

Spatial bias in sampling small rodents in the Atlantic Forest: a landscape and accessibility perspective

Lana Pavão Candelária (✉ lanapavao@gmail.com)

Universidade Federal de Mato Grosso

Nicolas Silva Bosco

Universidade Federal de Mato Grosso

Viviane Maria Guedes Layme

Universidade Federal de Mato Grosso

Thadeu Sobral-Souza

Universidade Federal de Mato Grosso

Domingos Jesus Rodrigues

Universidade Federal de Mato Grosso

Research Article

Keywords: Biodiversity, Atlantic Forest, Inventory completeness, Macroecology, Wallacean Shortfall

Posted Date: May 18th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1654418/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Understanding the effects of habitat loss on species spatial distribution is a challenge mainly because the species occurrences knowledge is biased. Some sites are more sampled than others causing spatial sampling biases that are normally neglected. The needs to assess the sampling biases urge particularly in threatened regions, as the global hotspots. Here, we aimed to assess the spatial biases in the samplings of small rodents in the Atlantic Forest using an integrative approach toward accessibility and landscape metrics. We hypothesized that the well-sampled sites are spatially aggregated and closed to roads, urban centers, and larger fragments with a major percentage of habitat cover. First, we found 9910 species occurrence records for the Atlantic Forest including information from digital platforms and local communities databases. We also calculated the inventory completeness of small rodents achieving 507 sites as well-sampled sites that cover less than 1% of the Atlantic Forest spatial surface (in 1km² cell-size resolution). Spatially the well-sampled sites are aggregated, biased toward roads, urban centers, larger fragments and habitat with major forest cover percentage. Our findings suggest that even though the Atlantic Forest is highly diverse the spatial sampling knowledge of small rodents remains absent for most of its spatial extension. These results may be important to biome conservation tasks and sampling planning in still unknown regions.

Introduction

Human-induced activities have modified landscapes directly affecting the spatial species distribution worldwide (Ceballos and Ehrlich 2002; Chase et al. 2020). Currently, the natural habitat modification is recognized as a factor causing defaunation (Dirzo et al. 2014), species population decreasing and/or species extinction (Dirzo and Raven 2003; Hanski 2011). Understanding how human-induced habitat changes drive biodiversity loss is still a challenge for science, because there are many gaps in the biodiversity knowledge which need to be better understood and quantified (Hortal et al. 2015).

The most studied biodiversity knowledge gap refers to spatial biases in the species distribution pattern (Wallacean Shortfall) since the primary biodiversity data on species occurrence are still scarce or biased (Girardello et al. 2018). The Wallacean Shortfall is characterized by uneven variation of the species sampling effort in time and space. Some localities are temporally more sampled and very sites are unsampled, generating a spatial aggregated pattern in the biodiversity knowledge (Almeida et al. 2021; Hortal et al. 2015; Sobral-Souza et al. 2021a). Like any other shortfalls, it is dependent on the spatial scale and resolution at which it is evaluated (Hortal et al. 2015, Bosco et al. 2022 in press). Normally, the bias in local scale occur because the sampling tends to be realized closed to access routes, either terrestrial or aquatic ecosystems (Oliveira et al. 2016; Sousa-Baena et al. 2014) and close to urban centers (Almeida et al. 2021; Lessa et al. 2019). Beyond accessibility-related biases there are also landscape-related biases, including more samples towards large and connected forest fragments (Sobral-Souza et al. 2021a) and/or in protected areas (Almeida et al. 2021). These biases are explained by the interest of researchers in sampling a greater number of species (Reddy and Dávalos 2003; Sastre and Lobo 2009) in a short period of time, especially in intensively modified landscapes.

Global initiatives emerged to solve the problems of limited biodiversity data, such as making biodiversity data available on accessible digital platforms (for example: Global Biodiversity Information Facility – GBIF e SpeciesLink). The digital biodiversity knowledge has been important to researchers' understanding of patterns in the spatial biodiversity distribution. However, gaps and biases on biodiversity digital data can compromise descriptions of large-scale biodiversity patterns (Hortal et al. 2007) problematizing conservation tasks (Brooks et al. 2004; Sousa-Baena et al. 2014). The spatial bias can also distort the estimated patterns (Hortal et al. 2008; Boakes et al. 2010; Yang et al. 2013) and impairs the use of niche-based modeling techniques (Hortal et al. 2008), that also emerged as a technique to solve the problem to use limited biodiversity data (Beck et al. 2014; Rocchini et al. 2011). The limitations of biodiversity data are most prominent for impacted and biodiverse e regions, such as the global hotspots.

The Atlantic Forest is a biodiversity hotspot (Myers et al. 2000) that even with a high biodiversity index still lacks biodiversity knowledge (Collen et al. 2008; Dirzo et al. 2014). Due to its history of colonization and natural resources exploitation, the number of endemic species and the presence of large urban centers with many research institutions, the Atlantic Forest is the most studied brazilian biome (Candelária et al. 2021; Teixeira et al. 2020). Even so, the Atlantic Forest's biodiversity knowledge is biased for some regions (Sobral-Souza et al. 2021a). Testing the biases and gaps of Atlantic Forest's biodiversity knowledge using well resolved taxonomic groups, with extensive information of species distribution and inventories sampling, such as the non-volant small mammals, can be an interesting step to improve the Atlantic Forest conservation strategies.

The population reduction and/or species loss of large mammals induced by the modification of natural habitats (Dirzo et al. 2014) increases the taxon vulnerability (Galetti et al. 2017). Small rodents are widely distributed taxa (Wilson and Reeder 2005) and in the Atlantic Forest is considered a diverse group, with 94 species (Bovendorp et al. 2017; Galindo-Leal and Câmara 2003). Ordinarily, rodent species are generalists and can occupy different habitat types (Bonvicino et al. 2008). The taxa also present different responses to local habitat changes (Bergallo et al. 2005; Rosa et al. 2018), because they are dependent on habitat quality (Delciellos et al. 2016). Even though the Atlantic Forest is the most sampled biome for small rodents (Candelária et al. 2021), the sampling bias and spatial gaps have not yet been evaluated.

In fact, problems related to biodiversity data quality and data limitation occur in all domains of ecology (Hortal et al. 2015). Therefore, identifying biodiversity data gaps and biases are important to indicate priority areas for sampling that can help us to understand large-scale biodiversity patterns. Here, we aimed to evaluate the biases and gaps in the spatial distribution of Atlantic Forest's small rodent samples. Specifically, we quantified the biodiversity knowledge of small rodents for the Atlantic Forest; we also evaluated the pattern of spatial distribution of rodent samplings to test how accessibility and landscape metrics can explain the sampling biases and gaps. We expect that the rodent samples are spatially biased towards more accessible regions, closed to urban centers and roads. The small rodents show different responses to landscape modification being used as good models to test hypotheses related to biodiversity and species loss. Based on this, we expect sampling bias regarding the fragment

size (ha) and forest cover in which larger fragments with a higher forest cover will be more sampled than others.

Methods

Study area

The Atlantic Forest extends from the northeastern to the southern of the Brazilian coast. To the southwestern it extends into a small part of Paraguay and Argentina (Ribeiro et al. 2009; Tabarelli et al. 2010). The historical and geological context of the Atlantic Forest explain the environmental heterogeneity that contributed to a high species richness and high rates of endemism (Ribeiro et al. 2009). Silva and Casteleti (2003) proposed some biogeographic subdivisions for the Atlantic Forest, namely: Serra do Mar, Pernambuco, Bahia, Northeastern Brejos, Diamantina, Interior Forests, Araucaria Forests and São Francisco. These subregions are characterized by the presence of two or more endemic species with overlapping spatial distributions and/or by the specificities of some habitats (Silva and Casteleti 2003). As for their phytophysionomies, the Atlantic Forest is mainly composed of Dense, Mixed and Open Ombrophilous Forests and Seasonal Semideciduous and Deciduous Forests in the ecotonal areas (Joly et al. 2014).

The Atlantic Forest is also known for its high biodiversity and long history of deforestation (Joly et al. 2014). Since its colonization (more than 500 years) it has been historically threatened by the expansion of human occupation and land use conversion (Cincotta et al. 2000; Metzger 2009; Tabarelli et al. 2010) causing landscape fragmentation. Currently the Atlantic Forest is located in one of the most populous regions of Brazil where we find the three main Brazilian urban centers (Tabarelli et al. 2010). Even today, part of the Atlantic Forest is lost due to population growth and the expansion of logistic infrastructure such as roads (Joly et al. 2014). Currently, 28% of the Atlantic Forest is composed of small and isolated forest remnants (Rezende et al. 2018). Among the biogeographic subregions, Serra do Mar is the most conserved (about 36,5% from their area), while the Interior and São Francisco sub-regions are the most degraded (< 8% of remaining forest area - Ribeiro et al. 2009).

Species dataset

We compile occurrences of small rodent species from the Families Caviidae, Cricetidae, Ctenomyidae, Echimyidae, Erethizontidae, Muridae and Sciuridae. Occurrence records of small rodent species were obtained from different digital available databases. In addition, we also obtained occurrence records from two databases published by Bovendorp et al. (2017) and Figueiredo et al. (2017). These datasets compiled information on richness and abundance of local communities of non-volant small mammals, from scientific articles, thesis and unpublished data. We also use the occurrence data obtained from the *Global Biodiversity Information Facility* (GBIF 2021- DOI <https://doi.org/10.15468/dl.bnbeqo>), speciesLink (2021), Information System on Brazilian Biodiversity (*Sistema de Informação sobre a Biodiversidade Brasileira* - SiBBR 2021) and Biodiversity Portal (*Portal da Biodiversidade* - PortalBio 2021). We also included the occurrence records of six invasive species from the Neotropical Invasive Mammals

database (Rosa et al. 2020). After compiling the database, we filtered the records considering: i) the accuracy of taxonomic identification (records that were not identified at species level were removed), ii) the geographic location within the limits of the Atlantic Forest (Muylaert et al. 2018), excluding occurrences outside this, and iii) deleting duplicate records. All species found on digital platforms were searched for current taxonomy and had their nomenclatures corrected, if necessary. For that, we use the ASM Mammals diversity database da American Society of Mammalogists (ASM, <https://www.mammaldiversity.org/>), Wilson and Reeder's Mammal Species of the World, 3rd edition (MSW3, <http://www.departments.bucknell.edu/biology/resources/msw3/>), and Patton et al. (2015) as a way to obtain the synonymies of the species.

Inventory Completeness

The inventory completeness is used to identify well-sampled sites at a given spatial resolution and extension (Lobo et al. 2018). To assess the inventory completeness, we calculated the completeness for cells with 1 km² resolution in the Atlantic Forest delimitation (Muylaert et al. 2018). We chose this spatial resolution because landscape (re)structure and (re)configuration have been shaped by changes in land-use that occur and affect ecological processes at narrow spatial scales (Benítez-López et al. 2010, Sobral-Souza et al. 2021a).

To calculate the inventory completeness for each 1km² cell-size resolution we subdivided our database in two folding: i) digital databases with records of species occurrence and; ii) the information compiled based on the richness and abundance of local communities. As the information from the communities already measured the sampling effort, this database was once considered to have a high completeness index. Thus, completeness was calculated by only considering individual occurrence records from the GBIF, SpeciesLink, SiBBR, PortalBio and Neotropical Aliens database.

We calculated the inventory completeness through the ratio between the number of observed and predicted species where the number of predicted species for each 1km² resolution cell was obtained through the species accumulation curve using the "Clench" function of the *KnowBr* package (Lobo et al. 2018) in the R program (R Development Core Team 2021). The inventory completeness metric ranges from 0 to 1 with 0 being no inventory and 1 the maximum inventory completeness for each cell. We considered the as well-sampled site when i) it contained more than 20 occurrence records and an inventory completeness equal to or greater than 0.7 and ii) sites with communities sampling according to Bovendorp et al. (2017) and Figueiredo et al. (2017).

Spatial bias

We assess the spatial bias on the well-sampled based on accessibility and landscape metrics. We use as accessibility metrics the i) roads distances (m) and ii) nearest distance to urban center (m). The raster with the Euclidean roads distances was produced through the shapefile of the federal and state roads of the National Department of Transport Infrastructure (*Departamento Nacional de Infraestrutura de Transportes - DNIT*). The raster of Euclidean distances from urban centers was produced by compiling shapes from the Brazilian Foundation for Sustainable Development (*Fundação Brasileira para o*

Desenvolvimento Sustentável - FBDS), SOS Atlantic Forest Foundation (*Fundação SOS Mata Atlântica*) and the Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Estatística - IBGE*).

We also used fragment size (ha) and percentage of habitat cover (%) to assess how landscape configuration can influence spatial biases and gaps in small rodent samplings. To calculate the fragment size each fragment was given an ID (fragment ID) and an area (fragment size or area), calculated in hectares as the sum of the area of each cell. In the final map, each cell of the same fragment presented the same value of the area of the whole fragment. Non-fragment (= non-vegetation or vegetation corridors) cells were given a NULL value. To infer the habitat cover (%) each cell presented a value of the proportion of habitat (forest) within a square window with a given size, centered in the focal cell (= amount of vegetation cells/total number of cells in the window). It varies between 0 and 100%. This metrics were calculated by variables taken from *Fundação Brasileira para o Desenvolvimento Sustentável (FBDS)*, *Fundação SOS Mata Atlântica*, *Instituto Nacional de Pesquisas Espaciais (INPE)* and University of Maryland (Global Forest Change Project, Hansen et al. 2013). All metrics were calculated for resolution of 30m x 30m and resampled to for 1km².

To test the accessibility bias on the well-sampled sites we compared the frequency distribution between the road distance for the entire Atlantic Forest extension with the road distance of the well-sampled sites. We did the same procedure to compare the distribution of frequency of distances for urban centers (between the entire Atlantic Forest and the well-sampled sites). For the landscape metrics we compared the frequency of fragment sizes for the entire Atlantic Forest with the frequency of fragment sizes of the well-sampled sites. We used the same procedure to habitat cover (%) inferring the difference between the frequency of habitat percentage of the entire Atlantic Forest and the well-sampled sites. We used the Komolgorov-Smirnov analysis to test for differences between frequencies of entire Atlantic Forest and well-sampled sites. All analyzes were performed in R (R Development Core Team 2021).

Results

The dataset of small rodents compiled by Bovendorp et al. (2017) resulted in 1499 records, the Figueiredo et al. (2017) dataset has 1115 occurrence records and Rosa et al. (2020) 1186 occurrence records. For digital platforms GBIF, SpeciesLink, SibBr and Portal da Biodiversidade we also found 69657 small rodent occurrence records (see Table s1 and s2 in the supplementary material). Of these, only 9910 were within the Atlantic Forest limits (Fig. 1).

Considering the spatial resolution of 1 km² only 507 sites of Atlantic Forest (out of 1,649,932 cells) were well-sampled. Thus, only 0.03% of the spatial surface of the Atlantic Forest is well sampled for small rodents (Fig. 2). Considering the Atlantic Forest biogeographic subregions, proposed by Ribeiro et al. (2009), Serra do Mar is the best known with 177 well-sampled sites (0.15% of the sub-region area). The Pernambuco and Bahia sub-regions have between 0.7 and 0.9% of their areas well sampled. While the Araucaria, Florestas do Interior and Diamantina sub-regions are known less than 0.05% of their areas. For

the São Francisco sub-region, there are no records of inventories with high completeness of small rodents (Figure s1).

The sampling of small non-volant mammals are biased towards sites close to roads ($D = 0.1234$, $p = 3.948^{-07}$) and urban centers ($D = 0.10877$, $p = 1.236^{-05}$; Fig. 3). In addition to accessibility, the landscape configuration also interferes in the sampling bias. Well-sampled sites are biased towards large fragments ($D = 0.28178$, $p < 2.2^{-16}$) with a higher percentage of habitat cover ($D = 0.41727$, $p\text{-value} < 2.2^{-16}$) (Fig. 4).

Discussion

Our findings suggest a spatial bias on the sampling sites of small rodents in the Atlantic Forest. The well-sampled sites are close to urban centers, roads, large fragments and with greater habitat (forest) cover. These results also indicate a sampling gap in small fragments with low forest cover and sites far from cities and roads, precluding the inferences of relationship between fragmentation and small mammal biodiversity on broad-scale perspectives. In addition, based on data available on digital platforms the spatial aggregation of sampling efforts generated a 99% of spatial gap in the knowledge of small mammals. The well-sampled sites cover only 0.03% of the biome's extension. Our results demonstrate that the small rodent's knowledge in a broad-scale perspective is spatially limited which impairs inferences about the factors that govern and change the spatial distribution, a key factor to propose conservation tasks.

Although the aim of digital accessible platforms is to provide more data on biodiversity to increase the biodiversity knowledge (Devictor and Bensaude-Vincent 2016) and helping to fill some gaps, the current available data is still limited (Hortal et al. 2007). The limitation of biodiversity information occurs because the available data were collected in different ways, by different researchers and with different purposes (Wüest et al. 2020). The imprecision in the Atlantic Forest's small rodent data can be explained by the lack of taxonomic expertise in identifying specimens in the field, by the retention of biodiversity data in museums or undigitized private collections and by the temporal degradation of information (Ladle and Hortal 2013; Meyer et al. 2015; Tessarolo et al. 2017). Additionally, there are imprecisions on species spatial information. The lack of geographic coordinates or their imprecision contributes to the increase of biogeographic ignorance. This imprecision makes part of the information "unusable" for some analyzes (Hortal et al. 2008) such as inventory completeness. Therefore, digitally accessible data, but with inaccurate information, also increases the biogeographic ignorance (Tessarolo et al. 2021) instead of decreasing them. It would be interesting if there was a validation, monitoring and mapping of biodiversity digital information so that it becomes accurate, useful and reliable. The correct digitization of biodiversity information would allow a better use of available data, cost reduction and promotion of new analyzes and/or research/inventories increasing the possibility of macroecological pattern inferences and conservation efforts based on the relationship between fragmentation and biodiversity loss.

Our findings also suggest a spatial bias from well-sampled sites to more accessible regions. The proximity to access routes and urban centers are well-known factors in generating bias in the biodiversity

knowledge from different taxonomic groups (Almeida et al. 2021; Correia et al. 2019; Oliveira et al. 2016). The historic natural exploitation of the Atlantic Forest resources causes forest loss resulting in only 28% in forest original remnants (Rezende et al. 2018). The modification in the Atlantic Forest landscape resulted in a fragmented landscape with a high density of access routes and many cities (Lapola et al. 2014; Oliveira et al. 2016), including also the Protected Areas that are easily accessible by roads. The Atlantic Forest protected areas are spatially joint embracing the largest and conserved forest remnants that being easily accessible to biodiversity sampling results on broad-scale spatial biases on both accessibility and landscape perspectives. In addition, historically, the promotion of biological research in the Atlantic Forest has favored projects on a local scale and with long-term sampling (for example PELDs). These projects were able to clarify and describe diverse ecological and biodiversity patterns on a local scale, however failed to clarify patterns on a broad-scale which can also be explained by the recent shift in the macroecology view (hierarchical approaches - Guisan and Rahbek (2011), Sobral-Souza et al. (2021b). There are still a density of sites with a distance greater than 15 km from urban centers and access routes that have not been sampled, becoming priority sites for future sampling efforts.

The disadvantage of biodiversity sampling close to roads and cities is that it can be underestimated mainly because these sites may not represent regional biodiversity (Benítez-López et al. 2010; Laurance et al. 2009) and may have the effect of biological invasion (Hobbs et al. 2009). Access routes facilitate habitat degradation (Freitas et al. 2010) and for small rodent changes on landscape configuration such as edge effect, change the species composition (Rosa et al. 2018). On the other hand, sampling sites close to large centers and roads offers better infrastructure, more qualified human resources and reduces costs for researchers with few financial support (Lessa et al. 2019; Meyer et al. 2015).

We also found bias for larger fragments and with greater habitat (forest) cover. These biases may be explained by the access facility, the greater environmental heterogeneity of larger fragments and species-area ratio which are more likely to record a greater number of species in a few sampling times (Sobral-Souza et al. 2021a). Additionally, in the Atlantic Forest, larger fragments are more temporally stable and have less effects caused by human-induced landscape modification (Hansen et al. 2020). The human colonization and habitat fragmentation result in a spatial configuration where the main forest remnants are on sites closed to high human density (Ribeiro et al. 2009). Thus the largest forest fragments are spatially joint explaining the spatial aggregation pattern of well-sampled sites. In the Atlantic Forest the main forest fragments are also protected areas generating long term habitat stability and providing long-lasting, temporal samplings with little human action.

The Serra do Mar is the most conserved biogeographic region (Ribeiro et al. 2009) with 35% of Atlantic Forest well-sampled sites within this region, covering 0.15% of the spatial surface of this region. The Serra do Mar are located in the main large protected forest massifs (such as Serra do Mar State Park and Carlos Botelho State Park), however, it is a geographically extensive region that has large urban centers, such as São Paulo city. Our findings highlight that the Floresta do Interior is six times larger than the Serra do Mar and the biodiversity knowledge is also smaller, with 0.02% of surface well studied. In the Floresta do Interior, the sampling sites are spatially joint to the Serra do Mar bioregion, probably due to

the presence of large urban centers and researchers Institutes (Almeida et al. 2021; Candelária et al. 2021; Lessa et al. 2019). The biogeographic regions of the northeast are practically unknown, especially São Francisco, which does not have well-sampled sites. These northeast regions are classified by the degree of endemism, which attracts taxonomists to discover new species (Silva-Soares et al. 2021) but even so areas with high rates of endemism are poorly sampled, such as Diamantina, Bahia and Pernambuco. This can be explained by the low number of universities and Protected areas in the region. The northeast is also known as an ecotone, having a vegetational mosaic with Caatinga and Cerrado enclaves which can increase the complexity to sampling Atlantic Forest taxa.

Biased biodiversity data also hamper the building of species distribution maps because the maps will reflect the sampling bias and not the true species distribution (Rocchini et al. 2011; Ladle and Hortal 2013). Landscape configuration also affects species dispersion (Fahrig 2005) and not knowing the true species distribution based on their sampling bias may be to impairs inferences about the effects of landscape modifications on local and regional biodiversity (Sobral-Souza et al. 2021b, Santos et al. 2020). Like any other deficit on biodiversity knowledge, the Wallacean Shortfall is dependent on the spatial scale on which it is assessed. At high spatial resolution for a wide spatial scale the knowledge of biodiversity can be considered complete as all grid cells would have enough occurrences to reach a high degree of completeness (Lobo et al. 2018). But at low spatial resolution, as here, we infer that there are many spatial gaps in the knowledge of the biodiversity of small mammal in the Atlantic Forest.

Currently, the biodiversity knowledge of small rodents in the Atlantic Forest is insufficient to understand how landscape modification affects the spatial species distribution on a large scale. The bias highlighted here demonstrated a biodiversity knowledge gap for small fragments, with little habitat (forest) coverage and far from roads and cities. Most of the Atlantic Forest area is represented by small fragments, with little forest cover, as they are composed of secondary forests or in forest recovery (Ribeiro et al. 2009). In addition, the well-sampled sites are spatially aggregated, not covering the entire range of accessibility and landscape conditions of the Atlantic Forest. Therefore, samplings of small rodents need to be directed to sites of difficult accessibility, away from cities, in small fragments and with little forest habitat cover in order to contribute to the increase of the inventory completeness through the landscape configuration spectrum mainly due to the dependence on responses that small rodents can present with the landscape human-induced modification.

Declarations

Acknowledgements

We thanks Milton Cezar Ribeiro, Marcos Penhacek, Jessie Pereira Santos and Rodrigo Pirilo for manuscript suggestions. We also thanks the MacrEco Lab students for the valuable discussions. LPC and NSB thanks the CAPES for fellowship.

Author contributions

L.P.C; V.M.G.L and D.J.R concern the manuscript idea

L.P.C; N.S.B and T.S-S building the dataset and run the analysis

All authors reviewed the manuscript

Funding

The Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) granted PhD scholarships to LPC.

Competing Interests

Not declare.

References

1. de Almeida TC, Tessarolo G, Nabout JC, Teresa FB (2021) Non-stationary drivers on fish sampling efforts in Brazilian freshwaters. *Divers Distrib* 27:1224–1234. <https://doi.org/10.1111/ddi.13269>
2. Beck J, Böller M, Erhardt A, Schwanghart W (2014) Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecol Inf* 19:10–15. <https://doi.org/10.1016/j.ecoinf.2013.11.002>
3. Benítez-López A, Alkemade R, Verweij PA (2010) The impacts of roads and other infrastructure on mammal and bird populations: A meta-analysis. *Biol Conserv* 143:1307–1316. <https://doi.org/10.1016/j.biocon.2010.02.009>
4. Bergallo HG, Luz JL, Raíces DS, Hatano FH, Martins-Hatano F (2005) Habitat use by *Oryzomys subflavus* (Rodentia) in an open shrubland formation in Restinga de Jurubatiba National Park, RJ, Brazil. *Braz J Biol* 65:583–588. <https://doi.org/10.1590/s1519-69842005000400004>
5. Boakes EH, McGowan PJK, Fuller RA, Chang-Qing D, Clark NE, O'Connor K, Mace GM (2010) Distorted views of biodiversity: Spatial and temporal bias in species occurrence data. *PLoS Biol* 8. <https://doi.org/10.1371/journal.pbio.1000385>
6. Bonvicino CR, Oliveira JAde, D'Andrea OS (2008) Guia dos roedores do Brasil, com chaves para gêneros baseadas em caracteres externos. Centro Pan-Americano de Febre Aftosa-OPAS/OMS, Rio de Janeiro
7. Bovendorp RS, Villar N, de Abreu-Junior EF, Bello C, Regolin AL, Percequillo AR, Galetti M (2017) Atlantic small-mammal: a dataset of communities of rodents and marsupials of the Atlantic forests of South America. *Ecology* 98:2226. <https://doi.org/10.1002/ecy.1893>
8. Brooks TM, Da Fonseca GAB, Rodrigues ASL (2004) Protected areas and species. *Conserv Biol* 18:616–618. <https://doi.org/10.1111/j.1523-1739.2004.01836.x>
9. Candelária LP, Zucchetto M, Gonçalves SRA, Braga GDSF, Izzo TJ, Teixido AL (2021) Biogeographic and fragmentation-related research biases on antbirds and non-flying small mammals in Brazil. *J*

- Trop Ecol 37:175–184. <https://doi.org/10.1017/S0266467421000250>
10. Ceballos G, Ehrlich PR (2002) Mammal population losses and the extinction crisis. *Science* 296:904–907. <https://doi.org/10.1126/science.1069349>
 11. Correia RA, Ruete A, Stropp J, Malhado ACM, dos Santos JW, Lessa T, Alves JA, Ladle RJ (2019) Using ignorance scores to explore biodiversity recording effort for multiple taxa in the Caatinga. *Ecol Indic* 106:105539. <https://doi.org/10.1016/j.ecolind.2019.105539>
 12. Chase JM, Blowes SA, Knight TM, Gerstner K, May F (2020) Ecosystem decay exacerbates biodiversity loss with habitat loss. *Nature* 584:238–243. <https://doi.org/10.1038/s41586-020-2531-2>
 13. Cincotta RP, Wisnewski J, Engelman R (2000) Human population in the biodiversity hotspots. *Nature* 404:990–992. <https://doi.org/10.1038/35010105>
 14. Collen B, Ram M, Zamin T, McRae L (2008) The Tropical Biodiversity Data Gap: Addressing Disparity in Global Monitoring. *Trop Conserv Sci* 1:75–88. <https://doi.org/10.1177/194008290800100202>
 15. Devictor V, Bensaude-Vincent B (2016) From ecological records to big data: the invention of global biodiversity. *Hist Philos Life Sci* 38:1–23. <https://doi.org/10.1007/s40656-016-0113-2>
 16. Delciellos AC, Vieira MV, Grelle CEV, Cobra P, Cerqueira R (2016) Habitat quality versus spatial variables as determinants of small mammal assemblages in Atlantic Forest fragments. *J Mammal* 97:253–265. <https://doi.org/10.1093/jmammal/gyv175>
 17. Dirzo R, Raven PH (2003) Global state of biodiversity and loss. *Annu Rev Environ Resour* 28:137–167. <https://doi.org/10.1146/annurev.energy.28.050302.105532>
 18. Dirzo R, Young HS, Galetti M, Ceballos G, Isaac NJB, Collen B (2014) Defaunation in the Anthropocene. *Science* 345:401–406. <https://doi.org/10.1126/science.1251817>
 19. Fahrig L (2005) When is a landscape perspective important? In: Wiens J, Moss M (eds) *Issues and Perspectives in Landscape Ecology* (Cambridge Studies in Landscape Ecology). Cambridge University Press, Cambridge, pp 3–10. doi:10.1017/CBO9780511614415.002
 20. Figueiredo MSL, Barros CS, Delciellos AC, Guerra EB, Cordeiro-Estrela P, Kajin M, Alvarez MR, Asfora PH, Astúa D, Bergallo HG, Cerqueira R, Geise L, Gentile R, Grelle CEV, Iack-Ximenes GE, Oliveira LC, Weksler M, Vieira MV (2017) Abundance of small mammals in the Atlantic Forest (ASMAF): a data set for analyzing tropical community patterns. *Ecology* 98:2981–2981. <https://doi.org/10.1002/ecy.2005>
 21. Freitas SR, Hawbaker TJ, Metzger JP (2010) Effects of roads, topography, and land use on forest cover dynamics in the Brazilian Atlantic Forest. *For Ecol Manage* 259:410–417. <https://doi.org/10.1016/j.foreco.2009.10.036>
 22. Galetti M, Brocardo CR, Begotti RA, Hortenci L, Rocha-Mendes F, Bernardo CSS, Bueno RS, Nobre R, Bovendorp RS, Marques RM, Meirelles F, Gobbo SK, Beca G, Schmaedecke G, Siqueira T (2017) Defaunation and biomass collapse of mammals in the largest Atlantic forest remnant. *Anim Conserv* 20:270–281. <https://doi.org/10.1111/acv.12311>
 23. Galindo-Leal C, Câmara IG (2003) *The Atlantic Forest of South America: biodiversity status, threats, and outlook*. Center for Applied Biodiversity Science at Conservation International State of the

Hotspots series. Island Press, Washington

24. Girardello M, Martellos S, Pardo A, Bertolino S (2018) Gaps in biodiversity occurrence information may hamper the achievement of international biodiversity targets: Insights from a cross-taxon analysis. *Environ Conserv* 45:370–377. <https://doi.org/10.1017/S0376892918000115>
25. Guisan A, Rahbek C (2011) SESAM – a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. *J Biogeogr* 38:1433–1444. doi:10.1111/j.1365-2699.2011.02550.x
26. Hansen MC, Wang L, Song XP, Tyukavina A, Turubanova S, Potapov PV, Stehman SV (2020) The fate of tropical forest fragments. *Sci Adv* 6:11. <https://www.science.org/doi/full/10.1126/sciadv.aax8574>
27. Hanski I (2011) Habitat loss, the dynamics of biodiversity, and a perspective on conservation. *Ambio* 40:248–255. <https://doi.org/10.1007/s13280-011-0147-3>
28. Hobbs RJ, Higgs E, Harris JA (2009) Novel ecosystems: implications for conservation and restoration. *Trends Ecol Evol* 24:599–605. <https://doi.org/10.1016/j.tree.2009.05.012>
29. Hortal J, De Bello F, Diniz-Filho JAF, Lewinsohn TM, Lobo JM, Ladle RJ (2015) Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. *Annu Rev Ecol Evol Syst* 46:523–549. <https://doi.org/10.1146/annurev-ecolsys-112414-054400>
30. Hortal J, Jiménez-Valverde A, Gómez JF, Lobo JM, Baselga A (2008) Historical bias in biodiversity inventories affects the observed environmental niche of the species. *Oikos* 117:847–858. <https://doi.org/10.1111/j.0030-1299.2008.16434.x>
31. Hortal J, Lobo JM, Jiménez-Valverde A (2007) Limitations of biodiversity databases: Case study on seed-plant diversity in Tenerife, Canary Islands. *Conserv Biol* 21:853–863. <https://doi.org/10.1111/j.1523-1739.2007.00686.x>
32. Joly CA, Metzger JP, Tabarelli M (2014) Experiences from the Brazilian Atlantic Forest: Ecological findings and conservation initiatives. *New Phytol* 204:459–473. <https://doi.org/10.1111/nph.12989>
33. Ladle R, Hortal J (2013) Mapping species distributions: living with uncertainty. *Front Biogeogr* 5:4–6. <https://doi.org/10.21425/F5FBG12942>
34. Lapola DM, Martinelli LA, Peres CA, Ometto JPHB, Ferreira ME, Nobre CA, Aguiar APD, Bustamante MMC, Cardoso MF, Costa MH, Joly CA, Leite CC, Moutinho P, Sampaio G, Strassburg BBN, Vieira ICG (2014) Pervasive transition of the Brazilian land-use system. *Nat Clim Chang* 4:27–35. <https://doi.org/10.1038/nclimate2056>
35. Laurance W (2009) Conserving the hottest of the hotspots. *Biol Conserv* 142. doi:10.1016/j.biocon.2008.10.011
36. Lessa T, Dos Santos JW, Correia RA, Ladle RJ, Malhado ACM (2019) Known unknowns: Filling the gaps in scientific knowledge production in the Caatinga. *PLoS ONE* 14:1–12. <https://doi.org/10.1371/journal.pone.0219359>
37. Lobo JM, Hortal J, Yela JL, Millán A, Sánchez-Fernández D, García-Roselló E, González-Dacosta J, Heine J, González-Vilas L, Guisande C (2018) KnowBR: An application to map the geographical

- variation of survey effort and identify well-surveyed areas from biodiversity databases. *Ecol Indic* 91:241–248. <https://doi.org/10.1016/j.ecolind.2018.03.077>
38. Metzger JP (2009) Conservation issues in the Brazilian Atlantic forest. *Biol Conserv* 142:1138–1140. <https://doi.org/10.1016/j.biocon.2008.10.012>
 39. Meyer C, Kreft H, Guralnick R, Jetz W (2015) Global priorities for an effective information basis of biodiversity distributions. *Nat Commun* 6. <https://doi.org/10.1038/ncomms9221>
 40. Muylaert RL, Vancine MH, Bernardo R, Oshima JEF, Sobral-Souza T, Tonetti VR, Niebuhr BB, Ribeiro MC (2018) Uma nota sobre os limites territoriais da Mata Atlântica. *Oecologia Aust* 22:302–311. [10.4257/oeco.2018.2203.09](https://doi.org/10.4257/oeco.2018.2203.09)
 41. Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403:853–858. <https://doi.org/10.1038/35002501>
 42. Oliveira U, Paglia AP, Brescovit AD, de Carvalho CJB, Silva DP, Rezende DT, Leite FSF, Batista JAN, Barbosa JPPP, Stehmann JR, Ascher JS, de Vasconcelos MF, De Marco P, Löwenberg-Neto P, Dias PG, Ferro VG, Santos AJ (2016) The strong influence of collection bias on biodiversity knowledge shortfalls of Brazilian terrestrial biodiversity. *Divers Distrib* 22:1232–1244. <https://doi.org/10.1111/ddi.12489>
 43. Patton JL, Pardiñas UFJ, D’Elía G (2015) *Mammals of South America; Volume 2: Rodents*. University of Chicago Press
 44. da Rosa CA, Secco H, Carvalho N, Alves C, Maia AC, Bager A, Catarina S (2018) Edge effects on small mammals: Differences between arboreal and ground-dwelling species living near roads in Brazilian fragmented landscapes. *Austral Ecol* 43:117–126. <https://doi.org/10.1111/aec.12549>
 45. da Rosa CA, Ribeiro BR, Bejarano V et al (2020) NEOTROPICAL ALIEN MAMMALS: a data set of occurrence and abundance of alien mammals in the Neotropics. *Ecology* 101. <https://doi.org/10.1002/ecy.3115>
 46. R Development Core Team (2021) *R: a language and environment for statistical computing*. v 3.6. R Foundation for Statistical Computing, Vienna, Austria. URL:<https://www.Rproject.org>
 47. Reddy S, Dávalos LM (2003) Geographical sampling bias and its implications for conservation priorities in Africa. *J Biogeogr* 30:1719–1727. <https://doi.org/10.1046/j.1365-2699.2003.00946.x>
 48. Rezende CL, Scarano FR, Assad ED, Joly CA, Metzger JP, Strassburg BBN, Tabarelli M, Fonseca GA, Mittermeier RA (2018) From hotspot to hopespot: An opportunity for the Brazilian Atlantic Forest. *Perspect Ecol Conserv* 16:208–214. <https://doi.org/10.1016/j.pecon.2018.10.002>
 49. Ribeiro MC, Metzger JP, Martensen AC, Ponzoni FJ, Hirota MM (2009) The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. *Biol Conserv* 142:1141–1153. <https://doi.org/10.1016/j.biocon.2009.02.021>
 50. Rocchini D, Hortal J, Lengyel S, Lobo JM, Jiménez-Valverde A, Ricotta C, Bacaro G, Chiarucci A (2011) Accounting for uncertainty when mapping species distributions: The need for maps of ignorance. *Prog Phys Geogr* 35:211–226. <https://doi.org/10.1177/0309133311399491>

51. Santos JP, Sobral-Souza T, Brown KS, Vancine MH, Ribeiro MC, Freitas AVL (2020) Effects of landscape modification on species richness patterns of fruit-feeding butterflies in Brazilian Atlantic Forest. *Divers Distrib* 26:196–208. <https://doi.org/10.1111/ddi.13007>
52. Sastre P, Lobo JM (2009) Taxonomist survey biases and the unveiling of biodiversity patterns. *Biol Conserv* 142:462–467. <https://doi.org/10.1016/j.biocon.2008.11.002>
53. Silva JMC, Casteleti CHM (2003) Status of the biodiversity of the Atlantic Forest of Brazil. In: Galindo-Leal C, Câmara IG (eds) *The Atlantic Forest of South America: Biodiversity Status; Threats; and Outlook*. CABS and Island Press, pp 43–59
54. Silva-Soares T, Ferreira RB, Ornellas IA, Zocca CZ, Caramaschi U, Cruz CAG(2021) A new species of *Ischnocnema* (Anura: Brachycephalidae) from the mountainous region of Atlantic Forest, southeastern Brazil, with a new phylogeny and diagnose for *Ischnocnema parva* series. *Zootaxa* doi: 10.11646/ZOOTAXA.5082.3.1
55. Sobral-Souza T, Stropp J, Santos JP, Prasniewski VM, Szinwelski N, Vilela B, Freitas AVL, Ribeiro MC, Hortal J (2021a) Knowledge gaps hamper understanding the relationship between fragmentation and biodiversity loss: The case of Atlantic Forest fruit-feeding butterflies. *PeerJ* 1–23. <https://doi.org/10.7717/peerj.11673>
56. Sobral-Souza T, Santos JP, Maldaner ME, Lima-Ribeiro MS, Ribeiro MC (2021b) EcoLand: A multiscale niche modelling framework to improve predictions on biodiversity and conservation. *Perspect Ecol Conserv* 19:362–368. <https://doi.org/10.1016/j.pecon.2021.03.008>
57. Sousa-Baena MS, Garcia LC, Peterson AT (2014) Knowledge behind conservation status decisions: Data basis for “Data Deficient” Brazilian plant species. *Biol Conserv* 173:80–89. <https://doi.org/10.1016/j.biocon.2013.06.034>
58. Tabarelli M, Aguiar AV, Ribeiro MC, Metzger JP, Peres CA (2010) Prospects for biodiversity conservation in the Atlantic Forest: Lessons from aging human-modified landscapes. *Biol Conserv* 143:2328–2340. <https://doi.org/10.1016/j.biocon.2010.02.005>
59. Teixido AL, Gonçalves SRA, Fernández-Arellano GJ, Dáttilo W, Izzo TJ, Layme VMG, Moreira LFB, Quintanilla LG (2020) Major biases and knowledge gaps on fragmentation research in Brazil: Implications for conservation. *Biol Conserv* 251. <https://doi.org/10.1016/j.biocon.2020.108749>
60. Wilson DE, Reeder DM (2005) *Mammal Species of the World: A Taxonomic and Geographic Reference*. Johns Hopkins University Press, Baltimore and London
61. Wüest RO, Zimmermann NE, Zurell D, Alexander JM, Fritz SA, Hof C, Kreft H, Normand S, Cabral JS, Szekely E, Thuiller W, Wikelski M, Karger DN (2020) Macroecology in the age of Big Data – Where to go from here? *J Biogeogr* 47:1–12. <https://doi.org/10.1111/jbi.13633>
62. Yang W, Ma K, Kreft H (2013) Geographical sampling bias in a large distributional database and its effects on species richness-environment models. *J Biogeogr* 40:1415–1426. <https://doi.org/10.1111/jbi.12108>

Figures

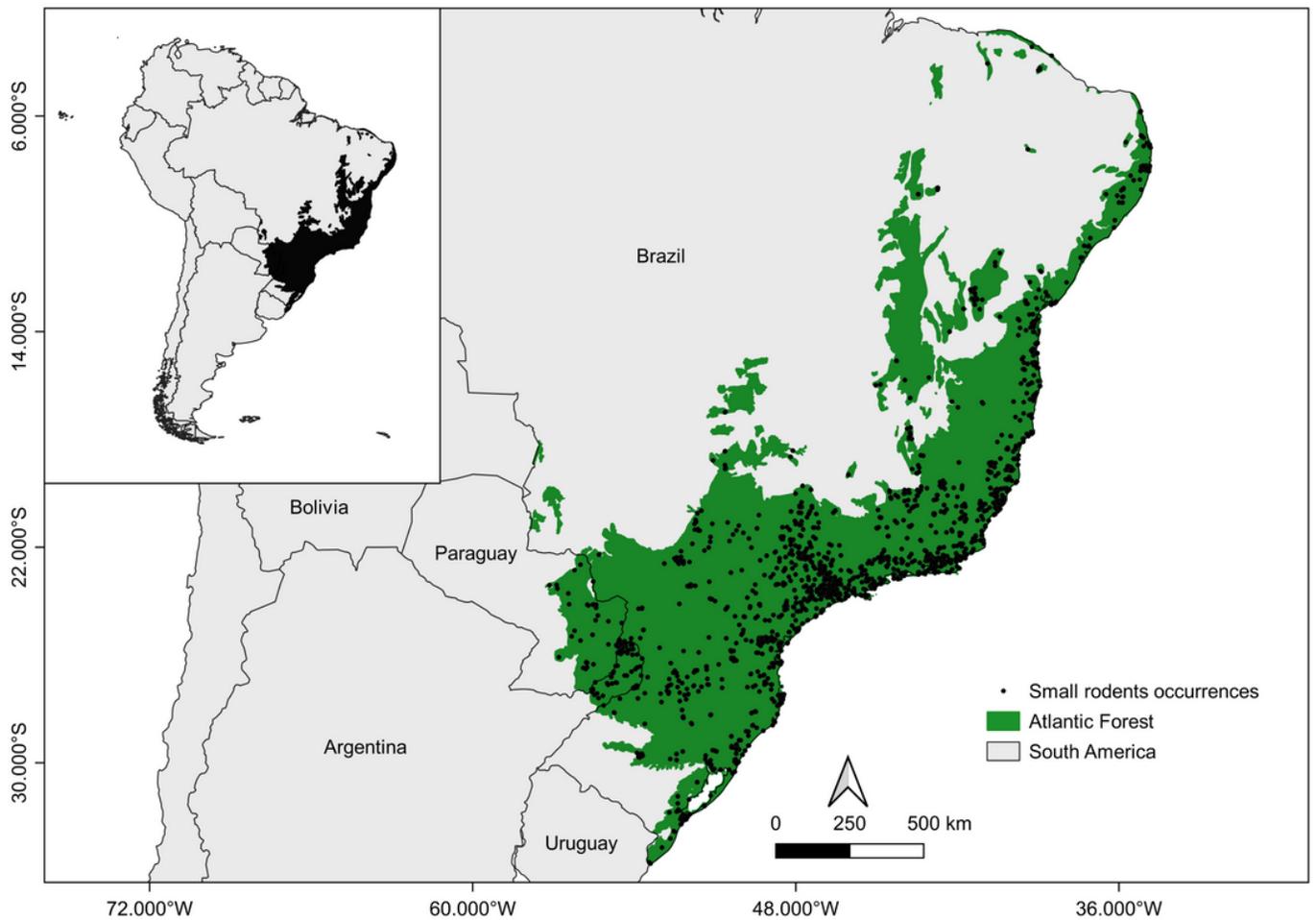


Figure 1

Spatial distribution of small rodent occurrences compiled from databases (Bovendorp et al., 2017 and Figueiredo et al., 2017) and digital platforms (GBIF, SpeciesLink, SibBr and Portal da Biodiversidade) for the Atlantic Forest, proposed by Muylaert et al. (2018).

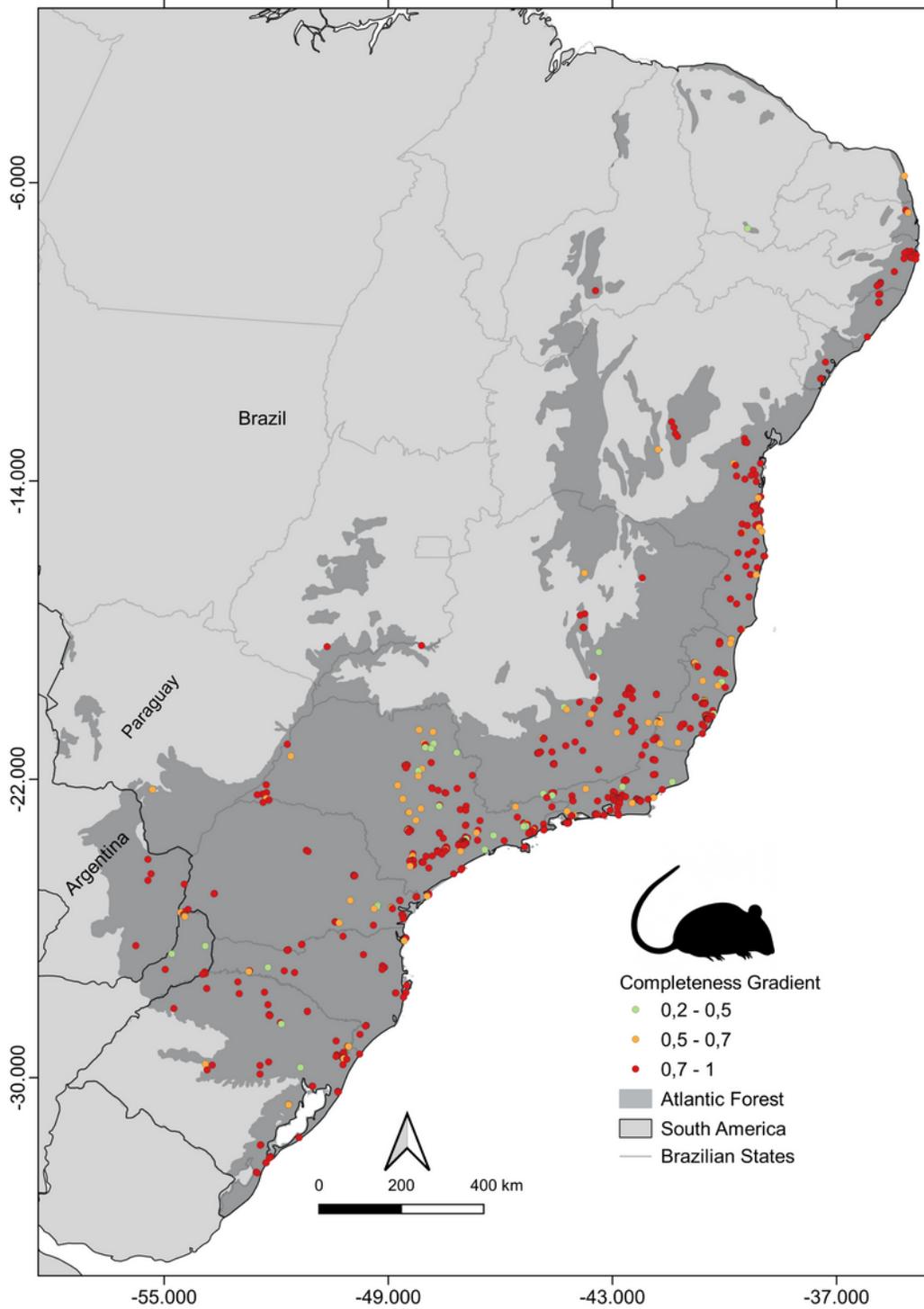


Figure 2

Well-sampled sites for small rodent inventories in the Atlantic Forest. Red dots indicate well-sampled sites (n = 507). Green and yellow dots indicate locations where less than 69% of the biodiversity is known (n = 41 and 91, respectively).

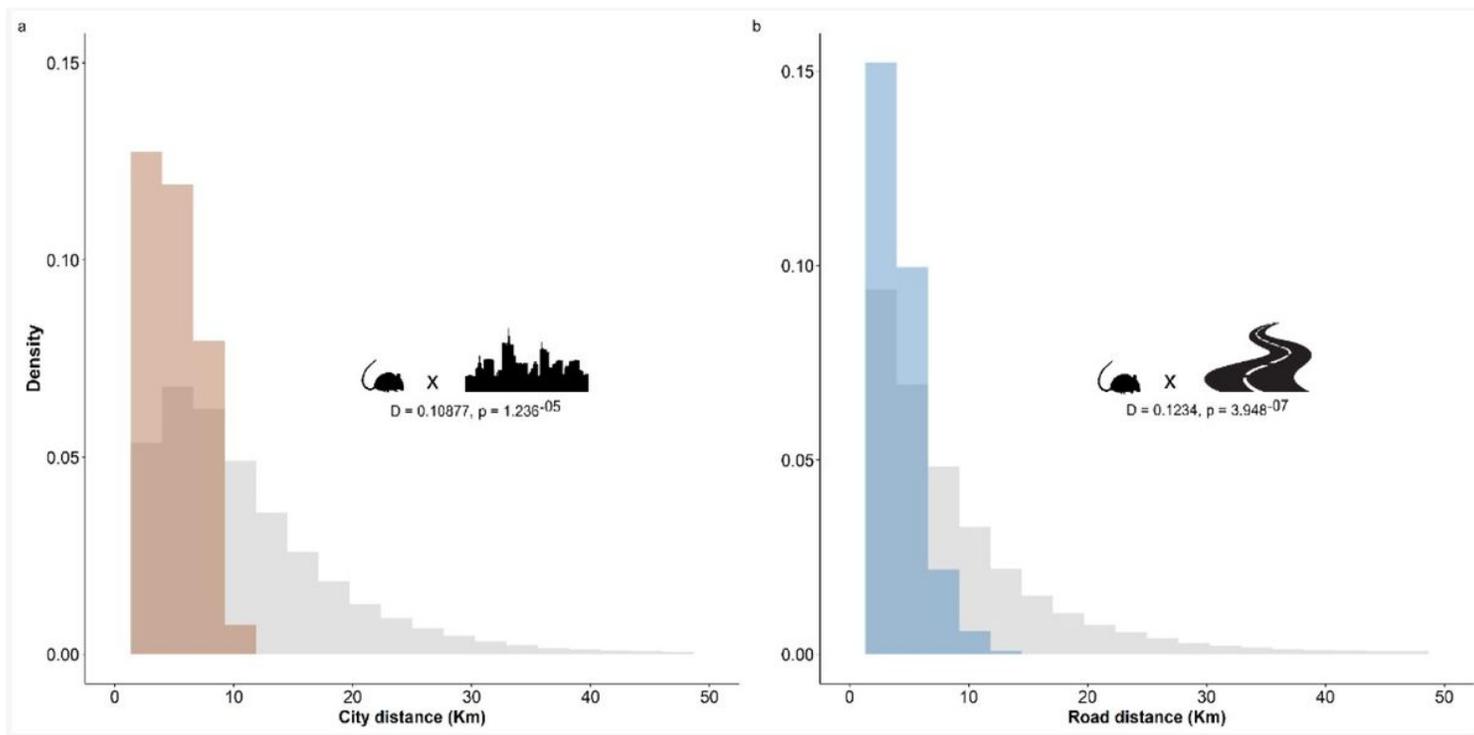


Figure 3

Spatial bias of the small rodents well-sampled sites related to city distance (m) and road distance (m). The well-sampled sites are biased to locations closed to cities and roads. The gray bars indicate the density of cities and road distances for extension Atlantic Forest (limit proposed by Muylaert et al., 2018). Colored bars indicate the density of distance of (a) cities centers and (b) roads of well-sampled sites.

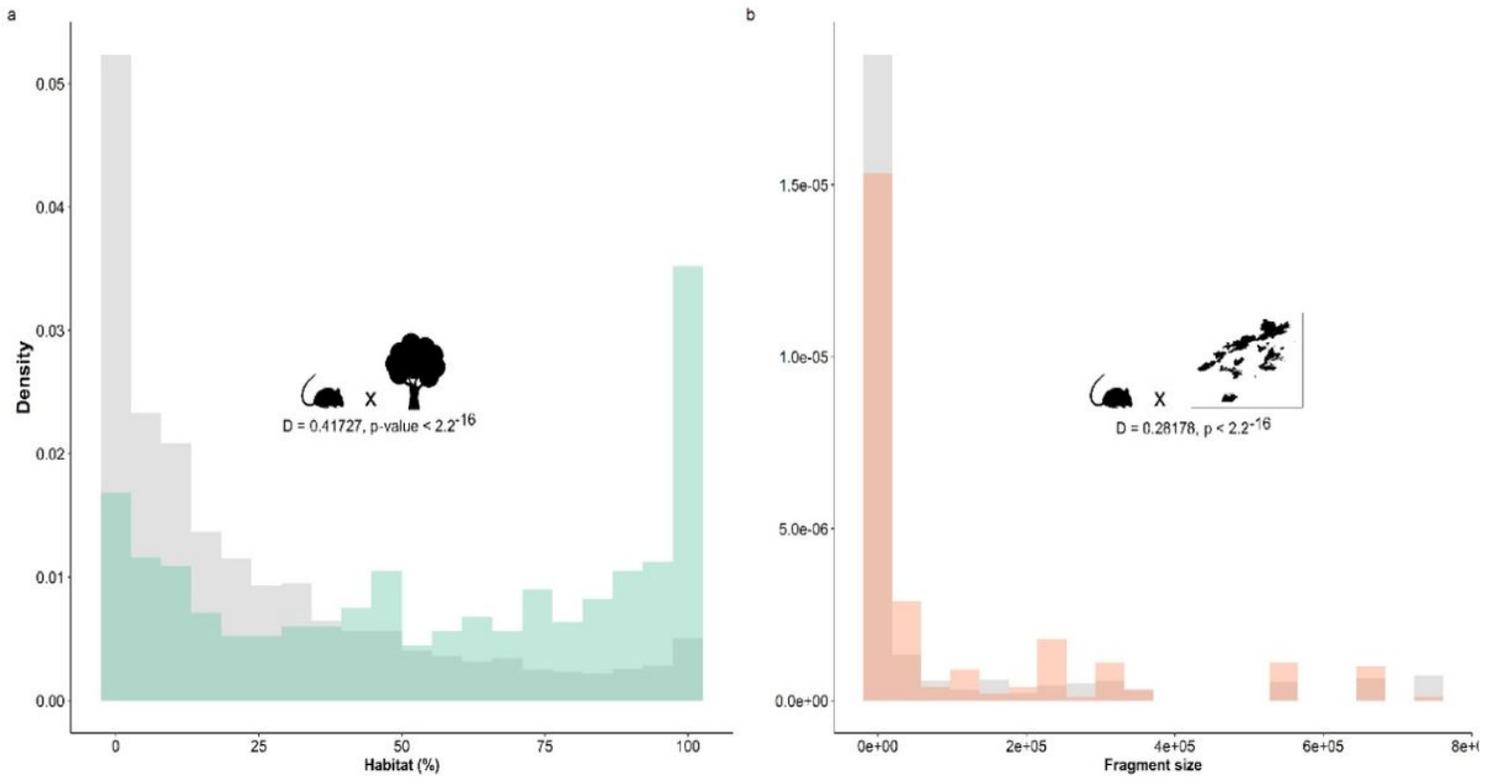


Figure 4

Spatial bias of the small rodents well-sampled site related to landscape. The well-sampled sites are biased to locations closed to large fragments and higher percentage of habitat. Colored bars indicate the percentage habitat (a) and density of fragment size (b) of well-sampled site.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [CandelariaetalSupplementarymaterial.docx](#)