression cluster samples by the three populations, but PC1 (explaining 20% of the variance) separates the samples by pigmentation phenotype (Figure 2.1B), with individuals from both pigmented populations grouped together and separated from depigmented individuals (Figure 2.1B).

To determine the contribution of population structure or biological processes on the variation of gene expression profiles, we performed enrichment analysis of the 100 genes with the most positive and most negative loadings on PC1 and PC2, and compared it to null models

constructed with 100 random permutations. We found that genes with the most negative loadings on PC1 were significantly enriched for the GO terms biological processes of pigmentation (P-value = $1.23e-05$) and oxidation-reduction process (P-value = $4.07e-03$), whereas genes with most positive loadings were significantly enriched for the immune response (P-value = $4.09e-06$) and defense response (P-value = $5.31e-03$) GO terms. No significant enrichment was detected for any of the GO terms for genes on PC2. Both null models for genes on PC2 and the model for genes with negative loadings on PC1 did not show enrichment for pigmentation or immune GO terms, whereas these terms were enriched just once in the null model for positive loadings of PC1. The enrichment of pigmentation genes on the negative loadings of PC1 is consistent with their complete separation by pigmentation type on axis 1 of the PCA (Figure 2.1B), in which individuals from pigmented populations have negative or low values on PC1, whereas enrichment of immune functions in genes with positive loadings may suggest they are also involved in driving this separation. The almost negligible enrichment of pigmentation and immune or defense response in the sets of random genes shows that the probability of that our results would be seen in a random set of genes is less than or equal to 0.01 and supports the notion that population structure plays a less significant role in the separation observed on PC1.

Differential expression and WGCNA

We compared the gene expression profiles between skin pigmentation groups and identified 1,186 genes differentially expressed between pigmented and depigmented populations, comprising 760 genes up-regulated and 426 down-regulated in opossums with depigmented skin. Of all the DE genes, WGCNA organized 1,045 genes into 10 modules of highly correlated genes. Four (*i.e.* purple, black, blue and pink) and six (*i.e.* brown, yellow, green, turquoise, magenta and

red) modules contain genes that are down-regulated and up-regulated, respectively, in depigmented opossums. Almost all the gene modules had significant correlation with the three skin pigmentation traits (Figure 2.2). In general, the proportion of pigmentation of the ear was the trait with lower correlation values, having non-significant correlation with the black, green, red and yellow modules (Figure 2.2). Table 2.2 presents the gene ontology (GO) biological process terms and biological pathways significantly enriched in each module. Notably, the pink module is functionally enriched for pigmentation, while the brown, yellow, and green modules are all enriched for GO terms related to immune or inflammatory responses (*i.e.* activation of NF-kappaB-inducing kinase activity, antigen processing and presentation, and humoral immune response, respectively; Table 2.2; Baldwin 1996, Bonizzi and Karin 2004). In addition, the blue module is enriched for DNA repair/cell cycle and the turquoise and magenta modules are both enriched for ion transport processes (Table 2.2).

In the network of co-expressed genes (Figure 2.3), the increasing color intensity indicates higher connectivity (coexpression similarity) among genes in the network, with red indicating genes with the greatest intramodular connectivity (or hub genes). This analysis reveals that the blue, pink, yellow, and turquoise modules have hub genes with higher connectivity with other genes within their modules (Figure 2.3). Because of the potential biological relevance and importance of hub genes in co-expression networks in driving phenotypic variation (Han *et al.* 2004; Filteau *et al.* 2013; Langfelder *et al.* 2013), we ranked the genes within each of these modules by intramodular connectivity to identify the top 25 module hubs. Among the top hub genes in the pink module are eight genes involved in melanogenesis, melanosome biogenesis, pigmentation, or production of melanin process including *MLANA*, *MLPH*, *EDNRB*, *PAX3*, *TYRP1*, *DCT*, *TCF7L2* and *TYR* (Bennett and Lamoreux 2003; Slominski *et al.* 2004; Hoeskstra

et al. 2006; Lin and Fisher 2007; Hou and Pavan 2008) and genes involve in transmembrane ion transport (*KCNGL1*, *CACNB3*, *TRPC3*, *KCNJ13*; Table 2.3; Su *et al.* 1998; Bros *et al.* 2011; Pattnaik *et al.* 2013; Sulk and Steinhoff 2015). In addition to the WGCNA results, we found that several genes with melanocytic functions that were down-regulated in depigmented opossums (*i.e.* *TRPM1*, *SLC45A2*, *GPR143*), including some of the hub genes in the pink module (*i.e.* *MLANA*, *TYRP1*, *DCT*, and *TYR*), are directly regulated by *MITF* (Figure 2.4), which is known as the “melanocyte master regulator,” because it plays a crucial role in melanocyte development by regulating the expression of several genes in the melanogenesis pathway (Bennett and Lamoreux 2003; Baxter *et al.* 2009; Lu *et al.* 2010). Although *MITF* itself did not show differences in expression, this result may suggests *MITF* is involved in regulating depigmentation in opossums. The presence of several hub genes with melanocytic functions in the pink (pigmentation) module, and the fact that this module is positively correlated with the skin pigmentation traits and down-regulated in depigmented individuals, suggests that these genes may be important in regulating the skin pigmentation differences observed between tropical and temperate opossum populations.

In the yellow module, the top five hub genes have immune functions including major histocompatibility complex genes *DQA1*, *DQA2*, and *DQB1*, participating in antigen recognition and presentation processes (Tsai and Santamaria 2013; Unanue *et al.* 2016). Similarly, *CTSH* (*Cathepsin H*) also participates in antigen presentation and is involved in apoptosis (Conus and Simon 2008; D’Angelo *et al.* 2010), while *IL4I1* (*Interleukin 4 Induced 1*) encodes an enzyme with immunoregulatory and antibacterial properties (Puiffe *et al.* 2013). Moreover, other genes implicated in apoptosis suppression, stress-induced immune or inflammatory response are among the top hub genes in this module including *STAT5A*, *STAT5B*, *MMP25*, *ELF4*, *TRAF1* and

PRKCB (Table 2.3; Wei *et al.* 2008; Shiryaev *et al.* 2009; Lu *et al.* 2004; Chung *et al.* 2002; Wang *et al.* 2009). Although the green module did not exhibit a strong signal for hub genes, several of the genes with the highest kME in this module have immune functions (Table 2.3). Specifically, these genes participate in the activation of innate immune response (*e.g.* *CD209*; Cox *et al.* 2015), the inflammatory response (*e.g.* *CDI63*; Moestrup and Møller 2004; Onofre *et al.* 2009) or are part of the complement system (*e.g.* *CIQA, CIQB, CIQC, CIR, CD59*; Nesargikar *et al.* 2012). These results imply that depigmented individuals are mounting a distinct immune response.

The turquoise module has hub genes involved in ion transmembrane transport including *VPS9D1, KCNN1, PRKAB2, and NDUFA4* (Table 2.3; Sugimoto *et al.* 1999; Boettger *et al.* 2002; Steinberg and Kemp 2009; Pitceathly *et al.* 2013); this observation together with the up-regulation of the Magenta module, which is enriched for ion transport processes (Table 2.2), suggests that ion transportation is activated in depigmented opossums. In contrast, several of the hub genes in the blue module participate in cell cycle processes including *TOP2B, NEK2, KNSTRN, CDK5RAP2, UBE2C* and *PBK* (Table 2.3), these genes act as positive signals for cell division (*i.e.* *NEK2, KNSTRN, CDK5RAP2, UBE2C* and *PBK*; Rape and Kirschner 2004; Fang *et al.* 2009; Zhang *et al.* 2009; Boekhout and Wolthuis 2015; Rizkallah *et al.* 2015) or for DNA replication (*i.e.* *TOP2B*; Sakaguchi and Kikuchi 2004). The down-regulation of these genes in depigmented individuals and of other genes in the blue module (Table 2.2; Table 2.3) may indicate the inhibition of cellular proliferation in depigmented skin.

Discussion

We investigated whether gene expression and gene networks are correlated with skin

pigmentation variation between tropical and temperate Virginia opossum populations to provide mechanistic and evolutionary insight into the origin of this pattern and to identify candidate genes for future targeted sequencing and functional studies. Using WGCNA we find that skin pigmentation traits of the Virginia opossum are correlated with genes and gene modules with melanocytic and immune functions, suggesting that these genes have functional importance in regulating skin pigmentation variation and differences in immune response of the skin.

Pigmentation and Immune genes drive gene expression variation

We cannot rule out the effect of population structure on our results because all samples with depigmented skin derive from one population. However, several observations suggest pigmentation differences are driving the pattern. First, all individuals from two different southern populations had pigmented phenotypes that are distinguished from those in the North with depigmented skin (Figure 2.1B, Table 2.1), and considerable distance (1,000-3,000 kilometers) separates all three populations from each other (Table 2.S2), indicating that the skin coloration differences observed might not be solely explained by genetic drift. Secondly, the enrichment of pigmentation and immune response GO terms among genes with the most negative and positive loadings in PC1 of the PCA (Figure 2.1B) was not seen in random sets of genes, suggesting that genes involved in those functions show differences in expression beyond what is expected from population structure alone. Finally, our observations are concordant with morphological studies suggesting a functional explanation for the skin and coat coloration variation in the Virginia opossum rather than effects of random drift (Gardner 1973; Nigenda *et al.* unpublished data). Specifically, our results are consistent with genetic changes in gene regulation driven by pathogen-related selection (Oleksiak *et al.* 2002; Alvarez *et al.* 2015) or developmentally

induced changes responding to temperature stimuli (Gilbert 2005; Espinosa-Soto *et al.* 2011).

Using differential expression analysis and WGCNA, we identified multiple genes and gene modules significantly correlated with skin pigmentation differences in tropical and temperate populations. As with PCA, the WGCNA of DE genes revealed gene modules enriched for pigmentation and immune response GO terms (Table 2.2; Figure 2.3). Many of the hub genes in the pigmentation (i.e. pink) module (Table 2.3) that were positively correlated with the skin pigmentation traits (Figure 2.2) and down-regulated in depigmented individuals, also are associated with pigmentation variation in wild mammal populations, (*e.g.* *TYRP1*; Gratten *et al.* 2007) and skin or coat coloration phenotypes in humans and laboratory mice strains (*e.g.* *TYR*, *DCT*, *EDNRB*, *PAX3*, *MLPH*, *MLANA*; Bennett and Lamoreux 2003; Beermann *et al.* 2004; Hoekstra *et al.* 2006). In addition, other down-regulated genes in depigmented opossums that are neither located within the pigmentation module nor are hub genes, are important transcription factors or genes that cause depigmentation in mice (*GPRI43*, *SOX10*, *ATRN*; Bennett and Lamoreux 2003; Hoekstra 2006), hooded crows, horses, polar bears and humans (*TRPM1*, *SLC45A2*, and *LEF1*; Kingo *et al.* 2008; Norton *et al.* 2007; Bellone *et al.* 2008; Nielsen *et al.* 2009; Miller *et al.* 2012; Poelstra *et al.* 2014). The consistent functional enrichment by different analyses for pigmentation GO terms suggests that these genes are important in the molecular mechanisms underlying skin pigmentation in opossums and may be driving the differences between tropical and temperate populations. Taken together, these results suggest that the genetic regulation of skin depigmentation in opossums is complex, including genes in the melanogenesis pathway, melanocyte development, and melanosome transport (Bennett and Lamoreux 2003; Dessinioti *et al.* 2009). Such down-regulation of many melanocytic genes may be caused by an upstream regulatory change or a low number of melanocytes (Spritz 2013; Poelstra *et al.* 2014).

The negative correlation of the pigmentation traits with the immune-related modules (*i.e.* yellow, brown and green; Figure 2.2), down-regulation of the pigmentation module and up-regulation of immunity modules (Table 2.2; Figure 2.3) in depigmented individuals, indicate that lower pigmentation is associated with higher levels of expression of immune genes and pathways. Several hub genes in the immune modules are important in vertebrate innate immunity (e.g. *TRAF1*, *STAT5B*, *C1QB*, *C1QC*, *C1R*, *CD209*, *CD103*; Chung *et al.* 2002; Wei *et al.* 2008; Onofre *et al.* 2009; Sarma and Ward 2011; Cox *et al.* 2015) and the adaptive immune response to extracellular pathogens (e.g. *DQA*, *DQB*; Bernatchez and Landry 2003; Piertney and Oliver 2006). These results imply that depigmented opossums have activated an immune response absent in their pigmented conspecifics.

The enrichment for ion transport processes found in the turquoise and magenta modules may be part of the response of the skin's neuroendocrine system in depigmented individuals to the stimuli caused by low temperatures, which involves the up-regulation of genes participating in ion transport to facilitate the communication of environmental signals to the central nervous system in order to provide appropriate physiological or behavioral responses (Slominski 2005; Slominski *et al.* 2012).

Skin pigmentation variation hypotheses

Two hypotheses may explain the gene and module expression patterns we found. An adaptive hypothesis of pathogen resistance, proposing that darker morphs are favored in tropical latitudes because the immune properties of melanin; or a developmental plasticity hypothesis in which thermal stress due to cold conditions in high latitudes causes an autoimmune reaction against melanocytes.

Pathogen resistance hypothesis

Finding direct evidence for the mechanisms underlying Gloger's rule in natural populations has proven to be difficult (Caro 2005, Lai *et al.* 2008). Here, we provide evidence that the skin pigmentation variation between opossums from tropical and temperate populations is correlated with expression of gene modules with melanocytic and immune functions. Our findings are consistent with the role of melanocytes, melanosomes and eumelanin as integral components of the innate immune system (Mackintosh 2001; Plonka *et al.* 2009) and with the hypothesis that both immunity and skin color represent reaction pathways to differences in pathogen prevalence (Wassermann 1965; Mackintosh 2001, Elias 2007). The pathogen resistance hypothesis to explain Gloger's rule proposes that higher pathogen load in humid tropical environments results in selection for increased pigmentation in tropical populations (Burt and Ichida 2004), presumably because more highly pigmented skin, hair or feathers confer better resistance to pathogenic infection (Wassermann 1965; Mackintosh 2001; Burt and Ichida 2004; Elias 2007). Natural selection related to immune functions has been implicated in driving the pigmentation polymorphism in some wild vertebrates populations, including Greenfinches (*Carduelis chloris*; Männiste and Hõrak 2014) and gray wolves (*Canis lupus*; Anderson *et al.* 2009; Coulson *et al.* 2011), with darker individuals being more resistant to pathogens. Potential mechanisms by which highly pigmented skin confer increased pathogen resistance in individuals include: 1) having a more competent membrane barrier function, causing a drier skin surface adverse for pathogenic colonization (Elias 2007; Elias and Williams 2013); 2) lower pH (acidic environment), which is hostile to microbial pathogen growth (Korting *et al.* 1990; Elias 2007); 3) antimicrobial activity of melanin intermediates (Mackintosh, 2001; Drake *et al.* 2008; Elias *et al.* 2009); and 4) antibiotic properties of eumelanin (Montefiori and Zhou 1991; Mackintosh 2001).

Therefore, the potential effect of pigmentation on pathogen resistance may be an important mechanism explaining conformance to Gloger's rule in the Virginia opossum (Mackintosh 2001; Burt and Ichida 2004; Elias 2007). We hypothesize that opossums in the temperate zone have reduced innate protection from pathogens due to their lower pigmentation, and must therefore maintain a more active (*i.e.* highly expressed) immune system in their skin compared to opossums in the tropics.

Our findings imply that high skin pigmentation has an important function in tropical environments to protect opossums from the higher pathogens diversity in those habitats (Guernier *et al.* 2004; Elias *et al.* 2009; Lafferty 2009). In contrast, as opossums expanded northwards 15,000 years ago colonizing temperate environments (Graham *et al.* 1996; Morgan 2008; Graham and Lundelius 2010), a decreased pathogen diversity or less virulent pathogen array would have relaxed selective pressures on skin pigmentation and favored less pigmentation (Mackintosh 2001; Elias 2007). Specifically, there may be trade-offs between maintaining high pigmentation, which has a relatively high metabolic cost, and other energetically expensive physiological requirements of living in colder environments (*e.g.* more energy allocation to maintain body temperature or increase body size; Stoehr 2006; Moreno and Møller 2006; Gasparini *et al.* 2009). Nevertheless, we observe a heightened immune response in temperate opossums that evolved rapidly and may be in response to novel pathogens and the decrease in pigmentation. Opossums show high frequency of frostbite injuries (*i.e.* lost or lesions in ears, toes and tail) in high latitudes locations (Gardner and Sunquist 2003), these injuries may facilitate pathogen infection of the body parts affected. In addition, pathogens affecting the skin have caused severe manifestations in temperate populations of mammals such as the white-nose syndrome in bats (Blehert *et al.* 2009; Lorch *et al.* 2011) and mange in wolves (Jimenez *et al.*

2010) and bobcats (Riley *et al.* 2007), therefore, an effective immune response against skin pathogens is important in temperate environments. Our results suggest that the pathogens causing this immune response may not be susceptible to the effects of melanin or do not impose a selection pressure that is greater than that imposed by the energetic costs of melanin production.

Environmental stress hypothesis

Alternatively, depigmentation and increased immune response in individuals from temperate populations may be in response to environmental stress. Skin depigmentation in humans is caused by environmental stress on melanocytes (Richmond *et al.* 2013) triggering an autoimmune response, a disease condition called vitiligo (Figure 2.5; Rezaei *et al.* 2007; Spritz 2012). Environmental stressors induce the production of reactive oxygen species (ROS) and damage-associated molecule patterns (DAMPs), which provide the signal to activate the innate immune system and stimulate an adaptive immune response leading to melanocytes destruction and consequent depigmentation (Ezzedine *et al.* 2015). Low temperatures in high latitudes might produce the initial stress in the opossum. Melanocytes are sensitive to cold (Gage 1979; Page and Shear 1988), and at about 8°C skin cells start producing ROS (Awad *et al.* 2013). Since the Virginia opossum is a marsupial species of tropical origins (with lower body temperature than Eutherian mammals; McNab 1978; Tyndale-Biscoe 2005) the thermal stress may be even higher. We found that three chaperones of heat shock proteins (*i.e.* *DNAJA2*, *DNAJB6*, *DNAJC28*), which are common DAMPs (Chen and Nuñez 2010), were up-regulated in depigmented opossums and may have triggered the innate and adaptive immune responses observed in our data. Skin transcriptome studies in humans discovered that depigmented vitiligo skin shows

down-regulation of several melanocytic hub genes we found under-expressed in depigmented opossums (*i.e.* *TYR*, *TYRP1*, *MLANA*, *DCT*, *GPR143* and *PAX3*; Yu *et al.* 2012; Dey-Rao and Sinha 2016; Regazzetti *et al.* 2015; Table 2.2; Table 2.3), and suggest that such down-regulation may indicate a reduction in melanocytes numbers (Spritz 2013). Concordant with our results, these studies also found innate immunity genes up-regulated in depigmented skin (Yu *et al.* 2012; Dey-Rao and Sinha 2016; Regazzetti *et al.* 2015; Table 2.2). In addition, genome wide association studies (GWAS) have implicated MHC class I and II genes with vitiligo (Quan *et al.* 2010; Jin *et al.* 2010; Jin *et al.* 2012). This is consistent with the enrichment we observed in the yellow module for antigen processing and presentation processes and up-regulation of MHC genes in depigmented opossums, including *DQA1*, *DQA2* and *DQB* for the class II, and *MODO-UT3* and *MODO-UT5* for class I (Table 2.2; Table 2.3). *MODO-UT3* and *MODO-UT5* are part of a recently discovered group of MHC class I genes that is unique for marsupials (Papenfuss *et al.* 2015). Also, the down-regulation of cell cycle genes in depigmented opossums (Table 2.2; Table 2.3) suggests an inhibition of cellular reproduction of melanocytes. The inhibition of cellular reproduction has been described in other autoimmune condition of the skin (Subramanya *et al.* 2010). Therefore, our results may reflect an epigenetic autoimmune reaction against skin melanocytes triggered by the low temperatures that opossums encounter in northern latitudes. Such developmental changes are common when populations colonize new environments (Price and Sol 2008; Lande 2009; Wund 2012), which is true for the Virginia opossum. In addition, both adaptive and non-adaptive plasticity seem to favor adaptive evolution (Ghalambor *et al.* 2007; Ghalambor *et al.* 2015).

Wnt signaling pathway regulates opossum depigmentation

The melanogenesis pathway is a transcriptional activation process that results in the production of eumelanin (Poelstra *et al.* 2015). Our results show that regardless of the evolutionary mechanism underlying skin depigmentation in opossums, the cAMP-dependent pathway might not be involved in regulating skin pigmentation in the opossum, as might be predicted based on research in other vertebrate species (Hoekstra 2006; Hubbard *et al.* 2010; Poelstra *et al.* 2014; Poelstra *et al.* 2015), since the up-regulation of one of the genes in this pathway (*i.e.* *PRKACB*) suggests the activation of melanin production, which is inconsistent with depigmentation (Figure 2.4). Instead, the Wnt/ β -catenin signaling pathway, which plays critical roles in several biological process, including embryonic development, cell proliferation and tissue homeostasis (Logan and Nuse 2004), may be controlling the depigmentation in opossums since we found several genes down-regulated in this pathway in depigmented individuals (Figure 2.4), consistent with transcriptional suppression inhibiting melanin production (Poelstra *et al.* 2015). *WNT3* expression promotes the differentiation and reproduction of melanocytes in cultured mouse neural crest cells (Hou and Pavan 2008), whereas *DVL3* is an important regulator of β -catenin, which interacts with two important transducers of the Wnt pathway, *LEF1* and *TCF7L2* (Gao and Chen 2010), which were found underexpressed in depigmented opossums (Table 2.2; Figure 2.4; Saito *et al.* 2003). *LEF1* down-regulation has been described as characteristic of depigmented skin (Kingo *et al.* 2008; Regazzetti *et al.* 2015) and *LEF1/TCF7L2* directly induce the expression of *MITF* (Saito *et al.* 2003; Hou and Pavan 2008). Although *MITF* was not itself differentially expressed, possibly because its wide expression in several cell types dilutes the signal from melanocytes (Poelstra *et al.* 2015), the fact that most melanocytic genes down-regulated in depigmented opossums are directly regulated by this gene (Figure 2.4) suggests that *MITF*, or genes closely upstream in the melanogenesis pathway, might play

important roles regulating skin depigmentation (Poelstra *et al.* 2015). In addition, *LEF1* can also regulate melanogenesis through *SOX10* transcriptional induction that then regulates the expression of *EDNRB* (Yokoyama *et al.* 2006; Pingault *et al.* 2010). Further, the transcription factor *PAX3* transactivates *MITF* and can directly regulate the TYRP1 promoter (Figure 2.4; Galibert *et al.* 1999; Baxter *et al.* 2009). Finally, the involvement of Wnt/ β -catenin signaling pathway in vitiligo skin depigmentation has been previously described in humans (Dey-Rao and Sinha 2016; Regazzeti *et al.* 2015).

Summary and conclusions

To our knowledge, this study is the first to investigate the genetic mechanisms underlying the variation of skin pigmentation in a marsupial species. Using an integrative gene network analysis we demonstrate that the skin pigmentation variation between opossums from tropical and temperate populations is associated with gene expression levels of genes involved in pigmentation and immunity. We present several candidate (hub) genes related to these functions that may be driving the variation observed in opossums. Also, we postulate that the Wnt/ β -catenin signaling pathway might be implicated in regulating the skin pigmentation in the opossum, contrary to that found in other vertebrate species. Finally, we show that skin pigmentation geographic variation in the Virginia opossum is related to evolutionary processes in temperate environments involving immune functions and given the recent expansion of the species in temperate latitudes this phenotypic evolution may have occurred rapidly, we present two feasible hypotheses to explain this phenotypic variation that merit further investigation. To test which hypothesis explains this phenotypic variation, immunohistochemical analysis to determine if depigmentation is due to low numbers of melanocytes (depleted by an autoimmune

response) would be needed. Resequencing the candidate genes identified in our study, including up- and downstream genomic regions, in a genome capture array design would facilitate the discovery of sequence variation associated with changes in gene expression (*e.g.* Schweizer *et al.* 2016b) and help to determine if the generalized down-regulation of melanocytic genes is caused by an upstream regulatory change. Finally, common garden experiments including the rearing of southern and northern opossums under temperate and tropical conditions may be used to test the stress hypothesis and better identify the development factors affecting pigmentation and immunity.

We provide evidence that Gloger's ecogeographical pattern of skin pigmentation variation in the Virginia opossum is associated with changes in the expression of genes with melanocytic and immune response functions, and that differences in pathogenic selective pressure or in thermal stress associated with living in temperate environments could be driving the phenotypic variation in this marsupial species. Our study postulates a new model for investigating the complex mechanistic, evolutionary and genetic basis of trait differences in non-model vertebrates using a system biology gene network approach (especially when previous genomic information is absent) to develop testable hypotheses of causation that could be the subject of future better-controlled studies. We show that gene expression quantification in wild populations with defined phenotypic variation offer natural experiments to test how individuals can respond rapidly to environmentally varying conditions.

Tables and Figures

Table 2.1. Geographic and phenotypic information on the sampled opossums used for the transcriptome analysis. LAT= Latitude, LONG= Longitude, TPP=Tail pigmentation proportion, EPP= ear pigmentation proportion, TOP= toe pigmentation proportion. We provide the mean for the phenotypic traits per population.

ID	Sex	Population	LAT	LONG	TPP	EPP	TOP	RL	CL	TL
MOR1	M	Morelos	18.985	-99.234	0.685	1	1	40.22	70.95	49.08
MOR2	M	Morelos	18.982	-99.235	0.501	1	1	33.18	69.14	26.68
MOR3	M	Morelos	18.981	-99.237	0.544	1	1	34.15	62.96	40.77
MOR4	M	Morelos	18.984	-99.239	0.508	1	1	39.76	66.48	36.24
<i>Mean</i>					<i>0.559</i>	<i>1</i>	<i>1</i>	<i>36.83</i>	<i>67.38</i>	<i>38.19</i>
YUC1	M	Yucatan	21.034	-89.616	0.475	1	1	39.63	58.55	52.37
YUC2	M	Yucatan	20.964	-89.576	0.606	1	1	35.02	61.08	38.64
YUC3	M	Yucatan	21.035	-89.618	0.698	1	1	34.68	62.82	35.84
YUC4	F	Yucatan	21.036	-89.617	0.589	1	1	39.99	62.47	44.25
<i>Mean</i>					<i>0.592</i>	<i>1</i>	<i>1</i>	<i>37.33</i>	<i>61.23</i>	<i>42.78</i>
SCA1	M	South Carolina	33.287	-81.563	0.184	0.92	0	60.32	66.77	59.2
SCA2	M	South Carolina	33.292	-81.701	0.153	0.71	0	64.41	72.11	65.06
SCA3	M	South Carolina	33.295	-81.692	0.241	0.96	0	63.47	68.62	63.43
SCA4	M	South Carolina	33.294	-81.544	0.179	0.98	0	41.74	65.52	56.71
SCA5	F	South Carolina	33.310	-81.691	0.144	0.57	0	44.64	60.88	61.19
<i>Mean</i>					<i>0.180</i>	<i>0.828</i>	<i>0</i>	<i>54.92</i>	<i>66.78</i>	<i>61.12</i>

Table 2.2. Gene ontology terms and KEEG pathways enriched in the gene modules of differentially expressed genes. The direction of the gene expression of the modules in depigmented opossums is shown as Down (down-regulated) or Up (up-regulated). Genes that belong to the enriched terms and pathways are shown for all modules. N/S= No significant.

Module	Direction of expression	# genes	Enriched biological process (Enrichment p-value) (# of genes)	Genes	Enriched biological pathway (Enrichment p-value) (# of genes)	Genes
Purple	Down	47	N/S		N/S	
Black	Down	73	Connective tissue development (0.05) (6)	<i>GLI2, SFRP2, TGFB2, MKX, DLX2, NKX3-2</i>	TGF-beta signaling pathway (0.04) (3)	<i>TGFB2, RHOA, SKP1</i>
Blue	Down	160	DNA repair (1.8e-03) (12)	<i>REC8, NVL, MGME1, SWI5, MBD4, TRPC2, RAD51C, BRCA1, MSH5, NHEJ1, WRN, NSMCE2</i>	N/S	
			Regulation of cell cycle process (1.40e-04) (14)	<i>KNSTRN, BUB1B, PSRC1, CEP131, NEK2, PKP4, KIF20B, FGFR2, RNASEH2B, RAD51C, BRCA1, UHRF2, CDK5RAP2, NSMCE2</i>		
Pink	Down	64	Pigmentation (6.30e-07) (7)	<i>EDNRB, TYR, DCT, TYRP1, GPR143, LEF1, PAX3</i>	Melanogenesis (5.99e-06) (6)	<i>EDNRB, TYR, DCT, TYRP1, TCF7L2, LEF1</i>
			Epithelium development (3.64e-03) (11)	<i>EDNRA, B9D1, EDNRB, TCF7L2, TYRP1, TBC1D32, ERCC3, TFAP2B, LEF1, PAX3, APCDD1</i>	Tyrosine metabolism (2.58e-03) (3)	<i>TYR, DCT, TYRP1</i>
Brown	Up	125	Activation of NF-kappaB-inducing kinase activity (2.47e-04) (4)	<i>TRAF2, TNFSF15, TRAF6, MALTI</i>	Small cell lung cancer (0.05) (5)	<i>E2F3, TRAF2, PIK3CB, TRAF6, LAMA1</i>
Yellow	Up	104	Antigen processing and presentation (0.05) (5)	<i>MODO-UT3, MODO-UT5, DQB, DQA1, DQA2</i>	Leishmaniasis (8.86e-03) (4)	<i>C3, DQB1, DQA2, PRKCB</i>
Green	Up	94	Humoral immune response (3.48e-03) (5)	<i>CD59, SPNS2, CIS, CIR, CIQA</i>	Complement and coagulation cascades (1.47e-09) (9)	<i>CD55, CD59, F13A1, C4A, CIQC, CIS, CIR, CIQB, CIQA</i>

					Systemic lupus erythematosus (6.61e-07) (7)	<i>C4A, CIQC, CIS, CIR, CIQB, C1QA, H3F3C</i>
Turquoise	Up	244	Inorganic ion transmembrane transport (4.86e-02) (15)	<i>MICU1, KCNC4, SLC11A2, ABCC8, SLC26A7, ATP5G1, UQCERS1, SLC17A7, NDUFA4, KCNK3, COX5B, ANK3, ABCC9, KCNN1, HCN1</i>	N/S	
Magenta	Up	54	Ion transport (3.74e-02) (8)	<i>RAB3B, SLC9A2, EGF, KCNC1, GRIK4, SLC4A3, KCNN2, ATP5E</i>	N/S	
Red	Up	80	N/S		N/S	

Table 2.3. Top 25 hub genes in the pink, yellow, turquoise, blue and green modules. Hub genes were identified with WGCNA as the genes with higher connectivity and module membership (KME) in that particular module. For genes that were not annotated in the opossum genome their corresponding Ensembl ID numbers are shown. In bold are the genes that have their function concordant to the biological processes enriched in their respective modules (Table 2.2): Melanocytic genes in the pink module, immune-related genes in the yellow and green modules, ion transport genes in the turquoise module, and cell cycle genes in the blue module.

Pink module genes	Pink module KME	Yellow module genes	Yellow module KME	Turquoise module genes	Turquoise module KME	Blue module genes	Blue module KME	Green module genes	Green module KME
<i>KCNKG1</i>	0.9646	<i>HLA-DQA2</i>	0.9731	<i>GOT1</i>	0.9627	<i>TRIM9</i>	0.9597	<i>SYT9</i>	0.9455
<i>ZSCAN23</i>	0.9622	<i>HLA</i>	0.9649	<i>MTCL1</i>	0.9526	<i>KNSTRN</i>	0.9476	<i>MKS1</i>	0.9430
<i>ZFYVE28</i>	0.9612	<i>HLA-DQB</i>	0.9626	<i>TP53INP2</i>	0.9518	<i>HSD17B7</i>	0.9462	<i>FAM47E-STBD1</i>	0.9347
<i>MLANA</i>	0.9351	<i>HLA-DQA1</i>	0.9602	<i>EIF3L</i>	0.9445	<i>PGM3</i>	0.9407	<i>ENDOG</i>	0.9310
<i>MLPH</i>	0.9266	<i>CTSH</i>	0.9587	<i>TDRD10</i>	0.9443	<i>PACRGL</i>	0.9395	<i>PNKD</i>	0.9292
<i>CACNB3</i>	0.9236	<i>DIAPH1</i>	0.9584	<i>MRPS7</i>	0.9437	<i>NHEJ1</i>	0.9349	<i>MCU</i>	0.9291
<i>ZNF420</i>	0.9194	<i>SLC28A1</i>	0.9577	<i>VPS9D1</i>	0.9417	<i>CDK5RAP2</i>	0.9248	<i>CD209-2</i>	0.9250
<i>PROM1</i>	0.9163	<i>DPYD</i>	0.9517	<i>TRNP1</i>	0.9404	<i>ZC RBI</i>	0.9220	<i>C19ORF67</i>	0.9247
<i>ENSMODG22842</i>	0.9155	<i>STAT5B</i>	0.9385	<i>TMEM127</i>	0.9341	<i>KIAA0586</i>	0.9151	<i>CD209</i>	0.9188
<i>TRPC3</i>	0.9120	<i>TRAF1</i>	0.9373	<i>PRKAB2</i>	0.9333	<i>NEK2</i>	0.9116	<i>NUDT9</i>	0.9176
<i>EDNRB</i>	0.9083	<i>MMR25</i>	0.9348	<i>TREH</i>	0.9328	<i>MES2</i>	0.9072	<i>H3FC</i>	0.9173
<i>TLE1</i>	0.9075	<i>CCDC69</i>	0.9300	<i>KCEN1</i>	0.9327	<i>THY1</i>	0.9070	<i>ADCK3</i>	0.9168
<i>SLC5A2</i>	0.9049	<i>EIF3J</i>	0.9290	<i>PHOSPHO1</i>	0.9300	<i>TYC23</i>	0.9033	<i>SHE</i>	0.9153
<i>ZNF184</i>	0.9047	<i>C2CD4A</i>	0.9236	<i>TCF15</i>	0.9287	<i>GALK2</i>	0.9028	<i>IDH3B</i>	0.9121
<i>PAX3</i>	0.9016	<i>RAP1GAP2</i>	0.9230	<i>C10L2</i>	0.9270	<i>ATL1</i>	0.9019	<i>CIR</i>	0.9118
<i>C14ORF159</i>	0.8955	<i>PRKCB</i>	0.9223	<i>CITTD4</i>	0.9264	<i>UBE2C</i>	0.9007	<i>CD59</i>	0.9106
<i>TBC1D2B</i>	0.8944	<i>STAT5A</i>	0.9127	<i>ANKRD13B</i>	0.9233	<i>PBR</i>	0.8994	<i>CNTN1</i>	0.9084
<i>ERCC3</i>	0.8899	<i>UPFI</i>	0.9115	<i>SLX1A</i>	0.9226	<i>ENSMODG29436</i>	0.8964	<i>PABPC1</i>	0.9065
<i>TYRPI</i>	0.8849	<i>CYSTM1</i>	0.9112	<i>ENSMODG28635</i>	0.9191	<i>RADIL</i>	0.8952	<i>CD163</i>	0.9063
<i>FAM103A1</i>	0.8848	<i>ADRBK2</i>	0.9091	<i>MVP</i>	0.9189	<i>MFSID8</i>	0.8915	<i>ABCC8</i>	0.9055
<i>DCT</i>	0.8831	<i>MED12L</i>	0.9087	<i>GAPDH</i>	0.9162	<i>ATRN</i>	0.8907	<i>C1QC</i>	0.9026
<i>CHN2</i>	0.8804	<i>ELF4</i>	0.9070	<i>NDEFA4</i>	0.9147	<i>CPIIE3</i>	0.8903	<i>C1QB</i>	0.9003
<i>KCNJ13</i>	0.8788	<i>B3GNT7</i>	0.9069	<i>NCAM1</i>	0.9141	<i>KIAA1107</i>	0.8890	<i>SSUH2</i>	0.8977
<i>TCF7L2</i>	0.8775	<i>CASS4</i>	0.9056	<i>DUSP14</i>	0.9136	<i>SSX2IP</i>	0.8887	<i>AKR1C1</i>	0.8974
<i>TYR</i>	0.8769	<i>CE7P</i>	0.9053	<i>SPATAS</i>	0.9135	<i>TOP2B</i>	0.8881	<i>BICD1</i>	0.8972

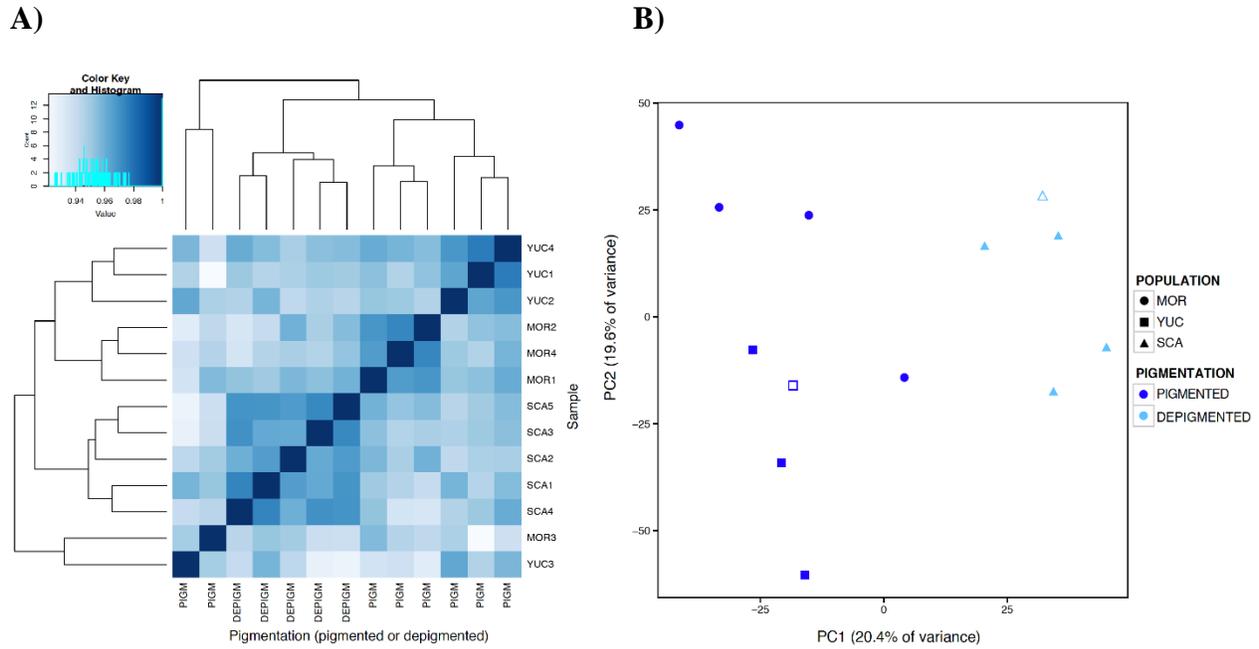


Figure 2.1. Clustering of the 13 individuals based on their gene expression profiles from 14,289 genes with at least three mapped reads in all individuals. A) Heatmap plot based on pairwise Pearson correlation of the expression data. B) Two-dimensional PCA plot of the gene expression profile of all the samples. Different colors identify different pigmentations, each shape indicates a particular population, while filling of the shape shows the sex of the individual, filled for males and not-filled for females.

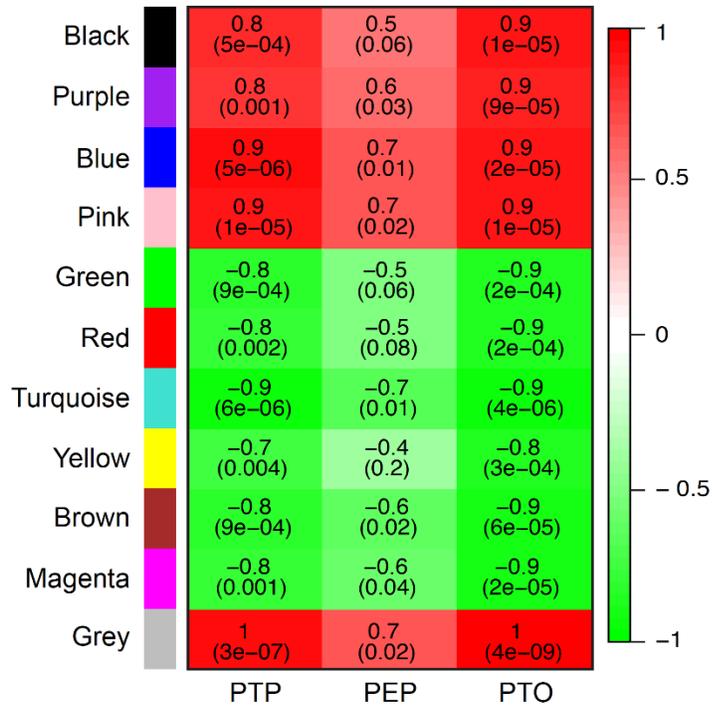


Figure 2.2. Correlation between skin gene modules and the pigmentation traits. Each row denotes a module, identified by its color, and each column a trait (PTP= proportion of tail pigmentation, PEP= proportion of ear pigmentation, PTO= proportion of toe pigmentation). Each cell shows the Pearson correlation coefficient (at the top) between the module eigengene and the skin pigmentation traits with the correlation p-value in parenthesis. The color of the cells corresponds to correlation values based on the color scale. Positive correlations are represented in red and negative correlations in green.

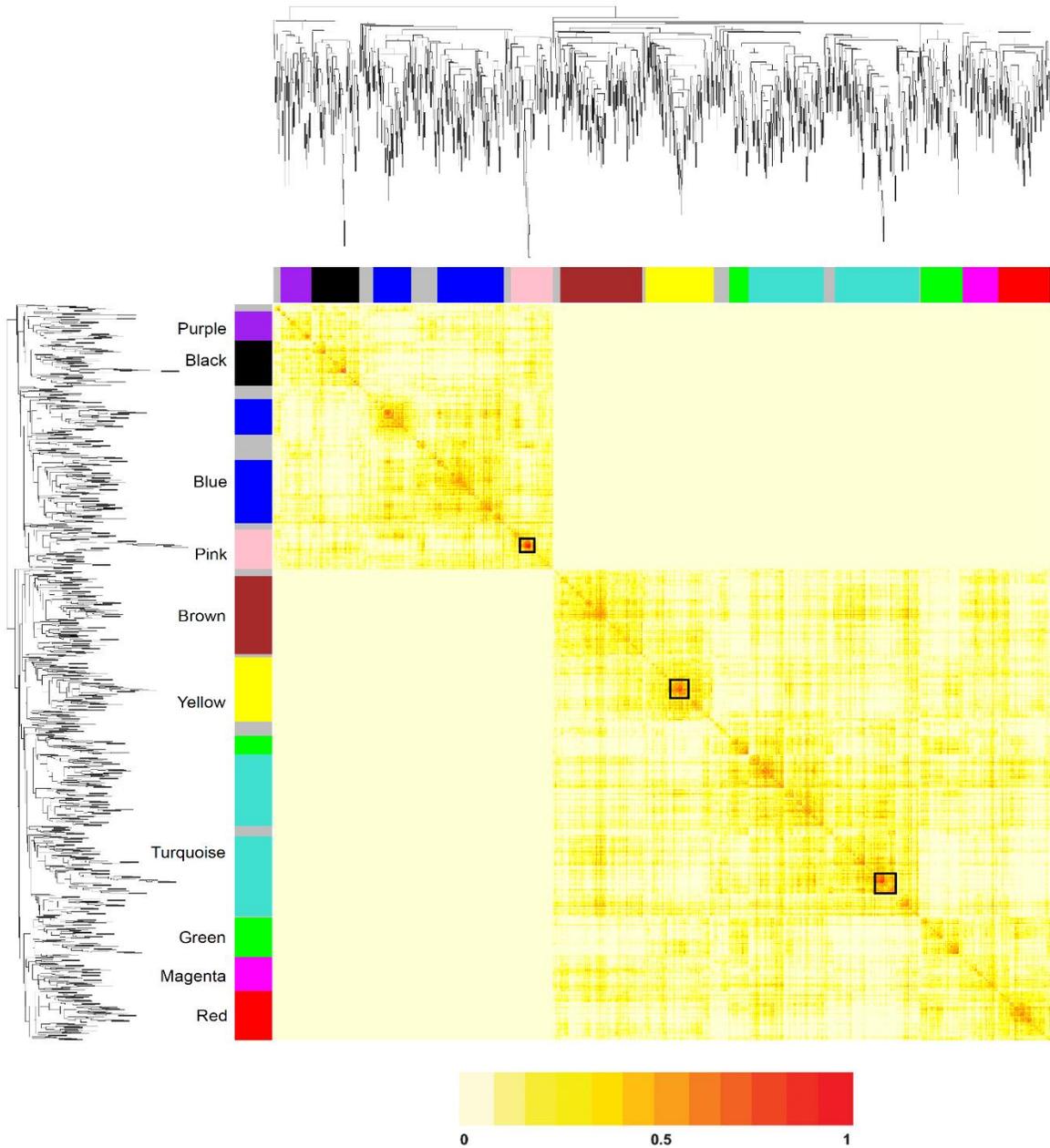


Figure 2.3. Gene coexpression network of the 1,186 genes differentially expressed between pigmented and depigmented opossums and module detection. Pearson correlations between expression profiles of all pairs of genes were transformed into network connection strengths (denoted by intensity of red color) by using a power function (see *Methods*). Average linkage hierarchical clustering with the topological overlap dissimilarity measure was used to identify gene coexpression modules. The modules correspond to branches of the tree, and each module was assigned a unique color. Grey denotes genes outside of proper modules. Rows and columns are symmetric and represent genes. The modules with more intramodular hub genes (those with higher correlation with other genes within their module) were the pink, yellow and turquoise modules; these groups of hub genes are highlighted in black squares.

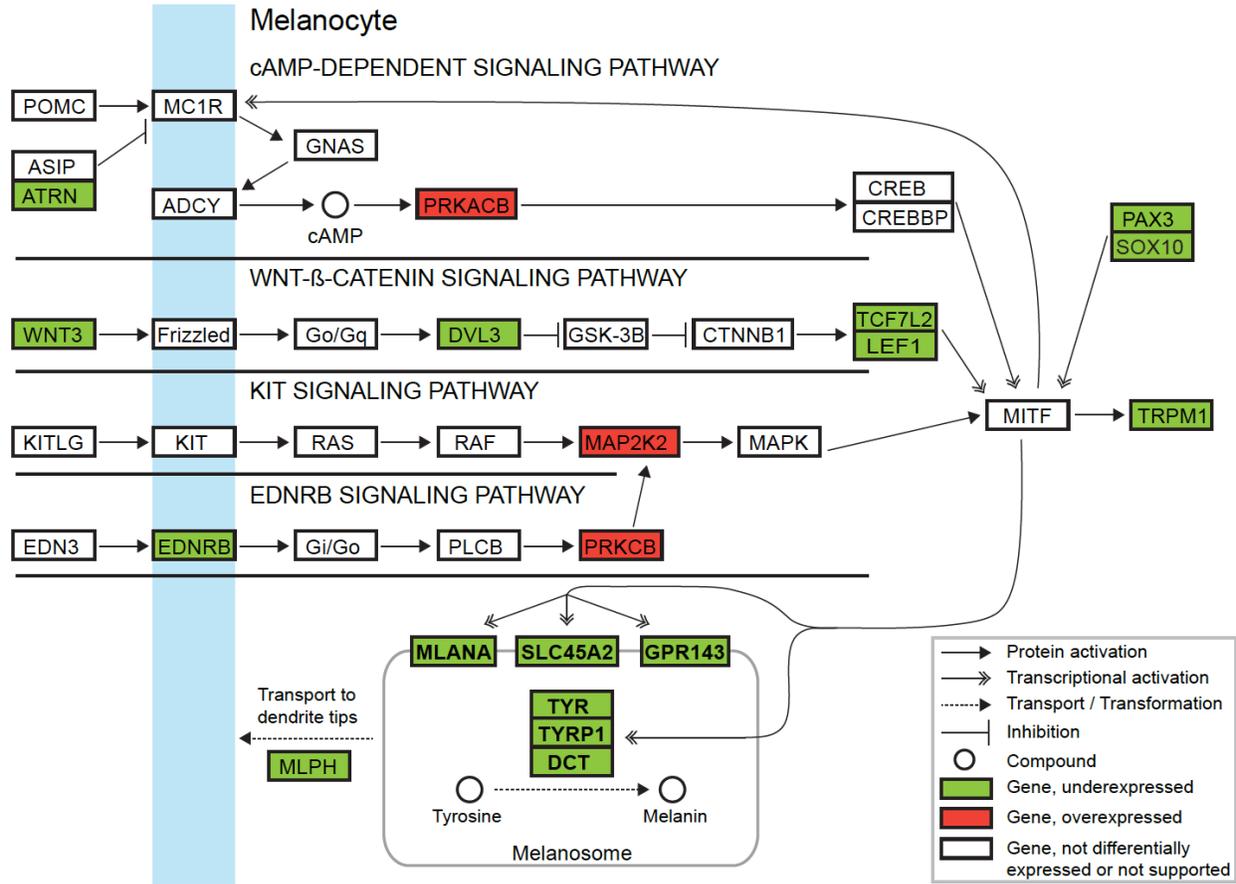


Figure 2.4. Schematic overview of the signaling pathways involved in melanocyte development and melanogenesis. The light blue vertical line represents the cellular membrane of a melanocyte. Genes that were found down-regulated in depigmented opossum skin the ear shown in green, genes up-regulated shown in red, and genes that were not differentially expressed in white. Genes participating in development and differentiation of melanocytes are shown in bold letters. Pathways relationships are shown for protein activation (black single arrows), transcriptional activation (black double arrow), transport or transformation (black arrow with dotted line) and inhibition (black line with blunt end). The Wnt/ β -catenin pathway is the most likely to be affecting the depigmentation of the skin in opossums since the down-regulation of several genes within this pathway, including WNT3, DVL3, TCF7L2 and LEF1, and of the downstream melanocytic genes TRPM1, MLANA, SLC45A2, GPR143, TYR, TYRP1 and DCT, which are regulated by MITF, is consistent with the depigmented phenotype. All the other pathways showed up-regulation of some genes, which is inconsistent with depigmentation. [Modified from the melanogenesis pathway map of Kyoto Encyclopedia of Genes and Genomes (KEEG) based on information from Hoekstra 2006; Hou and Pavan 2008; Yamaguchi and Hearing 2009; Lu *et al.* 2010; Schiaffino 2010; Poelstra *et al.* 2014; Poelstra *et al.* 2015].

a)



b)



Figure 2.5. Depigmented skin phenotypes in humans and Virginia opossums. a) Depigmentation in human hands due to vitiligo. b) Depigmentation observed in Virginia opossum individuals from northern populations on the superior and inferior aspect of their forefeet, hindfeet and the tip of the ear. (Photographies from opossums taken by Aislinn Sarnack and from creative commons. Photography from human hands taken from creative commons)

Supplementary Material

Table 2.S1. Analysis of variance for the skin pigmentation traits between the three populations sampled. The p-values for each population pairwise comparison are shown. The F-value for each trait is within parenthesis. MOR: Morelos, YUC: Yucatan, SCA: South Carolina.

Population comparisons	ANOVA		
	Tail pigmentation (F=45.7)	Ear pigmentation (F=3.51)	Toe pigmentation (F≥ 2.2e+31)
MOR vs YUC	0.8	1.0	0.47
MOR vs SCA	≤ 0.0001	0.11	≤ 0.0001
YUC vs SCA	≤ 0.0001	0.11	≤ 0.0001

Table 2.S2. Distances in Kilometers separating the three populations analyzed. MOR: Morelos, YUC: Yucatan, SCA: South Carolina

Population comparisons	Distance
MOR - YUC	1,190 Km
MOR - SCA	2,600 Km
YUC - SCA	3,550 Km

References

- Allen JA. 1901. A preliminary study of the North American opossums of the genus *Didelphis*. *Bulletin of the American Museum of Natural History* 14:149-195.
- Alvarez M, Schrey AW, Richards CL. 2015. Ten years of transcriptomics in wild populations: what have we learned about their ecology and evolution? *Molecular Ecology* 24(4):710-725.
- Anders S, Pyl PT, Huber W. 2014. HTSeq- a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166-169.
- Anderson TM, vonHoldt BM, Candille SI, Musiani M, Greco C, Stahler DR, Smith DW, Padhukasahasran B, Randi E, Leonard JA, Bustamante CD, Ostrander EA, Tang H, Wayne RK, Barsh GS. 2009. Molecular and evolutionary history of melanism in North American gray wolves. *Science* 323:1339-1343.
- Andrews S. 2010. FastQC. Available from <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Awad EM, Khan SY, Sokolikova B, Brunner PM, Olcaydu D, Wojta, J, Breuss JM, Uhrin P. 2013. Cold induces reactive oxygen species production and activation of the NF-kappa B response in endothelial cells and inflammation in vivo. *Journal of Thrombosis and Haemostasis* 11:1716-1726.
- Baldwin AS Jr. 1996. The NF-kappa B and I kappa B proteins: new discoveries and insights. *Annual Review of Immunology* 14:649-83.
- Bartlein PJ, Anderson KH, Anderson PM, Edwards ME, Mock CJ, Thompson RS, Webb RS, Webb III T, Whitlock C. 1998. Paleoclimate simulations for North America over the past 21,000 years: features of the simulated climate and comparisons with paleoenvironmental data. *Quaternary Science Reviews* 17:549-585.
- Baxter LL, Loftus SK, Pavan WJ. 2009. Networks and pathways in pigmentation, health, and disease. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 1:359-371.
- Beermann F, Orlov SJ, Lamoreux ML. 2004. The Tyr (albino) locus of the laboratory mouse. *Mammalian Genome* 15:749-758.
- Bellone R, Brooks S, Sandmeyer L, Murphy BA, Forsyth G, Archer S, Bailey E, Grahn B. 2008. Differential gene expression of TRPM1, the potential cause of congenital stationary night blindness and coat spotting patterns (LP) in the Appaloosa horse (*Equus caballus*). *Genetics* 179:1861-1870.
- Bennett DC, Lamoreux ML. 2003. The color loci of mice – a genetic century. *Pigment Cell*

Research 16:333-344.

- Bernatchez L, Landry C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology* 16:363-377.
- Blehert DS, Hicks AC, Behr M, Meteyer CU, Berlowski-Zier BM, Buckles EL, Coleman JT, Darling SR, Gargas A, Niver R, Okoniewski JC, Rudd RJ, Stone WB. 2009. Bat white-nose syndrome: an emerging fungal pathogen? *Science* 323:227.
- Boekhout M, Wolthuis R. 2015. Nek2A destruction marks APC/C activation at the prophase-to-prometaphase transition by spindle-checkpoint-restricted Cdc20. *Journal of Cell Science* 128:1639-1653.
- Boettger MK, Till S, Chen MX, Anand U, Otto WR, Plumpton C, Trezise DJ, Tate SN, Bountra C, Coward K, Birch R, Anand P. 2002. Calcium-activated potassium channel SK1 and IK1-like immunoreactivity in injured human sensory neurones and its regulation by neurotrophic factors. *Brain* 125:252-263.
- Bonizzi G, Karin M. 2004. The two NF- κ B activation pathways and their role in innate and adaptive immunity. *TRENDS in Immunology* 25:280-288.
- Bros M, Dexheimer N, Ross R, Trojandt S, Höhn Y, Tampe J, Sutter A, Jährling F, Grabbe S, Reske-Kunz AB. 2011. Differential gene expression analysis identifies murine *Cacnb3* as strongly upregulated in distinct dendritic cell populations upon stimulation. *Gene* 472:18-27.
- Burt EH. 1981. The adaptiveness of animal colors. *Bioscience* 31:723-729.
- Burt EH, Ichida JM. 2004. Gloger's rule, feather-degrading bacteria, and color variation among song sparrows. *The Condor* 106:681-686.
- Caro T. 2005. The adaptive significance of coloration in mammals. *BioScience* 55:125-136.
- Chen GY, Nuñez G. 2010. Sterile inflammation: sensing and reacting to damage. *Nature Reviews Immunology* 10:826-837.
- Chung JY, Park YC, Ye H, Wu H. 2002. All TRAFs are not created equal: Common and distinct molecular mechanisms of TRAF-mediated signal transduction. *Journal of Cell Science* 115:679-688.
- Conus S, Simon HU. 2008. Cathepsins: Key modulators of cell death and inflammatory responses. *Biochemical Pharmacology* 76:1374-1382.
- Coulson T, MacNulty DR, Stahler DR, vonHoldt B, Wayne RK, Smith DW. 2011. Modeling effects of environmental change on wolf population dynamics, trait evolution, and life history. *Science* 334:1275-1278.

- Cox N, Pilling D, Gomer RH. 2015. DC-SIGN activation mediates the differential effects of SAP and CRP on the innate immune system and inhibits fibrosis in mice. *Proceedings of the National Academy of Sciences U.S.A* 112:8385-8390.
- D'Angelo ME, Bird PI, Peters C, Reinheckel T, Trapani JA, Sutton VR. 2010. Cathepsin H is an additional convertase of pro-granzyme B. *Journal of Biological Chemistry* 285:20514–20519.
- Dawson TJ, Maloney SK. 2004. Fur versus feathers: the different roles of red kangaroo fur and emu feathers in thermoregulation in the Australian arid zone. *Australian Mammalogy* 26:145-151.
- Dawson TJ, Webster KN, Maloney SK. 2014. The fur of mammals in exposed environments; do crypsis and thermal needs necessarily conflict? The polar bear and marsupial koala compared. *Journal of Comparative Physiology B* 184:273-284.
- Dessinioti C, Stratigos AJ, Rigopoulos D, Katsambas AD. 2009. A review of genetic disorders of hypopigmentation: lessons learned from the biology of melanocytes. *Experimental Dermatology* 18:741-749.
- Dey-Rao R, Sinha AA. 2016. Interactome analysis of gene expression profile reveals potential novel key transcriptional regulators of skin pathology in vitiligo. *Genes and Immunity* 17:30-45.
- Drake DR, Brogden KA, Dawson DV, Wertz PW. 2008. Thematic review series: skin lipids. Antimicrobial lipids at the skin surface. *Journal of Lipid Research* 49:4-11.
- Eklblom R, Galindo J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* 107:1-15.
- Elias PM. 2007. The skin barrier as an innate immune element. *Seminars in Immunopathology* 29:3-14.
- Elias PM, Williams ML. 2013. Re-appraisal of current theories for the development and loss of epidermal pigmentation in hominins and modern humans. *Journal of Human Evolution* 64:687- 692.
- Elias PM, Menon G, Wetzel BK, Williams JW. 2009. Evidence that stress to the epidermal barrier influenced the development of pigmentation in humans. *Pigment Cell and Melanoma Research* 22:420-434.
- Endler JA. 1990. On the measurement and classification of colour in studies of animal colour patterns. *Biological Journal of the Linnean Society* 41:315-352.
- Espinosa-Soto C, Martin OC, Wagner A. 2011. Phenotypic plasticity can facilitate adaptive

- evolution in gene regulatory circuits. *BMC Evolutionary Biology* 11:5.
- Ezzedine K, Eleftheriadou V, Whitton M, Geel N. 2015, Vitiligo. *The Lancet* 386:74–84.
- Fang L, Seki A, Fang G. 2009. SKAP associates with kinetochores and promotes the metaphase-to-anaphase transition. *Cell Cycle* 8:2819-2827.
- Filteau M, Pavey SA, St-Cyr J, Bernatchez L. 2013. Gene coexpression networks reveal key drivers of phenotypic divergence in lake whitefish. *Molecular Biology and Evolution* 30:1384-1396.
- Fraser HB. 2013. Gene expression drives local adaptation in humans. *Genome Research* 23:1089–1096.
- Gage AA. 1979. What temperature is lethal for cells? *Journal of Dermatologic Surgery and Oncology* 5:459-464.
- Galibert MD, Yavuzer U, Dexter TJ, Goding CR. 1999. Pax3 and regulation of the melanocyte-specific tyrosinase-related protein-1 promoter. *The Journal of Biological Chemistry* 274:26894-26900.
- Gao C, Chen Y. 2010. Dishevelled: The hub of Wnt signaling. *Cellular Signalling* 22:717-727.
- Gardner AL. 1973. The systematics of the genus *Didelphis* (Marsupialia: Didelphidae) in North and Middle America. *Museum, Texas Tech University* 4:1-81.
- Gardner AL. 1982. Virginia opossum, pp. 3-36. In: Chapman JA, Feldhamer GA (eds). *Wild mammals of North America: biology, management, and economics*. Johns Hopkins University Press, Baltimore, MD.
- Gardner AL, Sunquist ME. 2003. Opossum, *Didelphis virginiana*, pp. 3-29. In: Feldhamer GA, Thompson BC, Chapman JA (eds.). *Wild Mammals of North America*, 2nd ed. Johns Hopkins University Press, Baltimore, MD.
- Gasparini J, Bize P, Piau R, Wakamatsu K, Blount JD, Ducrest A-L, Roulin A. 2009. Strength and cost of an induced immune response are associated with a heritable melanin-based colour trait in female tawny owls. *Journal of Animal Ecology* 78:608-616.
- Gaston KV, Chown SL, Evans KL. 2008. Ecogeographical rules: elements of a synthesis. *Journal of Biogeography* 35:483-500.
- Ghalambor CK, McKay JK, Carroll SP, Reznick DN. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* 21:394-407.
- Ghalambor CK, Hoke KL, Ruell EW, Fischer EK, Reznick DN, Hughes KA. 2015. Non-

- adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature* 525:372-375.
- Ghazalpour A, Doss S, Zhang B, Wang S, Plaisier C, Castellanos R, Brozell A, Schadt EE, Drake TA, Lusis AJ, Horvath S. 2006. Integrating genetic and network analysis to characterize genes related to mouse weight. *PLoS Genetics* 2(8):e130.
- Gilad Y, Pritchard JK, Thornton K. 2009. Characterizing natural variation using next-generation sequencing technologies. *Trends in Genetics* 25:463-471.
- Gilbert SF. 2005. Mechanisms for the environmental regulation of gene expression: ecological aspects of animal development. *Journal of Biosciences* 30:65-74.
- Gloger CL. 1833. *Das Abändern der vögel durch einfluss des klimas*. August Schulz, Breslau, Poland.
- Gratten J, Beraldi D, Lowder BV, McRae AF, Visscher PM, Pemberton JM, Slate J. 2007. Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proceedings of the Royal Society B: Biological Sciences* 274:619–626.
- Graham RW, and Lundelius Jr EL. 2010. FAUNMAP II: New data for North America with a temporal extension for the Blancan, Irvingtonian and early Rancholabrean. FAUNMAP II Database, version 1.0.
- Graham RW, Lundelius Jr EL, Graham MA, Schroeder EK, Toomey III RS, Anderson E, Barnosky AD, Burns JA, Churcher CS, Grayson DK, Guthrie D, Harington CR, Jefferson JT, Martin LD, McDonald HG, Morlan RE, Semken Jr HA, Webb D, Werdelin L, Wilson M. 1996. Spatial response of mammals to late Quaternary environmental fluctuations. *Science* 272:1601-1606.
- Guernier V, Hochberg ME, Guégan JF. 2004. Ecology drives the worldwide distribution of human diseases. *PLoS Biology* 2:e141.
- Guilday JE. 1958. The prehistoric distribution of the opossum. *Journal of Mammalogy* 39:39-43.
- Han JD, Bertin N, Hao T, Goldberg DS, Berriz GF, Zhang LV, Dupuy D, Walhout AJ, Cusick ME, Roth FP, Vidal M. 2004. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* 430:88–93.
- Hansen KD, Irizarry RA. 2012. Removing technical variability in RNA-seq data using conditional quantile normalization. *Biostatistics* 13:204-216.
- Hedges SB, Dudley J, Kumar S. 2006. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* 22:2971-2972.

- Hedges SB, Marin J, Suleski M, Paymer M, Kumar S. 2015. Tree of life reveals clock-like speciation and diversification. *Molecular Biology and Evolution* 32:835–845.
- Hoekstra HE. 2006. Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity* 97:222-234
- Hoekstra HE, Hirschmann R, Bunday R, Insel PA, Crossland JP. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* 313:101-104.
- Hope RM, Godfrey GK. 1988. Genetically-determined variation of pelage color and reflectance in natural and laboratory populations of the marsupial *Sminthopsis-Crassicaudata* (Gould). *Australian Journal of Zoology* 36:441-456.
- Horvath S, Zhang B, Carlson M, Lu KV, Zhu S, Felciano RM, Laurence MF, Zhao W, Qi S, Chen Z, Lee Y, Scheck AC, Liao LM, Wu H, Geschwind DH, Febbo PG, Kornblum HI, Cloughesy TF, Nelson SF, Mischel PS. 2006. Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proceedings of the National Academy of Sciences U. S. A.* 103:17402-17407.
- Horvath S and Dong J. 2008. Geometric interpretation of gene co-expression network analysis. *PLoS Computational Biology* 4(8):e1000117.
- Hou L, Pavan WJ. 2008. Transcriptional and signaling regulation in neural crest stem cell-derived melanocyte development: do all roads lead to Mitf?. *Cell Research* 18:1163-1176.
- Hubbard JK, UY JA, Hauber ME, Hoekstra HE, Safran RJ. 2010. Vertebrate pigmentation: from underlying genes to adaptive function. *Trends in Genetics* 26:231-239.
- Jablonski NG, Chaplin G. 2010. Human skin pigmentation as an adaptation to UV radiation. *Proceedings of the National Academy of Sciences U. S. A.* 107:8962-8968.
- Jacobs GH, Williams GA. 2010. Cone pigments in a North American marsupial, the opossum (*Didelphis virginiana*). *Journal of Comparative Physiology A* 196:379-384.
- Jimenez MD, Bangs EE, Sime C, Asher VJ. 2010. Sarcoptic mange found in wolves in the Rocky Mountains in western United States. *Journal of wildlife diseases* 46:1120-1125.
- Jin Y, Birlea SA, Fain PR, Gowan K, Riccardi SL, Holland PJ, Mailloux CM, Sufit AJ, Hutton SM, Amadi-Myers A, Bennett DC. 2010. Variant of TYR and autoimmunity susceptibility loci in generalized vitiligo. *New England Journal of Medicine* 362:1686-1697.
- Jin Y, Birlea SA, Fain PR, Gowan K, Riccardi SL, Holland PJ, Bennett DC, Herbstman DM, Wallace MR, McCormack WT, Kemp EH, Gawkrödger DJ, Anthony PW, Picardo M, Leone G, Taïeb A, Jouary T, Ezzedine K, Geel N, Lambert J, Overbeck A, Spritz RA.

2012. Genome-wide analysis identifies a quantitative trait locus in the MHC class II region associated with generalized vitiligo age of onset. *Journal of Investigative Dermatology* 131:1308-1312.
- Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* 14:R36.
- Kingo K, Aunin E, Karelson M, Rätsep R, Silm H, Vasar E, Kõks S. 2008. Expressional changes in the intracellular melanogenesis pathways and their possible role the pathogenesis of vitiligo. *Journal of Dermatological Science* 52:39-46.
- Kogenaru S, del Val C, Hotz-Wagenblatt A, Glatting KH. 2010. TissueDistributionDBs: a repository of organism-specific tissue-distribution profiles. *Theoretical Chemistry Accounts* 123:651-658.
- Lafferty KD. 2009. The ecology of climate change and infectious diseases. *Ecology* 90:888-900.
- Lai Y, Shiroishi T, Moriwaki K, Motokawa M, Yu H. 2008. Variation of coat color in house mice throughout Asia. *Journal of Zoology* 274:270-276.
- Lande R. 2009. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of Evolutionary Biology* 22:1435–1446.
- Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559.
- Langfelder P, Zhang B, Horvath S. 2008. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* 24:719-720.
- Langfelder P, Mischel PS, Horvath S. 2013. When is hub gene selection better than standard meta-analysis. *PLoS ONE* 8(4): e61505.
- Larison B, Harrigan RJ, Thomassen HA, Rubenstein DI, Chan-Golston AM, Li E, Smith TB. 2015. How the zebra got its stripes: a problem with too many solutions. *Royal Society open science* 2:140452.
- Law CW, Chen Y, Shi W, Smyth GK. 2014. Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology* 15:R29.
- Linnen CR, Kingsley EP, Jensen JD, Hoekstra HE. 2009. On the Origin and Spread of an Adaptive Allele in Deer Mice. *Science* 325:1095-1098.
- Linnen C, Poh Y, Peterson BK, Barrett RDH, Larson JG, Jensen JD, Hoekstra HE. 2013. Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* 339:1312-1316.

- Liu X, Yu X, Zack DJ, Zhu H, Qian J. 2008. TiGER: a database for tissue-specific gene expression and regulation. *BMC Bioinformatics* 9:271.
- Logan CY, Nuse R. 2004. The Wnt signaling pathway in development and disease. *Annual Review of Cell and Developmental Biology* 20:781-810.
- Lorch JM, Meteyer CU, Behr MJ, Boyles JG, Cryan PM, Hicks AC, Ballmann AE, Coleman JT, Redell DN, Reeder DM, Blehert DS. 2011. Experimental infection of bats with *Geomyces destructans* causes white-nose syndrome. *Nature* 480:376-378.
- Lu Z, Kim KA, Suico MA, Shuto T, Li JD, Kai H. 2004. MEF up-regulates human beta-defensin 2 expression in epithelial cells. *FEBS Letters*. 561:117–121.
- Lu S, Slominski A, Yang S, Sheehan C, Ross J, Carlson JA. 2010. The correlation of TRPM1 (Melastatin) mRNA expression with microphthalmia-associated transcription factor (MITF) and other melanogenesis-related proteins in normal and pathological skin, hair follicles and melanocytic nevi. *Journal of Cutaneous Pathology* 37:26-40.
- Mackintosh JA. 2001. The antimicrobial properties of melanocytes, melanosomes and melanin and the evolution of black skin. *Journal of Theoretical Biology* 211:101-113.
- Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE. 2010. Convergence in pigmentation at multiple levels: mutations, genes and function. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365:2439-2450.
- Männiste M, Hõrak P. 2014. Emerging infectious disease selects for darker plumage coloration in greenfinches. *Frontiers in Ecology and Evolution* 2:4.
- McManus JJ. 1974. *Didelphis virginiana*. *Mammalian Species* 40:1-6.
- McNab BK. 1978. The comparative energetics of Neotropical marsupials. *Journal of Comparative Physiology* 125:115-128.
- Mikkelsen TS, Wakefield MJ, Aken B, Amemiya CT, Chang JL, Duke S, Garber M, Gentles AJ, Goodstadt L, Heger A, Jurka J, *et al.* 2007. Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature* 447:167-177.
- Miller JA, Oldham MC, Geschwind DH. 2008. A systems level analysis of transcriptional changes in Alzheimer's disease and normal aging. *The Journal of Neuroscience* 28:1410-1420.
- Miller W, Schuster SC, Welch AJ, Ratan A, Bedoya-Reina OC, Zhao F, Kim HL, Burhans RC, Drautz DI, Wittekindt NE, Tomsho LP, Ibarra-Laclette E, Herrera-Estrella L, Peacock E, Farley S, Sage GK, Rode K, Obbard M, Montiel R, Bachmann L, Ingolfsson O, Aars J, Mailund T, Wiig Ø, Talbot SL, Lindqvist C. 2012. Polar and brown bear genomes reveal

- ancient admixture and demographic footprints of past climate change. *Proceedings of the National Academy of Sciences U. S. A.* 109:E2382-E2390.
- Millien V, Lyons SK, Olson K, Smith FA, Wilson AB, Yom-Tov Y. 2006. Ecotypic variation in the context of global climate change: revisiting the rules. *Ecology Letters* 9:853-869.
- Moestrup S, Møller H. 2004. CD163: a regulated hemoglobin scavenger receptor with a role in the anti-inflammatory response. *Annals of Medicine* 36:347-354.
- Montefiori DC, Zhou JY. 1991. Selective antiviral activity of synthetic soluble L- tyrosine and L-dopa melanins against human immunodeficiency virus in vitro. *Antiviral Research* 15:11-25.
- Moreno J, Møller AP. 2006. Are melanin ornaments signals of antioxidant and immune capacity in birds? *Acta Zoologica Sinica* 52:202-208.
- Morgan GS. 2008. Vertebrate fauna and geochronology of the great american biotic interchange in North America, 93-140 pp. In: Lucas SG, Morgan GS, Spielmann JA, Prothero DR (eds.). *Neogene Mammals*. New Mexico Museum of Natural History and Science Bulletin 44.
- Mullen LM, Hoekstra HE. 2008. Natural selection along an environmental gradient: a classic cline in mouse pigmentation. *Evolution* 62:1555-1570.
- Nesargikar PN, Spiller B, Chavez R. 2012. The complement system: history, pathways, cascade and inhibitors. *European Journal of Microbiology and Immunology* 2:103–111.
- Nielsen R, Hubisz MJ, Hellmann I, Torgerson D, Andres AM, Albrechtsen A, Gutenkunst R, Adams MD, Cargill M, Boyko A, Indap A, Bustamante CD, Clark AG. 2009. Darwinian and demographic forces affecting human protein coding genes. *Genome Research* 19:838-849.
- Norton HL, Kittles RA, Parra E, McKeigue P, Mao X, Cheng K, Canfield VA, Bradley DG, McEvoy B, Shriver MD. 2007. Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Molecular Biology and Evolution* 24:710-722.
- Oldham MC, Horvath S, Geschwind DH. 2006. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proceedings of the National Academy of Sciences U. S. A.* 103:17973-17978.
- Oleksiak MF, Churchill GA, Crawford DL. 2002. Variation in gene expression within and among natural populations. *Nature Genetics* 32:261-266.
- Onofre G, Kolackova M, Jankovicova K, Krejsek, J. 2009. Scavenger receptor CD163 and its biological functions. *Acta Medica (Hradec Kralove)* 52:57-61.

- Page EH, Shear NH. 1988. Temperature-dependent skin disorders. *Journal of the American Academy of Dermatology* 18:1003-1019.
- Papenfuss AT, Feng ZP, Krasnec K, Deakin JE, Baker ML, Miller RD. 2015. Marsupials and monotremes possess a novel family of MHC class I genes that is lost from eutherian lineage. *BMC Genomics* 16:535.
- Pardo-Diaz C, Salazar C, Jiggins CD. 2015. Towards the identification of the loci of adaptive evolution. *Methods in Ecology and Evolution* 6:445-464.
- Pattnaik BR, Tokarz S, Asuma MP, Schroeder T, Sharma A, Mitchell JC, Edwards AO, Pillers DA. 2013. Snowflake vitreoretinal degeneration (SVD) mutation R162W provides new insights into Kir7.1 ion channel structure and function. *PLoS ONE* 8(8):e71744.
- Piertney SB, Oliver MK. 2006. The evolutionary ecology of the major histocompatibility complex. *Heredity* 96:7-21.
- Pingault V, Ente D, Moal DL, Goossens M, Marlin S, Bondurand N. 2010. Review and update of mutations causing Waardenburg syndrome. *Human mutation* 31:391-406.
- Pitceathly RD, Rahman S, Wedatilake Y, Polke JM, Cirak S, Foley AR, Sailer A, Hurles ME, Stalker J, Hargreaves I, Woodward CE, Sweeney MG, Muntoni F, Houlden H, UK10K Consortium, Taanman JW, Hanna MG. 2013. NDUFA4 mutations underlie dysfunction of a cytochrome c oxidase subunit linked to human neurological disease. *Cell Reports* 3:1795-1805.
- Plaisier CL, Horvath S, Huertas-Vazquez A, Cruz-Bautista I, Herrera MF, Tusie-Luna T, Aguilar-Salinas C, Pajukanta P. 2009. A systems genetics approach implicates USF1, FADS3, and other causal candidate genes for familial combined hyperlipidemia. *PLoS Genetics* 5:e1000642.
- Plonka PM, Passeron T, Brenner M, Tobin DJ, Shibahara S, Thomas A, Slominski A, Kadekaro AL, HersHKovitz D, Peters E, Nordlund JJ, Abdel- Malek Z, Takeda K, Paus R, Ortonne JP, Hearing VJ, Schallreuter KU. 2009. What are melanocytes really doing all day long...? *Experimental Dermatology* 18:799– 819.
- Poelstra J, Vijay N, Bossu C, Lantz H, Ryll B, Müller I, Baglione V, Unneberg P, Wikelski M, Grabherr MG, Wolf JBW. 2014. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* 344:1410-1414.
- Poelstra J, Vijay N, Hoepfner M, Wolf JBW. 2015. Transcriptomics of colour patterning and coloration shifts in crows. *Molecular Ecology* 24:4617-4628.
- Price TD, Sol D. 2008. Introduction: Genetics of colonizing species. *American Naturalist*. 172: S1–S3.

- Puiffe M-L, Lachaise I, Molinier-Frenkel V, Castellano F. 2013. Antibacterial Properties of the Mammalian L-Amino Acid Oxidase IL4I1. *PLoS ONE* 8:e54589.
- Quan C, Ren YQ, Xiang LH, Sun LD, Xu AE, Gao XH, Chen HD, Pu XM, Wu RN, Liang CZ, Li JB. 2010. Genome-wide association study for vitiligo identifies susceptibility loci at 6q27 and the MHC. *Nature Genetics* 42:614-618.
- R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rape M, Kirschner MW. 2004. Autonomous regulation of the anaphase-promoting complex couples mitosis to S-phase entry. *Nature* 432:588-595.
- Reimand J, Arak T, Vilo J. 2011. g:Profiler - a web server for functional interpretation of gene lists (2011 update). *Nucleic Acids Research* 39:307-315.
- Regazzetti FJ, Marty C, Rivier M, Mehul B, Reiniche P, Mounier C, Rival Y, Piwnica D, Cavalié M, Chignon-Sicard B, Ballotti R, Voegel J, Passeron T. 2015. Transcriptional analysis of vitiligo skin reveals the alteration of WNT pathway: a promising target for repigmenting vitiligo patients. *Journal of Investigative Dermatology* 135:3105-3114.
- Rezaei N, Gavalas NG, Weetman AP, Kemp EH. 2007. Autoimmunity as an aetiological factor in vitiligo. *Journal of the European Academy of Dermatology and Venereology* 21:865-876.
- Richmond JM, Frisoli ML and Harris JE. 2013. Innate immune mechanisms in vitiligo: danger from within. *Current Opinion in Immunology* 25:676-682.
- Riley SP, Bromley C, Poppenga RH, Uzal FA, Whited L, Sauvajot RM. Anticoagulant exposure and notoedric mange in bobcats and mountain lions in urban southern California. 2007. *The Journal of Wildlife Management* 71:1874-1884.
- Rizkallah R, Batsomboon P, Dudley GB, Hurt MM. 2015. Identification of the oncogenic kinase TOPK/PBK as a master regulator of C2H2 zinc finger proteins. *Oncotarget* 6:1446-1461.
- Saito H, Yasumoto KI, Takeda K, Takahashi K, Yamamoto H, Shibahara S. 2003. Microphthalmia-associated transcription factor in the Wnt signaling pathway. *Pigment Cell Research* 16:261-265.
- Sakaguchi A, Kikuchi A. 2004. Functional compatibility between isoform α and β of type II DNA topoisomerase. *Journal of Cell Science* 117:1047-1054.
- Santana SE, Lynch-Alfaro J, Alfaro M. 2012. Adaptive evolution of facial colour patterns in Neotropical primates. *Proceedings of the Royal Society of London B: Biological Sciences* 279: 2204–2211.

- Sarma JV, Ward PA. 2011. The complement system. *Cell and Tissue Research* 343:227-235.
- Schiaffino MV. 2010. Signaling pathways in melanosome biogenesis and pathology. *The International Journal of Biochemistry and Cell Biology* 42:1094-1104.
- Schweizer RM, Vonholdt BM, Harrigan R, Knowles JC, Musiani M, Coltman D, Novembre J, Wayne RK. 2016a. Genetic subdivision and candidate genes under selection in North American grey wolves. *Molecular Ecology* 25:380-402.
- Schweizer RM, Robinson J, Harrigan R, Silva P, Galaverni M, Musiani M, Green RE, Novembre J, Wayne RK. 2016b. Targeted capture and resequencing of 1040 genes reveal environmentally driven functional variation in grey wolves. *Molecular Ecology* 25:357-379.
- Shiryaev SA, Remacle AG, Savinov AY, Chernov AV, Cieplak P, Radichev IA, Williams R, Shiryaeva TN, Gawlik K, Postnova TI, Ratnikov BI, Eroshkin AM, Motamedchaboki K, Smith JW, Strongin AY. 2009. Inflammatory proprotein convertase-matrix metalloproteinase proteolytic pathway in antigen-presenting cells as a step to autoimmune multiple sclerosis. *The Journal of Biological Chemistry* 284:30615-30626.
- Slominski A. 2005. Neuroendocrine system of the skin. *Dermatology* 211:199-208.
- Slominski AT, Tobin DJ, Shibahara S, Wortsman J. 2004. Melanin pigmentation in mammalian skin and its hormonal regulation. *Physiological Reviews* 84:1155-1228.
- Slominski AT, Zmijewski MA, Skobowiat C, Zbytek B, Slominski M, Steketee JD. 2012. Sensing the environment: regulation of local and global homeostasis by the skin's neuroendocrine system. *Advance in Anatomy Embriology and Cell Biology* 212:1-115.
- Smyth GK. 2005. limma: Linear models for microarray data, pp. 397-420. In: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W (eds.). *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Statistics for Biology and Health. Springer, New York.
- Spritz RA. 2012 Six decades of vitiligo genetics: genome-wide studies provide insights into autoimmune pathogenesis. *Journal of Investigative Dermatology* 132:268-273.
- Spritz RA. 2013. Modern vitiligo genetics sheds new light on an ancient disease. *The Journal of Dermatology* 40:310-314.
- Steinberg GR, Kemp BE. 2009. AMPK in Health and Disease. *Physiological Reviews* 89: 1025-1078.
- Stoehr AM. 2006. Costly melanin ornaments: the importance of taxon? *Functional Ecology* 20:276-281.

- Stoner C, Caro T, Graham C. 2003. Ecological and behavioral correlates of coloration in artiodactyls: systematic analyses of conventional hypotheses. *Behavioral Ecology* 14:823-840.
- Storey JD, Tibshirani R. 2003. Statistical significance for genome-wide experiments. *Proceedings of the National Academy of Sciences U. S. A.* 100:9440-9445.
- Su K, Kyaw H, Fan P, Zeng Z, Shell BK, Carter KC, Li Y. 1998. Isolation, characterization, and mapping of two human potassium channels. *Biochemical and Biophysical Research Communications* 241: 675–681.
- Subramanya RD, Coda AB, Sinha AA. 2010. Transcriptional profiling in alopecia areata defines immune and cell cycle control related genes within disease-specific signatures. *Genomics* 96:146-153.
- Sulk M, Steinhoff M. 2015. Role of TRP channels in skin diseases, pp. 293-323. In: Arpad Szallasi (ed.). *TRP Channels as Therapeutic Targets, From Basic Science to Clinical Use.* Academic Press, San Diego, CA.
- The GTEx consortium. 2013. The Genotype-tissue expression (GTEx) project. *Nature Genetics* 45:580-585.
- Todd EV, Black MA, Gemmell NJ. 2016. The power and promise of RNA-Seq in ecology and evolution. *Molecular Ecology* 25:1224-1241.
- Tyndale-Biscoe H. 2005. *Life of marsupials.* CSIRO Publishing, Australia. 464 pp.
- Ungerer MC, Johnson LC, Herman MA. 2008. Ecological genomics: understanding gene and genome function in the natural environment. *Heredity* 100:178-183.
- Vignieri SN, Larson JG, Hoekstra HE. 2010. The selective advantage of crypsis in mice. *Evolution* 64:2153-2158.
- Wang F, Liu HM, Irwin MG, Xia ZY, Huang Z, Ouyang J, Xia Z. 2009. Role of protein kinase C β 2 activation in TNF- α -induced human vascular endothelial cell apoptosis. *Canadian Journal of Physiology and Pharmacology* 87:221-229.
- Wassermann HP. 1965. Human pigmentation and environmental adaptation. *Archives of Environmental Health: An International Journal* 11:691-694.
- Wei L, Arian L, O'Shea JJ. 2008. New insights into the roles of Stat5a/b and Stat3 in T cell development and differentiation. *Seminars in Cell and Developmental Biology* 19:394-400.
- Weiss JN, Karma A, MacLellan WR, Deng M, Rau CD, Rees CM, Wang J, Wisniewski N, Eskin E, Horvarth S, Qu Z, Wang Y, Lusis AG. 2012. “Good enough solutions” and the

- genetics of complex diseases. *Circulation Research* 111:493-504.
- West-Eberhard MJ. 2005. Developmental plasticity and the origin of species differences. *Proceedings of the National Academy of Sciences U. S. A.* 102:6543-6549.
- Whitehead A. 2012. Comparative genomics in ecological physiology: toward a more nuanced understanding of acclimation and adaptation. *Journal of Experimental Biology* 215:884–891.
- Wund MA. 2012. Assessing the impacts of phenotypic plasticity on evolution. *Integrative and Comparative Biology* 52:5-15.
- Xiao SJ, Zhang C, Zou Q, Ji ZL. 2010. TiSGeD: a database for tissue-specific genes. *Bioinformatics* 26:1273-1275.
- Yamaguchi Y, Hearing VJ. 2009. Physiological factors that regulate skin pigmentation. *Biofactors* 35:193-199.
- Yokoyama S, Takeda K, Shibahara S. 2006. SOX10, in combination with Sp1, regulates the endothelin receptor type B gene in human melanocyte lineage cells. *FEBS Journal* 273:1805-1820.
- Yu R, Broady R, Huang Y, Wang Y, Yu J, Gao M, Levings M, Wei S, Zhang S, Xu A, Su M. 2012. Transcriptome analysis reveals markers of aberrantly activated innate immunity in vitiligo lesional and non-lesional skin. *PLoS ONE* 7(12): e51040.
- Zhang B, Horvath S. 2005. A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology* 4:17.
- Zhang X, Liu D, Lv S, Wang H, Zhong X, Liu B, Wang B, Liao J, Li J, Pfeifer GP, Xu X. 2009. CDK5RAP2 is required for spindle checkpoint function. *Cell Cycle* 8:1206-1216.

CHAPTER 3

Phylogeography and diversification pattern of the white-nosed coati (*Nasua narica*)

Abstract

Middle America is a region with complex orography, dynamic climatic and geological history that has enabled the development of great biological diversity. Within Middle America, Central America played an important role in shaping this diversity due to the rise of the Isthmus of Panama that facilitated the Great American Biotic Interchange (GABI). Procyonids are thought to have colonized South America twice, once before and once during the GABI, specifically, the genus *Nasua* apparently arrived from North to South America in the late Pleistocene after the opening of the land bridge in Central America. To test this hypothesis we performed a phylogeographic analysis of the White-nosed coati (*Nasua narica*), which is mainly distributed in North and Central America but also inhabits Northern South America. Using several mitochondrial and nuclear loci we determined the phylogeographic pattern, population structure, divergence times, and direction of migration rates between White-nosed coati populations throughout its distribution range. We found a high degree of population structure and that the most southerly distributed population (Panama) branched off earlier, diverging before the GABI took place. The migration rates were low and mostly northwards and westwards. This phylogeographic structure is associated to geographic barriers and habitat shifts caused by Pliocene-Pleistocene climate oscillations. Our findings suggest the dispersion of *Nasua narica* was south-to-north beginning in the Pliocene, not in the opposite direction during the Pleistocene. The species might have crossed overwater to Southern Central America before the beginning of the GABI or on a land bridge due to an earlier uplift of Isthmus of Panama

triggering an earlier GABI. Regardless which GABI scenario is accepted, our study implies the possibility that the diversification of *Nasua* species, and other extant procyonid lineages, may have occurred in South America.

Introduction

Phylogeographic studies illuminate the historical geography of genetic lineages and are essential to better understand the evolutionary process of lineage diversification (Riddle *et al.* 2008; Hickerson *et al.* 2010). Unexpectedly, one of the regions with a paucity of phylogeographic studies is the Neotropics that contains seven of the world's 25 biodiversity hotspots (Myers *et al.* 2000; Beheregaray 2008; Bagley and Johnson 2014). Among these seven hotspots is Middle America, the area lying between the United States and South America (Winker, 2011), which is the second most important area for vertebrate species richness and endemism worldwide (Myers *et al.* 2000). 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 extinction, lineage diversification, dispersion and speciation occurs at high rates (Savage 1982; Marshall 2007; Castoe *et al.* 2009; Bryson *et al.* 2011a; Hardy *et al.* 2013; Gutiérrez-García and Vázquez-Domínguez 2013). Within Middle America, Central America has been of paramount importance 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).

The traditional model of the GABI assumes that the total closure of the Isthmus of Panama occurred circa 3.0-3.5 million years ago (Mya) in the Middle of the Pliocene, resulting in the closing of the Central American Seaway (CAS) (Coates *et al.* 1992; Coates and Obando 1996; Haug and Tiedemann 1998; Haug *et al.* 2001; Coates and Stallard 2013; O’Dea *et al.* 2016). After the emergence of the Isthmus, the fossil record indicates that mammalian lineages predominantly migrated south, to colonize South America 2.4-2.8 Mya (Simpson 1980; Webb 2006; Woodburne 2010), while South American lineages of invertebrate, plant, freshwater fishes, reptiles and birds mainly moved northwards to colonize Central and North America (Bagley and Johnson 2014). However, recent geological and biological studies have challenged the traditional model and suggested that the formation of the Isthmus of Panama and the GABI have a more complex and protracted history (Bacon *et al.* 2015; Erkens 2015). The new model proposes that a land bridge first appeared ca. 25-23 Ma with the collision of the Panama Block and South America in the Oligocene–Miocene transition, exposing a large portion of land, shoaling and reducing the CAS and allowing some terrestrial migration. This event was followed by the complete closure of the CAS, representing the almost full emergence of the land bridge between 13-15 Ma, instead of a Pliocene event (Farris *et al.* 2011; Montes *et al.* 2012a, Montes *et al.* 2012b; Montes *et al.* 2015). Concordant with the new model, a recent study using extensive fossil record and molecular data of vertebrate species, plants, insects and marine invertebrates, proposed that the GABI occurred over a much longer span, with pulses of biotic migration around 40, 20, 9 and 5 Ma (Bacon *et al.* 2015). This study suggests that the most significant increments in migration of terrestrial taxa, including mammals, occurred at ca. 20 and 6 Ma, with

migration rates between North and South America being similar through time and asymmetric differences in migration appearing only after the last pulse of migration (ca. 6 Ma), when more taxa migrated from South to North America. This model also suggests environmental processes, not geological or orographic features, might have been responsible for preventing the faunal dispersal (Montes *et al.* 2015; Bacon *et al.* 2016).

Although limited, phylogeographic research in Central America has allowed the juxtaposition of the evolutionary and biogeographic history of a variety of taxa with the complex geological events that occurred in the region over the last several million years, and helped to identify geological and climatic features that have shaped similar phylogeographic and genetic structure patterns in different taxa (Daza *et al.* 2010; Gutierrez-Garcia and Vazquez-Dominguez 2013; Bagley and Johnson 2014; Leigh *et al.* 2014). Therefore, phylogeographic studies can allow a test of the contrasting models of flora and fauna evolution due to the development of the Isthmus of Panama and the GABI. Ultimately, these data can contribute to a better understanding of the origins of the high biodiversity observed in the Neotropics (Bagley and Johnson 2014).

Procyonids (Mammalia: Carnivora: Procyonidae) are a group of mammals comprised of 14 species that include the olingos (*Bassaricyon* spp.), ringtails (*Bassariscus* spp.) raccoons (*Procyon* spp.), coatis (*Nasua* spp. and *Nasuella* spp.) and kinkajou (*Potos flavus*.) (Nowak 2005). Their diversification, evolution and biogeographic history in the New World is controversial (Koepfli *et al.* 2007, Eizirik 2012; Soibelzon and Prevosti 2013; Forasiepi *et al.* 2014). The fossil record indicates that procyonids dispersed from North America into South America two separate times, and were among the very first group of North American mammals to invade South America. The first dispersion event occurred during the Late Miocene with the

appearance of fossil genus *Cyonasua* in South America, long before the closure of the Isthmus of Panama and the major migration events of the GABI (traditional view; Marshall *et al.* 1979; Webb 1985; Webb 2006; Woodburne 2010; Forasiepi *et al.* 2014). 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 around or after 0.125 Ma (sensu Woodburne 2010; Webb 2006). Because of the gap in the fossil record of >500 thousand years between the extinct procyonids of the first dispersion event and the appearance of the extant species (Soibelzon 2011; 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; Soiblezon 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 extinct genera in South America (Koepfli *et al.* 2007; Eizirik *et al.* 2010). This result suggests the possibility that the extant genera may have entered and/or diversified in South America as part of the first procyonid dispersion event during the Middle to Late Miocene, or that extant genera are descendants of those first migrant taxa (Koepfli *et al.* 2007; Eizirik 2012). Clearly, additional fossil and molecular studies are needed to help resolve the discrepancies in the evolutionary and biogeographic history of procyonids.

Among coatis, four extant species are recognized: the white-nosed or Central American coati (*Nasua narica*), the South American coati (*Nasua nasua*), the Western mountain coati (*Nasuella olivacea*) and the Eastern mountain coati (*Nasuella meridensis*). *Nasua narica* is the only coati species distributed in North, Central and South America, from Southern Arizona, New

Mexico and Texas in the United States, to Northern Colombia (U.S.; Gompper 1995). Across its range, four subspecies are currently recognized based on body size, cranial features and coat coloration differences (Hall 1981; Decker 1991; Gompper 1995). The distribution of the subspecies are 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 flatlands of the Yucatan Peninsula and *N. n. nelsoni* is confined to Cozumel Island in Mexico (Hall 1981; Decker 1991; Gompper 1995). However, the validity of these subspecies, and the genetic structure among white-nosed coati populations, have never been formally tested using genetic data. Given the morphological, ecological and geographical differences observed across the wide distribution of *N. narica*, phylogeographic analysis could help illuminate not only the patterns of temporal and spatial diversification of this species, but also could help shed light on the patterns of faunal dispersion associated with procyonids during the GABI.

Here, we performed the first phylogeographic study of the white-nosed coati throughout its distribution range using a combination of 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 *Nasua 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 much earlier than 3.0 Mya better explains the biogeographic history of the species. Additionally, we assessed whether ecological, biological or geological processes have driven the population structure. In order to test this hypotheses we: 1) determined the population structure and phylogenetic relationships of white-nosed coati populations throughout most of its geographic 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; and 3) calculated recent and ancient migration between populations to evaluate possible directions of historical population movements. Our results have important novel implications for the diversification and evolutionary history of procyonids in South, Central and North America, and suggest a different perspective on the faunal dynamics previous to and during the GABI.

Materials and methods

Tissue collection and DNA extraction

We assembled a collection of 85 white-nosed coati samples from throughout most of the distribution range of the species, except for localities in northwestern Colombia (Figure 3.1A). Our sample collection includes representatives of the four recognized species of *Nasua narica*. Samples were collected from a variety of tissues including whole blood, ear punches and muscle obtained from animals caught in the field, roadkills, zoo and museum specimens (Appendix 3.1). In addition, we obtained samples from 10 South American coatis (*N. nasua*), two Western lowland olingos (*Bassaricyon medius*) and one Eastern lowland olingo (*B. alleni*) (Appendix 3.1). Genomic DNA was extracted from all samples using a Qiagen DNEasy Kit (Valencia, CA) following the manufacturer's protocol.

Mitochondrial sequences

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 majority of the sequences of the three genes were

newly generated. However, we also included four *CYTB* sequences used in a previous study: *Bassaricyon alleni* (Genbank accession DQ660299), *B. medius* (DQ660300; originally classified as *B. gabbii* but see Helgen *et al.* 2013), *Nasua narica* (DQ660302), and *N. nasua* (DQ660303). Furthermore, the orthologous sequences of the three gene segments were extracted from the mitochondrial genome of a *Nasua nasua* deposited in Genbank (HM106331) and used in the molecular dating analyses (see below). The *CYTB* gene was amplified in two overlapping segments in two separate polymerase chain reactions (PCRs) using primers L14724 and H151513 (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 (Sorenson *et al.* 1999), respectively. All PCRs were performed in a total volume of 50 μ L containing 1 μ L of DNA (~0.5 μ g), 5 μ L 10x PCR buffer, 5 μ L of 25 mM MgCl₂, 1 μ L 10 mM dNTP mix, 1 μ L of 25 pM/uL forward and reverse primers, 0.3 μ L *Taq* polymerase (Sigma-Aldrich, St. Louis, MO), and 35.7 μ L 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; followed by hold at 4°C. All PCRs were run with a negative control (1 μ L sterile ddH₂O instead of DNA). Electrophoresis was used to visualize a 5 μ L aliquot of the PCR products in a 1% agarose/Tris-acetic acid-EDTA 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, Affymetrix, Santa Clara, CA). Purified PCR products were cycle sequenced in a 10 μ L volume using the forward and reverse PCR primers, the BigDye Terminator v3.1 Cycle Sequencing Kit (Life

Technologies, Grand Island, NY) and SeqSaver Sequencing Premix Dilution Buffer (Sigma-Aldrich, St. Louis, MO), following the manufacturer's protocol. We used the following thermal cycling conditions: one cycle at 96°C for 1 min; 45 cycles at 96°C for 10 sec, 50°C for 5 sec, 60°C for 4 min; followed by incubation at 4°C. Cycle sequencing products were purified and sequenced on a 96-capillary 3730xl DNA Analyzer (Life Technologies, Grand Island, NY) at the UCLA DNA and Genotyping Core Facility. We visually inspected, edited and assembled forward and reverse sequence reads (chromatograms) using Sequencher 3.1 (Gene Codes Corporation, Ann Arbor, MI) or Geneious Pro versions v3.7-3.8 (BioMatters Ltd., Auckland, New Zealand). *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.

Alignment and phylogenetic analyses

Sequences from each of the mitochondrial loci were aligned using MAFFT v7.017 (Kato *et al.* 2002) in the Geneious v7.1.4 package with the following parameter settings: scoring matrix = 200PAM/k=2, gap open penalty = 1.53, offset value = 0.123. The resulting alignments were adjusted by eye. The alignments were then concatenated to generate a data matrix containing the three mitochondrial gene segments (*CYTB+NADH5+16S*), totaling 2,201 bp. We converted this alignment into a fasta file in Geneious, exported it, and then applied the DNA to haplotype collapser tool in FaBox v1.41 (Villesen, 2007) so that only unique haplotypes were represented for *Nasua narica* (21 haplotypes). We used jModelTest v2.1.4 (Darriba *et al.* 2012) to select the best-fitting models of DNA substitution for this data matrix from among 24 models using the Bayesian information criterion (BIC), and a starting tree estimated with BIONJ (Gascuel, 1997). We then estimated maximum likelihood (ML) trees using raxmlGUI 1.3.1

(Silvestro and Michalak, 2012), a front-end application that runs RAxML v7.4.2 (Stamatakis, 2006). ML trees were estimated with the GTR+G model (the only model available in RAxML). Branch support was evaluated using 1000 thorough-search bootstrap replicates. A total of three different ML+bootstrap searches were conducted. Output trees and their branch supports were visualized in FigTree v1.4.2 (Rambaut, 2014). Trees were rooted using sequences of the western lowland oringo (*Bassaricyon medius*) and eastern lowland oringo (*Bassaricyon alleni*) as outgroups.

Divergence time estimation

We used BEAST v2.3.1 (Bouckaert *et al.* 2014), which employs a Markov chain Monte Carlo (MCMC) method to average over tree space, to estimate the divergent times among the phylogroups within *Nasua narica*. We set up and generated XML files for different MCMC runs using the BEAUti application. We generated two different XML files, 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 as the best-fitting model of DNA substitution (estimated with jModelTest v2.1.4 as above) with empirical base frequencies, and the proportion of invariable sites set to Estimate; clock model set to Strict; tree prior set to coalescent constant population; gamma distribution (0.001-1000) set for the clock rate prior; a 1/X distribution set for the population size prior; a calibration prior with a truncated normal mean of 7.2 Mya and standard deviation of 1.7 Mya, based on the estimated time of the split between the *Bassaricyon* and *Nasua* lineages from an analysis of 14 concatenated nuclear genes (Eizirik *et al.* 2010; also see Helgen *et al.* 2013); and the auto optimize setting was 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 a mean of 1.75 Mya and a standard deviation of 0.15 Mya, to calibrate the split between *Nasua 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; Gilmore, 2013). For both XML files, MCMC analyses were run for 20,000,000 generations, trees and parameters sampled every 2,000 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 v1.6 (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 independent run of the Mitochondrial-1, Mitochondrial-2 XML files. The post-burn-in samples of the posterior distribution from the two independent runs of each of the three data sets were merged using LogCombiner (Bouckaert *et al.* 2014) and FigTree v1.4.2 (Rambaut, 2014) was used to visualize the maximum clade credibility topology and mean node heights.

Haplotype genetic diversity and haplogroups structure

We estimated the genetic diversity among the concatenated mtDNA haplotypes by calculating the nucleotide and haplotype diversity indices with ARLEQUIN v3.5.2.2 (Excoffier and Lischer 2010). The genetic structure of the mtDNA data was tested using an analysis of molecular variance (AMOVA) using the Tamura-Nei model of molecular evolution (which is the closest to the GTR+G) implemented in ARLEQUIN (Excoffier and Lischer 2010), to estimate the amount of genetic variation among and within populations. To further investigate the relationships between all haplotypes and haplogroups identified with the AMOVA test, we constructed unrooted networks using two methods: a median-joining haplotype network of the

concatenated mitochondrial haplotypes with the software NETWORK v4.6.1.1 (Bandelt *et al.* 1999) using default parameters, and a network based on uncorrected patristic distances using the Neighbor-Net method with 1000 bootstrap replicates implemented in the SPLITSTREE4 program (Huson and Bryant 2006).

Number of migrants estimation

We estimated potential past migration events and determined the number of immigrants 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 data from the three mitochondrial loci. 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. We then used the estimates of these parameters as priors for the final Bayesian inference analysis that used the Brownian mutation model, mutation rate estimated from the data, uniform data distribution, a long chain of 5,000,000 steps with four replicates, a burn-in length of 10,000 steps and a static heating scheme with four chains (1, 1.5, 3 and 1×10^6).

Microsatellites

Microsatellite amplification and genotyping

We tested 15 putatively neutral dinucleotide microsatellite loci previously described for the South American Coati (*N. nasua*; Almany *et al.* 2009). Two of these loci amplified inconsistently (*i.e.* in less than in 50% of our samples) and two others were monomorphic, precluding us from using them in our analysis. We were able to successfully amplify 11 microsatellite loci (Table 3.S1) in 85 white-nosed coati individuals. Loci amplification was

performed by a touchdown PCR 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 10mg/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 Biosystem (ABI) 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).

Genetic diversity, Hardy-Weinberg equilibrium and linkage disequilibrium

Genetic diversity of the microsatellite data was measured as allelic diversity in GENALEX v6.5 (Peakall and Smouse 2012) and as the observed (H_o) and expected (H_e) 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 using GENEPOP v4.5.1 (Raymond and Rousset 1995; Rousset 2008). GENEPOP 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.

Genetic structure

We performed several analyses to test for population structure in our sample. First, 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) without *a priori* assumptions about sample location and to assign individuals to clusters based on their multilocus genotypes, assuming correlated allele frequencies and admixture ancestry between clusters. We assessed K values from 1 through 10 using a burn-in period of 50,000 iterations and 500,000 sampling iterations for each value of K. The stability of the inferred clusters was evaluated using 10 independent runs per K value. The log likelihood [$\ln P(X | K)$] values and ΔK parameter of Evanno *et al.* (2005) were calculated with the software STRUCTURE HARVESTER v0.6.94 (Earl and vonHoldt, 2012) and used to determine the most probable number of genetic clusters. We used the CLUMPP v1.1.2 program (Jakobsson and Rosenberg 2007) with the greedy algorithm to account for the admixture variation of individuals over the 10 replicate runs per K in STRUCTURE.

Second, 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 F_{ST} (Weir and Cockerham 1984) between populations using ARLEQUIN (Excoffier and Lischer 2010). AMOVA and F_{ST} estimations were calculated using permutation tests of 10,000 randomized runs. In addition, 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) in the program POPULATIONS v1.2.30 (Langella 1999). We utilized the program GENECLASS2 (Piry *et al.* 2004) to perform Bayesian assignments tests (Rannala

and Mountain 1997) in order to determine the likelihood of each individual being assigned to its population of origin based on its multilocus genotype. We tested different clustering schemes and the assignment probabilities were estimated using 10,000 simulations of multilocus genotypes (Paetkau et. al. 2004) and an assignment threshold level of 0.01.

Migration rates estimates

Recent potential migration rates among the identified populations were estimated using a Bayesian MCMC analysis of microsatellite genotypes in BAYESASS v3.0.4 (Wilson and Rannala 2003). Individuals were pre-assigned to five populations based on the population structure analysis described above. 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 for migration rate, allele frequency and level of inbreeding of 0.25, 0.65 and 0.80, respectively, which were adjusted to allow correct mixing of the chain. To assess convergence of the results, four runs under the same parameters were performed using different initial seed numbers.

Results

Mitochondrial sequences

Phylogenetic and divergence time analyses

Among the 85 white-nosed coati samples sequenced for 2,201 bp across three mitochondrial gene regions, we identified 21 haplotypes that differed by 1 to 237 substitutions (0.046 – 10.78% uncorrected p-distance). Four haplotypes differing from 2 to 122 substitutions (0.092 – 5.55% uncorrected p-distance) were observed among the 10 samples of South American

coatis. The 21 white-nosed coati haplotypes were assorted into three primary clades based on ML phylogenetic analyses (Figure 3.1B). Three haplotypes from samples derived from several locations in Panama comprised the earliest branching lineage within mitochondrial gene tree and these haplotypes were highly divergent from the remaining haplotypes (9.92 – 10.78 % uncorrected p-distance). The middle clade, which we denoted the Yucatan Peninsula-Guatemala clade, was comprised of 11 haplotypes from samples collected in Belize, the Yucatan Peninsula region of Mexico, Cozumel Island, Guatemala and Costa Rica. Lastly, seven haplotypes defined a clade containing samples from central Mexico (*i.e.* state of Morelos), western Mexico (*i.e.* state of Jalisco) and the southwestern United States (*i.e.* states of Arizona and New Mexico). Eight samples from central Mexico, contained two haplotypes and defined a distinct subclade separated from two other subclades north and west of the Sierra Madre Occidental mountain range comprised of samples from the southwestern United States (U.S.) and from western Mexico (Figure 3.1A; Figure 3.1B). Bootstrap and posterior probability support was high (>95% and 0.95, respectively) for most of the clades and sub-clades across the mitochondrial gene tree (Figure 3.1B).

Divergence times estimated from the two data sets (Mitochondrial-1 and Mitochondrial-2) employing different sets of calibration priors were generally congruent. The molecular dating analyses showed that the clade consisting of the Panama haplotypes first diverged approximately 4 Mya (95% highest posterior density [HPD] = 2.0 – 6.7 Ma and 2.6 – 5.1 Mya for the Mitochondrial-1 and Mitochondrial-2 analyses, respectively), whereas the other two primary clades split around 1.3 Mya (95% HPD = 0.59 – 2.1 Ma and 0.78-1.6 Ma for the Mitochondrial-1 and Mitochondrial-2 analyses, respectively; Figure 3.2). The split separating the subclade of the central Mexico haplotypes from the subclades of the southwestern U.S. and western Mexico,

haplotypes occurred about 0.89 to 1.0 Mya (Figure 3.2). Interestingly, the timing of the splits within white-nosed coatis were similar to those observed within the limited samples obtained from South America coatis. It is clear in both the phylogenetic and molecular dating analyses that the order of splits within white-nosed coatis proceeds in a south to north pattern rather than the reverse direction.

Population genetic analyses

The AMOVA analysis confirmed our phylogenetic results by identifying five haplogroups corresponding to the main clades and subclades in the phylogenetic tree as the grouping scheme with the highest genetic variance among groups of populations (Table 3.S2). In addition, both median-joining and patristic distance networks are consistent with the phylogenetic tree, and high bootstrap values of the patristic distance network supported the five haplogroups associated with the deepest *Nasua narica* clades (Figure 3.1B; Figure 3.3; Figure 3.S1). The haplotype and nucleotide diversity were low in Central Mexico and Panama populations, whereas higher diversity was found in the Southwestern U.S., Western Mexico and Yucatan Peninsula–Guatemala populations (Table 3.1). Also, the F_{ST} values for the mitochondrial sequences were statistically significant and very high between all the five genetic groups, ranging from 0.910 to 0.998 (Table 3.2).

The estimated number of migrants per generation based on the concatenated mitochondrial sequences was low among all populations, ranging from 0.084 to 0.239 (Table 3.S3), which indicate less than one individual per generation migrate between populations. The number of migrants between the Panama population and all other populations were consistently among the lowest we observed (0.084231- 0.176656) and had asymmetric patterns, with

migration from Panama into Northern populations usually greater than in the opposite direction (Figure 3.4A; Table 3.S3), suggesting that this population has been isolated for a considerable time, which is consistent with the deepest divergence time of this clade and the south-to-north pattern found for the cladogenesis of the species (Figure 3.1B). All other populations showed higher migration between them, with the Western Mexico population sending more immigrants out than receiving them from other populations (Figure 3.4A; Table 3.S3), which suggests a northeastern migration from this population.

Microsatellite analyses

Genetic diversity

We found that all 11 microsatellite loci were variable in white-nosed coatis with the number of alleles per loci ranging from 3 to 17 (Table 3.S4). The loci were in Hardy-Weinberg equilibrium, all genetic populations had from one to three loci with heterozygote deficit, but no loci showed deviation from Hardy-Weinberg equilibrium across all populations (Table 3.S5). Also, no evidence of linkage disequilibrium, null alleles, allelic dropout or stuttering was found in our data.

Population structure

The Bayesian cluster analysis showed that $K=5$ had the maximum likelihood (-2601.14; Figure 3.5A). The Evanno ΔK parameter showed two peaks, at $K=2$ and $K=5$ (Figure 3.5B). At $K=2$ the group including samples from the southwestern United States and western Mexico was separated from the group containing individuals from Central Mexico southwards (Figure 3.5C). 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) combined with the extensive geographic range of our taxon sampling, we tested 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; Figure 3.5C). Although K=6 showed substructure within the Yucatan Peninsula-Guatemala cluster (Figure 3.5C), when we ran this cluster alone in STRUCTURE no evidence of substructure was found (Figure 3.S2). Therefore, we decided that K=5 is the most probable number of genetic clusters in our sample. In addition, when assuming these five genetic clusters, the Bayesian assignment test performed in GENECLASS2 correctly assigned the highest number of individuals (76 individuals; 89.4%) to their populations of origin ($p < 0.01$) compared to alternative clustering schemes (Table 3.S6). Furthermore, the neighbor-joining tree of allele-sharing distance clearly defined five clusters, with only three individuals mixed between them (Figure 3.5D), two of which may represent second generation migrants according to the structure plot (Figure 3.5C).

The differences in allele frequency (F_{ST}) between the five genetic groups was highly significant in all pairwise comparisons ranging from 0.101 to 0.328 (Table 3.2), suggesting low gene flow between these populations. Surprisingly, the lowest F_{ST} value was between the populations from Morelos and the Yucatan Peninsula, which are more distantly separated than from the Jalisco and Panama populations, respectively. The AMOVA of the microsatellite loci following this 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 3.S7). The cluster comprised of samples from Central Mexico showed the highest level of heterozygosity whereas the cluster of

individuals from Panama had the lowest (Table 3.3). The Yucatan-Guatemala cluster showed an intermediate level of heterozygosity (Table 3.3)

Concordant with the results from the mitochondrial sequence data, the estimates of recent migration rates based on the microsatellite data are low (Table 3.S8), ranging from 0.0079 to 0.0671, suggesting that very small numbers of individuals have migrated between these populations in recent generations. In general, recent migration has occurred northwards and westwards (Figure 3.4B), with the Yucatan-Guatemala and Panama populations migrating at a higher rate into the populations in Central and Western Mexico than vice versa. Although the population in the Southwestern U.S. have migrated south into Central and Western Mexico, the migration rates into these populations are not as high as the migration from southern populations (Figure 3.4B; Table 3.S8). The proportion of migration is almost symmetrical between Central and Western Mexico (Figure 3.4B; Table 3.S8).

Discussion

Genetic diversity and structure

We found remarkable genetic differentiation between white-nosed coati populations characterized by high F_{ST} values (Table 3.2), high population structure (Table 3.S2; Table 3.S7; Figure 3.5), and low migration between populations with less than one migrant per generation (Table 3.S3; Table 3.S8; Figure 3.4). In the analysis of mtDNA sequences, we detected the presence of three main clades: Panama, Yucatan Peninsula-Guatemala and Southwestern U.S.-Western Mexico-Central Mexico. The latter clade is subdivided in three subclades: central Mexico, western Mexico and Southwestern U.S. (Figure 3.1B), and Bayesian clustering based on

11 microsatellite nuclear loci strongly support the presence of these five populations (Figure 3.5). These results reflect the history of high diversification in taxa inhabiting Middle America, molded by the complex topographical features of the region combined with environmental changes resulting from Pliocene-Pleistocene glacial cycles in the temperate zone (Dansgaard *et al.* 1993; Hewitt 1996; Bryson *et al.* 2011a; Gutierrez-Garcia and Vazquez-Dominguez 2013; Hardy *et al.* 2013; Bagley and Johnson *et al.* 2014).

The Panama population is the earliest branching and most divergent lineage (Figure 3.1B; Figure 3.3; Figure 3.S1) with more than 10% mitochondrial sequence divergence from other populations. Furthermore, this population had the highest F_{ST} values (Table 3.2) and lowest migration rates for both mitochondrial and nuclear markers (Table 3.S3; Table 3.S8 Figure 3.4). It is important to note that the observed level of mitochondrial sequence divergence of the Panama population is twice the amount of that which typically defines recognized species of mammals based on the divergence of the cytochrome-*b* gene (>5%; Baker and Bradley 2006). Similar levels of genetic divergence in cytochrome-*b* sequences are found between other procyonid species: olingos (*Bassaricyon alleni*, *B. medius*, *B. gabbii*) and olinguito (*B. neblina*) 9.6-11.3% (Helgen *et al.* 2013), raccoons *Procyon lotor* and *P. cancrivorus* 10-11% (Helgen and Wilson 2005; Helgen *et al.* 2013), and coatis *Nasua narica* and *Nasuella olivacea* 9-13% (Helgen *et al.* 2009). Additional data from autosomal and sex chromosome sequences as well as morphology will be required to further explore whether white-nosed coatis from Panama and neighboring regions represent a distinct species. Also, both haplotype and nucleotide diversity are reduced in the Panama population, which is the opposite of what would be expected in a population founded several million years ago (DeGiorgo *et al.* 2009). This discrepancy may be due to a recent population bottleneck or sampling artifact. In addition, the social behavior of the

species favor a low diversity in the mitochondrial DNA. Studies including more thorough sampling of this population and the analysis of additional nuclear genes would help to determine the cause of the low diversity we observed.

The other white-nosed coati populations show high F_{ST} values (Table 3.2) and a sequence divergence of 2-4% (except between Southwestern U.S. and western Mexico populations, which is 0.5%) a level that is more in accordance with the recognition of subspecies (Avice and Walker 1999; Helgen *et al.* 2009). Our results are only marginally concordant with the 4-subspecies scheme proposed for white-nosed coatis based on morphological data (Decker 1991; Gompper 1995; Wozencraft 2005). Our phylogeographic and population genetic analyses indicate that populations from the Southwestern US (Arizona and New Mexico) and Western Mexico (Jalisco) are genetically differentiated yet define a subclade in the ML and Bayesian phylogenetic trees that generally conforms to the distribution ascribed to *N. n. molaris*, north of the Trans-Mexican Volcanic Belt (TMVB). However, south of the TMVB, the populations or phylogroups we identified did not coincide with the ranges suggested for the other three subspecies (*N. n. narica*, *N. n. nelsoni*, *N. n. yucatanica*; Gompper 1995). Although the number of samples of putative *N. n. nelsoni* was low, we did not find evidence of significant differentiation indicating individuals from Cozumel island represent a different subspecies from those in mainland Yucatan peninsula (Table 3.S2; Table 3.S5; Table 3.S7; Figure 3.1B; Figure 3.3; Figure 3.5). Nonetheless, the presence of private haplotypes and the star-shape haplotypes distribution (Figure 3.3) may indicate a recent expansion and imminent diversification. Samples from Cozumel Island were clustered with white-nosed coatis sampled from different areas of the Yucatan Peninsula, which in turn grouped with individuals from Guatemala and one sample from Costa Rica. These latter two areas, along with localities in Panama (and possibly further south to

northern Colombia) encompass the range of *N. n. narica* (Gompper 1995) but our results clearly show this range is comprised of multiple, deeply divergent phylogroups with a polyphyletic history (according to the mitochondrial gene tree).

The strong differences observed in the mitochondrial sequences may be in part facilitated by the social behavior of white-nosed coatis. Individuals form phylopatric groups consisting of only adult females and their offspring, and the home ranges of these groups are relatively small, 0.34 – 3 Km² (Gompper *et al.* 1995; Valenzuela and Ceballos 2000), which may result in a limited dispersion of mitochondrial maternal lineages. In contrast, males disperse from the groups when they are adults (Valenzuela and Ceballos 2000). This could explain the instances of admixture between populations detected with the microsatellite loci and thus represent dispersion events by males (Figure 3.5C; Figure 3.5D).

Phylogeographic pattern and drivers of divergence.

The divergence among clades and population structure observed in *Nasua narica* tightly corresponds with several geographic barriers and environmental changes associated to Pliocene-Pleistocene climate oscillations, whose combined actions may have shaped the phylogeographic pattern of the species (Castoe *et al.* 2009; Bryson *et al.* 2011a; Gutierrez-Garcia and Vazquez-Dominguez 2013; Hardy *et al.* 2013). The pattern of diversification and estimated divergence time of the Panama population (~4 Ma, Figure 3.1B; Figure 3.2) coincides with the final uplift of the Cordillera de Talamanca mountain range 3- 5 Mya (Abratis and Wörner 2001; MacMillan *et al.* 2006; Mann *et al.* 2007) that separates northwestern Panama from southern Costa Rica, suggesting that this mountain range may have played an important role in the isolation of this population. Genetic studies have identified the Cordillera de Talamanca range as a significant

barrier involved in the diversification of a broad range of taxa including amphibians (Wang *et al.* 2008; Hauswaldt *et al.* 2011), reptiles (Castoe *et al.* 2003; Daza *et al.* 2010), birds (Cadena *et al.* 2007; Arbeláez-Cortés *et al.* 2010) and small mammals (Hardy *et al.* 2013; Arellano *et al.* 2005; Bradley *et al.* 2008). Interestingly, the only individual from Costa Rica included in our analysis is more closely related to the geographically more distant Yucatan Peninsula-Guatemala clade than to the Panama clade (Figure 3.1A; Figure 3.1B; Figure 3.3), although it represents the earliest branching haplotype within the former clade (ca. 0.6 Mya; Figure 3.2). This suggests that white-nosed coatis from this location likely represent a distinct population. If this is the case, the Nicaragua depression, which involved several marine transgressions during the Pliocene and probably the Pleistocene (Coates and Obando 1996; Bagley and Johnson 2014), may have separated areas of Costa Rica from those further north. This geologic event would have facilitated the differentiation of populations in this region as it apparently did for other mammals and carnivore species including mice, ocelots and margays (Eizirik *et al.* 1998; Gutiérrez-García and Vázquez-Domínguez 2012). More exhaustive sampling of white-nosed coatis from Costa Rica as well as from both sides of the Nicaragua depression is needed to test the hypothesis that coatis south of the Nicaragua depression define a distinct population.

The Yucatan Peninsula-Guatemala clade has been identified in other co-distributed species of reptiles, birds and mammals, 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 and probably drive the diversification of this clade (Castoe *et al.* 2009; Daza *et al.* 2010; Bryson *et al.* 2011a; Gutiérrez-García and Vázquez-Domínguez 2012; Gutiérrez-García and Vázquez-Domínguez 2013 and references there in). However, these barriers were formed well before the split of the white-nosed coati Yucatan Peninsula-Guatemala

clade 1.3 Mya (Figure 3.2). The last marine incursion that inundated the Isthmus of Tehuantepec was 2.5 Mya and retreated by 1.8 Mya (Barrier *et al.* 1998; Gutiérrez-García and Vázquez-Domínguez 2012), whereas the main uplift of the Motagua-Polochic-Jocotan fault occurred in the late Miocene-early Pliocene (6 – 2.5 Mya; Ortega-Gutiérrez *et al.* 2007). Nonetheless, the split of this clade closely coincides with the sixth North American glaciation that occurred 1.10-1.30 Mya (Barendregt and Duk-Rodkin 2011; Rutter *et al.* 2012). Similarly, the mountains in northwestern and western Mexico and those of the TMVB are also too old for their emergence to have caused diversification of white-nosed coati populations 0.1-1.2 Mya (Figure 3.2). The Sierra Madre del Sur range, which separates the central from the western Mexico population in the west, and the Sierra Madre Occidental range, that isolates the western Mexico from the Southwestern U.S. population in the North, were almost completely formed in the Oligocene-early Miocene (20-33 Mya) and late Eocene- early Oligocene (30-35 Mya), respectively (Ferrari *et al.* 1999; Ferrari *et al.* 2000; Ferrusquía-Villafranca *et al.* 2005; Nieto-Samaniego *et al.* 2006). Whereas the TVBM reached most of its current form in the late Miocene-early Pliocene, 5-10 Mya (Ferrari *et al.* 2000; Ferrusquía-Villafranca *et al.* 2005). Yet, the split of the Central Mexico population ~ 1.0 Mya (Figure 3.2) is concordant with the glacial episode of the seventh North America glaciation 0.99-1.07 Mya, whereas the divergence of the western Mexico and southwestern U.S. populations in the Middle Pleistocene is consistent with the Reid glaciation 0.13-0.28 Mya (Barendregt and Duk-Rodkin 2011; Rutter *et al.* 2012; Figure 3.2). Multiple taxa of reptiles, birds and mammals show phylogeographic patterns that are congruent to the one we found for the white-nosed coati in this region of Middle America, and have been found to be associated with geographical barriers and/or habitat changes caused by glacial climatic oscillations (Sullivan *et al.* 2000; McCormack *et al.* 2008; Bryson *et al.* 2011a; Bryson *et*

al.2011b; Hardy *et al.* 2013; Castañeda-Rico *et al.* 2014; Suarez-Atilano *et al.* 2014). This observation suggests that these high mountain ranges have likely prompted phylogeographical breaks due to altitude limitation of the constituent taxa. Further, our findings are consistent with the hypothesis that vicariant events cause species level differentiation in vertebrate species, whereas habitat shifts provoked by climate change may be responsible for population-level structure (Castoe *et al.* 2009; Bryson *et al.* 2011a). Although our phylogenetic results are highly supported, the results based solely on mitochondrial DNA analysis should be taken with caution since they represent one locus and may only indicate the evolutionary history of maternal lineages. Genome-wide studies would help to confirm our results without bias toward any sex lineage and considering information from thousands of loci across the whole genome.

Patterns of diversification of Nasua narica and implications for the evolution of procyonids in the context of the Great American Biotic Interchange

Our results clearly indicate that the direction of cladogenesis in the white-nosed coati was in a south-to-north direction (Figure 3.1B) and the migration between populations has been mostly northwards and westwards (Figure 3.4). In addition, our estimates locate the earliest branching clade of the species in Panama around 4 Mya (Figure 3.2), previous to the first episode of the GABI (*sensu* Woodburne 2010), which occurred 2.4 -2.8 Mya (Figure 3.6). These results challenge the hypothesis that ancestors of living procyonid species, specifically *Nasua* spp., migrated from North America to South America during the last episode of the GABI in the late Pleistocene 0.125 Mya (Webb 2006; Woodburne 2010; Figure 3.6). The earliest fossils of the genus *Nasua* have been identified in North America, from the late Hemphillian (NALMA; 6.8-4.9 Mya) to early Irvingtonian (2.5-1.0 Mya; Baskin 1982; Dalquest 1978; Cassiliano 1999; Gilmore

2013). Based on these remains and other North American procyonid fossils, and the late appearance (*i.e.* Pleistocene) of extant species in the fossil record of South America, it was proposed that extant procyonid species, including *Nasua* spp., are descendant from North American lineages that migrated to South America during the Pleistocene following the closure of the Panamanian isthmus (Baskin 1982; Baskin 1989; Baskin 2003; Forasiepi *et al.* 2014; Figure 3.6). Conversely, phylogenetic studies indicate that the diversification of *Nasua* and *Procyon* species and South American extinct species (*Cyonasua* spp. and *Chapalmalania* spp.) occurred around the same time 5- 7 Mya (Koepfli *et al.* 2007; Eizirik *et al.* 2010; Eizirik 2012, Helgen *et al.* 2013; Forasiepi *et al.* 2014), predating the GABI and concordant with our finding of the mean divergence time between *N. narica* and *N. nasua* estimated about 6 Mya (Figure 3.2; Figure 3.6). In addition, most of the 14 extant currently recognize procyonid species inhabit South and Central America (*i.e.* nine and eight species, respectively), with all coati species (*i.e.* *Nasua* spp. and *Nasuella* spp.) at least partially distributed in South America (Helgen *et al.* 2013). Concordantly, in a study on the taxonomic revision of olingos (*Bassaricyon* spp.), a dispersal-extinction-cladogenesis analysis to infer geographic range evolution based on a molecular phylogeny identified Central America, not North America, as the most probable region of origin of dispersion for all extant procyonid genera including *Nasua* (Helgen *et al.* 2013). These findings are consistent with our results that the Panama population is the earliest branching lineage of *N. narica* and migration between populations has been mostly northwards (Figure 3.2; Figure 3.4). Molecular phylogenies, distribution of extant species and the fossil record have led to the suggestion that extant *Procyon*, *Nasua* and the South American extinct lineages *Cyonasua* and *Chapalmalania* may have been part of a temporally concordant diversification event in South America 5-7.3 Mya, previous to the GABI as suggested in the

traditional model (Koepfli *et al.* 2007; Eizirik 2012; Forasiepi *et al.* 2014; Figure 3.6). Also, because of the age of the North American *Nasua* and *Procyon* fossils (6.8 -1.0 Mya) it is possible that they may actually represent South American lineages that migrated into North America and went extinct in the late Pliocene and Pleistocene due to climate change related to glacial cycles, as it has been documented for several mammal species (Martin 1984; Stuart 1991; Guthrie 2003; Figure 3.6).

The fact that *N. narica* is almost exclusively distributed in Central and North America inhabiting only the most northern part of South America west of the Andes, combined with our results of diversification and migration northwards suggests that the most probable location for the initial diversification of *Nasua* species might have been the northern Andes. The northern Andes fast uplift 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* (Figure 3.2). Furthermore, this region has played an important role for the diversification of other procyonid species (Helgen *et al.* 2013), including the divergence of two species of mountain coatis that only inhabit the northern Andean range, *Nasuella olivacea* and *Nasuella meridensis* (Helgen *et al.* 2009; Helgen *et al.* 2013). Recent analyses of cytochrome-b sequences from both *Nasua* and *Nasuella* species showed that *Nasuella meridensis* and *N. olivacea* comprise a clade sister to *Nasua narica*, rendering the genus *Nasua* paraphyletic (Helgen *et al.* 2009). *N. nasua* represents the deepest split and sister taxon to these two lineages, which suggests that the all coati species should be included within the genus *Nasua* (Helgen *et al.* 2009). These findings suggest the Northern Andes was both the origin of diversification of currently recognized coati species and a barrier to dispersal, restricting *N. narica* to the western side of this mountain range, *N. nasua* to the lowlands of South America in the eastern side, and isolated the mountain coatis to the

highlands of this range (Helgen *et al.* 2009; Helgen *et al.* 2013). Subsequent to these diversification events, *Nasua narica* eventually reached southern Central America and dispersed northwards (Figure 3.6). A more comprehensive phylogeographic analysis including samples of *N. narica* from northern South America populations (*i.e.* northern Colombia) are required to further test our hypothesis for the evolution and diversification of *Nasua* in South America, along with the discovery of older *Nasua* fossils in tropical Central and South America.

In the context of the GABI, our results have a more parsimonious explanation under the new GABI model since the traditional model would require procyonids to cross twice overwater, first during the Miocene dispersion to South America and then *N. narica* crossing back to Central America ~4 Mya (Figure 3.6). Under the new model both dispersion events would occur over the land bridge which formed 13-15 Mya and climatic or environmental changes, not geological events, would be responsible for preventing or facilitating the faunal dispersion (Molnar 2008; Bacon *et al.* 2015; Montes *et al.* 2015; Bacon *et al.* 2016; Figure 3.6). As most of mammals migrant taxa in the fossil record that participated in the GABI were open-country species that do not thrive in forested environments (Webb 2006; Woodburne 2010), the new GABI model proposes that the moist and warm climate that existed in northern South America and Central America before 3.5 Mya favored tropical forest environments and prevented the faunal interchange from occurring earlier, even when the land bridge was already present (Molnar 2008; Montes *et al.* 2012b; Leigh *et al.* 2014). Then, climate change initiated by cooling glacial periods in the middle Pliocene (3-3.5 Mya) caused the evolution of dry savanna-like habitats replacing the tropical forest and facilitating the dispersion of taxa across the isthmus of Panama (Webb 2006; Molnar 2008; Bacon *et al.* 2016). The white-nosed coati is considered a tropical woodland species adapted to forested habitats (Gompper 1995) and would have easily dispersed through

tropical forests before savanna-like habitats surged 3 Mya. This scenario is consistent with two of our previous hypotheses: 1) that the white-nosed coati may have migrated north from South into Central America through forested habitats before the GABI, under the traditional model, started 2.4 -2.8 Mya (Woodburne 2010; Figure 3.6); and 2) that habitat changes have been important drivers of diversification events in this species. However, the dating of the GABI and the final uplift of the Panamanian isthmus remain contentious (Coates and Stallard 2013; Montes *et al* 2015; Bacon *et al.* 2015; Bacon *et al.* 2016; O’Dea *et al.* 2016). If the traditional model is accepted, it would not affect our conclusion on the direction of diversification and migration, and would indicate that *Nasua narica* dispersed overwater from South to Central America before the final closure of the isthmus (Figure 3.6). In fact, there is evidence for overwater dispersion of *Nasua* and *Procyon* species to Caribbean islands (McFadden *et al.* 2008), which renders this explanation plausible and suggests a complex dispersal and evolutionary history of procyonids in South and Central America.

Our study reveals that there is strong population structure through the distribution range of the white-nosed coati. The finding that the Panama population, which diverged 4 Mya, may represent a distinct species, warrants further analysis. We demonstrate, contrary to what was previously thought, that the migration and diversification of the species has been in a south-to-north direction and that the diversification have probably been driven by a combined action of geographic barriers and habitat shifts associated with glacial periods. Finally, we show that our results better fit the new model of the evolution of the GABI, although the alternative explanation invoking the traditional model is also reasonable. In general, our findings imply that most of the evolution of extant procyonids species may have occurred in South and Central America, instead of North America, previous to the episodes of the Great American Biotic

Interchange under the traditional model. This implies a more complex evolutionary history than previously acknowledge of procyonids in general, and *Nasua* species in particular.

Tables and Figures

Table 3.1. Mitochondrial genetic diversity. Haplogroups as identified in the AMOVA, sample sizes (N), number of haplotypes, haplotype and nucleotide diversity indices per sampling locality and haplogroup.

Population groups (haplogroups)	Population/locality	N	# haps	<i>h</i>	π
Southwestern U.S. (SWUS)		16	3	0.4250 ± 0.1326	0.0011 ± 0.0007
	New Mexico (NM)	2	2	1.0000 ± 0.5000	0.000460 ± 0.000650
	Arizona (AZ)	14	2	0.3626 ± 0.1302	0.001179 ± 0.000754
Western Mexico (WMEX)	Jalisco, Mexico (JAL)	11	2	0.5455 ± 0.0722	0.000249 ± 0.000249
Central Mexico (CMEX)	Morelos, Mexico (MOR)	8	2	0.2500 ± 0.1802	0.000229 ± 0.000247
Yucatan Peninsula and Guatemala (YUCP-GUAT)		37	11	0.6817 ± 0.0759	0.005024 ± 0.002596
	Yucatan, Mexico (YUC)	5	4	0.9000 ± 0.1610	0.000639 ± 0.000543
	Cozumel, Mexico (COZ)	9	5	0.7222 ± 0.1592	0.001875 ± 0.001168
	Belize (BLZ)	2	2	1.0000 ± 0.5000	0.000456 ± 0.000645
	Guatemala (GUA)	20	1	0.0000 ± 0.0000	0.0000 ± 0.0000
	Costa Rica (CRA)	1	1	1.0000 ± 0.0000	0.0000 ± 0.0000
Panama (PAN)		13	3	0.2949 ± 0.1558	0.000353 ± 0.000309

Table 3.2. Pairwise F_{ST} values for the five white-nosed coati populations identified in the AMOVA and Bayesian clustering analysis. Below the diagonal are the F_{ST} values calculated from the concatenated mtDNA sequences and above the diagonal are the values obtained from the analysis of 11 microsatellite loci. All F_{ST} p-values are significant (< 0.0001).

	SWUS	WMEX	CMEX	YUCP-GUAT	PAN
SWUS	-	0.15560	0.20061	0.26192	0.32829
WMEX	0.91038	-	0.10573	0.16536	0.28029
CMEX	0.97841	0.99403	-	0.10155	0.19840
YUCP-GUAT	0.92574	0.92276	0.91909	-	0.23303
PAN	0.99616	0.99847	0.99847	0.97913	-

Table 3.3. Average genetic diversity for all five white-nosed coati populations based on analysis of 11 microsatellite loci. N= number of individuals sampled in each population, Ho= Observed heterozygosity, He= Expected heterozygosity. Asterisks indicate the level of significance: (*) P < 0.05, (**) P<0.01.

Population	N	Average # of alleles	Ho	He	Fis	Allelic range
SWUS	21	4.091	0.552	0.579	0.03619	13.636
WMEX	11	5.364	0.621	0.732	0.12664*	18.636
CMEX	8	5.636	0.682	0.776	0.12797**	18.182
YUCP-GUAT	37	7.000	0.639	0.674	0.05085	18.909
PAN	13	3.636	0.456	0.492	0.02234	12.636

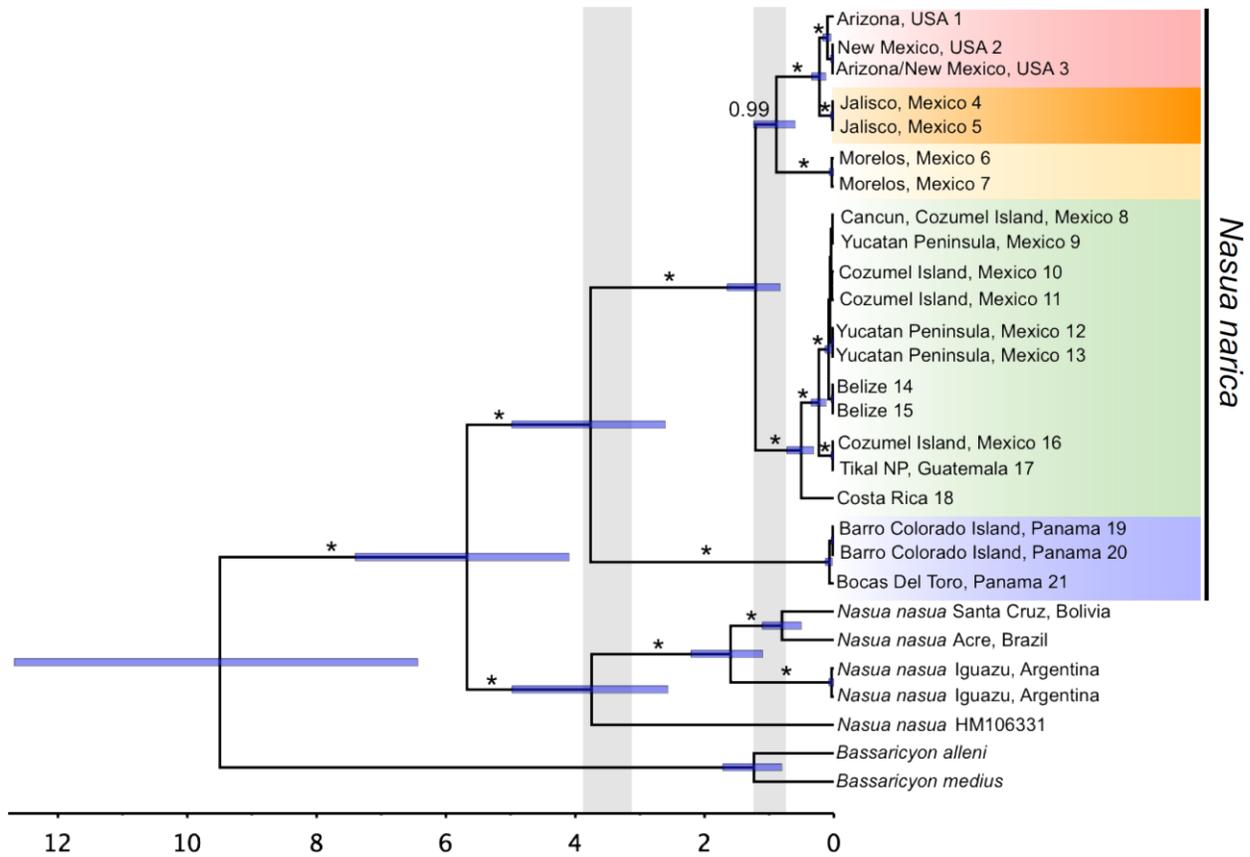


Figure 3.2. Divergence times among *Nasua narica* lineages estimated from the concatenated sequences of three mitochondrial loci. Asterisks and numbers shown at the nodes indicate posterior probabilities (PP), with asterisk representing PP=1.0. Node bars denote 95% highest posterior density for divergence times. The tree was rooted using the olingos *Bassaricyon medius* and *Bassaricyon alleni*. The timescale is in million years ago (Mya). The colors represent the main lineages found in the Maximum Likelihood phylogenetic analysis (see Figure 3.1B).

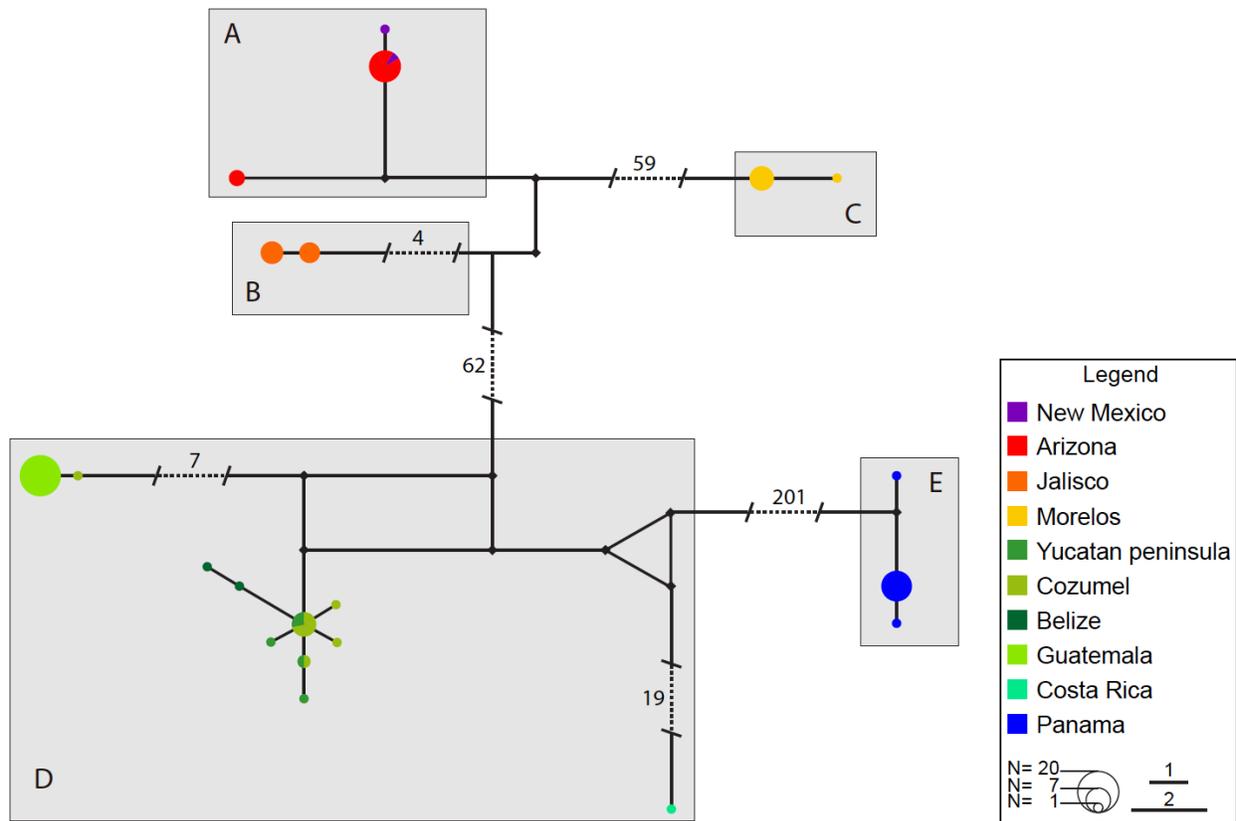


Figure 3.3. Median-joining network showing the phylogeographic structure of the 21 haplotypes of white-nosed coati. The haplotype network is based on 2,201 bp of concatenated mitochondrial sequence data from three genes (*CYTB*, *NADH5*, *16S*) in 85 individuals. Concordant with the AMOVA results, five haplogroups were identified, indicated by the gray rectangles A – E. The color code indicates the geographic localities from which the samples were collected (see legend). Small black squares represent median vectors corresponding to homoplasies or missing haplotypes. The size of the circles indicates the number of individuals having that particular haplotype and the length of the lines represents the number of nucleotide mutations separating the haplotypes (See legend). Dotted lines represent larger numbers of nucleotide mutations indicated by the number above or by the sides of these lines.

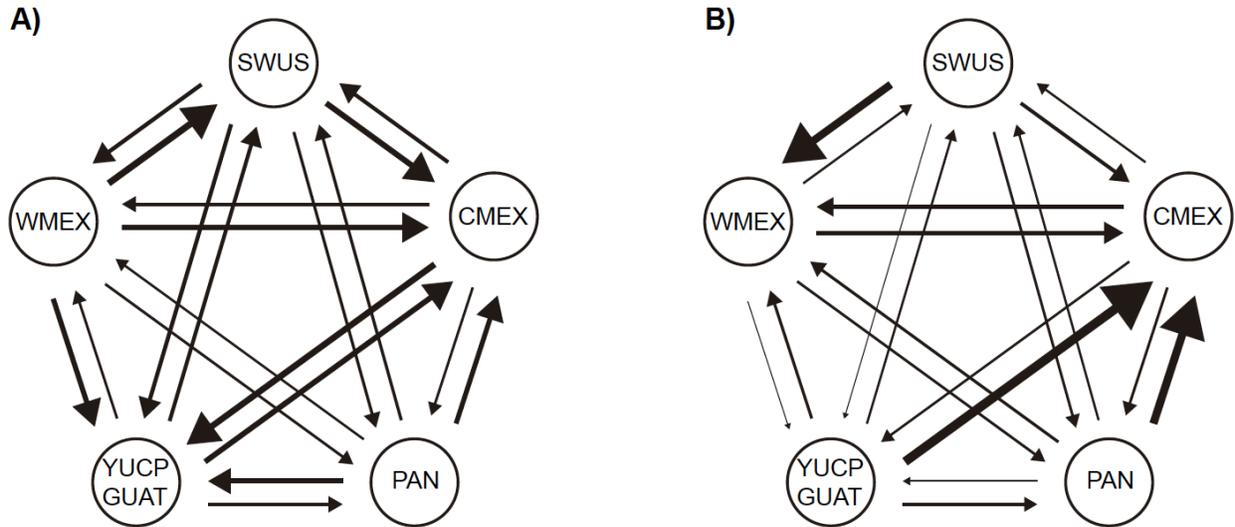


Figure 3.4. Schematic summary of gene flow among white-nosed coati populations. A) Ancient migration estimates based on concatenated mitochondrial sequences using MIGRATE-N. B) Recent migration estimates based on microsatellites loci using BAYESASS. The arrows represent the direction of the migration and the thickness is proportional to the numbers of migrants in A) and the migration rate per generation in B).

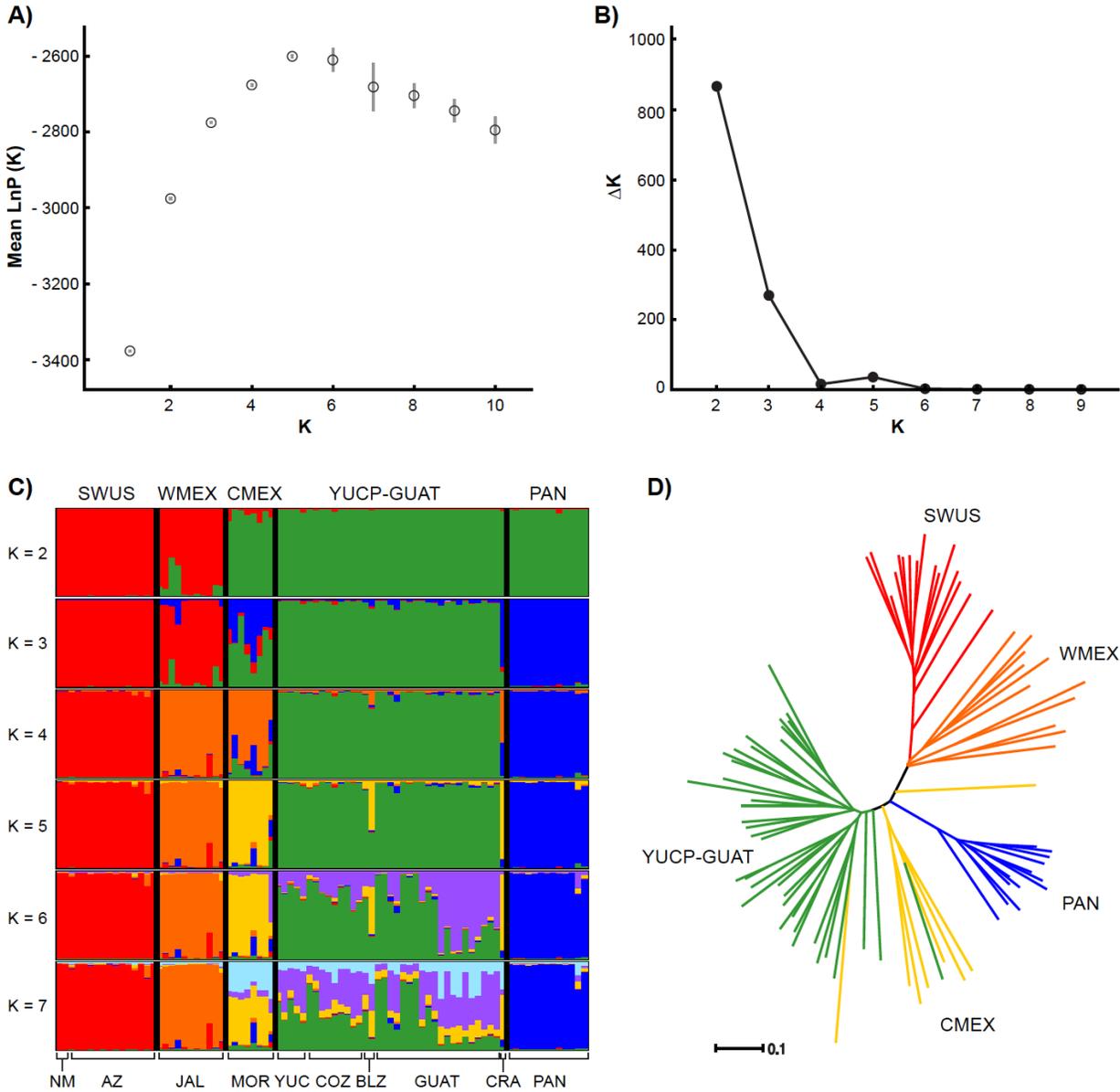


Figure 3.5. Genetic subdivision structure among white-nosed coati populations based on the analysis of 11 microsatellites in 85 individuals. Plots showing A) the log likelihood probability and B) delta K values for K=1 to K=10 based on the STRUCTURE analysis. In B) the standard deviation of the mean likelihood value for each K is shown as a gray vertical line. C) Assignment bar plots of the STRUCTURE analysis for K=2 to K=7. The five most probable genetic clusters (K= 5) are delimited in all assignment plots. The name of the five clusters are shown at the top of the figure, whereas the sampling locations of the individuals within the clusters are indicated at the bottom of the figure. D) Neighbor-joining tree based on DA distance of microsatellite loci. Colors in D) represent the five clusters identified in the STRUCTURE analysis. SWUS: Southwestern U.S., WMEX: Western Mexico, CMEX: Central Mexico, YUCP-GUAT: Yucatan Peninsula - Guatemala, PAN: Panama, NM: New Mexico, AZ: Arizona, JAL: Jalisco, MOR: Morelos, YUC: Yucatan, COZ: Cozumel, BLZ: Belize, GUAT: Guatemala, CRA: Costa Rica.

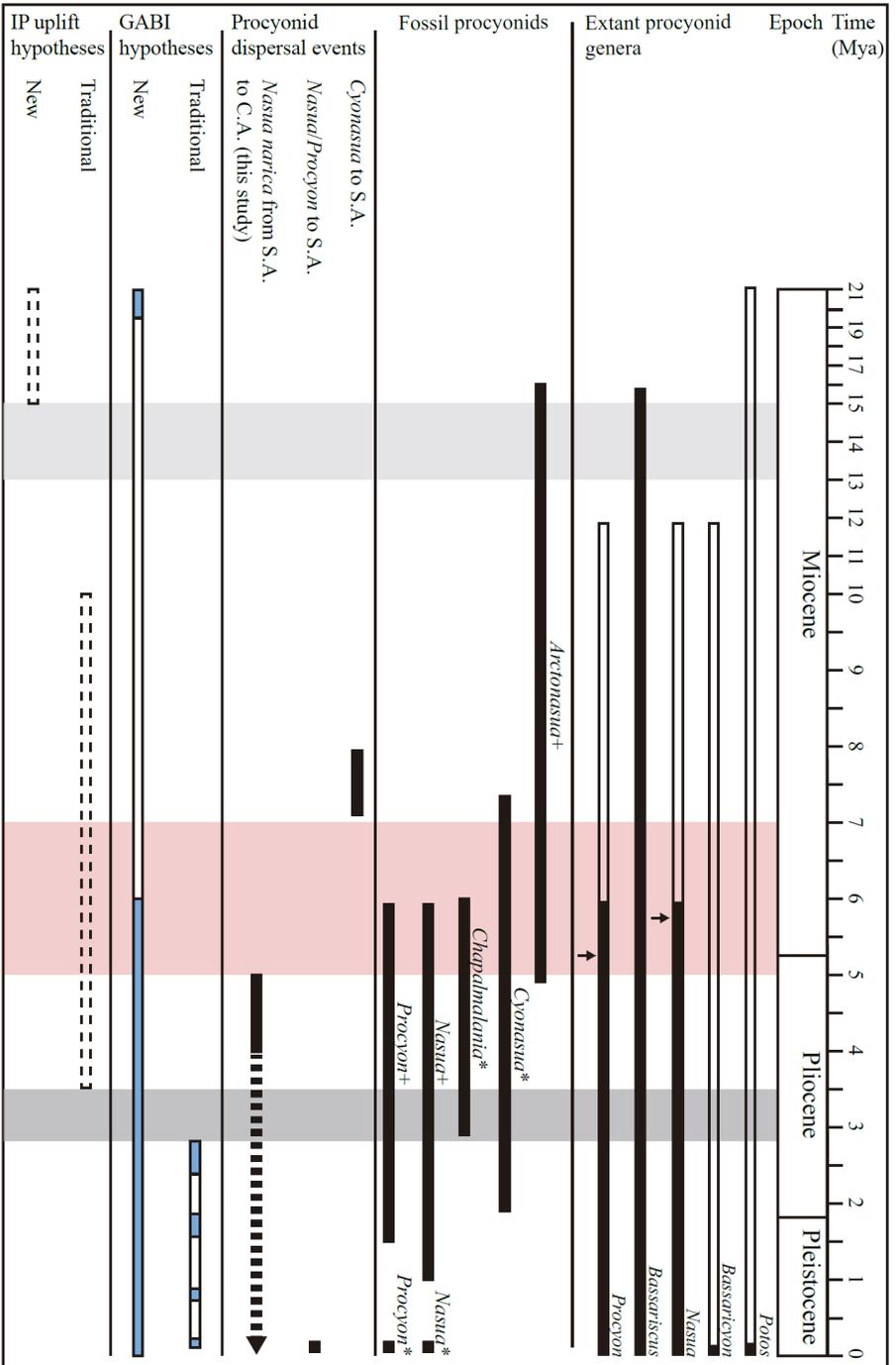


Figure 3.6. Temporal range for extinct and extant procyonid genera, procyonid dispersal events, GABI hypotheses and Isthmus of Panama uplift hypotheses. For extant procyonid genera black lines indicate temporal range based on the fossil record (Baskin 1982, 1998, 2003, 2004, Gilmore 2013, Forstner *et al.* 2014), whereas appended white lines indicate temporal range estimations based on molecular data (Koepfli *et al.* 2007; Eizirik *et al.* 2010; Eizirik 2012). This part of the figure is based on Figure 6 of Koepfli *et al.* 2007. The arrows below *Nasua* and *Procyon* signal the mean divergence time between different species of those genera based on our data and Koepfli *et al.* 2007, respectively. For procyonid fossils taxa with a plus sign (+) and an asterisk (*) represent remains found in North and South America, respectively. *Nasua* and *Procyon* have remains in both of this locations. The time range for the different proposed dispersal event of procyonid is marked by black lines. The dashed arrow after the dispersion event of *Nasua narica* indicate the time range of the colonization of Central and North America by this species proposed in this study. For the GABI hypotheses the line segments colored in light blue represent episodes of high migration, whereas the white segments are relative pauses in migration (Woodburne 2010; Bacon *et al.* 2015). For the hypotheses on the uplift of the Isthmus of Panama the dashed rectangles represent the range in which the Central America Seaway was shallow. The vertical light and dark gray rectangles represent the time estimates in which the final uplift of the Isthmus occurred under the new and traditional models respectively (Coates and Obando 1996; Montes *et al.* 2012b; Montes *et al.* 2015; O'Dea *et al.* 2016). The vertical red rectangle indicates the time frame that we proposed the diversification of procyonids may have occurred in South America.

Supplementary Material

Table 3.S1. Primer sequences and repeat motifs of the 11 microsatellite loci used in this study. All the loci were previously identified in the brown-nosed coati (*Nasua nasua*; Almany *et al.* 2009).

Locus	Primer sequence	Repeat motif
NanaSTR-A02	F: CACGACGTTGTAAAACGAC AGG AAC GCT CAA ACC AAA GA R: TGT GAT GCA GCA GCC TAA TC	(GT)17
NanaSTR-A09	F: CACGACGTTGTAAAACGAC TCA AGA TCC TCT GCA ACT TGT G R: ACC CTT AAG TCT TGA GTG GAA GAA	(TG)18
NanaSTR-B09	F: CACGACGTTGTAAAACGAC GCT TTT GCT GGC CAT AGT TT R: TCA CTA ATT ACA ACT AAA AAC CCT GA	(TG)19
NanaSTR-C09	F: CACGACGTTGTAAAACGAC GCT GCG TCC TCT GCT TAG AA R: TGA TCC ATG TAC TGG TGG AGA A	(TG)20
NanaSTR-D03	F: CACGACGTTGTAAAACGAC AGG CTT GAA TTT GTC CAG CTA R: CCA AGA ATC CTG TGG CAA A	(CA)14
NanaSTR-D07	F: CACGACGTTGTAAAACGAC TGT TAT CTC TGC TTC TTC GGT CT R: TGG TCA CAA GTT TCT CAA ATG C	(TG)18
NanaSTR-E04	F: CACGACGTTGTAAAACGAC CCT GCA GCT TTA CTG AAT TTG A R: TTT GCA AAT GAC ATG ATA TGA TAG G	(TG)22
NanaSTR-E05	F: CACGACGTTGTAAAACGAC CCC AAT CCT GAT AGC CCT TC R: TAT TTT TGT TGG GCC CGA GT	(CA)18
NanaSTR-F02	F: CACGACGTTGTAAAACGAC CAT TTG AGT GAA AAT CCA GTG A R: GCT CTT GAT AAA GCA AGC ACA A	(TG)15
NanaSTR-G03	F: CACGACGTTGTAAAACGAC AAT CCG TGG CTG AAA CAT TC R: GCA AAA TGC AGG GTA AGG TG	(GT)20
NanaSTR-H03	F: CACGACGTTGTAAAACGAC GCC CCT GAG CCA ATT CTT R: TTC TCC TGT ATT AGG GTT CTC CA	(TC)17 (AC)12

Table 3.S2. AMOVA results of the concatenated mtDNA sequence data based on different grouping schemes. We grouped sampling localities to maximize the among-group variance component (Φ_{ct}). The grouping scheme with the highest among group variance value was the one with five groups that included the Cozumel site within the YUCP-GUAT group. The populations or sites are designated as: SWUS (Arizona-Nuevo Mexico), WMEX (Jalisco), CMEX (Morelos), YUC (Yucatan), COZ (Cozumel), BLZ (Belize), GUA (Guatemala), CRA (Costa Rica), PAN (Panama), YUCP-GUAT (Yucatan-Guatemala, including Cozumel, Costa Rica and Belize). Asterisks indicate the level of significance: (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$, (****) $P < 0.0001$.

Grouping scheme	df	SS	Variance component	% of variance component
[SWUS+WMEX] [CMEX] [YUCP-GUAT] [PAN]				
Among Groups [Φ_{ct}]	3	6375.156	106.21689	92.87****
Among Populations [Φ_{sc}]	6	292.562	7.54196	6.59****
Within Populations [Φ_{st}]	75	46.287	0.61717	0.54****
[SWUS] [WMEX] [CMEX] [YUCP-GUAT] [PAN]				
Among Groups [Φ_{ct}]	4	6487.563	100.15544	93.25****
Among Populations [Φ_{sc}]	5	180.155	6.63449	6.18****
Within Populations [Φ_{st}]	75	46.287	0.61717	0.57****
[SWUS+WMEX] [CMEX] [COZ] [YUC+BLZ+GUA+CRA] [PAN]				
Among Groups [Φ_{ct}]	4	6418.307	94.77806	90.90**
Among Populations [Φ_{sc}]	5	249.411	8.87532	8.51****
Within Populations [Φ_{st}]	75	46.287	0.61717	0.59****
[SWUS] [WMEX] [CMEX] [YUC+COZ+BLZ] [GUA+CRA] [PAN]				
Among Groups [Φ_{ct}]	5	6611.017	91.80321	95.48**
Among Populations [Φ_{sc}]	4	56.702	3.73255	3.88**
Within Populations [Φ_{st}]	75	46.287	0.61717	0.64****
[SWUS] [WMEX] [CMEX] [COZ] [YUC+BLZ+GUA+CRA] [PAN]				
Among Groups [Φ_{ct}]	5	6530.714	89.55586	90.91*
Among Populations [Φ_{sc}]	4	137.004	8.33405	8.46****
Within Populations [Φ_{st}]	75	46.287	0.61717	0.63****

Table 3.S3. Number of migrants per generation estimated with MIGRATE-N based on the concatenated mitochondrial sequences.

	SWUS	WMEX	CMEX	YUC-GUA	PAN
SWUS	-	0.239522	0.138557	0.137668	0.108077
WMEX	0.145873	-	0.098441	0.091875	0.084231
CMEX	0.207156	0.189588	-	0.199714	0.176656
YUCP-GUAT	0.145803	0.186613	0.226681	-	0.161385
PAN	0.104834	0.095188	0.089888	0.101654	-

Table 3.S4. Summary statistics for the 11 microsatellite loci used in this study. Average for allelic diversity, observed heterozygosity (Ho) and expected heterozygosity (He) among all 85 white-nosed coati individuals.

LOCUS	Allelic diversity	Ho	He
A02	10	0.468	0.539
A09	10	0.698	0.716
B09	9	0.382	0.497
C09	11	0.791	0.749
D03	10	0.595	0.559
D07	13	0.631	0.687
E04	7	0.593	0.588
E05	3	0.335	0.426
F02	9	0.629	0.669
G03	12	0.656	0.723
H03	17	0.713	0.715

Table 3.S5. Observed and expected heterozygosity, and deviations from Hardy-Weinberg equilibrium for 11 microsatellite loci in five white-nosed coati populations. Minus sign (-) indicates significant heterozygote deficiency, $p < 0.05$, Bonferroni multiple tests correction.

Locus	SWUS			WMEX			CMEX			YUCP-GUAT			PAN		
	Ho	He	HW	Locus	Ho	He	HW	Locus	Ho	He	HW	Locus	Ho	He	HW
A02	0.250	0.365	-	A02	0.273	0.533	-	A02	0.750	0.734		A02	0.838	0.732	
A09	0.667	0.682		A09	0.636	0.851	-	A09	0.875	0.758		A09	0.541	0.637	-
B09	0.067	0.064		B09	0.400	0.645	-	B09	0.625	0.672		B09	0.568	0.605	
C09	0.750	0.721		C09	0.727	0.711		C09	0.875	0.766		C09	0.757	0.807	
D03	0.400	0.480		D03	0.800	0.705		D03	0.750	0.750		D03	0.486	0.469	
D07	0.688	0.652		D07	0.909	0.814		D07	0.625	0.813	-	D07	0.703	0.846	-
E04	0.438	0.432		E04	0.636	0.649		E04	0.625	0.664		E04	0.730	0.757	
E05	0.563	0.537		E05	0.273	0.483		E05	0.250	0.531	-	E05	0.514	0.506	
F02	0.688	0.742		F02	0.727	0.698		F02	0.750	0.797		F02	0.595	0.663	
G03	0.813	0.785		G03	0.636	0.731		G03	0.750	0.797		G03	0.541	0.545	
H03	0.750	0.705		H03	0.818	0.860		H03	0.625	0.719		H03	0.757	0.748	

Table 3.S6. GENECLASS2 results. The highest number of correctly assigned individuals to their population of origin was obtained for the grouping scheme consisting of five populations which is concordant with the STRUCTURE analysis.

Grouping scheme	No. populations	% correctly assigned individuals (No. of individuals)	Quality Index
[SWUS+WMEX][CMEX][YUCP-GUAT][PAN]	4	85.9 (73)	76.06%
[SWUS] [WMEX][CMEX][YUCP-GUAT][PAN]	5	89.4 (76)	76.64%
[SWUS+WMEX][CMEX][COZ][YUC+BLZ+GUA+CRA] [PAN]	5	80 (68)	65.11%
[SWUS] [WMEX][CMEX][YUC+COZ+BLZ] [GUA+CRA] [PAN]	6	77.6 (66)	65.12%
[SWUS][WMEX][CMEX][COZ][YUC+BLZ+GUA+CRA] [PAN]	6	80 (68)	65.20%

Table 3.S7. AMOVA results for the microsatellite data based on different grouping schemes. The grouping scheme with the highest among group variance value was the one identified with STRUCTURE of five groups comprised by [SWUS] [WMEX] [CMEX] [YUCP-GUAT] [PAN]. Asterisks indicate the level of significance: (*) P < 0.05, (**) P<0.01, (***) P<0.001, (****) P<0.0001. SWUS (Arizona-Nuevo Mexico), WMEX (Jalisco), CMEX (Morelos), YUC (Yucatan), COZ (Cozumel), BLZ (Belize), GUA (Guatemala), CRA (Costa Rica), PAN (Panama), YUCP-GUAT (Yucatan-Guatemala, including Cozumel, Costa Rica and Belize).

Grouping scheme	df	SS	Variance component	% of variance
[SWUS+WMEX][CMEX][YUCP-GUAT][PAN]				
Among Groups [Φ_{ct}]	3	103.573	0.80364	19.25****
Among Populations [Φ_{sc}]	81	297.833	0.30612	7.33****
Within Populations [Φ_{st}]	85	260.500	3.06471	73.42****
[SWUS] [WMEX][CMEX][YUCP-GUAT][PAN]				
Among Groups [Φ_{ct}]	4	122.708	0.88121	21.21****
Among Populations [Φ_{sc}]	80	278.698	0.20951	5.04**
Within Populations [Φ_{st}]	85	260.500	3.06471	73.75****
[SWUS+WMEX][CMEX][COZ][YUC+BLZ+GUA+CRA] [PAN]				
Among Groups [Φ_{ct}]	4	106.989	0.72645	17.72****
Among Populations [Φ_{sc}]	80	294.417	0.30775	7.51****
Within Populations [Φ_{st}]	85	260.500	3.06471	74.77****
[SWUS] [WMEX][CMEX][YUC+COZ+BLZ] [GUA+CRA] [PAN]				
Among Groups [Φ_{ct}]	5	127.336	0.78989	19.47****
Among Populations [Φ_{sc}]	79	274.070	0.20227	4.99**
Within Populations [Φ_{st}]	85	260.500	3.06471	75.54****
[SWUS][WMEX][CMEX][COZ][YUC+BLZ+GUA+CRA] [PAN]				
Among Groups [Φ_{ct}]	5	126.124	0.80344	19.70****
Among Populations [Φ_{sc}]	79	275.282	0.20994	5.15**
Within Populations [Φ_{st}]	85	260.500	3.06471	75.15****

Table 3.S8. Inferred migration rates among white-nosed coati populations based on the microsatellite data. Means of the posterior distributions of the migration rate into each population are shown. The migration rate values should be interpreted as the fraction of individuals in populations listed in the rows that are migrants derived from the populations in the columns. Values along the diagonal are the proportions of individuals derived from the source populations each generation.

	AZ-NM	JAL	MOR	YUC-GUA	PAN
SWUS	0.9364	0.0159 (0.0153)	0.0159 (0.0151)	0.0159 (0.0152)	0.0159 (0.0151)
WMEX	0.0564 (0.0357)	0.8776 (0.0428)	0.0242 (0.0224)	0.0208 (0.0196)	0.0210 (0.0197)
CMEX	0.0258 (0.0238)	0.0277 (0.0255)	0.8206 (0.0551)	0.0671 (0.0466)	0.0588 (0.0410)
YUCP-GUAT	0.0079 (0.0078)	0.0081 (0.0078)	0.0173 (0.0128)	0.9561 (0.0179)	0.0106 (0.0099)
PAN	0.0185 (0.0176)	0.0186 (0.0175)	0.0188 (0.0177)	0.0186 (0.0176)	0.9256 (0.0320)

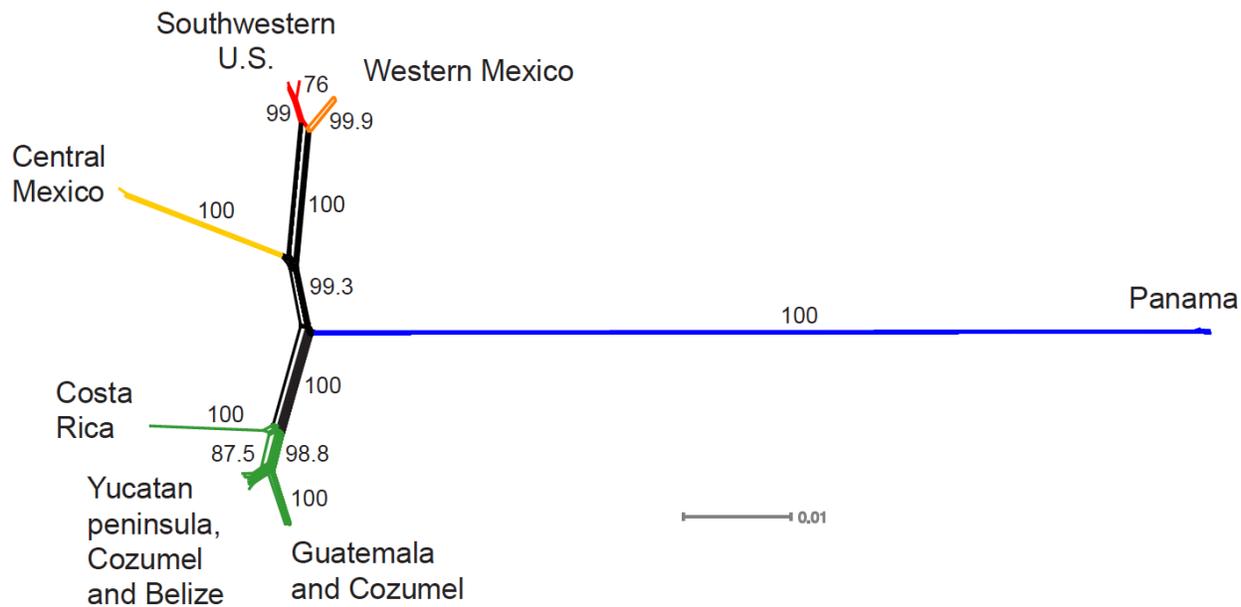


Figure 3.S1. Haplotype Neighbor-Net network based on the patristic distance between 21 haplotypes of the concatenated mitochondrial sequences. This network is consistent with the phylogenetic trees and the Median-Joining haplotype network supporting the five haplogroups found in our data with high bootstrap values (1000 replicates). The five haplogroups are color coded according to the phylogenetic trees and the names of the general localities from which the samples were collected are shown.

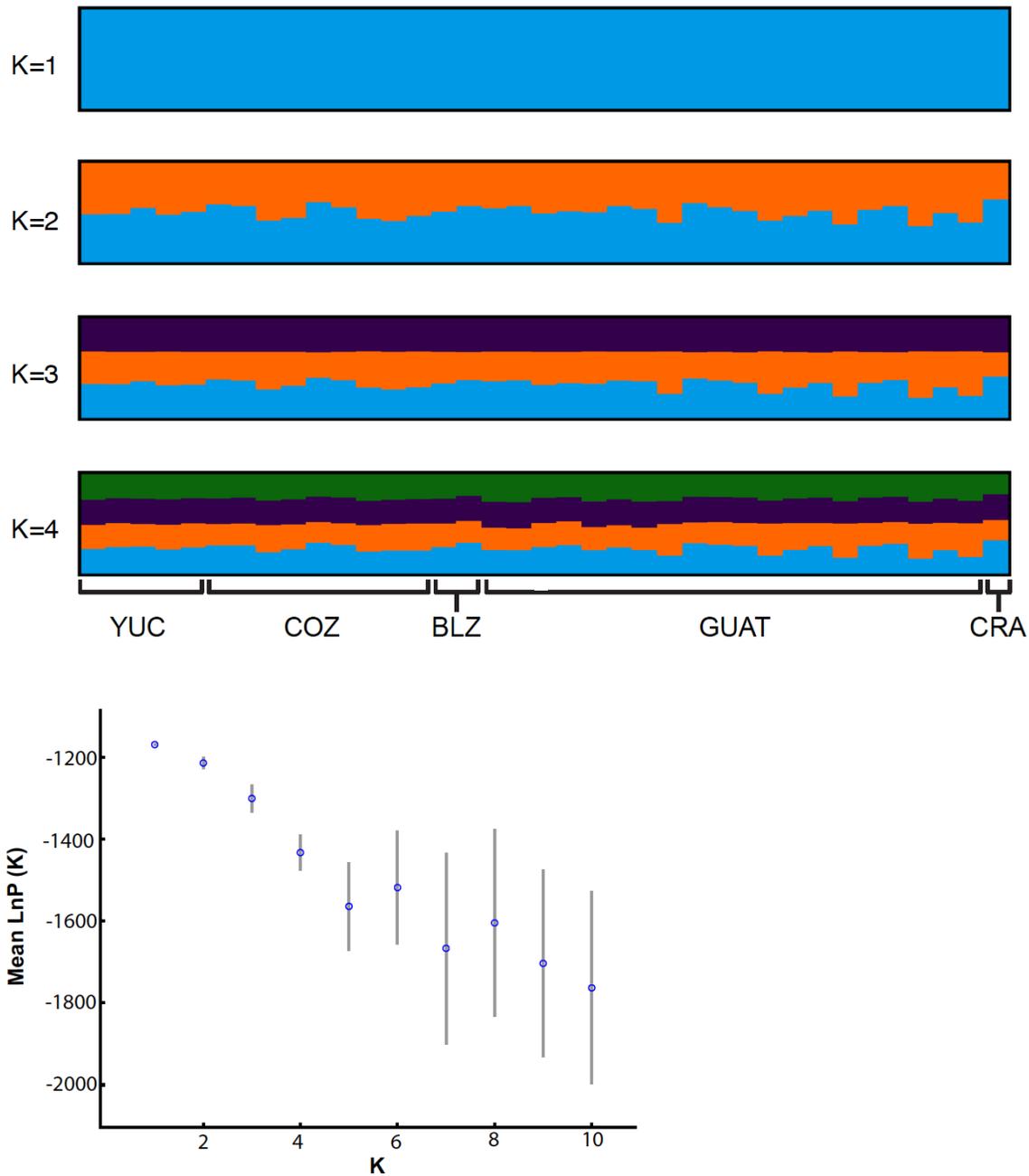


Figure 3.S2. Structure analysis results for the YUCP-GUAT cluster. The assignment bar plots show that all individuals within this cluster are assigned to only one cluster ($K=1$) and further substructure was not found at higher K values ($K=2 - K=4$). Accordingly, the plot of the likelihood probability for $K=1$ to $K=10$ shows the higher probability of the data is at $K=1$.

Appendix 3.1. Species names, sample identification numbers, source locality, collector information, and Genbank accession numbers for the samples used in this study. MSB = Museum of Southwestren Biology, University of New Mexico; MVZ = Museum of Vertebrate Zoology, University of Californina, Berkeley; USNM = National Museum of Natural History, Washington, D.C.

Species	Sample ID	Locality	Source
<i>Nasua narica</i>	AZ_14	Arizona, USA	Katherine McFadden
<i>Nasua narica</i>	AZ_U1	HWY 82, Patagonia, Santa Cruz Co., Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_U3	Parker Canyon Lake, Santa Cruz Co., Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_U8	Guindani Canyon, Whetstones, Pima Co., Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_C257	Chiricahua Mountains, Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_C151	Chiricahua Mountains, Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_C206	Chiricahua Mountains, Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_C213	Chiricahua Mountains, Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_C225	Chiricahua Mountains, Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_C234	Chiricahua Mountains, Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_606	Chiricahua Mountains, Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_611	Chiricahua Mountains, Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_614	Chiricahua Mountains, Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_628	Chiricahua Mountains, Arizona, USA	Christine Hass
<i>Nasua narica</i>	AZ_C257	Chiricahua Mountains, Arizona, USA	Christine Hass
<i>Nasua narica</i>	NM_140072	Big Buuro Mountain, Grant Co., New Mexico, USA	MSB
<i>Nasua narica</i>	NM_283633	Iron Knot Ranch, Grant Co., New Mexico, USA	MSB
<i>Nasua narica</i>	JAL_9	Jalisco, Mexico	David Valenzuela
<i>Nasua narica</i>	JAL_14	Jalisco, Mexico	David Valenzuela
<i>Nasua narica</i>	JAL_25	Jalisco, Mexico	David Valenzuela
<i>Nasua narica</i>	JAL_27	Jalisco, Mexico	David Valenzuela
<i>Nasua narica</i>	JAL_38	Jalisco, Mexico	David Valenzuela
<i>Nasua narica</i>	JAL_43	Jalisco, Mexico	David Valenzuela
<i>Nasua narica</i>	JAL_A2	Jalisco, Mexico	David Valenzuela
<i>Nasua narica</i>	JAL_B2	Jalisco, Mexico	David Valenzuela
<i>Nasua narica</i>	JAL_C2	Jalisco, Mexico	David Valenzuela
<i>Nasua narica</i>	JAL_D2	Jalisco, Mexico	David Valenzuela
<i>Nasua narica</i>	JAL_E2	Jalisco, Mexico	David Valenzuela
<i>Nasua narica</i>	MOR_A1	Morelos, Mexico	David Valenzuela
<i>Nasua narica</i>	MOR_C1	Morelos, Mexico	David Valenzuela
<i>Nasua narica</i>	MOR_D1	Morelos, Mexico	David Valenzuela
<i>Nasua narica</i>	MOR_E1	Morelos, Mexico	David Valenzuela
<i>Nasua narica</i>	MOR_F2	Morelos, Mexico	David Valenzuela
<i>Nasua narica</i>	MOR_G2	Morelos, Mexico	David Valenzuela
<i>Nasua narica</i>	MOR_H2	Morelos, Mexico	David Valenzuela
<i>Nasua narica</i>	MOR_I2	Morelos, Mexico	David Valenzuela
<i>Nasua narica</i>	MOR_J2	Morelos, Mexico	David Valenzuela
<i>Nasua narica</i>	COZ_NAM02	Cozumel Island, Mexico	Katherine McFadden
<i>Nasua narica</i>	COZ_NAF03	Cozumel Island, Mexico	Katherine McFadden
<i>Nasua narica</i>	COZ_NEF04	Cozumel Island, Mexico	Katherine McFadden
<i>Nasua narica</i>	COZ_NAF07	Cozumel Island, Mexico	Katherine McFadden
<i>Nasua narica</i>	COZ_NAM09	Cozumel Island, Mexico	Katherine McFadden
<i>Nasua narica</i>	COZ_NAF10	Cozumel Island, Mexico	Katherine McFadden
<i>Nasua narica</i>	COZI_NAF36	Cozumel Island, Mexico	Katherine McFadden
<i>Nasua narica</i>	COZI_NAF37	Cozumel Island, Mexico	Katherine McFadden
<i>Nasua narica</i>	COZI_NEM61	Cozumel Island, Mexico	Katherine McFadden
<i>Nasua narica</i>	YUC_NAM05	Yucatan, Peninsula, Mexico	Katherine McFadden
<i>Nasua narica</i>	YUC_NAF08	Yucatan, Peninsula, Mexico	Katherine McFadden

<i>Nasua narica</i>	YUC_NEF20	Yucatan, Peninsula, Mexico	Katherine McFadden
<i>Nasua narica</i>	YUC_NAM092	Yucatan, Peninsula, Mexico	Katherine McFadden
<i>Nasua narica</i>	CANC_95	Cancun, Mexico	Katherine McFadden
<i>Nasua narica</i>	BZ_1	Belize	Katherine McFadden
<i>Nasua narica</i>	BZ_2	Belize	Katherine McFadden
<i>Nasua narica</i>	GUAT_003	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_017	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_019	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_020	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_021	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_024	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_028	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_032	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_034	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_037	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_043	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_047	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_055	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_072	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_073	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_079	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_080	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_083	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_089	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	GUAT_091	Tikal National Park, Petén, Guatemala	Susan Booth-Binczik, Gerry Booth-Binczik
<i>Nasua narica</i>	CR_JWD9	Costa Rica	Jerry Drago
<i>Nasua narica</i>	CR_M3010	Costa Rica	Matthew Gompper
<i>Nasua narica</i>	PAN_449556	Bocas Del Toro, Panama	USNM
<i>Nasua narica</i>	PAN_FK10	Barro Colorado Island, Panama	Matthew Gompper
<i>Nasua narica</i>	PAN_FK14	Barro Colorado Island, Panama	Matthew Gompper
<i>Nasua narica</i>	PAN_FK16	Barro Colorado Island, Panama	Matthew Gompper
<i>Nasua narica</i>	PAN_SAFK18	Barro Colorado Island, Panama	Matthew Gompper
<i>Nasua narica</i>	PAN_SM8	Barro Colorado Island, Panama	Matthew Gompper
<i>Nasua narica</i>	PAN_SM11	Barro Colorado Island, Panama	Matthew Gompper
<i>Nasua narica</i>	PAN_38	Gigante, Panama	Matthew Gompper
<i>Nasua narica</i>	PAN_39	Gigante, Panama	Matthew Gompper
<i>Nasua narica</i>	PAN_40	Gigante, Panama	Matthew Gompper
<i>Nasua narica</i>	PAN_42	Gigante, Panama	Matthew Gompper
<i>Nasua narica</i>	PAN_60	Bohio Peninsula, Panama	Matthew Gompper
<i>Nasua narica</i>	PAN_Limbo	Limbo, Panama	Matthew Gompper
<i>Nasua nasua</i>	BOL_56101	Santa Cruz, Bolivia	MSB
<i>Nasua nasua</i>	BRZ_195089	Acre, Brazil	MVZ
<i>Nasua nasua</i>	ARG_30B	Iguazu, Argentina	Ben Hirsch

<i>Nasua nasua</i>	ARG_71B	Iguazu, Argentina	Ben Hirsch
<i>Nasua nasua</i>	ARG_73B	Iguazu, Argentina	Ben Hirsch
<i>Nasua nasua</i>	ARG_82B	Iguazu, Argentina	Ben Hirsch
<i>Nasua nasua</i>	ARG_93B	Iguazu, Argentina	Ben Hirsch
<i>Nasua nasua</i>	ARG_105A	Iguazu, Argentina	Ben Hirsch
<i>Nasua nasua</i>	ARG_116A	Iguazu, Argentina	Ben Hirsch
<i>Nasua nasua</i>	ARG_126B	Iguazu, Argentina	Ben Hirsch
<i>Bassaricyon alleni</i>	PER_155219	Amazonas, Peru	MVZ
<i>Bassaricyon medius</i>	PAN_BOG	Limbo, Panama	Roland Kays
<i>Bassaricyon medius</i>	PAN_EDG	Limbo, Panama	Roland Kays

References

- Abratis M, Wörner G. 2001. Ridge Collision, Slab-Window Formation, and the Flux of Pacific Asthenosphere into the Caribbean Realm. *Geology* 29:127–130.
- Almany GR, De Arruda MP, Arthofer W *et al.* 2009. Permanent genetic resources added to molecular ecology resources database 1 May 2009–31 July 2009. *Molecular Ecology Resources* 9:1460-1466.
- Almendra AL, Rogers DS. 2012. Biogeography of Central American Mammals, pp. 203-229. In: Patterson BD, Costa LP (eds.). *Bones, Clones and Biomes. The History and Geography of Recent Neotropical Mammals*. University of Chicago Press, Chicago, IL, USA.
- Arbeláez-Cortés E, Nyári AS, Navarro-Sigüenza AG. 2010. The differential effect of lowlands on the phylogeographic pattern of a Mesoamerican montane species (*Lepidocolaptes affinis*, Aves: Furnariidae). *Molecular Phylogenetics and Evolution* 57: 658–668.
- Arellano EF, González-Cozátl X, Rogers DS. 2005. Molecular systematics of middle American harvest mice *Reithrodontomys* (Muridae), estimated from mitochondrial cytochrome b gene sequences. *Molecular Phylogenetics and Evolution* 37:529–540.
- Avise JC, Walker D. 1999. Species realities and numbers in sexual vertebrates: perspectives from an asexually transmitted genome. *Proceedings of the National Academy of Sciences of the United States of America* 96:992-995.
- Bacon CD, Silvestro D, Jaramillo C, Smith BT, Chakrabarty P, Antonelli A. 2015. Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proceedings of the National Academy of Sciences of the United States of America* 112:6110-6115.
- Bacon CD, Molnar P, Antonelli A, Crawford AJ, Montes C, Vallejo-Pareja MC. 2016. Quaternary glaciation and the Great American Biotic Interchange. *Geology* 44:375-378.
- Bagley JC, Johnson JB. 2014. Phylogeography and biogeography of the lower Central American Neotropics: diversification between two continents and between two seas. *Biological Reviews* 89:767-790.
- Baker RJ, Bradley RD. 2006. Speciation in mammals and the genetic species concept. *Journal of Mammalogy* 87: 643–662.
- Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16:37-48.
- Barber BR, Klicka J. 2010. Two pulses of diversification across the isthmus of Tehuantepec in a montane Mexican bird fauna. *Proceedings of the Royal Society B* 277:2675–2681.

- Barendregt RW, Duk-Rodkin A. 2011. Chronology and extent of Late Cenozoic ice sheets in North America: a magnetostratigraphical assessment. *Developments in Quaternary Science* 15:419-426.
- Barrier E, Velasquillo L, Chavez M, Gaulon R. 1998. Neotectonic evolution of the Isthmus of Tehuantepec (southeastern Mexico). *Tectonophysics* 287:77–96.
- Baskin JA. 1982. Tertiary Procyoninae (Mammalia: Carnivora) of North America. *Journal of Vertebrate Paleontology* 2:71-93.
- Baskin JA. 1989. Comments on new world tertiary procyonidae (Mammalia: Carnivora). *Journal of Vertebrate Paleontology* 9:110-117.
- Baskin JA. 2003. New Procyonines from the Hemingfordian and Barstovian of the Gulf Coast and Nevada, Including the First Fossil Record of the Potosini. *Bulletin of the American Museum of Natural History* 279:125-146.
- Baskin JA. 2004. *Bassariscus* and *Probassariscus* (Mammalia, Carnivora, Procyonidae) from the early Barstovian (Middle Miocene). *Journal of Vertebrate Paleontology* 24:709-720.
- Beerli P. 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22:341-345.
- Beerli P, Felsenstein J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America* 98:4563-4568.
- Beerli P, Palczewski M. 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics* 185:313-326.
- Beheregaray LC. 2008. Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology* 17:3754-3774.
- Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* 10:e1003537.
- Bradley RD, Henson DD, Durish ND. 2008. Re-evaluation of the geographic distribution and phylogeography of the *Sigmodon hispidus* complex based on mitochondrial DNA sequences. *Southwestern Naturalist* 53:301–310.
- Bryson RW, García-Vázquez UO, Riddle BR. 2011a. Phylogeography of the Middle American gopher snakes: mixed responses to biogeographical barriers across the Mexican transition zone. *Journal of Biogeography* 38:1570–1584.

- Bryson RW, Murphy RW, Lathrop A, Lazcano-Villareal D. 2011b. Evolutionary drivers of phylogeographical diversity in the highlands of Mexico: a case study of the *Crotalus triseriatus* species group of montane rattlesnakes. *Journal of Biogeography* 38:697-710.
- Cadena CD, Klicka J, Ricklefs RE. 2007. Evolutionary differentiation in the Neotropical montane region: molecular phylogenetics and phylogeography of Buarremon brush-finches (Aves-Emberizidae). *Molecular Phylogenetics and Evolution* 44, 993–1016.
- Cassiliano ML. 1999. Biostratigraphy of Blancan and Irvingtonian mammals in the Fish Creek-Vallecito section, southern California, and a review of the Blancan–Irvingtonian boundary. *Journal of Vertebrate Paleontology*. 19:169-186.
- Castañeda-Rico S, León-Paniagua L, Vázquez-Domínguez E, Navarro-Sigüenza AG. 2014. Evolutionary diversification and speciation in rodents of the Mexican lowlands: the *Peromyscus melanophrys* species group. *Molecular Phylogenetics and Evolution* 70:454-463.
- Castoe TA, Daza JM, Smith EN, Sasa MM, Kuch U, Campbell JA, Chippindale PT, Parkinson CL. 2009. Comparative phylogeography of pitvipers suggests a consensus of ancient Middle American highland biogeography. *Journal of Biogeography* 36:88-103.
- Coates AG, Obando JA. 1996. The geologic evolution of the Central American Isthmus, pp. 21-56. In: Jackson JBC, Budd AF, Coates AG (eds.). *Evolution and Environment in Tropical America*, University of Chicago Press, Chicago, IL, USA.
- Coates AG, Stallard RF. 2013. How old is the Isthmus of Panama? *Bulletin of Marine Science* 89:801-813.
- Coates AG, Jackson JB, Collins LS, Cronin TM, Dowsett HJ, Bybell LM, Jung P, Obando JA. 1992. Closure of the Isthmus of Panama: the near-shore marine record of Costa Rica and Panama. *Geological Society of America Bulletin* 104:814-828.
- Dalquest WW. 1978. Early Blancan Mammals of the Beck Ranch Local Fauna of Texas. *Journal of Mammalogy* 59:269-298.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest2: more models, new heuristics and parallel computing. *Nature Methods* 9:772.
- Dansgaard W, Johnson SJ, Clausen HB, Dahl-Jensen D, Gundestrup NS, Hammer CU, Hvidberg CS, Steffensen JP, Sveinbjörnsdóttir AE, Jouzel J, Bond G. 1993. Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature* 364: 218–220.
- Daza JM, Castoe TA, Parkinson CL. 2010. Using regional comparative phylogeographic data from snake lineages to infer historical processes in Middel America. *Ecography* 33:343-354.
- Decker DM. 1991. Systematics of the coatis, genus *Nasua* (Mammalia:Procyonidae). *Proceedings of the Biological Society of Washington* 104:370-386.

- DeGiorgio M, Jakobsson M, Rosenberg NA. Out of Africa: modern human origins special feature: explaining worldwide patterns of human genetic variation using a coalescent-based serial founder model of migration outward from Africa. *Proceedings of the National Academy of Sciences of the United States of America* 106:16057-16062.
- Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4:359-361.
- Eizirik E. 2012. A molecular view on the evolutionary history and biogeography of Neotropical carnivores (Mammalia, Carnivora), pp. 123-142. In: Patterson BD, Costa LP (eds.). *Bones, Clones and Biomes. The History and Geography of Recent Neotropical Mammals*. University of Chicago Press, Chicago, IL, USA.
- Eizirik E, Bonatto SL, Johnson WE, Crawshaw Jr PG, Vié JC, Brousset DM, O'Brien SJ, Salzano FM. 1998. Phylogeographic patterns and evolution of the mitochondrial DNA control region in two Neotropical cats (Mammalia, Felidae). *Journal of Molecular Evolution* 47:613–624.
- Eizirik E, Murphy WJ, Koepfli KP, Johnson WE, Dragoo JW, Wayne RK, O'Brien SJ. 2010. Pattern and timing of diversification of the mammalian order Carnivora inferred from multiple nuclear gene sequences. *Molecular Phylogenetics and Evolution* 56:49-63.
- Erkens RHJ. 2015. The less-splendid isolation of the South American continent. *Frontiers of Biogeography* 7:89-90.
- Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611-2620.
- Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10:564-567.
- Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567-1587.
- Farris DW, Jaramillo C, Bayona G, Restrepo-Moreno SA, Montes C, Cardona A, Mora A, Speakman RJ, Glascock MD, Valencia V. 2011. Fracturing of the Panamanian Isthmus during initial collision with South America. *Geology* 39: 1007-1010.
- Ferrari L, López-Martínez M, Aguirre-Díaz G, Carrasco-Núñez G. 1999. Space-time patterns of Cenozoic arc volcanism in central Mexico: From the Sierra Madre Occidental to the Mexican Volcanic Belt. *Geology* 27:303-306.

- Ferrari L, Conticelli S, Vaggelli G, Petrone CM, Manetti P. 2000. Late Miocene volcanism and intra-arc tectonics during the early development of the Trans-Mexican Volcanic Belt. *Tectonophysics* 318:161-185.
- Ferrusquía-Villafranca I, González-Guzmán LI, Cartron JE. 2005. Northern Mexico's landscape, part I: the physical settings and constraints on modelling biotic evolution, 11-38 pp. In: Cartron JE, Ceballos G, Felger RS (eds.). *Biodiversity, ecosystems and conservation in northern Mexico*. Oxford University Press, New York, NY.
- Forasiepi AM, Soibelzon LH, Gomez CS, Sánchez R, Quiroz LI, Jaramillo C, Sánchez-Villagra MR. 2014. Carnivorans at the Great American Biotic Interchange: new discoveries from the northern neotropics. *Naturwissenschaften* 101:965-974.
- Gascuel O. 1997. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Molecular Biology and Evolution* 14:685–695.
- Gilmore LS. 2013. Analysis of the Blancan Procyonids of Florida. Master of Science Thesis, East Tennessee State University. Electronic Theses and Dissertations. Paper 1127. <http://dc.etsu.edu/etd/1127>.
- Gompper ME. 1995. *Nasua narica*. *Mammalian Species* 487:1-10.
- Guthrie RD. 2003. Rapid body size decline in Alaskan Pleistocene horses before extinction. *Nature* 426:169-171.
- Gutiérrez-García TA, Vázquez-Domínguez E. 2012. Biogeographically dynamic genetic structure bridging two continents in the monotypic Central American rodent *Otodylomys phyllotis*. *Biological Journal of the Linnean Society* 107:593-610.
- Gutiérrez-García TA, Vázquez-Domínguez E. 2013. Consensus between genes and stones in the biogeographic and evolutionary history of Central America. *Quaternary Research* 79:311-324.
- Hall ER. 1981. *The mammals of North America*. John Wiley & Sons, New York, USA.
- Hardy DK, González-Cózatl FX, Arellano E, Rogers DS. 2013. Molecular phylogenetics and phylogeographic structure of Sumichrast's harvest mouse (*Reithrodontomys sumichrasti*: Cricetidae) based on mitochondrial and nuclear DNA sequences. *Molecular Phylogenetics and Evolution* 68:2822-92.
- Haug GH, Tiedemann R. 1998. Effect of the formation of the Isthmus of Panama on Atlantic Ocean thermohaline circulation. *Nature* 393:673-676.
- Haug GH, Tiedemann R, Zahn R, Ravelo AC. 2001. Role of Panama uplift on oceanic freshwater balance. *Geology* 29:207-210.

- Hauswaldt JS, Ludewig A-K, Vences M, Pröhl H. 2011. Widespread co-occurrence of divergent mitochondrial haplotype lineages in a Central American species of poison frog (*Oophaga pumilio*). *Journal of Biogeography* 38:711–726.
- Helgen KM, Kays R, Helgen LE, Tsuchiya-Jerep MTN, Pinto CM, Koepfli K, Eizirik E, Maldonado JE. 2009. Taxonomic boundaries and geographic distributions revealed by an integrative systematic overview of the mountain coatis, *Nasuella* (Carnivora: Procyonidae). *Small Carnivore Conservation* 41: 65–74.
- Helgen KM, Pinto CM, Kays R, Helgen LE, Tsuchiya MTN, Quinn A, Wilson DE, Maldonado JE. 2013. Taxonomic revision of the olingos (*Bassaricyon*), with description of a new species, the olinguito. *ZooKeys* 324:1-83.
- Hewitt GM. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* 58: 247–276.
- Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, Johnson JB, Rissler L, Victoriano PF, Yoder AD. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution* 54, 291-301.
- Horn C, Wesselingh FP, Ter Steege H, Bermudez MA, Mora A, Sevink J, Sanmartín I, Sanchez-Meseguer A, Anderson CL, Figueiredo JP, Jaramillo C, Riff D, Negri FR, Hooghiemstra H, Lundberg J, Stadler T, Särkinen T, Antonelli A. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* 330:927-931.
- Housley DJE, Zalewski ZA, Beckett SE, Venta PJ. 2006. Design factors that influence PCR amplification success of cross-species primers among 1147 mammalian primer pairs. *BMC Genomics* 7:253.
- Huson DH, Bryant D. 2006. Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution* 23:254-267.
- Iturralde-Vinent MA. 2006. Meso-Cenozoic Caribbean Paleogeography: Implications for the historical biogeography of the region. *International Geology Review* 48: 791-827.
- Irwin DM, Kocher TD, Wilson AC. 1991. Evolution of the cytochrome b gene of mammals. *Journal of Molecular Evolution* 32:128–144.
- Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801-1806.
- Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* 30:3059–3066.

- Koepfli K-P, Gompper ME, Eizirik E, Ho CC, Linden L, Maldonado JE, Wayne RK. 2007. Phylogeny of the Procyonidae (Mammalia: Carnivora): Molecules, morphology and the great American interchange. *Molecular Phylogenetics and Evolution* 43:1076-1095.
- Langella M. 1999. Populations 1.2.30: Population genetic software (individuals or population distances, phylogenetic trees). Available from <http://bioinformatics.org/~tryphon/populations/> (accessed 15 May 2016).
- Leigh EG, O’Dea A, Vermeij GJ. 2014. Historical biogeography of the Isthmus of Panama. *Biological Review* 89:148-172.
- MacMillan I, Gans PB, Alvarado G. 2006. Middle Miocene to present plate tectonic history of the southern Central American volcanic Arc. *Tectonophysics* 392:325-348.
- Mann P, Rogers RD, Gahagan L. 2007. Overview of plate tectonic history and its unresolved tectonic problems, pp. 205–241. In: Bundschuh J, Alvarado GE. (eds.). *Central America: Geology, Resources and Hazards*. Taylor and Francis, Philadelphia, PA, USA.
- Marshall LG. 1985. Geochronology and land-mammal biochronology of the transamerican faunal interchange, pp. 357-386. In: Stehli FG, Webb SD (eds.). *The Great American Biotic Interchange*. Plenum Press, New York, NY, USA.
- Marshall LG, Butler RF, Drake RE, Curtis GH, Tedford RH. Calibration of the great American interchange. *Science* 204:272-279.
- Marshall LG, Webb SD, Sepkoski JJ, Raup DM. 1982. Mammalian evolution and the great american interchange. *Science* 215:1351-1357.
- Marshall JS. 2007. The geomorphology and physiographic provinces of Central America, pp. 1-51. In: Bundschuh J, Alvarado GE (eds.). *Central America: Geology, Resources and Hazards*. Taylor & Francis, Philadelphia, PA, USA.
- Martin PS. 1984. Prehistoric overkill: The global model, pp. 354 – 403. In: Martin PS, Klein RG (eds.). *Quaternary Extinction: A prehistoric revolution*. University of Arizona Press, Tucson, AZ, USA.
- McCormack JE, Peterson AT, Bonaccorso E, Smith, TB. 2008. Speciation in the highlands of Mexico: genetic and phenotypic divergence in the Mexican jay (*Aphelocoma ultramarina*). *Molecular Ecology* 17:2505–2521.
- McFadden KW, Gompper ME, Valenzuela D, Morales JC. 2008. Evolutionary history of the critically endangered Cozumel dwarf carnivores inferred from mitochondrial DNA analyses. *Journal of Zoology* 276: 176–186.

- Molnar P. 2008. Closing of the Central American Seaway and the Ice Age: A critical review. *Paleoceanography* 23:PA2201.
- Montes C, Bayona G, Cardona A, Buchs DM, Silva CA, Morón S, Hoyos N, Ramírez DA, Jaramillo CA, Valencia V. 2012a. Arc-continent collision and orocline formation: Closing of the Central American seaway. *Journal of Geophysical Research*: 117: B04105.
- Montes C, Cardona A, McFadden R, Morón SE, Silva CA, Restrepo-Moreno S, Ramírez DA, Hoyos N, Wilson J, Farris D, Bayona GA, Jaramillo CA, Valencia V, Bryan J, Flores JA. 2012b. Evidence for middle Eocene and younger land emergence in central Panama: Implications for Isthmus closure. *Geological Society of America Bulletin* 124:780-799.
- Montes C, Cardona A, Jaramillo C, Pardo A, Silva JC, Valencia V, Ayala C, Pérez-Angel LC, Rodríguez-Parra LA, Ramirez V, Niño H. 2015. Middle Miocene closure of the Central American Seaway. *Science* 348: 226-229.
- Mora A, Baby P, Roddaz M, Parra M, Brusset S, Hermoza W, Espurt N. 2010. Tectonic history of the Andes and sub-Andean zones: implications for the development of the Amazon drainage basin, 38-60 pp. In: Hoorn C, Wesselingh FP (eds.). *Amazonia, landscape and species evolution, a look into the past*. Wiley-Blackwell, West Sussex, UK.
- Morgan GS. 2008. Vertebrate fauna and geochronology of the great American biotic interchange in North America, 93-140 pp. In: Lucas SG, Morgan GS, Spielmann JA, Prothero DR (eds.). *Neogene Mammals*. New Mexico Museum of Natural History and Science Bulletin 44.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GA, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403, 853-858.
- Nei M, Tajima F, Tateno Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution* 19:153-170.
- Nieto-Samaniego AF, Alaniz-Álvarez SA, Silva-Romo G, Eguiza-Castro MH, Mendoza-Rosales CC. 2006. Latest Cretaceous to Miocene deformation events in the eastern Sierra Madre del Sur, Mexico, inferred from the geometry and age of major structures. *Geological Society of America Bulletin* 118:238-252.
- Nowak RN. 2005. *Walker's carnivores of the world*. Johns Hopkins University Press. Baltimore, MD, USA.
- Ortega-Gutiérrez F, Solari LA, Ortega-Obregón C, Elías-Herrera M, Martens U, Morán-Icál S, Chiquín M, Keppie JD, De León RT, Schaaf P. 2007. The Maya-Chortís boundary: a tectonostratigraphic approach. *International Geology Review* 49:996-1024.
- O'Dea A, Lessios HA, Coates AG, Eytan RI, Restrepo-Moreno SA, Cione AL, Collins LS, de Queiroz A, Farris DW, Norris RD, Stallard RF, Woodburne MO, Aguilera O, Aubrey M-P,

- Berggren WA, Budd AF, Cozzuol MA, Coppard SE, Duque-Caro H, Finnegan S, Gasparini GM, Grossman EL, Johnson KG, Keigwin LD, Knowlton N, Leigh EG, Leonard-Pingel JS, Marko PB, Pyenson ND, Ravello-Dolme PG, Soibelzon E, Soibelzon L, Todd JA, Vermeij GJ, Jackson JBC. 2016. Formation of the Isthmus of Panama. *Science Advances* 2:e1600883.
- Paetkau D, Slade R, Burden M, Estoup A. 2004. Direct, real-time estimation of migration rate using assignment methods: a simulation-based exploration of accuracy and power. *Molecular Ecology* 13:55-65.
- Peakall R, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28:2537-2539.
- Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A. 2004. GeneClass2: A Software for Genetic Assignment and First-Generation Migrant Detection. *Journal of Heredity* 95:536-539.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.
- Rambaut A. 2014. FigTree v1.4.2. Available from <http://tree.bio.ed.ac.uk/software/figtree/>.
- Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. Tracer v1.6, Available from <http://beast.bio.ed.ac.uk/Tracer>.
- Rannala B, Mountain JL. 1997. Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* 94:9197-9201.
- Raymond M, Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86:248-249.
- Riddle BR, Dawson MN, Hadly EA, Hafner DJ, Hickerson MJ, Mautooth SJ, Yoder AD. 2008. The role of molecular genetics in sculpting the future of integrative biogeography. *Progress in Physical Geography* 32, 173-202.
- Rousset F. 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources* 8:103-106.
- Rutter N, Coronato A, Helmens K, Rabassa J, Zárata M. 2012. Glaciations in North and South America from the Miocene to the last Glacial Maximum, comparisons, linkages and uncertainties. Springer, New York, NY, USA.
- Sato JJ, Woslan M, Minami S, Hosoda T, Sinaga MH, Hiyama K, Yamaguchi Y, Suzuki H. 2009. Deciphering and dating the red panda's ancestry and early adaptive radiation of Musteloidea. *Molecular Phylogenetic and Evolution* 53:907-922.

- Savage JM. 1982. The enigma of the Central American herpetofauna: dispersals or vicariance? *Annals of the Missouri Botanical Garden* 69: 464–547.
- Silvestro D, Michalak I. 2011. raxmlGUI: a graphical front end for RAxML. *Organisms Diversity and Evolution* 12:335–337.
- Simpson GG. 1980. *Splendid Isolation: The curious history of South American mammals*. Yale University Press, New Haven, CT, USA.
- Soibelzon LH, Prevosti F. 2013. Fossils of South American Land Carnivores (Carnivora, Mammalia), pp 509–527. In: Ruiz M, Shostell J (eds.). *Molecular population genetics, evolutionary biology and biological conservation of neotropical carnivores*. Nova Science Publisher, New York, NY, USA.
- Sorenson MD, Ast JC, Dimcheff DE, Yuri T, Mindell DP. 1999. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Molecular Phylogenetics and Evolution* 12:105–114.
- Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.
- Stehli FG, Webb SD (eds.). 1985. *The Great American Biotic Interchange*. Plenum Press, New York, NY, USA.
- Stuart AJ. 1991. Mammalian extinctions in the late Pleistocene of Northern Eurasia and North America. *Biological Reviews* 66:453-562.
- Suárez-Atilano M, Burbrink F, Vázquez-Domínguez E. 2014. Phylogeographical structure within *Boa constrictor imperator* across the lowlands and mountains of Central America and Mexico. *Journal of Biogeography* 41:2371-2384.
- Sullivan J, Arellano E, Rogers DS. 2000. Comparative phylogeography of Mesoamerican highland rodents: concerted versus independent response to past climatic fluctuations. *American Naturalist* 155:754–768.
- Takezaki N, Nei M. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* 144:389-399.
- Trigo TC, Freitas TRO, Kunzler G, Cardoso L, Silva JC, Johnson WE, O'Brien SJ, Bonatto SL, Eizirik E. 2008. Inter-species hybridization among Neotropical cats of the genus *Leopardus*, and evidence for an introgressive hybrid zone between *L. geoffroyi* and *L. tigrinus* in southern Brazil. *Molecular Ecology* 17:4317–4333.
- Valenzuela D, Ceballos G. 2000. Habitat selection, home range, and activity of the White-nosed coati (*Nasua narica*) in a Mexican tropical dry forest. *Journal of Mammalogy* 81:810-819.

- Van Oosterhout C, Hutchinson W, Wills D, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4:535-538.
- Villesen P. 2007. FaBox: an online toolbox for fasta sequences. *Molecular Ecology Notes* 7:965–968.
- Wang IJ, Crawford AJ, Bermingham E. 2008. Phylogeography of the pygmy rain frog (*Prismantis ridens*) across the lowland wet forest of isthmian Central America. *Molecular Phylogenetics and Evolution* 47:992–1004.
- Webb SD. 1985. Late Cenozoic mammal dispersal between the Americas, pp. 357-386. In: Stehli FG, Webb SD (eds.). *The Great American Biotic Interchange*. Plenum Press, New York, NY, USA.
- Webb SD. 2006. The Great American Biotic Interchange: patterns and processes. *Annals of the Missouri Botanical Garden* 93: 245–257.
- Weir B, Cockerham C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358-1370.
- Wilson GA, Rannala B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163:1177-1191.
- Woodburne MO. 2010. The great american biotic interchange: Dispersals, tectonics, climate, sea level and holding pens. *Journal of Mammalian Evolution* 17:245-264.
- Woodburne MO, Cione AL, Tonni EP. 2006. Central American provincialism and the Great American Biotic Interchange, pp. 73-101. In: Carranza-Castañeda O, Lindsay EH (eds.). *Advances in Late Tertiary vertebrate paleontology in Mexico and the Great American Biotic Interchange*. Universidad Nacional Autonoma de Mexico, Instituto de Geologia y Centro de Geociencias, Publicacion Especial 4.
- Wozencraft WC. 2005. Order Carnivora, pp. 279-348. In: Wilson DE, Reeder DM (eds.). *Mammal species of the world: a taxonomic geographic reference*. The Johns Hopkins University Press, Baltimore.