NeoBat Interactions: a data set of bat-plant interactions in the Neotropics

Guillermo L. Florez-Montero[1*] Cullen Geiselman[2] Marcelo R. Nogueira[3]
Renata de Lara Muylaert[4] Sharlene E. Santana[5] Richard D. Stevens[6]
Marco Tschapka[7] Francisco A. Rodrigues[8] Marco A. R. Mello[9]

1 Universidade Federal do ABC (UFABC). Centro de Ciências Humanas e Naturais. 09210-580 Santo André, SP, Brazil.

2 Institution. Departament. Postal Code. City, Country XXXXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXX XXXXXX

3 Universidade Federal Rural do Rio de Janeiro (UFRRJ). Instituto de Biologia, Laboratório de Mastozoologia. 23890-000 Seropédica, RJ, Brazil.

4 Institution. Departament. Postal Code. City, Country

5 Department of Biology and Burke Museum of Natural History and Culture, University of Washington, Seattle, WA, 98195, USA

6 Institution. Departament. Postal Code. City, Country XXXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXX XXXXXX

7 Institution. Departament. Postal Code. City, Country

8 Universidade de São Paulo. Departamento de Matemática Aplicada e Estatística, Instituto de Ciências Matemáticas e de Computação. 13560-970 São Carlos, SP, Brazil

9 Universidade de São Paulo (USP), Instituto de Biociências, Departamento de Ecologia Geral. 05508-090 São Paulo, SP, Brazil.

*Corresponding author: Guillermo L. Florez-Montero (gflorezmontero@gmail.com)

INTRODUCTION

Open access to primary scientific data is fundamental to assure the social contract that governs scientific publishing (Vision 2010). Making primary data available improves the transparency, reproducibility, and progress of science by allowing independent verification and reuse of published data. (Costello 2009). This access has been facilitated in the information age through important frameworks for the production, storage, curation, and sharing of ecological data. Those frameworks which aim to preserve data in the long term, even beyond the life of the initial compilers and curators (Chavan and Ingwersen 2009).

One of those frameworks are data papers, which optimize efforts in the discovery, organization, and availability of ecological data (Chavan and Penev 2011). They offer a highly reliable source of data, as they have been subjected to high-quality control measures, such as peer review and editorial control of data and metadata (Costello et al. 2013). This new kind of publication has revolutionized contemporary science by making decades of naturalistic information widely available in highly accessible and comprehensive formats. The revolution has also reached mammalogy. Data papers on mammal communities are growing in numbers. There are, for instance, data papers on the global non-volant mammal communities (Thibault et al. 2011). Another important source is the BioTIME database of biodiversity time series for the Anthropocene

(Dornelas et al. 2018). Others have larger scope, such as the global database for metacommunity ecology, integrating species, traits, environment and space (Jeliazkov et al. 2020).

In Brazil, the most famous mammalogical data papers were produced in the ATLANTIC Series, which contains information about the biodiversity of the Atlantic Forest of South America. This series includes data papers on plant-frugivore interactions (Bello et al. 2017), rodents and marsupials: (Bovendorp et al. 2017), bats (Muylaert et al. 2017), primates (Culot et al. 2019), and medium-and large-sized mammals (Souza et al. 2019). Many other data papers were produced by other research groups outside the ATLANTIC Series, focusing on groups such as small mammals (Figueiredo et al. 2017). In addition, only a few covered larger spatial scales, such as Neotropical xenarthrans (Santos et al. 2019). Nevertheless, most of them are based on abundance-incidence species data whereas species interactions received much less attention.

Mutualistic interactions between animals and plants are a cornerstone of terrestrial ecosystems. Almost 94% of plants in neotropical communities are pollinated by animals (Ollerton, Winfree, and Tarrant 2011), while 70 – 94% have their seeds dispersed by vertebrates (Jordano 2013). Bats are especially important in this context, as they represent the second largest group of seed dispersers in the Neotropics, after birds (Bello et al. 2017). On the other hand, even though insects pollinate most of the flowering plants, bats are also the second group of pollinating vertebrates since they pollinate about 2% of plant genera (Sekercioglu 2006). Bat-plant interactions also result in ecosystem services, such as the pollination of some economically important plants and the dispersal of seeds from pioneer plants that are key to habitat regeneration (Kunz et al. 2011).

In the present data paper,we compiled a georeferenced database of 2571 interaction records of frugivory and nectarivory between 93 bat species and 501 plant species. The data came from 169 studies covering 200 locations in 16 countries all over the Neotropical region. (Figure 1). The database compiled by Geiselman and Younger (2002) was used as a starting point and was filtered and updated. NeoBat Interactions is so far the most extensive bat-plant interaction database both in geographic and taxonomic terms. Most sampling sites are georeferenced with high coordinate accuracy. All records came from primary sources and were taxonomically verified and updated. Additionally, our database includes ecological information, such as a life form and successional stage of plants, and trophic guild of bats. The data are organized and standardized at different levels of ecological complexity and temporal and geographic scales, which allows using them in a variety of studies with different scopes.

METADATA

CLASS I. DATA SET DESCRIPTORS

A. Data set identity

Title: NeoBat Interactions: a data set of bat-plant interactions in the Neotropics

B. Data set identification code

Suggested data set identity codes:

NeoBat_Interactions_References.csv

NeoBat_Interactions_Sites.csv

NeoBat_Interactions_Records.csv

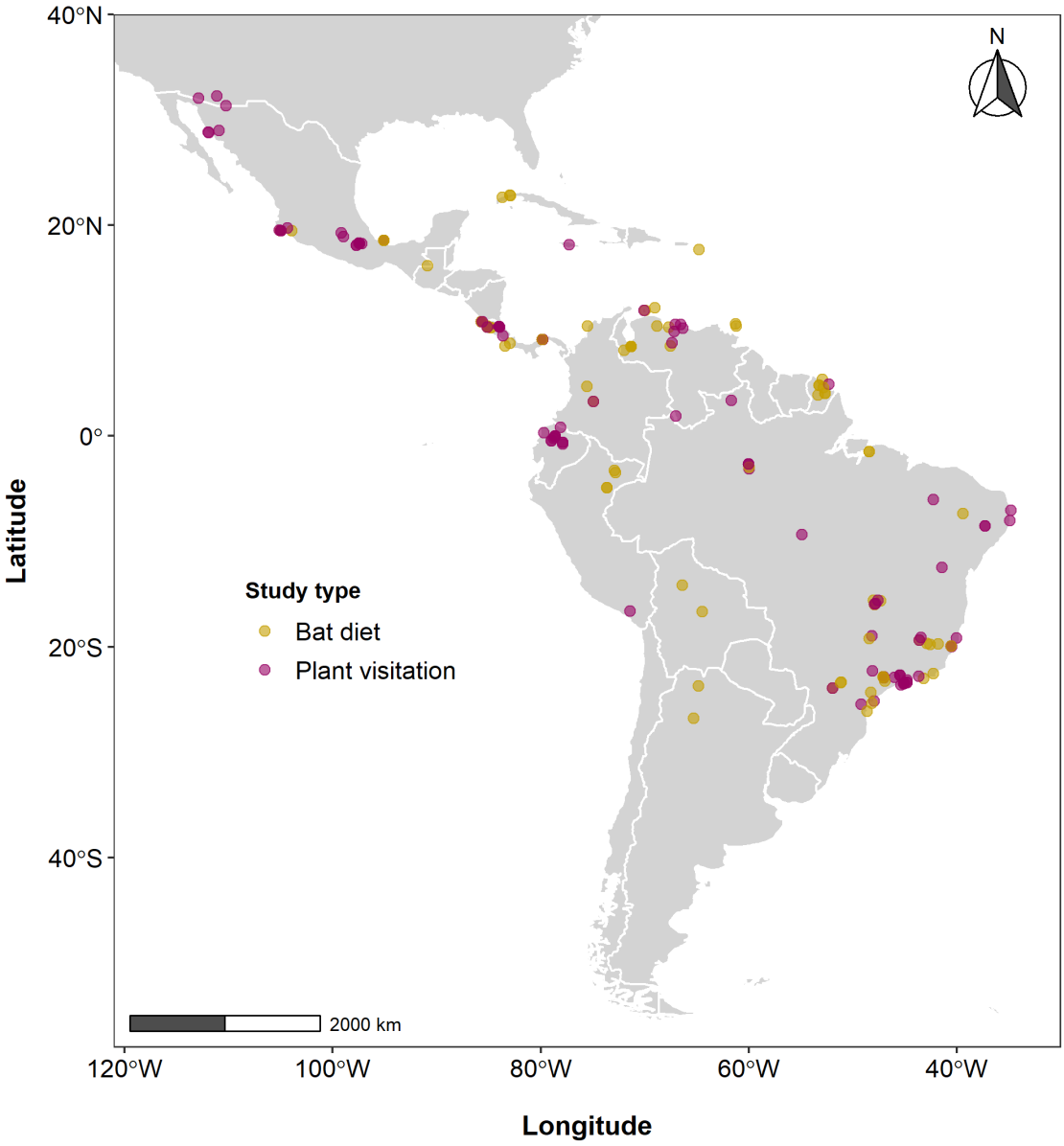

Figure 1: Distribution of sampling sites covered in NeoBat Interactions. Purple dots show the location of original studies reporting bat-plant interactions. White lines show country borders. We included only studies with records of bat-plant interactions that were confirmed either by indirect or direct observation.

C. Data set description

Principal Investigator(s):

1. Guillermo L. Florez-Montero. Universidade Federal do ABC. Centro de Ciências Humanas e Naturais. 09210-580 Santo André, SP, Brazil.
2. Marco A. R. Mello. Universidade de São Paulo, Instituto de Biociências, Departamento de Ecologia Geral. 05508-090 São Paulo, SP, Brazil.

Abstract:

Data papers and open databases revolutionized contemporary science, as they provide the long-needed incentive to collaborate in large international teams and make naturalistic information widely available. Nevertheless, most of them focus on occurrence or abundance, while species interactions received less attention. To help fill this gap, we compiled a georeferenced data set of interactions between 93 bat species of the family Phyllostomidae (Chiroptera) and 501 plant species of 68 families. Data were obtained from 169 studies published from 1957 to 2007 in the entire Neotropical Region, with most records coming from Brazil (34.5% of all study sites), Costa Rica (16%), and Mexico (14%). Our data set includes 2571 records of frugivory (75.1% of all records) and nectarivory (24.9%). The best represented bat genera are *Artibeus* (28% of all records), *Carollia* (24%), *Sturnira* (10.1%), and *Glossophaga* (8.8%). *Carollia perspicillata* (187), *Artibeus lituratus* (125), *Artibeus jamaicensis* (94), *Glossophaga soricina* (86), and *Artibeus planirostris* (74) are the bat species with the broadest diets recorded in number of plant species. Among plants, the best represented families are Moraceae (17%), Piperaceae (15.4%), Urticaceae (9.2%), and Solanaceae (9%). Plants of the genera *Cecropia* (46), *Ficus* (42), *Piper* (40), *Solanum* (31), and *Vismia* (27) hold the largest number of interactions. These data are stored as arrays (records, sites, and studies) organized by logical keys and rich metadata, which helps compile the information at different ecological and geographic scales, as required by different studies. Our data set on bat-plant interactions is so far the most extensive both in geographic and taxonomic terms, and also includes some ecological information of plants and bats. It has already helped us develop several studies and we hope it will stimulate novel analyses and syntheses, in addition to pointing out to important gaps in knowledge for future research.

D. Key words

Species interactions, mutualism, nectarivory, frugivory, pollination, seed dispersal, databases, networks.

E. Description

This data base includes 2571 records of interactions involving the consumption of nectar and fruit by bats, taken from studies carried out from the approach of the diet of bats and plant visitation across the Neotropics (Figure 1). The information came from 16 countries, from the southwestern United States to northwestern Argentina. We have collected 169 scientific papers carried out over 50 years, in 200 study locations. The spatial and temporal distribution of the studies is not at all homogeneous.Brazil (69), Costa Rica (32), Mexico (28) and Venezuela (16) are the countries with the most studies during that period (Figure 2A). Likewise, the number of published studies on frugivory and nectarivory were very scarce from the 1950s to the 1970s, having a peak of scientific production from the 1980s to the 2000s (Figure 2B).

We have recorded 2571 interaction events of frugivory (75.1% of all records) and nectarivory (24.9%) between 93 bat species and 501 plant species. Although most of the information comes from fruit- and nectar-feeding bats, we have also collected data of species with other trophic guilds that occasionally eat fruit or nectar. Likewise, this database includes information mainly on trees and shrubs, although it also contains other forms of life such as vines, herbs, and others. Many of species have no information on successional stage but many of the data cames from Early Successional species (Table 1). We have included information about the interaction strength, nevertheless 44.5% of the reccords have no data, mainly because it was not reported in the paper.

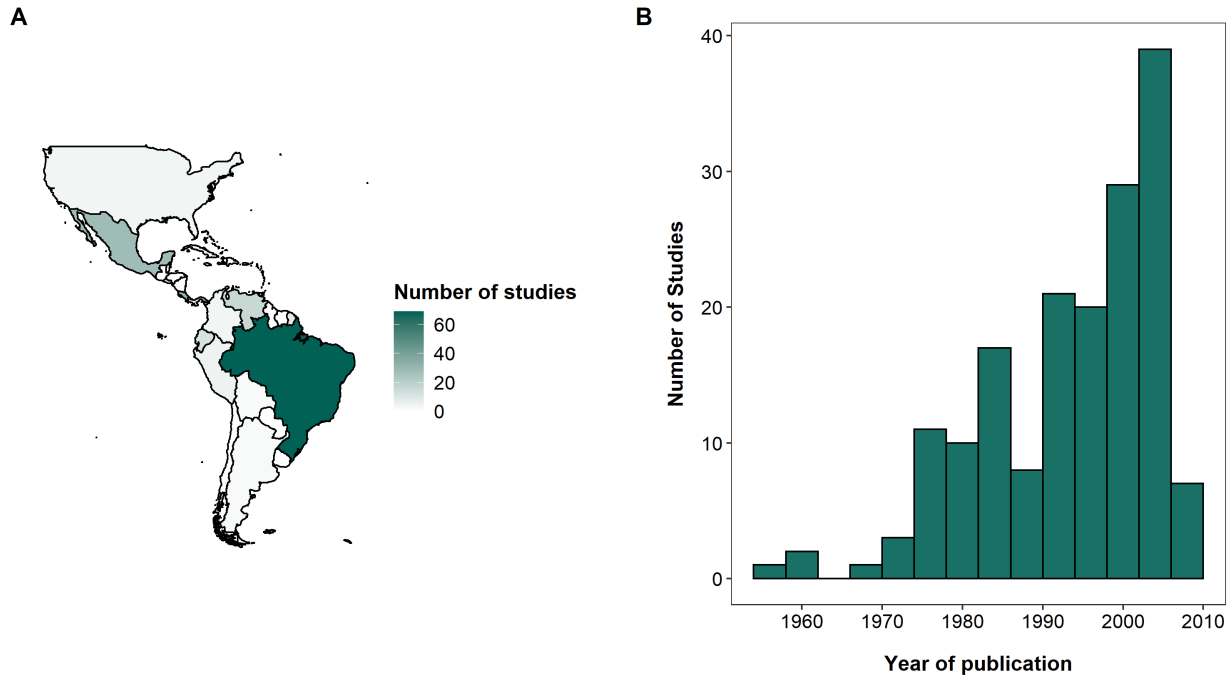

Figure 2: Distribution of the number of studies on frugivory and nectarivory by bats, published by country (A) and throughout the recorded period (B).

Table 1: Ecological information of bat and plant species recorded in the NeoBat Interactions database

Ecological trait	**Class**	**Number of species**	**%**
Trophic guild of bats	Foliage gleaner	5	5.4
	Frugivore	55	59.1
	Nectarivore	27	29.0
	Omnivore	6	6.5
Succesional stage of plants	Early	152	29.8
	Late	112	22.0
	Not information	246	48.2
Life form of plants	Herb	52	10.1
	Palm-tree	8	1.6
	Shrub	135	26.3
	Succulent	25	4.9
	Tree	231	45.0
	Vine	36	7.0
	Not information	26	5.1

Table 2: IUCN conservation status of animals and plants species reported in the NeoBat Interactions database

Group	IUCN Status	Number of species	%
Plants	Critically Endangered (CR)	1	0.2
	Endangered (EN)	4	0.8
	Vulnerable (VU)	5	1.0
	Near Threatened (NT)	4	0.8
	Least Concern (LC)	183	36.5
	Conservation Dependent (CD)	1	0.2
	Data Deficient (DD)	3	0.6
	Not Evaluated (NE)	300	59.9
Bats	Endangered (EN)	3	3.2
	Vulnerable (VU)	2	2.2
	Near Threatened (NT)	3	3.2
	Least Concern (LC)	67	72.0
	Data Deficient (DD)	2	2.2
	Not Evaluated (NE)	16	17.2

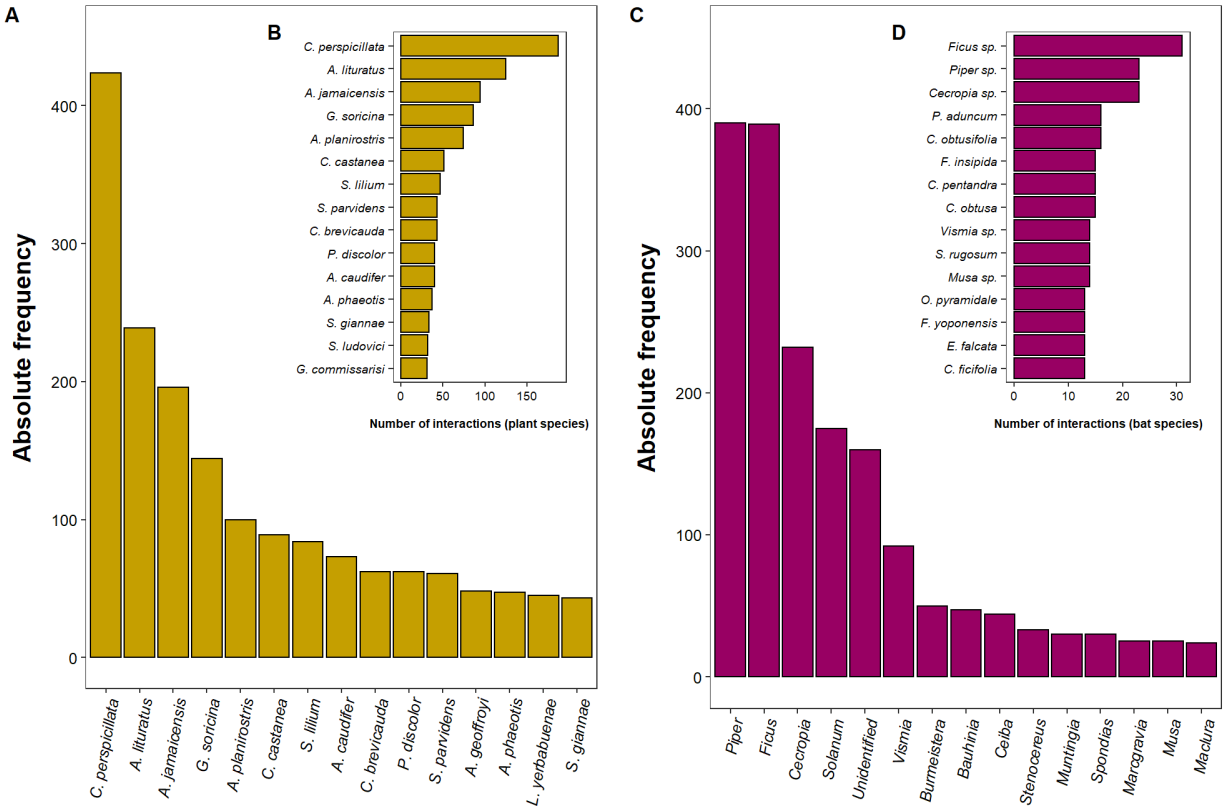

Figure 3: Ranking of frequency of the 15 most abbundant bat species (A) and plant genera (C). Internal plots represent the number of interaction of the 15 most abbundant species of bats (B) and plants (D).

Table 3: Information about the references in the NeoBat Interactions database

Variable	**Description**	**Levels**	**Example**
RefCode	Identification of each reference. This code links the reference matrix to the other matrices	BPR001 to BPR168	BPR066
Author	Short name of the author(s), if there are three or more authors, we use et al.	...	Hernández-Conrique et al.
Year	Year of publication.	1957 to 2007	1997
Reference	Extended reference.	...	Hernández-Conrique, D., L.I. Iñiguez-Dávalos & J.F. Storz. 1997. Selective feeding by phyllostomid fruit bats in subtropical montane cloud forest. Biotropica 29: 376-379

Table 4: Information about the samplig sites in the NeoBat Interactions database

Variable	**Description**	**Levels**	**Example**
SiteCode	Identification of each sampling site. This code links the site matrix to the record matrix. In many cases, a paper can have more than one sampling site.	BPA001 to BPA200	BPA074
RefCode	Identification of each reference. This code links the reference matrix to the other matrices.	BPR001 to BPR168	BPR066
Locality	Locality where fieldwork was carried out, based on information reported in the paper. We checked all names using the Google Earth database.	...	Sierra de Manantlan Biosphere Reserve
State	State, Department or Province of the study site based on the geographic coordinates.	...	Jalisco

Table 4: Information about the samplig sites in the NeoBat Interactions database *(continued)*

Variable	**Description**	**Levels**	**Example**
Country	Country where fieldwork was carried out (English name).	Argentina Bolivia Brazil Colombia Costa Rica Cuba Curacao Ecuador French Guiana Jamaica Mexico Panama Peru Trinidad and Tobago United States Venezuela	Mexico
Latitude	Corrected latitude in decimal degrees (Projection WGS84 EPSG:4326). In studies with two or more sampling sites with less than 5 km of linear distance between them we used the centroid coordinate.	Decimal degrees	19.485675
Longitude	Corrected longitude in decimal degrees. See Latitude for more information.	Decimal degrees	-103.950087
Precision	In some cases, papers reported the precise coordinates of each sampling site. Other papers reported only geographic references (basins, rivers, municipality, or distance from a village). In these cases, we validated these references with Google Earth satellite images. We consider Not Precise when coordinates mismatch the written information in the paper, or when the paper only reported the coordinates of the municipality or region.	Precise Not Precise	Precise
YearStart	The year in which sampling started.	1960 to 2006	1993

Table 4: Information about the samplig sites in the NeoBat Interactions database *(continued)*

Variable	**Description**	**Levels**	**Example**
YearEnd	The year in which sampling ended.	1960 to 2006	1994
Duration	Unstandardized duration of the sampling period (in months).	1 to 37	7
SamplingEffort	For studies based on feces collection: Total number of fecal samples. For studies based on the observation of plant visitation events: Total number of events recorded.	6 to 6809	68
StudyType	The type of study according to the focus reported in the reference paper. Bat diet refers to study focused on describing the diet of a bat species or assemblage. Plant visitation refers to studies aimed at describing the visitors of a plant species or assemblage.	Bat diet Plant visitation	Bat diet
EcologicalScale	The ecological scale studied. When there were more than one species of bat (when the Study type is Bat diet) or plant (when the study type is Plant visitation), we considered as Assemblage.	Population Assemblage	Assemblage
SamplingMethod	The sampling method as described in the reference paper. We have standardized the levels to five broad methods. Some studies have more than one sampling method	Direct observation Experimental Feces collection Pollen collection Roost inspection	Experimental
SamplingSeason	Climatic season in which sampling was performed. Some studies were conducted in both, dry and wet season.	Dry Wet	Dry and Wet
Vegetation	Vegetation type as described in the reference paper.	...	Subtropical montane cloud forest

Table 4: Information about the samplig sites in the NeoBat Interactions database *(continued)*

Variable	**Description**	**Levels**	**Example**
VegType	Vegetation type corrected according to Oliveira-Filho (2017).	Cloud forest Coastal sandy mosaic Deciduous forest Floodplain forest Limestone deciduous woodland Limestone Rainforest Mixed forest Montane woodland Rainforest Rocky woodland Savanna woodland Seasonal riverine forest Semi-arid thorny woodland Semi-desert and desert Semideciduous forest	Mixed forest
Ecoregion	Ecological region according to Olson et al. (2001).	...	Trans-Mexican Volcanic Belt pine-oak forests
Domain	Phytogeographic domain according to NatureServe (2013) and Oliveira-Filho (2017).	Amazonia Atlantic Forest Caatinga Caribbean Cerrado Los Llanos Mesoamerica Nearctic Mexico Northern Andes Southern Andes	Mesoamerica
Altitude	Meters above sea level reported in the reference paper.	2 to 2700	1900
X1kmAlt	Meters above sea level, from the Hydro-1K dataset (United States Geological Survey â€“ USGS, 2001. Global 30 arc-seconds Elevation (GTOPO30).	0 to 2686	2462
AnnRain	Annual Rainfall in mm from WorldClim 2.0 with 30 arc seconds resolution.	66 to 3912	1116

Table 4: Information about the samplig sites in the NeoBat Interactions database *(continued)*

Variable	**Description**	**Levels**	**Example**
MeaAnnTemp	Mean annual temperature in Celsius degrees from WorldClim 2.0 with 30 arc seconds resolution.	10.3 to 27.9	14.2
PET	Global Potential Evapo-Transpiration (annual average in mm) from CGIAR-CSI (Trabucco and Zomer 2009), with resolution of 30 arc seconds.	1021 to 2588	1577
GAI	Global Aridity Index model from CGIAR-CSI (Trabucco and Zomer 2009), with resolution of 30 arc seconds.	0 to 3	0.7072

Table 5: Information about the interaction records in the NeoBat Interactions database

Variable	**Description**	**Levels**	**Example**
IDCode	Identification code of each interaction record.	BPI0001 to BPI2574	BPI0077
SiteCode	Identification of each sampling site. This code links the site matrix to the record matrix. In many cases, a paper can have more than one sampling site.	BPA001 to BPA200	BPA008
RefCode	Identification of each reference. This code links the reference matrix to the other matrices.	BPR001 to BPR168	BPR008
BatGenus	Current scientific name of the bat genus.	...	Uroderma
BatSpecies	Scientific name of the bat as reported in the reference paper.	...	Uroderma bilobatum
CurrentBatSpecies	Current scientific name of the bat species	...	Uroderma bilobatum
TrophicGuild	Trophic guild of the bat species	Foliage gleaner Frugivore Nectarivore Omnivore	Frugivore

Table 5: Information about the interaction records in the NeoBat Interactions database *(continued)*

Variable	**Description**	**Levels**	**Example**
PlantFamily	Current scientific name of the plant family.	...	Piperaceae
PlantGenus	Current scientific name of the plant genus.	...	Piper
PlantSpecies	Scientific name of the plant as reported in the reference paper.	...	Photomorpha peltata
CurrentPlantSpecies	Current scientific name of the plant, validated with The Plant List database (http://www.theplantlist.org) and the REFLORA database (http://reflora.jbrj.gov.br).	...	Piper peltatum
LifeForm	Life form of the plant species.	Herb Palm-tree Shrub Succulent Tree Vine	Shrub
SuccessionalStage	Successional stage of the plant species.	Early Late	Early
Interaction	Type of interaction described.	Frugivory Nectarivory	Frugivory
Weight	Frugivory: Number of fecal samples containing the plant seed.Nectarivory: Number of visits per sampling unit.	1 to 219	1

REFERENCES

Bello, Carolina, Mauro Galetti, Denise Montan, Marco A. Pizo, Tatiane C. Mariguela, Laurence Culot, Felipe Bufalo, et al. 2017. "Atlantic Frugivory: A Plant–Frugivore Interaction Data Set for the Atlantic Forest." *Ecology* 98 (6): 1729–9. https://doi.org/10.1002/ecy.1818.

Bovendorp, Ricardo S., Nacho Villar, Edson F. de Abreu-Junior, Carolina Bello, André L. Regolin, Alexandre R. Percequillo, and Mauro Galetti. 2017. "Atlantic Small-Mammal: A Dataset of Communities of Rodents and Marsupials of the Atlantic Forests of South America." *Ecology* 98 (8): 2226–6. https://doi.org/10.1002/ecy.1893.

Chavan, Vishwas S., and Peter Ingwersen. 2009. "Towards a data publishing framework for primary biodiversity data: challenges and potentials for the biodiversity informatics community." *BMC Bioinformatic* 10 (14): S2. https://doi.org/10.1186/1471-2105-10-S14-S2.

Chavan, Vishwas S., and Lyubomir Penev. 2011. "The data paper: a mechanism to incentivize data publishing in biodiversity science." *BMC Bioinformatic* 12 (15): S2. https://doi.org/10.1186/1471-2105-12-S15-S2.

Costello, Mark J. 2009. "Motivating Online Publication of Data." *BioScience* 59 (5): 418–27. https://doi.org/10.1525/bio.2009.59.5.9.

Costello, Mark J., William K. Michener, Mark Gahegan, Zhi-Qiang Zhang, and Philip E. Bourne. 2013. "Biodiversity Data Should Be Published, Cited, and Peer Reviewed." *Trends in Ecology & Evolution* 28 (8): 454–61. https://doi.org/10.1016/j.tree.2013.05.002.

Culot, Laurence, Lucas Augusto Pereira, Ilaria Agostini, Marco Antônio Barreto de Almeida, Rafael Souza Cruz Alves, Izar Aximoff, Alex Bager, et al. 2019. "Atlantic-Primates: A Dataset of Communities and Occurrences of Primates in the Atlantic Forests of South America." *Ecology* 100 (1): e02525. https://doi.org/10.1002/ecy.2525.

Dornelas, Maria, Laura H. Antão, Faye Moyes, Amanda E. Bates, Anne E. Magurran, Dušan Adam, Asem A. Akhmetzhanova, et al. 2018. "BioTIME: A Database of Biodiversity Time Series for the Anthropocene." *Global Ecology and Biogeography* 27 (7): 760–86. https://doi.org/10.1111/geb.12729.

Figueiredo, Marcos S. L., Camila S. Barros, Ana C. Delciellos, Edú B. Guerra, Pedro Cordeiro-Estrela, Maja Kajin, Martin R. Alvarez, et al. 2017. "Abundance of Small Mammals in the Atlantic Forest (Asmaf): A Data Set for Analyzing Tropical Community Patterns." *Ecology* 98 (11): 2981–1. https://doi.org/10.1002/ecy.2005.

Geiselman, Cullen K., and Sarah Younger. 2002. "Bat Eco-Interactions Database." http://www.batbase.org.

Jeliazkov, Alienor, Darko Mijatovic, Stéphane Chantepie, Nigel Andrew, Raphaël Arlettaz, Luc Barbaro, Nadia Barsoum, et al. 2020. "A global database for metacommunity ecology, integrating species, traits, environment and space." *Scientific Data* 7 (6). https://doi.org/10.1038/s41597-019-0344-7.

Jordano, Pedro. 2013. "Fruits and Frugivory." In *Seeds: The Ecology of Regeneration in Plant Communities*, edited by Robert S. Gallagher, 3rd ed., 18–61. Wallingford, England: Commonwealth Agricultural Bureau International.

Kunz, Thomas H., Elizabeth Braun de Torrez, Dana Bauer, Tatyana Lobova, and Theodore H. Fleming. 2011. "Ecosystem Services Provided by Bats." *Annals of the New York Academy of Sciences* 1223 (1): 1–38. https://doi.org/10.1111/j.1749-6632.2011.06004.x.

Muylaert, Renata d. L., Richard D. Stevens, Carlos E. L. Esbérard, Marco A. R. Mello, Guilherme S. T. Garbino, Luiz H. Varzinczak, Deborah Faria, et al. 2017. "Atlantic Bats: A Data Set of Bat Communities from the Atlantic Forests of South America." *Ecology* 98 (12): 3227–7. https://doi.org/10.1002/ecy.2007.

Ollerton, Jeff, Rachael Winfree, and Sam Tarrant. 2011. "How Many Flowering Plants Are Pollinated by Animals?" *Oikos* 120 (3): 321–26. https://doi.org/10.1111/j.1600-0706.2010.18644.x.

Olson, David M, Eric Dinerstein, Eric D Wikramanayake, Neil D Burgess, George VN Powell, Emma C Underwood, Jennifer A D'amico, et al. 2001. "Terrestrial Ecoregions of the World: A New Map of Life on Eartha New Global Map of Terrestrial Ecoregions Provides an Innovative Tool for Conserving Biodiversity." *BioScience* 51 (11): 933–38.

Santos, Paloma Marques, Adriana Bocchiglieri, Adriano Garcia Chiarello, Adriano Pereira Paglia, Adryelle Moreira, Agnis Cristiane de Souza, Agustin Manuel Abba, et al. 2019. "Neotropical Xenarthrans: A Data Set of Occurrence of Xenarthran Species in the Neotropics." *Ecology* 100 (7): e02663. https://doi.org/10.1002/ecy.2663.

Sekercioglu, Cagan H. 2006. "Increasing Awareness of Avian Ecological Function." *Trends in Ecology & Evolution* 21 (8): 464–71. https://doi.org/10.1016/j.tree.2006.05.007.

Souza, Yuri, Fernando Gonçalves, Laís Lautenschlager, Paula Akkawi, Calebe Mendes, Mariana M. Carvalho, Ricardo S. Bovendorp, et al. 2019. "Atlantic Mammals: A Data Set of Assemblages of Medium- and Large-Sized Mammals of the Atlantic Forest of South America." *Ecology* 100 (10): e02785. https://doi.org/10.1002/ecy.2785.

Thibault, Katherine M., Sarah R. Supp, Mikaelle Giffin, Ethan P. White, and S. K. Morgan Ernest. 2011. "Species Composition and Abundance of Mammalian Communities." *Ecology* 92 (12): 2316–6. https://doi.org/10.1890/11-0262.1.

Vision, Todd J. 2010. "Open Data and the Social Contract of Scientific Publishing." *BioScience* 60 (5): 330–31. https://doi.org/10.1525/bio.2010.60.5.2.