Research Article

Population structure and genetic diversity of the giant anteater (*Myrmecophaga tridactyla*: Myrmecophagidae, Pilosa) in Brazil

Camila L. Clozato[1,2], Flávia R. Miranda[1,2], Paula Lara-Ruiz[1], Rosane G. Collevatti[3], and Fabrício R. Santos[1]

[1]*Laboratório de Biodiversidade e Evolução Molecular, Departamento de Biologia Geral, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brazil.*
[2]*Projeto Tamanduá, São Paulo, SP, Brazil.*
[3]*Laboratorio de Genética e Biodiversidade, Instituto de Ciências Biológicas, Universidade Federal de Goiás (UFG), Goiânia, GO, Brazil.*

Abstract

The giant anteater (*Myrmecophaga tridactyla*, Pilosa, Linnaeus 1758) belongs to the mammalian order Pilosa and presents a large distribution along South America, occupying a great variety of habitats. It is listed in the IUCN Red List of threatened species as Vulnerable. Despite threatened, there is a lack of studies regarding its genetic variability. The aim of this study was to examine the genetic diversity and patterns of genetic structure within remaining populations. We analyzed 77 individuals from seven different populations distributed in four biomes across Brazil: Cerrado, Pantanal, Atlantic Forest and Amazon Forest. We sequenced two mitochondrial markers (control region and Cyt-b) and two nuclear markers (AMELY and *RAG2*). We found high genetic diversity within subpopulations from National Parks of Serra da Canastra and Emas, both within the Cerrado biome, with signs of population expansion. Besides, we found a notable population structure between populations from the Cerrado/Pantanal and Amazon Forest biomes. This data is a major contribution to the knowledge of the evolutionary history of the species and to future management actions concerning its conservation.

Keywords: Giant Anteater, Xenarthra, Cerrado, genetic diversity, population structure.

Received: April 13, 2016; Accepted: September 30, 2016.

Introduction

The giant anteater, *Myrmecophaga tridactyla* Linnaeus 1758, is a mammal of the Myrmecophagidae family, order Pilosa (Gardner, 2005). It is the largest of all four anteater species and it occupies a great variety of habitats, such as rainforests, dry forests, wetlands and open fields (Fonseca and Aguiar, 2004). The species historical distribution corresponds to the area from Honduras in Central America to the Gran Chaco region of Bolivia, Paraguay and Argentina, and southern Pampas of Uruguay and Brazil in South America. The animals typically display solitary behavior and females give birth to a single young once a year after a 190 days of gestation (Eisenberg and Redford, 1999). They feed on ants and termites, and have a low metabolic rate and body temperature (McNab, 1985; Shaw *et al.*, 1985).

The giant anteater is the only member of its family listed in the World Conservation Union's 2014 IUCN Red List of Threatened Species as Vulnerable (VU) (Miranda *et al.*, 2014). The population numbers are declining along its range and the species is already extinct in some locations and countries, for instance, Uruguay, and possibly also Belize, Costa Rica and Guatemala (Fallabrino and Castiñeira, 2006; Miranda *et al.*, 2014). Moreover, within Brazil the species is critically threatened or even virtually extinct (not recorded for a long time, or rarely visualized) from a few states, such as Rio de Janeiro, Espírito Santo, Santa Catarina and Rio Grande do Sul (Bergallo *et al.*, 2000; Fontana *et al.*, 2003; Cherem *et al.*, 2004; Lorenzutti and Almeida, 2006) and appears listed as threatened in 19 Brazilian states according to the national red list of threatened species published by ICMBio (Miranda *et al.*, 2014).

The main causes for the observed population decline are the deterioration and reduction of natural habitats (Fonseca *et al.*, 1999), along with hunting for food, skin trade and pet purposes (Leeuwenberg, 1997; Peres, 2001; Ferreira *et al.*, 2013), frequent road kills (*e.g.* Cunha *et al.*, 2010; Diniz and Brito, 2013), and extensive wildfires in natural parks that usually kill hundreds of animals at once (Silveira *et al.*, 1999). In addition, the species' natural characteristics also contribute to increase its vulnerability, such

Send correspondence to Fabrício R. Santos. Laboratório de Biodiversidade e Evolução Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, Sala L3-244, 31270-010 Belo Horizonte, MG, Brazil. E-mail: fsantos@icb.ufmg.br

as the solitary habit with a long gestation time (Eisenberg and Redford, 1999), and the low metabolic rate that contributes to their slow movement (McNab, 1985) and makes them more susceptible to anthropic hunting.

These elements altogether are expected to make the remaining *M. tridactyla* populations progressively more isolated. It is known that reductions in size and range of populations increase their vulnerability to stochastic extinction, leading in several instances to local extinction (Gilpin and Soulé, 1986). Currently, it is largely accepted that genetic variability plays an important role in the persistence and adaptation of populations to changing environments (Lande and Shannon, 1996; Frankham *et al.*, 2002), and the loss of adaptive genetic diversity places wild populations in greater extinction risk (Frankham, 2005). Thus, the knowledge of population diversity within remaining populations of *M. tridactyla* is essential to aid in conservation management.

Regarding the giant anteater, Collevatti *et al.* (2007) performed a population genetic study with individuals from the National Park of Emas, one of the largest populations of the species in Brazil (Miranda *et al.*, 2006), using five microsatellite loci previously described by Garcia *et al.* (2005). The study revealed a marked inbreeding within the population, associated to low levels of polymorphism in all loci. Apart from that, no other population genetics or phylogeography studies have focused on *M. tridactyla.* The lack of scientific information concerning the species' diversity pattern and its population structure calls for detailed research studies on this threatened species.

In this study we present results on population genetic diversity and structure based on the analysis of mtDNA and nDNA markers in different remaining populations across the distribution range of *M. tridactyla* in Brazil. The aim of this study is to describe the genetic diversity within sampled populations by means of mtDNA and nuclear markers, and to evaluate the presence of genetic structure throughout the species distribution. When genetic structure was found, we searched for likely causes to explain the species' distribution along the landscape, commonly associated with either isolation by distance (IBD) or population structure driven by historical divergence. This information provides scientific resources for future management actions for the conservation of the species.

Materials and Methods

Sample collection

Seventy-seven individuals of *M. tridactyla* were collected between 1994 and 2007 for nine Brazilian federal states (Minas Gerais, Goiás, São Paulo, Mato Grosso, Mato Grosso do Sul, Paraná, Pará, Roraima, Amapá), in 20 different localities, except for seven samples that originated from a museum collection (Museu Paraense Emílio Goeldi, MPEG), which dated from 1957-1979. Sampled tissues

varied from dry skin (museum), to hair, bone and soft tissues. The original samples were collected from captive individuals of known origin, wild animals captured for ecological studies, road-killed individuals, and museum collection vouchers. All tissue samples were preserved in 95% ethanol, and stored in -20 ºC. The samples covered the populations of (i) Minas Gerais state (CEMG, n=21), comprising individuals from the National Park of Serra da Canastra and other localities, (ii) Goiás state (CEGO, n=28) comprising individuals from the National Park Emas, (iii) different localities in São Paulo state (CESP, n=7) and (iv) Mato Grosso state (CEMT, n=4), all representative of the biome Cerrado (CE). Additionally, a population from the Atlantic Forest of Paraná state (AF, n=5), from the Pantanal biome (PT, n=5) of Mato Grosso and Mato Grosso do Sul states, and individuals from the Amazon Forest (AM, n=8) were also sampled. Sampling localities are displayed in Figure 1, and details regarding samples are available in Table S1 from Supplementary Material.

All biological material was collected with authorization for activity with scientific purposes emitted by SISBIO/IBAMA under the accession number 15052-1.

Molecular Methods

Total genomic DNA was extracted from tissue samples using a standard phenol-chloroform protocol (Sambrook *et al.*, 1989). For extraction of dry skins and bone fragments a modified protocol from Holland *et al.* (2003) was used in combination with a DNA extraction Kit (DNA Tissue Kit, Qiagen). For mitochondrial DNA (mtDNA) analyses, two fragments were amplified by PCR: 450 base pairs (bp) of the first hypervariable segment from the control region (HVI), with two sets of primers, BrDi-L and BrDi-H (modified from original primers described in Arnason *et al.*, 2002 and Douzery and Randi, 1997, respectively), and Pro-L (Lara-Ruiz *et al.*, 2008) with H16498 (Ward *et al.*, 1991); and 555 bp of the Cytochrome *b* gene (*Cyt-b*) using primers CytB-L and CytB-H (Lara-Ruiz *et al.*, 2008) (or XL14733 from Kocher *et al.*, 1989, as an alternative reverse primer). For nuclear DNA (nDNA) analyses, a 700 bp fragment of the recombination activation gene (*RAG2*) was amplified with the set of primers RAG2-F220 and RAG2-R995 (Teeling *et al.*, 2001), and, finally, a 600 bp of the Y chromosome Amelogenin gene (AMELY) was amplified with primers AMELY-F2 and AMELY-R2 (Roca *et al.*, 2005), only in known male specimens.

PCR amplifications for mtDNA markers were done in 15 µL volumes containing 10X buffer, 200 µM dNTPs, 0.5 µM of each primer, and 1 unit of *Taq* DNA polymerase (Phoneutria Biotecnologia). Thermocycling conditions consisted of a denaturing step at 94 ºC for 3 min, 35 cycles at 94 ºC for 30 s, 50 ºC for 40 s, 72 ºC for 30 s, and a final extension at 72 ºC for 10 s. For nDNA markers, PCR amplification was done in 10 µL volumes containing 10X buffer, 1.5 µM MgCl$_2$, 200 µM dNTPs, 0.5 µM of each primer and
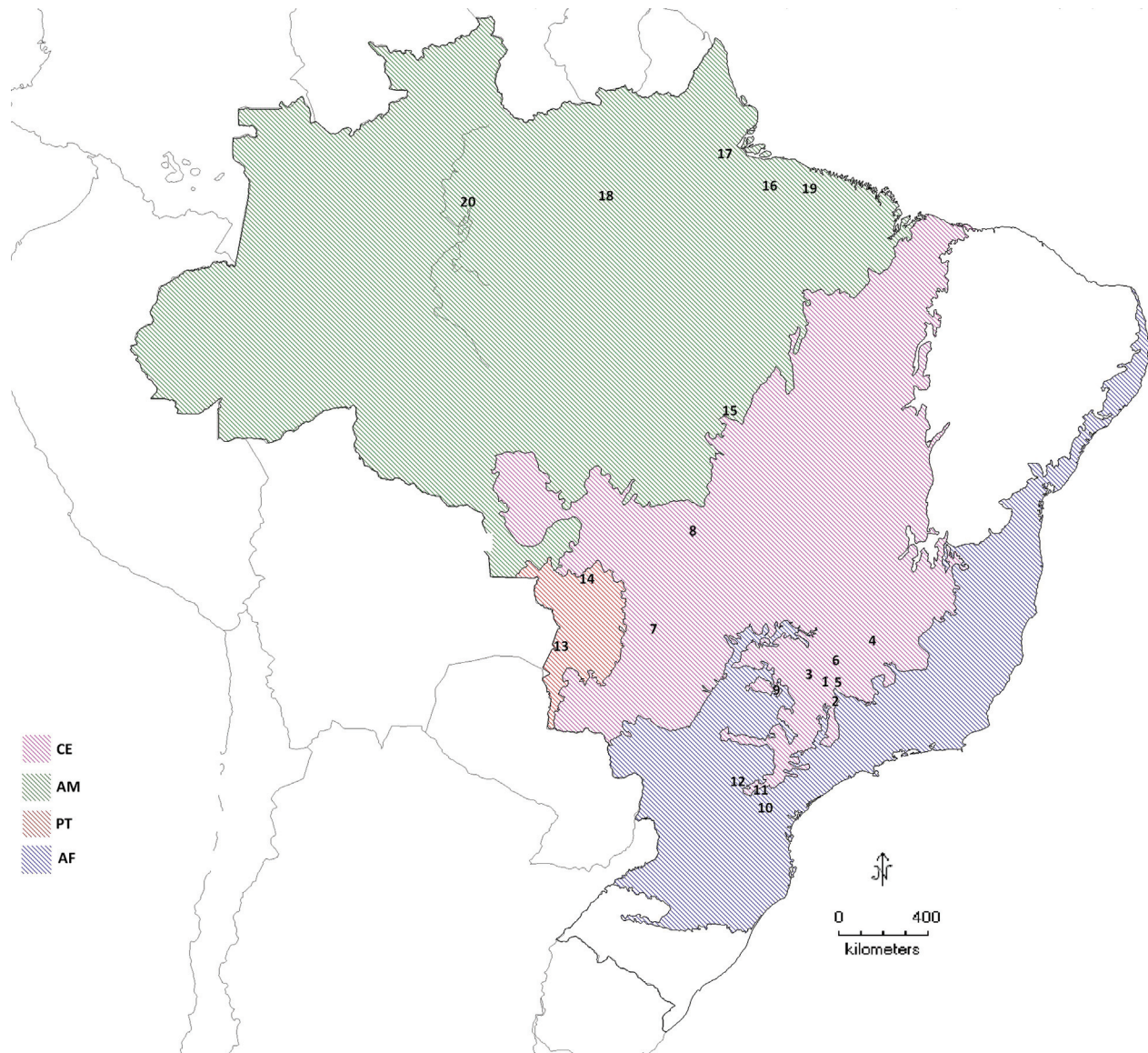
Figure 1 - Map with localities of *M. tridactyla* individuals sampled for this study. 1- National Park of Serra da Canastra, MG; 2- Piumhi, MG; 3- Araxá, MG; 4- Dores do Indaiá, MG; 5- Doresópolis, MG; 6- Uberlândia, MG; 7- National Park of Emas, GO; 8- Nova Xavantina, MT; 9- São José do Rio Preto, SP; 10- Jaguariaíva, PR; 11- Telêmaco Borba; 12- Piraí do Sul; 13- Corumbá, MS; 14- Poconé, MT; 15- Vila Rica, PA; 16- Ilha do Marajó, PA; 17- Mazagão, AP; 18- Oriximiná, PA; 19- Belém, PA; 20- Caracaraí, RR.

0.5 unit of *Platinum Taq*® polymerase (Thermo Fisher Scientific). In both cases, template DNA dilutions were used ranging between 20-100 ng/μL. Thermocycling consisted of a hotstart step at 95 ºC for 9 min 45 s, 5-10 (touchdown) cycles of a denaturing step at 94 ºC for 15 s, annealing at 49-54 ºC for 30 s, extension at 72 ºC for 80 s, and a final extension step at 72 ºC for 3 min. All products were examined on a 0.8% agarose gel stained with ethidium bromide, purified with a moidified polyethyleneglycol (PEG) protocol (Santos-Júnior *et al.*, 2015), and sequenced using a MegaBACE DNA Analysis System 1000 automatic sequencer (Amersham Biosciences). All samples were sequenced at least twice, in forward and reverse directions. Museum, hair and bone samples were doubled checked.

Data analysis

Sequence electropherograms were visually inspected using Phred v. 0.20425 (Ewing and Green, 1998), Phrap v. 0.99031 and Consed 12.0 (Gordon *et al.*, 1998), and aligned using Clustal W (Higgins and Sharp, 1988) algorithm implemented in MEGA 4.0 (Tamura *et al.*, 2007). Alignments were checked and edited by hand to account for artifacts. For autosomal data (*RAG2*), PolyPhred 5.04 (Nickerson *et al.*, 1997) was used for identifying heterozygote sites. In this case, PHASE 2.0 (Stephens *et al.*, 2001; Stephens and Donnely, 2003) was used to reconstruct haplotypes from genotypes, estimating the gametic phase. Software DNAsp 5.0 (Librado and Rozas, 2009) was used to obtain the

haplotypes and their polymorphic positions for both haploid and diploid dataset.

The relationship between haplotypes and their geographical distribution was visualized through a phylogenetic network, using the median-joining (MJ) algorithm in NETWORK v. 4.6 software (Bandelt *et al.*, 1999).

Software Arlequin v. 3.5 (Excoffier and Lischer, 2010) was used to calculate haplotype (h) and nucleotide diversity (π), ΘS values (a measure of the population nucleotide diversity), Tajimas D, a test of selective neutrality (Tajima, 1989), population pairwise F_{ST} values and analysis of molecular variance (AMOVA) computed with pairwise differences, with 1000 permutations to test for significance at the 0.05 level. Tajimas test of selective neutrality was used to distinguish between a random (neutrality) and non-random evolving DNA sequence dataset, which may be caused by positive/balancing selection or by demographic fluctuations (expansions and contractions). When there is an excess of low frequency polymorphisms in the dataset a negative D value results, which indicates population size expansions or positive selection (Tajima, 1989). Since we only used mtDNA data for the Tajimas D analysis, which is considered mainly a neutrally evolving marker (Saccone *et al.*, 2000), we assume that a negative significant D would be most likely indicative of demographic expansion. The fixation index F_{ST} measures population differentiation based on the population frequency of genetic polymorphisms. It was estimated for each population pair, and then in groups of populations, using AMOVA, to examine the level of genetic subdivision between localities. The analyses were grouped in three different ways: (i) populations from all four biomes in distinct groups [CE][PT][AF][AM]; (ii) forested versus open vegetation biomes [AM+AF][CE+PT]; and finally (iii), using a geographic distance criterion: [CE+PT+AF] [AM].

AMOVA was also used to test for sex biased dispersion, measured in the major sampled populations (CEMG and CEGO, individuals from National Parks only), grouped by population and gender (males, M, and females, F, from each park, and in both parks). For this purpose, mtDNA (HVI+Cyt-b) and nDNA (*RAG2*) was used. Groups were tested as follows: (i) HVI+Cyt-b [CEMG+CEGO][F]/[CEMG+CEGO][M] and (ii) RAG2 [CEMG+CEGO][F]/[CEMG+CEGO][M].

For the mismatch distribution analysis, observed and expected pairwise differences between alleles were calculated in Arlequin (Excoffier and Lischer, 2010). It is expected to show a unimodal distribution when populations have undergone a rapid expansion, and a bimodal distribution if populations are subdivided or in demographic equilibrium (Rogers and Harpending, 1992).

In order to infer about the hypothesis of isolation by distance (IBD) we conducted the non-parametric Mantel's test, which correlates genetic and geographical distances. To access the correlation coefficient reliability, 10,000 replicates were done. The test was performed in the software Alleles in Space, AIS vs. 1.0 (Miller, 2005). To explore the existence and location of barriers to gene flow, the software Barrier vs. 2.2 (Manni *et al.*, 2004) was used. The software uses Monmoniers maximum difference algorithm (Monmonier, 1973) designed to visualize on a geographic map (represented by geographical coordinates) the trend of data constrained in a matrix, in this case, a matrix of genetic distances between all populations sampled. The triangulation edge parameters were not modified. Genetic distance matrix input was calculated with MEGA 4.0 (Tamura *et al.*, 2007). We set the initial number of barriers to four, given that we sampled four different biomes.

Data Access

Sequence data for mtDNA and nDNA markers are publicly available at GenBank, under the accession numbers: KF543782-KF543820.

Results

Genetic diversity and haplotype distribution

All 77 individuals were successfully amplified and sequenced for mtDNA markers, HVI and *Cyt-b*. These fragments were analyzed jointly as one, totalizing 1005 bp. After alignment, the sequences showed 29 haplotypes distributed along four biomes (Table 1). Nuclear markers were successfully amplified and sequenced in 47 individuals for *RAG2* and in 34 individuals for AMELY. These markers were used to compare general results against mitochondrial data, once their data covered populations from CE, PT and AF, but excluded AM. *RAG2* showed eight haplotypes, and AMELY showed nine (Table 2). Sequences of haplotypes of all markers were deposited in GenBank (accession numbers KF543782-KF543820).

Regarding mitochondrial data, there were 36 polymorphic sites, being 19 singletons and 17 parsimony informative sites. Total haplotype diversity was 0.8267 ± 0.0416, ranging from 0.500 in CEMT to 1.000 in PT; total nucleotide diversity was 0.002163 ± 0.001366, ranging from 0.000498 in CEMT to 0.003383 in PT (Table 3). Values of ΘS ranged from 0.5454 in CEMT to the highest value of 4.3685 in CEGO. In general, the results showed high haplotype diversity and moderate nucleotide diversity. In terms of populations, the highest levels of diversity were encountered in PT, despite the few individuals sampled. No particular haplotype was shared by all populations, and only two haplotypes (H1, H9) were shared among three or more populations (Table 1). Haplotype H1 was the most frequent one, found in 31 of the 77 samples, and it was found mostly in CE populations (CEMG, CEGO, CEMT and CESP), shared by only one individual in the PT population. Some haplotypes were shared between CE and PT populations, and between AF and AM populations, and only one haplotype (H9) was shared among AM, CE and

Table 1 - Haplotypes for mtDNA (HVI and Cyt-b joined together), polymorphic sites and distribution per population.

Haplotype	56	573	787	269	256	262	264	265	267	271	274	281	294	296	482	501	528	535	553	635	751	757	792	855	872	890	903	945	CE MG	CE GO	CE SP	CE MT	PT	AF	AM
H1	T	C	G	C	A	G	T	T	G	C	A	C	C	A	T	G	A	C	A	A	G	A	C	A	A	C	C	A	10	15		1	1		
H2	G	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	1						
H3	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	1	1					
H4	.	.	T	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	1					
H5	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1						
H6	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	1	1					
H7	.	.	.	.	.	A	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	1	1					
H8	.	.	.	.	G	G	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1						
H9	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.				3		2	3
H10	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.						2	1
H11	.	.	.	.	.	.	T	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1						
H12	.	.	.	.	.	.	A	.	.	.	A	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.							2
H13	.	.	.	C	.	.	A	.	.	.	A	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.							1
H14	G	.	.	.	.	.	A	.	.	.	A	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.							1
H15	.	.	.	.	G	A	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1			1		
H16	.	.	.	.	G	A	.	.	.	.	A	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	1				1		
H17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G					1		
H18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	G	.	G	.	.	.	.		1			1		
H19	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.					1		
H20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.			1	1			
H21	.	.	.	.	.	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.				1			
H22	.	.	.	.	G	A	A	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		1			1		
H23	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.						1	
H24	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1						
H25	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	G	.	.	.		1					
H26	.	.	.	.	G	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	G	.		3					
H27	.	G	.	.	G	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		1					
H28	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		1					
H29	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1						

Table 2 - Haplotypes for nDNA (AMELY and *RAG2*), polymorphic sites and distribution per population. *RAG2* is based in 94 sequences from 47 individuals (autosomal data).

Haplotype		Polymorphic Sites in nDNA					Haplotype per Population						
			1	2	5		CEMG	CEGO	CESP	CEMT	PT	AF	AM
		3	7	6	1	1							
		8	1	0	0	2							
AMELY	A1	A	G	C	A	C		2	8		1			-
	A2	.	.	.	G	.		3		1		1		-
	A3	G	A	T	G	.						1		-
	A4	G	.	.	G	.		1						-
	A5	.	.	.	G	T				1				-
	A6	G	.	.	.	T		1	2			1		-
	A7	.	.	.	.	T		3	3	2		1		-
	A8	G	.	T	.	.							1	-
	A9	G	.	.	.	.			1					-
		2	2	3	4	7	7	CEMG	CEGO	CESP	CEMT	PT	AF	AM
		4	9	0	1	0	0							
		5	4	8	3	7	9							
RAG2	R1	C	C	A	A	G	C	26	28	7	1	6	2	-
	R2	.	.	.	A	.	T	2	4	2		2		-
	R3	.	.	.	A	A	.	1	4	1	1	1		-
	R4	T	.	.	A	.	.					1		-
	R5	.	G	.	A	A	.	1						-
	R6	.	G	.	A	.	.	1						-
	R7	.	.	G	A	.	.			1				-
	R8	.	.	.	G	.	.	1	1					-

Table 3 - Parameters of genetic diversity in each sampled population of *M. tridactyla*, and in the entire dataset. Significant Tajimas D values (p < 0.05) are marked with an asterisk (*).

Population	No. Individuals	Polymorphic Sites	Haplotype Diversity	Θ S	Nucleotide Diversity	Tajima's D	
						D	$P (Ds < Do)$
CEMG	21	12	0.7810 +/- 0.0943	3.3354	0.001838 +/- 0.001226	-1.58	0.04*
CEGO	28	17	0.7143 +/- 0.0929	4.3685	0.002022 +/- 0.001304	-1.85	0.02*
CESP	6	4	0.7143 +/- 0.1809	1.6326	0.001137 +/- 0.000952	-1.43	0.06
CEMT	4	1	0.5000 +/- 0.2652	0.5454	0.000498 +/-0.000617	-0.61	0.39
AF	5	3	0.8000 +/- 0.1640	1.4400	0.001393 +/- 0.001185	-0.17	0.49
PT	5	7	1.0000 +/- 0.1265	3.3000	0.003383 +/- 0.002425	0.08	0.57
AM	8	5	0.8571 +/- 0.1083	1.9283	0.001883 +/- 0.001363	-0.083	0.47
All	77	30	0.8267 +/- 0.0416	6.1043	0.002163 +/- 0.001366	-2.02	0.00*

AF populations. Five haplotypes were shared between CEMG and CEGO, and one between AF and AM (H10). The network of mitochondrial data exhibited a star-like shape pattern, with many haplotypes derived from the most common one (H1). Most haplotypes differ from one another by only one single mutational step (Figure 2a).

Nuclear DNA data revealed the same pattern regarding haplotype distribution. The Y chromosome marker, AMELY, showed four haplotypes (A1, A2, A6, A7) shared among three or more populations, and *RAG2* showed three haplotypes (R1, R2 and R3) shared among three or more populations. Of these, all haplotypes except A1 were shared with the PT population, and only one, R1, was shared with AF (Table 2, Figure 2b, c). Since only CE, PT and AF populations could be fully analyzed for nDNA, all the following analyses regarding population expansion and
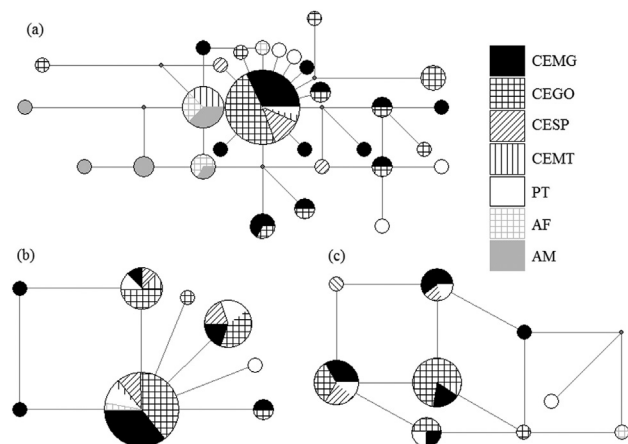
Figure 2 - Median-joining networks of *M. tridactyla* haplotypes. (a) mtDNA, HVI and Cyt-b; (b) nDNA, *RAG2;* (c) nDNA, iAMELY. Sizes of circles are proportional to the amount of individuals carrying the haplotype. Colors and patterns are representative of populations from which individuals originated (depicted in labels).

population structure and biogeographic analyses will focus solely in mtDNA data.

Population Expansion and population structure

Tajimas D for the entire sample set was -2.02 (p < 0.0001), indicating a likely demographic expansion. Only populations CEMG and CEGO also showed significant negative values of Tajimas D when analyzed separately (Table 3).

The mismatch distribution graphic (Figure 3b) presented a clear unimodal fashion between observed and expected average number of pairwise differences, indicating a scenario compatible with a recent population expansion, and corroborating the star-like shape of the mtDNA network (Figure 2a).
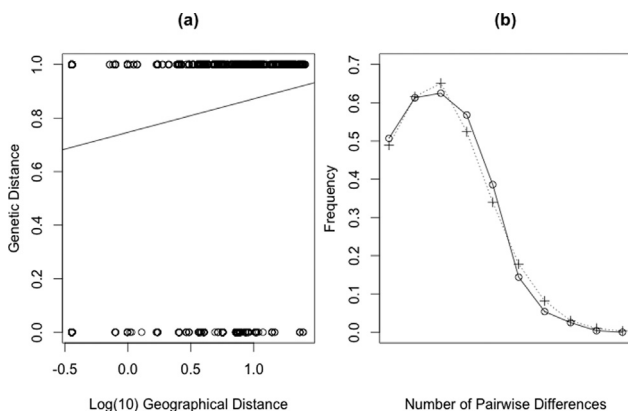
Figure 3 - Mantel test results. (a) Plot of Mantel test showing the relationship of genetic and geographic distances (r=0.18484, p=0.9995); (b) Mismatch distribution of mtDNA data for *M. tridactyla*. Observed (Obs, dashed line) and expected (Exp, solid line) average number of pairwise differences show a unimodal fashion, compatible with a recent population expansion scenario.

Population pairwise F_{ST} values were low and non-significant between CE populations (CEMG, CESP and CEGO) and PT. Values were also low between populations AF and AM. Among CE populations, F_{ST} was significant only between CEMT and two other populations, CEMG and CESP. All CE populations except CEMT showed significant F_{ST} values against AM and AF (the highest CESP/AM, F_{ST}=0.45252, p < 0.0001). PT showed high F_{ST} values against AF and AM as well, but was significant only against AM (Table 4).

The AMOVA test revealed the highest F_{ST} value when [AM] was separated from the group [CE+PT+AF], and the lowest when groups from all four biomes were separated. The percentage of genetic variation was always higher within groups and populations than between groups (Table 5). AMOVA tests between gender (groups [CEMG+CEGO][F]/[CEMG+CEGO][M] for each marker) did not show any significant evidence of sex biased dispersal, neither in mtDNA data (HVI+Cyt-b), nor in nDNA (*RAG2*). Considering mtDNA, only 2.99% of variation could be attributed to gender, and F_{ST} was not significant (0.3607, p = 0.3704). In nDNA, similarly, 0.093% of the variation was attributed to gender groups, but was not significant F_{ST} = 0.0635, p- = 0.0861.

Biogeographic analysis

The Mantel test of correlation between geographic and genetic distance was not significant (r = 0.18484, p = 0.9995) (Figure 3a). Barrier vs. 2.2 could place four different barriers between geographic regions: barrier *a* between CEGO and CEMT, barrier *b* between CEMG and CEMT, barrier *c* between three CE regions and AF and barrier *d* between PT and CEGO (Figure 4). When an additional barrier was requested in the software (five barriers), it appeared between CEMT and AM (*e*). CEMG, CEGO and CESP showed no evidence of barriers to gene flow among them.

Discussion

This work described the general patterns of genetic diversity variation in populations of *Myrmecophaga tridactyla* along Brazilian localities. We were able to sample populations from four biomes: Cerrado (CE), Pantanal (PT), Atlantic Forest (AF) and Amazon Forest (AM), constituting the largest description of giant anteater genetic diversity available in literature to date, with mitochondrial and nuclear markers. Despite low amplification efficiency and limited sampling, the nDNA dataset corroborated largely the patterns revealed by mtDNA regarding CE, PT and AF populations.

The population from Pantanal (PT) was the genetically most diverse, followed by CEMG and CEGO, as shown by genetic parameters (haplotype and nucleotide diversity), despite the few individuals sampled. The Pantanal biome is known to be composed by different elements gath-

Table 4 - Pairwise F_{ST} values for all populations analyzed. Significant values ($p < 0.05$) are displayed in parenthesis and highlighted in bold.

Population (N)	CEMG	CEMT	CEGO	CESP	PT	AF	AM
CEMG (N=21)	*						
CEMT (N=4)	**0.17818** **(0.04980)**	*					
CEGO (N=28)	0.00070 (0.41113)	0.15034 (0.07227)	*				
CESP (N=7)	-0.03304 (0.83789)	**0.33211** **(0.02344)**	0.03597 (0.82424)	*			
PT (N=6)	0.03674 (0.21094)	0.24950 (0.06836)	0.08851 (0.08691)	-0.03999 (0.29980)	*		
AF (N=5)	**0.21758** **(0.01074)**	-0.02601 (0.68555)	**0.22282** **(0.00195)**	**0.32280** **(0.00195)**	0.22078 (0.08105)	*	
AM (N=8)	**0.39636** **(0.00000)**	0.18410 (0.10938)	**0.39341** **(0.00000)**	**0.45252** **(0.00098)**	**0.36288** **(0.00098)**	0.05950 (0.29688)	*

Table 5 - AMOVA with groups of *M. tridactyla* populations, and percentage of variation within and between groups.

Source of Variation	AMOVA		% Variation	
	F_{ST} value	P-value	Within Groups/Populations	Between Groups
[CE] [PT] [AF] [AM]	0.28169	< 0.0001	73.59	26.41
[CE+PT] [AF+AM]	0.33571	< 0.0001	69.36	30.64
[CE+PT+AF] [AM]	0.37908	< 0.0001	67.36	32.64
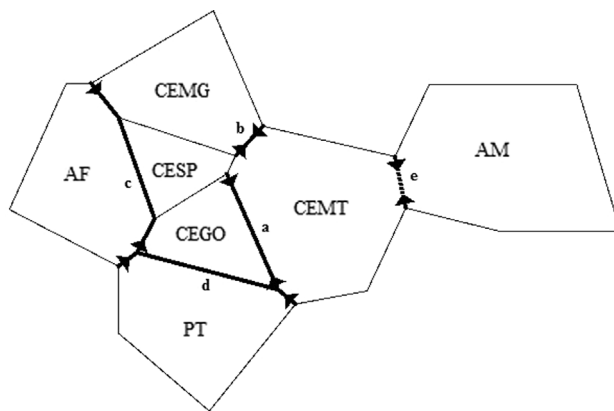
Figure 4 - Diagram of geographic location of populations (indicated by population codes) and the placement of genetic barriers detected. Solid bold lines represent the four barriers requested to software Barrier v 2.2 (*a*, *b*, *c* and *d*), and dashed bold line represent the extra barrier requested (*e*).

ered from other biomes, such as shared fauna and flora (Prance and Schaller, 1982). This miscellaneous nature of the biome itself was likely reflected in *M. tridactyla* genetic diversity. Despite the evidence of a barrier between CE and PT (Figure 4), estimates of F_{ST} between them were low and non-significant (Table 4, Table 5), and there was a high haplotype sharing between them (Figure 2). However, since PT showed some exclusive haplotypes, wider sampling across the biome may be necessary to confirm the nature of its relationship with adjacent populations.

The population with lowest overall diversity was CEMT (for instance, an 8-fold lower ΘS compared to CEGO). Despite being represented by only four individuals, the region where it is located (number 8 in Figure 1) was reported to present anteater hunting activities by indigenous tribes (Leeuwenberg, 1997), and, most importantly, the region constantly suffers from severe habitat loss (Buschbacher, 2000). This population was the only one to present significant F_{ST} values against other CE populations (CEMG and CESP). This differentiation was also evidenced by the barrier analysis, where CEMT showed to be separated from CE populations, but joined with AM when four barriers were considered (the fifth barrier is displayed between them, Figure 4). Moreover, it was the only CE population to share a haplotype with AM and AF (H9). This may reflect the geographic origin of this population, placed between grassland formations (CE) and forest vegetation biomes (AM), and representing, genetically, an intermediate population.

In addition to PT, the two most diverse populations were CEMG and CEGO. Even though they were the most sampled ones, which may cause a bias, they showed the highest numbers of segregating sites and ΘS values (Table 3). The ΘS estimates are made for non-recombining DNA from the relationship between infinite-site equilibrium number of segregating sites and sample size (Watterson, 1975), and, therefore, consider the different number of samples in each population studied. Even though this pa-

rameter showed a higher value for CEMG and CEGO, corroborating their higher genetic variability. CEMG and CEGO populations are composed mostly by specimens from the National Park Serra da Canastra and National Park Emas, respectively. These are Conservation Units in Brazil, and they may represent strongholds for the species diversification, once populations are kept protected from external anthropic disturbance. The Cerrado biome, where these populations are located, has been suffering a severe and accelerated decline in its range, due mostly to the spread in agriculture borders, especially for soy bean and sugarcane production (Ratter *et al.*, 1997). The Cerrado is also considered a biodiversity hotspot for global conservation, with less than 20% of its original range left (Myers *et al.*, 2000). Indeed, these populations are highly threatened by habitat loss, and even protected areas of the Cerrado suffer from wildfires at a regular basis (either natural fires or anthropically originated ones). At times, most of the park's vegetation coverage is burned. In 1994 the entire National Park of Emas was burned, and Silveira *et al.* (1999) estimated that about 332 anteaters were killed. This factor may play a major role in the populations' diversity, once such successive bottlenecks are responsible for genetic diversity loss. Nevertheless, when Collevatti *et al.* (2007) studied the CEGO population (most individuals overlap between this study and ours) they found a high level of inbreeding and low levels of polymorphism in microsatellite loci. This outcome was also attributed to the wildfire effects. The CEMG and CEGO populations also showed significant negative values of Tajima's D, corroborating a possible scenario of repeated bottlenecks over time, followed by expansions during population recovery. Many factors place these populations in constant threats, and they should be protected for long-term maintenance of their genetic diversity. They may be suitable sources of individuals for recolonization of other populations in the vicinities, as most populations in the Cerrado (CE) showed no significant genetic distance among them.

An evident genetic differentiation was detected between CE and PT populations and AM and AF. The result of Mantel's test (non-significant correlation) suggests that IBD may be not the main factor generating genetic structuration in *M. tridactyla*. The significant F_{ST} values suggest that there is a barrier preventing some level of gene flow between these groups of populations, a hypothesis supported by barrier analysis with the Monmoniers algorithm.

At the same time, individuals from the AM population were most closely related to individuals from AF, as demonstrated in the mtDNA haplotype network (Figure 2), and by population pairwise F_{ST} values (Table 4). Even though these individuals are separated by a large distance represented mainly by the dry diagonal (Caatinga, Cerrado, Pantanal and Chaco), they both come from forest formations, which could suggest some adaptive constraints af-

fecting population distribution. The Cerrado biome may have played a historical role as a barrier to connectivity between the two forest formations. At the same time, Cerrado vegetation also shows several fragments of deciduous and semi-deciduous forests, as well as gallery forest that constitutes a net of connections between Atlantic and Amazon biomes (Oliveira-Filho and Ratter, 1995; Vivo, 1997, Costa, 2003). Furthermore, Bigarella *et al.* (1975) suggested that both rainforests were possibly continuous in the past, and this ancient bridge can also explain the relationship between the individuals sampled from both forest biomes, as it has been reported in other studies (*e.g.* Cortés-Ortiz *et al.*, 2003; Costa, 2003; Martins *et al.*, 2007). To discuss this issue in more detail, a wider sampling covering these biomes and intermediate areas is needed.

The results presented in this study contributed to the understanding of the evolutionary history and population dynamics of the threatened giant anteater. Our data pointed out the importance of *M. tridactyla* populations of the Serra da Canastra and Emas National Parks as strongholds of diversity, an important source for future management actions for the species. Besides, it showed a marked genetic structure between Cerrado and Amazon and Atlantic forests populations, representing a historical break to gene flow, and high genetic similarity between Cerrado and Pantanal individuals. We encourage further studies with widespread populations of this species, including specimens from other biomes outside Brazil, in order to better understand its phylogeographic history and to be able to compare diversity indexes among such populations, providing useful information for conservation actions towards the species at a continental level.

Acknowledgments

References

Arnason U, Gullberg and Janke A (2002) Phylogenetic analyses of mitochondrial DNA suggest a sister group relationship between Xenarthra (Edentata) and Ferungulates. Mol Biol Evol 14:762-768.

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

Bergallo HG, Rocha CFD, Alves MAS and Sluys MV (2000) A Fauna Ameaçada de Extinção do Estado do Rio de Janeiro. Ed. UERJ, Rio de Janeiro, 648 p.

Bigarella JJ, Andrade-Lima D and Riehs PJ (1975) Considerações a respeito das mudanças paleoambientais na distribuição de

algumas espécies vegetais e animais no Brasil. An Acad Bras Ciênc 47:411-464.

Buschbacher R (2000) Expansão Agrícola e Perda da Biodiversidade no Cerrado. WWF Brasil, Brasília, 99 p.

Cherem JJ, Simões-Lopes SA and Graipel ME (2004) Lista dos mamíferos do Estado de Santa Catarina, sul do Brasil. Mastozool Neotrop 11:151-184.

Collevatti RG, Leite KCE, Miranda GHB and Rodrigues FHG (2007) Evidence of high inbreeding in a population of the endangered giant anteater, *Myrmecophaga tridactyla* (Myrmecophagidae), from Emas National Park, Brazil. Genet Mol Biol 120:112-120.

Cortés-Ortiz L, Bermingham E, Rico C, Rodríguez-Luna E, Sampaio I and Ruiz-García M (2003) Molecular systematics and biogeography of the Neotropical monkey genus, *Alouatta*. Mol Phylogenet Evo 26:64-81.

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

Cunha HF, Moreira FGA and Silva SS (2010) Roadkill of wild vertebrates along the GO-060 road between Goiânia and Iporá, Goiás State, Brazil. Acta Scient Biol Sci 32:257-263.

Diniz MF and Brito D (2013) Threats to and viability of the giant anteater, *Myrmecophaga tridactyla* (Pilosa: Myrmecophagidae), in a protected Cerrado remnant encroached by urban expansion in central Brazil. Zoologia 30:151-156.

Douzery E and Randi E (1997) The mitochondrial control region of Cervidae: Evolutionary patterns and phylogenetic content. Mol Biol Evol 14:1154-1166.

Eisenberg JF and Redford KH (1999) Mammals of the Neotropics. v. 3. The Central Neotropics: Ecuador, Peru, Bolivia, Brazil. University of Chicago Press, Chicago, 609 p.

Ewing B and Green P (1998) Base-calling of automated sequencer traces using Phred. II. error probabilities. Genome Res 8:186-194.

Excoffier L and Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564-567.

Fallabrino A and Castiñeira E (2006) Situacion de los Edentados en Uruguay. Edentata 7:1-3.

Ferreira FS, Fernandes-Ferreira H, Léo Neto NA, Brito SV and Alves RRN (2013) The trade of medicinal animals in Brazil: Current status and perspectives. Biodivers Conserv 22:839-870.

Fonseca GAB, Herrmann G and Leite YLR (1999) Macrogeography of Brazilian mammals. In: Eisenberg JF and Redford KH (eds) Mammals of the Neotropics: The Central Neotropics, Ecuador, Peru, Bolivia, Brazil. Vol. 3. The University of Chicago Press, Chicago, pp 549-563.

Fonseca GAB and Aguiar JM (2004) Species summaries and species discussions. Edentate Specialis Group. Edentata 6:3-26.

Fontana CS, Bencke GA and Reis RE (2003) Livro Vermelho da Fauna Ameaçada de Extinção no Rio Grande do Sul. EDIPUCRS, Porto Alegre, 632 p.

Frankham R (2005) Genetics and extinction. Biol Cons 126:131-140.

Frankham R, Ballou JD and Briscoe DA (2002) Introduction to Conservation Genetics. Cambridge University Press, Cambridge, 221 p.

Garcia JE, Vilas Boas LA, Lemos MVF, Macedo Lemos G and Contel EPB (2005) Identification of microsatellite DNA markers for the Giant Anteater *Myrmecophaga tridactyla*. J Hered 96:600-602.

Gardner AL (2005) Order Pilosa. In: Wilson DE and Reeder DM (eds) Mammal Species of the World: A Taxonomic and Geographic Reference. 3rd edition The John Hopkins University Press, Baltimore, pp 98-102.

Gilpin MS and Soulé ME (1986) Minimum viable populations: Processes of species extinction. In: Soule ME (ed) Conservation Biology: The Science of Scarcity and Diversity. Sinauer Associates Inc., Massachusetts, pp 19-34.

Gordon D, Abajian C and Green P (1998) Consed: A graphical tool for sequence finishing. Genome Res 8:195-202.

Higgins DG and Sharp PM (1988) CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. Gene 73:237-244.

Holland MM, Cave CA, Holland CA and Billie TW (2003) Development of a quality, high throughput DNA analysis procedure for skeletal samples to assist with the identification of victims from the World Trade Center attacks. Croat Med J 44:264-272.

IBAMA (2003) Lista das Espécies da Fauna Brasileira Ameaçadas de Extinção. Instrução Normativa n. 3, de 27 de maio de 2003, do Ministério do Meio Ambiente. Ministério do Meio Ambiente, Brasília.

Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablanca FX and Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. Proc Natl Acad Sci U S A 86:6196-6200.

Lande R and Shannon S (1996) The role of genetic variation in adaptation and population persistence in a changing environment. Evolution 50:434-437.

Lara-Ruiz P, Chiarello A and Santos FR (2008) Extreme population divergence and conservation implications for the rare endangered Atlantic Forest sloth, *Bradypus torquatus* (Pilosa: Bradypodidae). Biol Cons 141:1332-1342.

Leeuwenberg F (1997) Edentata as a food resource: Subsistence hunting by Xavante Indians, Brazil. Edentata 3:4-5.

Librado P and Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451-1452.

Lorenzutti R and Almeida AP (2006) A coleção de mamíferos do Museu Elias Lorenzutti em Linhares, Espírito Santo, Brasil. Bol Mus Biol Prof Mello Leitão 19:59-74.

Manni F, Guérard E and Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by using Monmonier's algorithm. Hum Biol 76:173-190.

Martins F, Ditchfield A, Meyer D and Morgante JS (2007) Mitochondrial DNA phylogeography reveals marked population structure in the common vampire bat, *Desmodus rotundus* (Phyllostomidae). J Zool Syst Evol Res 45:372-378.

McNab BK (1985) Energetics, population biology and distribution of Xenarthrans, living and extinct. In: Montgomery GG (ed) The Evolution and Ecology of Armadillos, Sloths, and Vermilinguas. Smithsonian Institution Press, Washington, pp 219-232.

Miller MP (2005) Alleles in space (AIS): Computer software for the joint analysis of interindividual spatial and genetic information. J Hered 96:722-724.

Miranda F, Bertassoni A and Abba AM (2014) *Myrmecophaga tridactyla*. The IUCN Red List of Threatened Species: e.T14224A47441961.

Miranda, GHB, Tomas W, Padua CV and Rodrigues FHG (2006) Giant anteater (*Myrmecophaga tridactyla*) population survey in Emas National Park, Brazil - A proposed monitoring program. Endangered Species Update 23:96.

Monmonier M (1973) Maximum-difference barriers: An alternative numerical regionalization method. Geogr Anal 3:245-261.

Myers N, Mittermeier RA, Mittermeier CG, Fonseca GAB and Kent J (2000) Biodiversity hotsposts for conservation priorities. Nature 403:853-858.

Nickerson DA, Tobe VO and Taylor SL (1997) PolyPhred: Automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. Nucleic Acids Res 25:2745-2751.

Oliveira-Filho AT and Ratter JA (1995) A study of the origin of central Brazilian forests by the analysis of plant species distribution patterns. Edinburgh J Bot 52:141-194.

Peres CA (2001) Synergistic effects of subsistence hunting and habitat fragmentation on amazonian forest vertebrates. Conserv Biol 15:1490-1505.

Prance TP and Schaller GB (1982) Preliminary study of some vegetation types of the Pantanal, Mato Grosso, Brazil. Brittonia 34:228-251.

Ratter JA, Ribeiro JF and Bridgewater S (1997) The Brazilian Cerrado vegetation and threats to its biodiversity. Ann Bot 80:223-230.

Roca AL, Georgiadis N and O'Brien SJ (2005) Cytonuclear genomic dissociation in African elephant species. Nat Genet 37:96-100.

Rogers AR and Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9:552-569.

Saccone C, Gissi C, Lanave C, Larizza A, Pesole G and Reyes A (2000) Evolution of the mitochondrial genetic system: An overview. Gene 261:153-159.

Sambrook E, Fritsch F and Maniatis T (1989) Molecular Cloning. Cold Spring Harbor Press, New York, 1626 p.

Santos-Junior JE, Santos FR and Silveira FA (2015) Hitting an Unintended Target: Phylogeography of Bombus brasiliensis Lepeletier, 1836 and the First New Brazilian Bumblebee Species in a Century (Hymenoptera: Apidae). Plos One 10:e0125847.

Shaw JH, Carter TS and Machado-Neto JC (1985) Ecology of the giant anteater *Myrmecophaga tridactyla* in Serra da Canastra, Minas Gerais, Brasil: A pilot study. In: Montgomery GG (ed) The Evolution and Ecology of Armadillos, Sloths, and Vermilinguas. Smithsonian Institution Press, Washington, pp 379-384.

Silveira L, Almeida Jacomo AT, Diniz Filho JAF and Rodrigues FHG (1999) Impact of wildfires on the megafauna of Emas National Park, Central Brazil. Oryx 33:108-114.

Stephens M and Donnely P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 73:1162-1169.

Stephens M, Smith NJ and Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978-989.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.

Tamura K, Dudley J, Nei M and Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software v. 4.0. Mol Biol Evol 24:1596-1599.

Teeling EC, Madsen O, Van Den Busshe RA, de Jong WW, Stanhope MJ and Springer MS (2001) Microbat paraphyly and the convergent evolution of a key innovation in Old World rhinolophoid microbats. Proc Natl Acad Sci U S A 99:1431-1436.

Vivo M (1997) Mammalian evidence of historical ecological change in the Caatinga semiarid vegetation of northeastern Brazil. J Comp Biol 2:65-73.

Ward RH, Frazier BL, Dew-Jager K and Pääbo S (1991). Extensive mitochondrial diversity within a single Amerindian tribe. Proc Natl Acad Sci U S A 88:8720-8724.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. Theor Popul Biol 7:256-276.

Internet Resources

Miranda FR, Chiarello AG, Röhe F, Braga FG, Mourão GM, Miranda GHB, Silva KFM, Faria-Corrêa MA, Vaz SM and Belentani SCS (2015) Avaliação do Risco de Extinção de *Myrmecophaga tridactyla* Linnaeus, 1758 no Brasil. Processo de avaliação do risco de extinção da fauna brasileira. ICMBio. http://www.icmbio.gov.br/portal/biodiversidade/fauna-brasileira/lista-de-especies/7049-mamiferos-myrmecophaga-tridactyla-tamandua-bandeira.html

Supplementary Material

The following online material is available for this article:

Table S1 – Municipalities (or national parks) in Brazil corresponding to numbers of localities indicated in the map (Figure 1 in text). The populations to which the localities belong are also indicated

Associate Editor: Guillermo Ortí