Journal of Insect Conservation
https://doi.org/10.1007/s10841-020-00262-y

ORIGINAL PAPER

How to prioritize areas for new ant surveys? Integrating historical data on species occurrence records and habitat loss

Raquel Divieso[1] · Ana Rorato[2] · Rodrigo M. Feitosa[1] · Andreas L. S. Meyer[3] · Marcio R. Pie[1]

Received: 8 April 2020 / Accepted: 7 August 2020

Abstract
Habitat loss is a leading cause of extinctions, which may occur even before species are recorded or formally described. On the other hand, limitations in species distribution data and sampling biases can hamper inferences about patterns of species richness that form the basis of conservation strategies. Insects, despite their crucial roles in terrestrial ecosystems, are still largely neglected when dealing with biological inventories. Among insects, ants are of unique importance because of their species richness, widespread distribution, and due to their key ecosystem functions such as seed dispersal, soil nutrient cycling, predation, and biological control. In this study, we prioritize different Brazilian biomes and ecoregions for new ant surveys based on information on the distribution of occurrence records and two estimates of habitat loss for the period between 2000 and 2016. We compiled nearly 8000 ant occurrence records, including a total of 1170 species. The Caatinga was the biome showing the greatest urgency for new inventories, whereas the Atlantic Forest had the lowest urgency. However, there were considerable differences of priority between ecoregions belonging to each biome, thus underscoring the need to pay special attention to these geographical units and their characteristic associated with ant species. The prioritization of poorly studied sites with an imminent risk of habitat loss can be a valuable starting point for filling knowledge gaps and can help in formulating new strategies of conservation. The dataset provided here may also be useful in studies on the distribution of ant diversity in Brazil.

Keywords Biotic surveys · Habitat loss · Conservation · Wallacean shortfall · Formicidae · Brazil

Introduction

After nearly a decade, the state-of-affairs identified by May (2011) still rings true: "[w]e are astonishingly ignorant about how many species are alive on earth today, and even more ignorant about how many we can lose yet still maintain ecosystem services that humanity ultimately depends upon". Alleviating this plight is particularly challenging given that time and financial resources can severely limit the accumulation of knowledge on biodiversity (Blackburn and Gaston 2003). As a consequence, our current understanding of biodiversity patterns might be biased by sampling disproportionately common species in more easily accessible areas irrespective of the actual underlying species distributions (Grand et al. 2007; Boakes et al. 2010). This is particularly important, given that sampling biases and knowledge gaps directly influence the interpretation of ecological studies (Brown 1995; Blackburn and Gaston 2003; Vale and Jenkins 2012; Sousa-Baena et al. 2014), especially regarding local and global diversity patterns (Boakes et al. 2010; Meyer et al. 2015), thus hampering the design of more effective conservation strategies (Boitani et al. 2011).

Habitat loss is the main cause of species extinctions (Brooks et al. 2002; Hanski 2011; Barlow et al. 2016; Giam 2017) and biotic homogenization (McKinney 2006). These impacts go beyond simple reductions in the composition of local communities, as they also affect the functional and

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10841-020-00262-y) contains supplementary material, which is available to authorized users.

✉ Raquel Divieso
raqueldivieso@gmail.com

[1] Departamento de Zoologia, Universidade Federal do Paraná, C.P. 19020, Curitiba, PR 81531-980, Brazil

[2] Centro de Ciência do Sistema Terrestre, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP, Brazil

[3] African Climate and Development Initiative, University of Cape Town, Cape Town, South Africa

Published online: 12 August 2020

phylogenetic aspects of biological diversity (Arnan et al. 2018). This decrease in species richness, sometimes even before the lost species can be formally described, is likely to be more intense in environments susceptible to deforestation, particularly in the case of megadiverse groups, such as many invertebrates.

Insects are rarely used in the design of conservation policies, despite performing essential ecosystem services. For instance, ants not only can account for 25% of the animal biomass in tropical rainforests (Schultz 2000), but also perform numerous ecosystem functions such as seed dispersal, soil nutrient cycling and interactions with other organisms (Folgarait 1998). These interactions include associations with fungi (e.g. Currie 2001; Schultz 2008) and plants (e.g. Frederickson et al. 2005; Dáttilo et al. 2013; Ward and Branstetter 2017), and biological control by predation (e.g. Gonthier et al. 2013; Offenberg 2015). Some studies have investigated worldwide data on ant occurrence records to address questions dealing mainly with diversity distribution patterns (Andersen 1995; Kaspari et al. 2000; Dunn et al. 2007, 2009). The distribution of ant diversity in Brazil has been explored in studies on specific Brazilian biomes, in Amazon forest (Vasconcelos et al. 2010), Atlantic Forest (Silva and Brandão 2010, 2014), Cerrado (Vasconcelos et al. 2018), Caatinga (Leal et al. 2017), and Pampas (Dröse et al. 2017), yet we are still far from a complete understanding of the mechanisms governing ant diversity patterns.

Given the double challenge of incomplete knowledge about ant distribution patterns and the pressing risk of habitat loss, it is important to establish priorities for future ant surveys to be carried out in areas that are both poorly studied and are under high risk of habitat loss. Indeed, Guénard et al. (2012), through diversity prediction models, showed that regions with high potential for new ant discoveries tend to be under greater threat of deforestation. In this study we compile a large-scale dataset of ant occurrence records in Brazil and use it as a proxy for sampling effort throughout the country. This information was compared to estimates of habitat loss at two spatial levels, namely ecoregions and biomes. We then provide a statistic—the urgency index—that incorporates both sources of information and then applied it to generate priority rankings for Brazilian biomes and ecoregions, which can be used to establish priorities for future surveys.

Materials and methods

Occurrence records

Ant occurrence records in Brazil were compiled between 2017 and 2018 in two ways, always including only records in which ants were identified at the species level. Literature ant records were compiled through searching in the AntCat.org database (Bolton 2020). This is the main online source of ant taxonomic publications and contains all the original articles related to each nomenclatural act performed in ant literature since the publication of Linneus' "Systema Naturae" (1758). Given that the articles in this database are authored by taxonomic authorities in the different ant taxa, all the papers collected in this platform were deemed credible. Records were then screened considering the occurrence of the ant species in Brazil. Regarding the ecological literature, since not all of the identifications could be verified, the occurrences were used only when deemed credible (i.e., confirming taxa that were already known to Brazil or removing the occurrence when the author specialist certify that the record would be improbable). Each entry in the list is backed by at least one published reference or data source. The validity and authority of species names follow Bolton (2020), as implemented on AntCat.org. Species described as morphospecies and specimens identified only to a level higher than species were not included in the checklist. In addition, RMF, as ant taxonomist and co-author of this paper, revised species occurrences for all the papers considered eligible. When only the city name was informed, we used the central coordinate of the municipality as a reference point for the record. The collection dates and corresponding citations were also included in the dataset (Supplementary Material, Table S1).

In addition, ant records were obtained from the Antweb platform (https://www.antweb.org/, last accessed on 09/05/2019). This is an online platform in which photos and additional information from a large number of ants collected worldwide are maintained. Most specimens in this repository were identified and curated by well-known researchers. Existing records for Brazil were filtered and downloaded using aw_data function in *AntWeb* package v. 0.6.7 (Ram 2014), implemented in R 3.4.4 (R Core Team 2018). Duplicate records of the same species within a distance of 10 km were excluded from the final dataset. Finally, all of our records that had not been included in AntWeb were submitted to their database and are now available in that platform. The distribution of each record type is presented in Fig. S5 in the Supplementary Material.

Habitat loss data

We compiled data on deforestation and loss of natural vegetation cover for Brazil using two datasets: the mapping of forest cover produced by Hansen et al. (2013) and the data generated by the Annual Land Use and Land Cover Mapping Project in Brazil—MapBiomas (http://mapbiomas.org/, last accessed 08/20/2019). Both datasets were chosen because they reflect complementary properties of habitat loss, differing mainly in the estimation of canopy loss. The loss of natural cover calculated by the MapBiomas project considers the replacement of natural areas by other types of unnatural vegetation, such as urban areas and agricultural activity, for example. Open natural areas, therefore, were better represented by the MapBiomas estimates, given that differences in canopy height are not considered. On the other hand, given that data from Hansen et al. (2013) reports deforestation as loss of vegetation cover above 5 m, the loss of large trees could be detected even within a matrix of preserved native vegetation. In both cases, vegetation loss mapping from 2000 to 2016 was used as a proxy for habitat loss. This period reflects well the current state of Brazilian environments considering the implementation of new policies to combat deforestation in recent years (Arima et al. 2014; Brancalion et al. 2016; Dupin et al. 2018), and therefore are suitable to compare regions according to their imminent danger of habitat loss. More information on these datasets is provided in the Supplementary Material.

Analyses

All files with information on global deforestation for the years 2000 to 2016 (Hansen et al. 2013) were joined using the *Mosaic* operation. After joining the files, the *Reproject* operation was used to convert the spatial resolution to about 1 km^2 per pixel, generating a raster of georeferenced deforestation information. Finally, a shapefile with the geographical limits of Brazil, made available by IBGE (Instituto Brasileiro de Geografia e Estatística), was used to truncate the produced raster. Thus, the high coverage loss values (HCL) proposed by Hansen et al. (2013) was calculated considering the percentage of deforestation cells recorded between 2000 and 2016.

MapBiomas natural vegetation cover loss data (MCL) per year was obtained by the difference in natural vegetation cover area between 2000 and 2016 generated directly from the Google Earth Engine platform (https://code.earthengine.google.com/?accept_repo=users/mapbiomas/user-toolkit, accessed on 10/15/2018), also at a resolution of 1 km per pixel. Information available in the MapBiomas Collection 3 was used in this work, considering as loss of natural vegetation cover the decrease in the area composed of natural forest formations, savannas, mangroves and non-forest natural formations, including non-forest wetlands, grasslands and *apicum*.

Finally, the boundaries of Brazil's biomes and ecoregions were used to calculate the habitat loss of these regions from both layers. We considered 45 Brazilian ecoregions (Olson et al. 2001) analyzed according to the updated limits in shapefiles available on the Ecoregions 2017 platform © Resolve (https://ecoregions2017.appspot.com/, last accessed 05/08/2018). These same Ecoregions limits are available on the official platform of the Brazilian Ministry of the Environment (http://mapas.mma.gov.br/i3geo/datadownload.htm, accessed on 06/01/2020). Ecoregion boundaries were used because the structure of the environment reflects the type of vegetation, which is an important determinant of faunal composition, especially for ants (Lassau and Hochuli 2004). More complex than biomes, ecoregion boundaries are also defined based on soil type, hydrology and geographic formations, so that the use of these boundaries can demonstrate variation in associated diversity with better refinement. We therefore assume that habitat loss in different ecoregions may lead to the loss of a different set of species. In addition, the limits of Brazilian biomes were also used, mainly because they are political delimitations used in designing conservation strategies or resource allocation for new research. Layers of the biomes were obtained through the governmental platform maps.mma.gov.br (last accessed 01/08/2018). All georeferencing operations were performed using the open QGIS software (version 2.18).

With the occurrence data of Brazilian ant species combined in the same dataset, a sampling pattern was generated for Brazil, identifying the sampling density (records/km²) for each area, between biomes and ecoregions. Using these sampling density data and habitat loss data new urgency indices (*UI*) were calculated for all biomes and ecoregions. As the two vegetation loss data have different underlying methodologies, the UI index was calculated separately for each one. This collection urgency index corresponds to the percentage of habitat loss in each region, weighted by the highest habitat loss value, divided by the corresponding density of ant species records also weighted by the highest calculated register density value. Thus, a region would have a high UI when under a severe deforestation regime and/or with a low sampling level. These weights were used in order to keep the values at a comparable scale for comparisons. To avoid negative or zero values, the habitat difference between 2000 and 2016 that was zero or below zero found by the MCL estimate was replaced by 0.01. This has not affected the order of priorities as we disregarded reforestation areas for all regions. The UI was calculated as

$$\mathrm{UI} = \frac{\mathrm{P}/\mathrm{P}_{max}}{D/D_{max}}$$

where: P = percentage of habitat loss in the area (biome or ecoregion), P_{max} = highest percentage of calculated habitat loss between areas; D = sampling density; D_{max} = highest calculated sampling density between areas.

Despite differences in magnitude, habitat loss values of the ecoregions from HCL and MLC were concordant, with a strong and positive correlation between the two UI estimates ($r^2 = 0.73$, $n = 45$, $P < 0.001$; Fig. S1). Therefore, although there were some differences in priorities between ecoregions according to the type of habitat loss estimate, overall the urgency of new collections followed a consistent pattern across different metrics. We therefore used the average between the HCL and MCL UI values to list the ecoregions according to their priority and categorize all ecoregions into three groups of high, medium and low priority for new collections according to the region's rank. Finally, the collection date information included in the compiled dataset was used to explore how occurrence records were distributed temporally among Brazilian ecoregions. In particular, we assessed whether there was a pattern of concentration or increase of sampling in past few decades and if there were fluctuations in the collection intensity in each ecoregion. It is important to note that qualitatively similar patterns were obtained omitting the AntWeb records (Table S2) and therefore will not be explored further here.

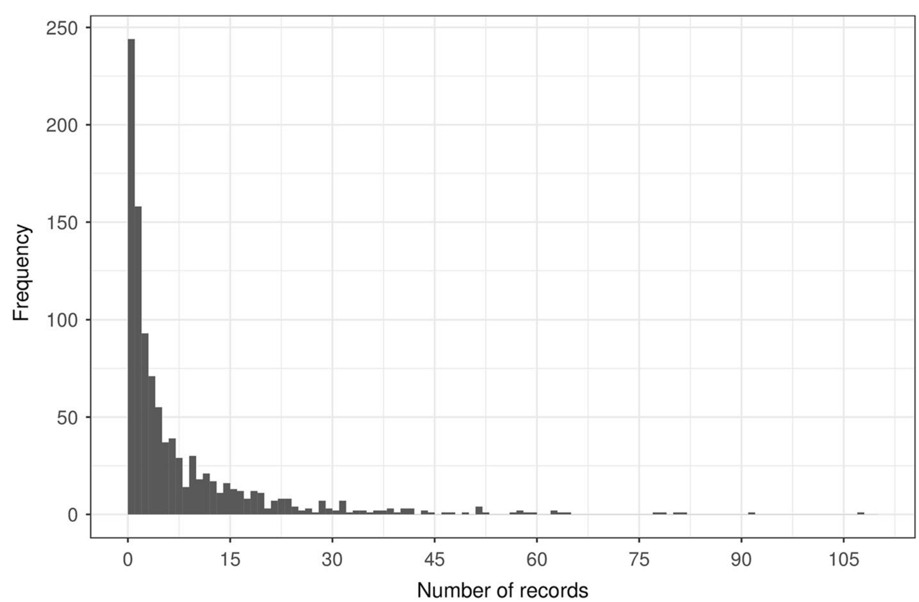
Results

The total set of compiled data involved 7804 occurrence records ($\bar{x} = 8.52 \pm 13.48$ records per species, median = 4, 1–32 95% quantiles, Fig. 1), of which 5447 from the literature and 2357 from AntWeb. There were 1170 species were from 113 genera sampled, of the approximately 1500 valid species with records confirmed for Brazil. This dataset involved occurrence data on approximately 78% of all ant species recorded for Brazil. All occurrence records used in our study have been made available on the AntWeb platform. By mapping occurrence data in a geographical context, the heterogeneous distribution of records across the Brazilian territory becomes clear (Fig. 2A). With the exception of the far south, there is a concentration of ant occurrences in the south and southeast regions, where the largest and oldest research centers in the country are located. There are large sampling gaps, mainly in the midwest region, and a low number of records in the northeast. Caatinga was the biome with the lowest number of records per area, whereas the Atlantic Forest was the best sampled biome (Table 1).

Habitat loss values calculated from Hansen et al. (2013) (HCL) and MapBiomas project (MCL) showed some discrepancies between biomes and ecoregions, as expected (Tables 1 and 2; Fig. 2b, Supplementary Material Fig. S2A). Overall, habitat loss values were higher when calculated using data from Hansen et al. (2013). However, for native vegetation regions characterized by open areas such as the Pampas, habitat loss values were higher when calculated from MapBiomas. When considering occurrences and habitat loss data for both used estimates, the UI values on the maps clearly highlight priority regions for future collection efforts (Table 1; Fig. 2c and d, Supplementary Material Figs. S2B, C). In terms of biomes, Caatinga represents the biome that should receive most special attention, whereas the Atlantic Forest would have the lowest priority given the considerable effort of collections already carried out along its extension. Priority among the other biomes changes according to the difference in HCL and MCL values.

When we evaluated UI values at the ecoregion scale for both categories of cover loss, we can notice considerable heterogeneity at the level of ecoregions within each biome (Figs. 2d and S2C), as observed in western Amazonia. The Guianan piedmont moist forests ecoregion has a high priority compared to other ecoregions with much lower UI, such as the Japurá-Solimões-Negro moist forests area (Table 2). The Gurupa várzea ecoregion has the highest collection priority for both habitat loss categories followed by the Maranhão Babaçu forests and Mato Grosso tropical dry forests. These three ecoregions are differentiated in terms of risk

Fig. 1 Frequency of occurrence records per species obtained from the literature and the AntWeb database



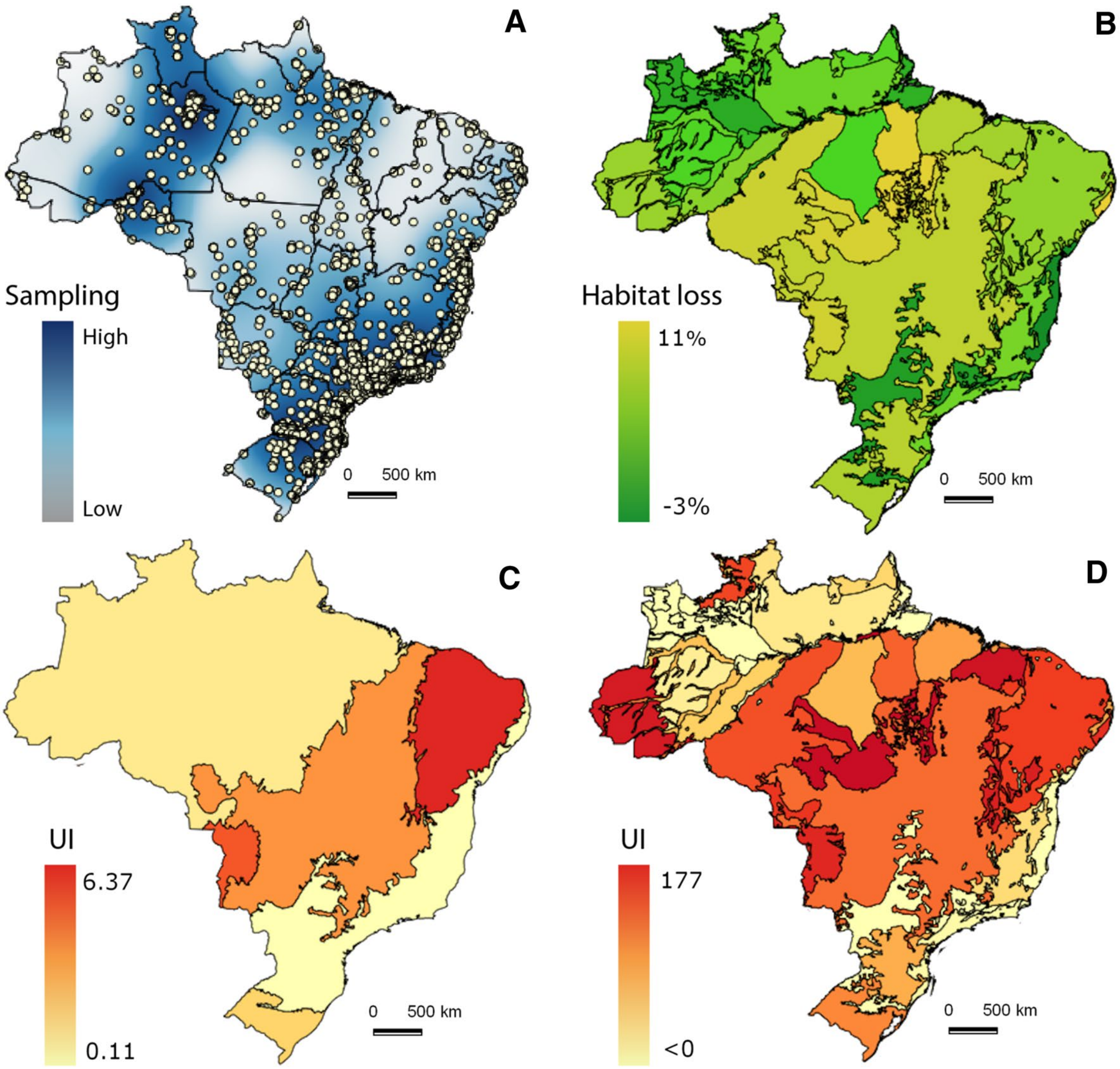


Fig. 2 Maps containing information on: **a** Density of ant records among Brazilian states containing all compiled occurrence points. The kernel smoothing methodology was used for record interpolation, with more intense blue color in regions with higher occurrence density; **b** Percentage of natural cover loss (MCL) calculated from MapBiomas land cover data between 2000 and 2016. **c** Urgency Index (IU) calculated for Brazilian Biomes based on record density and MCL values for each biome; **d** Urgency index calculated for Brazilian ecoregions also based on record density and MCL values

Table 1 Number of occurrence records per km^2, percentages of high coverage loss (PCA), loss of natural coverage (PCN) between 2000 and 2016 and their respective Urgency Index (UI) values and priority for new collections for each Brazilian biome

Biome	Nº records/km^2	HCL (%)	MCL (%)	IU_{HCL}	IU_{MCL}	Priority HCL	Priority MCL
Amazon	1.01	6.92	3.44	3.43	1.98	3º	5º
Caatinga	0.22	3.33	2.41	7.58	6.37	1º	1º
Cerrado	0.57	6.95	5.31	6.06	5.38	2º	3º
Atlantic Forest	3.51	5.01	0.67	0.71	0.11	6º	6º
Pampa	0.76	2.16	5.11	1.42	3.91	5º	4º
Pantanal	0.63	3.58	6.03	2.85	5.58	4º	2º

Table 2 Number of occurrence records per km^2, percentage of high coverage loss by Hansen et al. (2013; HCL), and natural coverage loss by MapBiomas project (MCL), their respective urgency index (UI) values, and Priority of collection for each Brazilian ecoregion

Ecoregion	Nº records /km	HCL (%)	MCL (%)	UI $_{HCL}$	UI $_{MCL}$	Priority HCL	Priority MCL	Priority level
Gurupa várzea	0.01	4.54	2.21	202.37	176.96	1º	1º	High
Maranhão Babaçu forests	0.1	13.16	4.02	58.66	32.19	2º	3º	High
Mato Grosso tropical dry forests	0.22	13.71	9.79	27.78	35.63	5º	2º	High
Tapajós-Xingu moist forests	0.11	8.44	3.06	34.20	22.27	4º	4º	High
Brazilian Atlantic dry forests	0.15	7.07	2.57	21.01	13.72	7º	5º	High
Guianan piedmont moist forests	0.01	1.08	0.12	48.14	9.61	3º	10º	High
Chiquitano dry forests	0.44	8.88	5.93	9.00	10.79	10º	8º	High
Caatinga	0.18	3.29	2.4	8.15	10.68	12º	9º	High
Madeira-Tapajós moist forests	0.49	9.68	5.99	8.81	9.79	11º	11º	High
Pantepui forests & shrublands	0.01	0.4	0.09	17.83	7.21	8º	15º	High
Northeast Brazil restingas	0.1	5.36	0.54	23.89	4.32	6º	19º	High
Pantanal	0.38	3.53	6.05	4.14	12.75	22º	6º	High
Xingu-Tocantins-Araguaia moist forests	1.06	15.09	11.09	6.35	8.38	17º	13º	High
Iquitos várzea	0.58	9.35	4.43	7.19	6.12	16º	16º	High
Pernambuco coastal forests	0.75	3.53	11.39	2.10	12.16	26º	7º	High
Cerrado	0.55	6.62	5.44	5.37	7.92	19º	14º	Medium
Tocantins Pindare moist forests	1.16	20.46	5.13	7.86	3.54	14º	20º	Medium
Southwest Amazon moist forests	0.2	2.81	1.09	6.26	4.36	18º	18º	Medium
Humid Chaco	0.86	5.25	10.25	2.72	9.54	25º	12º	Medium
Purus várzea	0.11	1.87	0.12	7.58	0.87	15º	24º	Medium
Guianan highlands moist forests	0.04	0.41	0.02	4.57	0.40	21º	25º	Medium
Uruguayan savanna	0.71	2.24	4.63	1.41	5.22	30º	17º	Medium
Amazon-Orinoco-Southern Caribbean mangroves	0.63	5.38	0.92	3.81	1.17	24º	23º	Medium
Pernambuco interior forests	1.47	5.23	3.8	1.59	2.07	29º	21º	Medium
Caqueta moist forests	0.01	0.35	-0.02	15.60	-1.60	9º	45º	Medium
Araucaria moist forests	3.42	7.07	5.27	0.92	1.23	32º	22º	Medium
Guianan lowland moist forests	0.09	0.38	0.02	1.88	0.18	27º	27º	Medium
Solimões-Japurá moist forests	0.01	0.18	0.01	8.02	0.80	13º	44º	Medium
Negro-Branco moist forests	0.08	0.84	0.01	4.68	0.10	20º	38º	Medium
Juruá-Purus moist forests	0.07	0.29	0.01	1.85	0.11	28º	31º	Medium
Bahia interior forests	2.97	5.98	0.52	0.90	0.14	33º	28º	Medium
Marajó várzea	0.25	2.23	0.01	3.98	0.03	23º	41º	Low
Uatumã-Trombetas moist forests	1.54	2.68	0.13	0.78	0.07	34º	30º	Low
Caatinga Enclaves moist forests	2.51	3.59	0.11	0.64	0.04	35º	32º	Low
Purus-Madeira moist forests	4.33	4.22	1.7	0.43	0.31	42º	26º	Low
Guianan savanna	3.99	4.55	0.41	0.51	0.08	40º	29º	Low
Monte Alegre várzea	3.17	4.03	0.01	0.57	0.00	37º	34º	Low
Bahia coastal forests	4.32	11.15	0.01	1.15	0.00	31º	42º	Low
Japurá-Solimões-Negro moist forests	0.66	0.81	0.14	0.55	0.17	38º	39º	Low
Campos Rupestres montane savanna	2.81	2.74	0.01	0.43	0.00	41º	37º	Low
Serra do Mar coastal forests	9.12	4.26	0.26	0.21	0.02	45º	33º	Low
Atlantic coast restingas	2.67	3.47	0.01	0.58	0.00	36º	43º	Low
Alto Parana Atlantic forests	1.67	1.91	0.01	0.51	0.00	39º	40º	Low
Southern Atlantic Brazilian mangroves	7.36	4.77	0.01	0.29	0.00	43º	36º	Low
Campinaranas de Alto Rio Negro	2.34	1.18	0.01	0.22	0.00	44º	35º	Low

Priority level categories were calculated by the average value of collection priorities between the two types of habitat loss status and knowledge of ant fauna. Although not under such a severe habitat loss regime, the highest priority ecoregion, Gurupa várzea, has no record of ants in our occurrence data set, which highlights the need for surveys in that area. On the other hand, the Maranhão Babaçu forests and Mato Grosso tropical dry forests have shown high levels of deforestation in recent years, where canopy cover fell by 13% between 2000 and 2016. Even though these areas do not have relatively low numbers of occurrence records, much of the vegetation is being lost. Thus, we notice that our index highlights equally worrying areas due to either the low knowledge of the fauna or the imminent loss of habitat.

Results relating to the temporal distribution of occurrence records are presented in Supplementary Material (Figs. S3 and S4). No period of consistent increase or decrease has been observed for all ecoregions for the last decades. However, even with fluctuations, some areas were consistently poorly sampled while others showed oscillations within a higher sampling margin, such as western Amazonia and southeastern Brazil, respectively.

Discussion

The extensive compilation of occurrence records carried out in this study is a breakthrough in our understanding of the large-scale distribution of ants in Brazil. While such extensive compilations provide important information about distribution patterns, they can be even more informative about data deficiency, directing strategies for addressing biodiversity knowledge gaps more efficiently (Brito 2010). Therefore, our data set not only identifies the regions with collection biases, but also indicates which regions have the least studies and the most limited knowledge on its ant fauna and possibly other invertebrate groups. This is particularly valuable, given that biological sampling is not evenly distributed across the globe (Grand et al. 2007), being concentrated in areas near research centers and large cities, as well as those with easier access to collection sites (Kadmon et al. 2004; Carneiro et al. 2008; Boakes et al. 2010; Oliveira et al. 2016), as seems to be the case in our study. Moreover, similar concentration of sampling in different Brazilian biomes has also been observed for other arthropods (see Oliveira et al. 2016), with the Atlantic Forest and the Caatinga being the most and least sampled biomes, respectively.

The numerous impacts of deforestation and fragmentation scenarios on biotas are a constant cause of concern (Laurance et al. 2009; Clark et al. 2010; Caro et al. 2014). For ants, many studies show changes and decreases in diversity along fragmented, disturbed or regenerating gradients (Floren and Linsenmair 2001; Ribas et al. 2005; Silva et al. 2007). According to both habitat loss estimates observed here, Caatinga is the biome with the greatest need for new ant inventories. Although considered as one of the main dry tropical forest areas in the world with a unique biological composition (Mittermeier et al. 2002) and high diversity, Caatinga is still poorly studied (Overbeck et al. 2015; Oliveira and Bernard 2017) and quite neglected in terms of taxonomic studies (Santos et al. 2011) and knowledge of biodiversity, especially for insects (Oliveira et al. 2016). As a major contribution to fill some of this gap, Leal et al. (2017) provide a multi-faceted approach on the composition of ant species in the Caatinga and its intrinsic relationship with anthropogenic expansion and environmental changes. As demonstrated in that study, (1) ants and their provided services are very sensitive to human disturbances in this biome, (2) remnants of native Caatinga vegetation are small and fragmented (see also Prado 2003), and (3) the recent habitat loss in this biome is also of concern, as protected areas account for only 7.5% of total area (data from the Brazilian Ministry of the Environment, https://www.mma.gov.br/areas-protegidas.html. Accessed on 01/03/2019). The work by Arnan et al. (2018) suggests a decrease in the functional and phylogenetic diversity of Caatinga ants with increasing of anthropogenic disturbance, emphasizing the sensitivity of these organisms and their intrinsic relationship with different areas, even within the same biome.

According to our prioritization rates, some areas of the Caatinga, when considered as more specific ecoregion domains, do not present such an alarming scenario, whereas areas located north of the biome are among the sites with the highest priority. The same is true in the Amazon, where intense sampling biases, especially near major research centers, dilute the priority of the biome as a whole. When one considers the ecoregions of the biome, it is possible to identify priority sites on a more refined scale. This shows that considering the total area of the biome in terms of sampling and habitat loss may not provide as accurate information as considering more heterogeneous environments such as ecoregions.

Recently, using new technologies and integration of various data types, Smith et al. (2018) demonstrated that ecoregions represent very robust biological units. It has been tested and confirmed that currently used limits delineate terrestrial biodiversity standards well. Thus, we can assume that the results based on ecoregion delimitations might better reflect the urgencies in the need for knowledge of the different ant communities. Still, we emphasize that the boundaries of Brazilian biomes are also important, as they reflect more political strategies in decision making, especially for resource allocation for conservation and research. In 2015, the Ministry of Environment, in partnership with the LIFE Institute, published a Technical Notebook (Available at: https://institutolife.org/material/caderno-tecnico-vol-3-ecorregioes-do-brasil-prioridades-terrestres-e-marinhas/, Accessed: 10/06/2018) which contains information about Brazilian ecoregions, including the cities belonging to each ecoregion, thus facilitating the knowledge of the characterization of the entire territory.

We also underscore that the prioritization of areas for new inventories was based on two important criteria - prior ant knowledge and current habitat loss - but these are not the only factors that highlight areas in need of further research. The determination of knowledge hotspots for ants, as argued by Guénard et al. (2012) indicates the choice of priority areas for further studies according to the ant diversity potential. They used climate modeling and diversity interpolation to estimate which sites would be most unknown in terms of diversity on a global scale. The results of this research showed that the Brazilian Northeast is among the world regions with the highest ant knowledge potential, especially the states of Piauí, Sergipe, Rio Grande do Norte, and Paraíba. Most of these areas belong to the Caatinga biome, which presented the highest priority for new samples according to our results. In addition, a very recent study (Jory and Feitosa 2020) presented a list of species of Piauí and showed that in two collecting expeditions in national parks, 58 new records were obtained for the state, eight for the northern region and one for Brazil. These results reaffirm the potential for diversity discoveries in the region.

Normally, strategies for conservation focus on more diverse environments and/or areas with high rates of endemism (Myers et al. 2000; Butchart et al. 2010). Despite great efforts in developing remediation strategies to reduce sampling bias in different types of studies (e.g. García 2006; Costa et al. 2010; Engemann et al. 2015), biodiversity databases still have limitations for many poorly studied sites and organisms (Soberón et al. 2000, 2007; Hortal et al. 2007). This lack of knowledge impacts the conservation of species, given that regions with less sampled diversity may also be environments with less focus on conservation strategies. Other works with the same approach proposed here may guide new sampling efforts for different groups and directly influence the delimitation of priority areas for conservation.

Although sampling in most ecoregions fluctuate without any obvious temporal pattern, the most sampled ecoregions had a more consistent increase in sampling from the 2000s onwards (Supplementary Material, Fig. S4). Even though these temporal data were only analyzed until 2015, due to the slow processing of collected material, much of this sampling effort of the last decade still awaits in collections to be properly analyzed. We expect this small increase to reflect more robust ant knowledge numbers in recent years. In particular, investing in the knowledge of ant faunas is a valuable tool for conservation planning, given that ants are important bioindicators of environmental disturbance (Andersen 1997; Andersen et al. 2002; Brandão et al. 2011) and can provide good monitoring of environmental conditions, mainly in their species composition (Kaspari and Majer 2000; Ribas et al. 2012; Underwood and Fisher 2006).

There is an important caveat for our use of the UI: by integrating two different variables into the same index, we are giving equal weight to sampling effort and habitat loss, which are very distinct types of data. Meanwhile, the weight given to each variable can be easily changed according to different emphases for prioritization. We fully acknowledge this limitation and we advise its use with the appropriate caution and good judgement. However, it is important to note that this is not a limitation that is exclusive of the UI. For the sake of argument, let us assume we were focusing exclusively on habitat loss. One could raise similar concerns despite that reduction in scope given that, for instance, habitat loss from logging would have different consequences from burning or mining. In addition, measuring habitat loss might not properly describe whether the impact is acute or chronic, or the existence of selective logging. Likewise, often conservation priorities are based largely on vertebrates, which might not reflect the remaining local biota (Donaldson et al. 2016). However, we still do rely on such summary statistics in spite of their limitations for two main reasons: (1) they are operational, as they provide a convenient way to rapidly compare and contrast different conditions of interest, and (2) they are likely to lead—more often than not—to informed decision making. We would not advise researchers to use our index as the sole criterion for new ant surveys, but we believe that it might provide important insight to design future surveys in a more strategic fashion.

Acknowledgements We thank Eduardo Carneiro dos Santos, Thiago J. Izzo, Mariane A. Nickele, Fernanda T. Brum, and Renattho Nitz Oliveira for their contributions and comments during the development of this work.

Funding This study was funded by Brazilian Coordination for the Improvement of Higher Education Personnel (CAPES). RMF was supported by the Brazilian Council of Research and Scientific Development (CNPq, Grant 302462/2016-3).

Data availability All records compiled in this study are available on the antweb.org platform.

Code availability Not applicable.

Compliance with ethical standards

Conflict of interests The authors declare no potential conflict of interests in this article.

References

Andersen A (1997) Using ants as bioindicators: multiscale issues in ant community ecology. Conserv Ecol 1(1):8. https://doi.org/10.5751/ES-00014-010108

Andersen AN (1995) Measuring more of biodiversity: genus richness as a surrogate for species richness in Australian ant faunas. Biol Conserv 73(1):39–43. https://doi.org/10.1016/0006-3207(95)90059-4

Andersen AN, Hoffmann BD, Muller WJ, Griffiths AD (2002) Using ants as bioindicators in land management: simplifying assessment of ant community responses. J Appl Ecol 39:8–17. https://doi.org/10.1046/j.1365-2664.2002.00704.x

Arima EY, Barreto P, Araújo E, Soares-Filho B (2014) Public policies can reduce tropical deforestation: lessons and challenges from Brazil. Land Use Policy. https://doi.org/10.1016/j.landusepol.2014.06.026

Arnan X, Arcoverde GB, Pie MR, Ribeiro-Neto JD, Leal IR (2018) Increased anthropogenic disturbance and aridity reduce phylogenetic and functional diversity of ant communities in Caatinga dry forest. Sci Total Environ 631–632:429–438. https://doi.org/10.1016/j.scitotenv.2018.03.037

Barlow J, Lennox GD, Ferreira J et al (2016) Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. Nature 535:144–147. https://doi.org/10.1038/nature18326

Blackburn TM, Gaston KJ (2003) Macroecology: concepts and consequences. Cambridge University Press, Cambridge

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(6):e1000385. https://doi.org/10.1371/journal.pbio.1000385

Boitani L, Maiorano L, Baisero D, Falcucci A, Visconti P, Rondinini C (2011) What spatial data do we need to develop global mammal conservation strategies? Philos Trans R Soc Lond B Biol Sci 366:2623–2632. https://doi.org/10.1098/rstb.2011.0117

Bolton B (2020) An online catalog of the ants of the world. http://antcat.org. Accessed 20 Apr 2020

Brancalion PHS, Garcia LC, Loyola R, Rodrigues RR, Pillar VD, Lewinsohn TM (2016) A critical analysis of the native vegetation protection law of Brazil 2012: updates and ongoing initiatives. Nat Conserv 14:1–15. https://doi.org/10.1016/j.ncon.2016.03.003

Brandão CRF, Silva RR, Feitosa RM (2011) Brazilian Savanna ground-dwelling ants (Hymenoptera: Formicidae) as indicators of edge effects. Zoologia 28:379–387. https://doi.org/10.1590/S1984-46702011000300012

Brito D (2010) Overcoming the Linnean shortfall: data deficiency and biological survey priorities. Basic Appl Ecol 11(8):709–713. https://doi.org/10.1016/j.baae.2010.09.007

Brooks TM, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Rylands AB, Konstant WR, Flick P, Pilgrim J, Oldfield S, Magin G, Hilton-Taylor C (2002) Habitat loss and extinction in the hotspots of biodiversity. Conserv Biol 16(4):909–923. https://doi.org/10.1046/j.1523-1739.2002.00530.x

Brown JH (1995) Macroecology. University of Chicago Press, Chicago

Butchart SHM, Walpole M, Collen B et al (2010) Global biodiversity: indicators of recent declines. Science 328:1164–1168. https://doi.org/10.1126/science.1187512

Carneiro ES, Mielke O, Casagrande MM (2008) Butterfly inventories in Brazil: the state of the art and the priority-areas model for research aiming at conservation. Nat Conserv 6(2):176–198

Caro T, Dobson A, Marshall AJ, Peres CA (2014) Compromise solutions between conservation and road building in the tropics. Curr Biol 24(16):722–725. https://doi.org/10.1016/j.cub.2014.07.007

Clark RW, Brown WS, Stechert R, Zamudio KR (2010) Roads, interrupted dispersal, and genetic diversity in timber rattlesnakes: roads and population genetics. Conserv Biol 24(4):1059–1069. https://doi.org/10.1111/j.1523-1739.2009.01439.x

Costa GC, Nogueira C, Machado RB, Colli GR (2010) Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. Biodivers Conserv 19:883–899. https://doi.org/10.1007/s10531-009-9746-8

Currie CR (2001) A community of ants, fungi, and bacteria: a multilateral approach to studying symbiosis. Annu Rev Microbiol 55:357–380. https://doi.org/10.1146/annurev.micro.55.1.357

Dáttilo W, Guimarães PR, Izzo TJ (2013) Spatial structure of ant-plant mutualistic networks. Oikos 122(11):1643–1648. https://doi.org/10.1111/j.1600-0706.2013.00562.x

Donaldson MR, Burnett NJ, Braun DC, Suski CD, Hinch SG, Cooke SJ, Kerr JT (2016) Taxonomic bias and international biodiversity conservation research. FACETS 1(1):105–113. https://doi.org/10.1139/facets-2016-0011

Dröse W, Podgaiski LR, Cavalleri A, Feitosa RM, Mendonca JM (2017) Ground-dwelling and vegetation ant fauna in southern Brazilian grasslands. Sociobiology 64(4):381–392. https://doi.org/10.13102/sociobiology.v64i4.1795

Dunn RR, Agosti D, Andersen AN et al (2009) Climatic drivers of hemispheric asymmetry in global patterns of ant species richness. Ecol Lett 12:324–333. https://doi.org/10.1111/j.1461-0248.2009.01291.x

Dunn RR, Sanders NJ, Fitzpatrick MC et al (2007) Global ant (Hymenoptera: Formicidae) biodiversity and biogeography: a new database and its possibilities. Myrm News 10:77–83

Dupin MGV, Espírito-Santo MM, Leite ME, Silva JO, Rocha AM, Barbosa RS, Felisa AC (2018) Land use policies and deforestation in Brazilian tropical dry forests between 2000 and 2015. Environ Res Lett 13(3):035008. https://doi.org/10.1088/1748-9326/aaadea

Engemann K, Enquist BJ, Sandel B, Boyle B, Jørgensen PM, Morueta-Holme N, Peet RK, Violle C, Svenning JC (2015) Limited sampling hampers "big data" estimation of species richness in a tropical biodiversity hotspot. Ecol Evol 5(3):807–820. https://doi.org/10.1002/ece3.1405

Floren AD, Linsenmair KE (2001) The influence of anthropogenic disturbances on the structure of arboreal arthropod communities. Plant Ecol 153:153–167. https://doi.org/10.1023/A:1017510312462

Folgarait PJ (1998) Ant biodiversity and its relationship to ecosystem functioning: a review. Biodiver Conserv 7:1221–1244. https://doi.org/10.1023/A:1008891901953

Frederickson ME, Greene MJ, Gordon DM (2005) 'Devil's gardens' bedevilled by ants: ecology. Nature 437:495–496. https://doi.org/10.1038/437495a

García A (2006) Using ecological niche modelling to identify diversity hotspots for the herpetofauna of Pacific lowlands and adjacent interior valleys of Mexico. Biol Conserv 130:25–46. https://doi.org/10.1016/j.biocon.2005.11.030

Giam X (2017) Global biodiversity loss from tropical deforestation. Proc Natl Acad Sci 114(23):5775–5777. https://doi.org/10.1073/pnas.1706264114

Gonthier DJ, Ennis KK, Philpott SM, Vandermeer J, Perfecto I (2013) Ants defend coffee from berry borer colonization. BioControl 58:815–820. https://doi.org/10.1007/s10526-013-9541-z

Grand J, Cummings MP, Rebelo TG, Ricketts TH, Neel MC (2007) Biased data reduce efficiency and effectiveness of conservation reserve networks. Ecol Lett 10(5):364–374. https://doi.org/10.1111/j.1461-0248.2007.01025.x

Guénard B, Weiser MD, Dunn RR (2012) Global models of ant diversity suggest regions where new discoveries are most likely are under disproportionate deforestation threat. Proc Natl Acad Sci 109(19):7368–7373. https://doi.org/10.1073/pnas.1113867109

Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman SV, Goetz SJ, Loveland TR, Kommareddy A, Egorov A, Chini L, Justice CO, Townshend JRG (2013) High-resolution global maps of 21st-century forest cover change. Science 342(6169):850–853. https://doi.org/10.1126/science.1244693

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

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(3):853–863. https://doi.org/10.1111/j.1523-1739.2007.00686.x

Jory TT, Feitosa R (2020) First survey of the ants (Hymenoptera, Formicidae) of piauí: filling a major knowledge gap about ant diversity in Brazil. Pap Avuls Zool 60:e20206014. https://doi.org/10.11606/1807-0205/2020.60.14

Kadmon R, Farber O, Danin A (2004) Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. Ecol Appl 14:401–413. https://doi.org/10.1890/02-5364

Kaspari M, Majer JD (2000) Using ants to monitor environmental change. In: Agosti D, Majer JD, Alonso LE, Schultz TR (eds) Ants: standard methods for measuring and monitoring biodiversity. Smithsonian Institution Press, Washington, DC, pp 89–98

Kaspari M, O'Donnell S, Kercher JR (2000) Energy, density, and constraints to species richness: ant assemblages along a productivity gradient. Am Nat 155:280–293. https://doi.org/10.2307/3078948

Lassau SA, Hochuli DF (2004) Effects of habitat complexity on ant assemblages. Ecography 27:157–164. https://doi.org/10.1111/j.0906-7590.2004.03675.x

Laurance WF, Goosem M, Laurance SGW (2009) Impacts of roads and linear clearings on tropical forests. Trends Ecol Evol 24:659–669. https://doi.org/10.1016/j.tree.2009.06.009

Leal IR, Ribeiro-Neto JD, Arnan X, Oliveira FMP, Arcoverde GB, Feitosa RM, Andersen AN (2017) Ants of the Caatinga: diversity, biogeography, and functional responses to anthropogenic disturbance and climate change. In: da Silva JMC, Leal IR, Tabarelli M (eds) Caatinga: the largest tropical dry forest region in South America, Switzerland. Springer, New York, pp 65–95

May RM (2011) Why worry about how many species and their loss? PLoS Biol 9:e1001130. https://doi.org/10.1371/journal.pbio.1001130

McKinney ML (2006) Urbanization as a major cause of biotic homogenization. Biol Conserv 127(3):247–260. https://doi.org/10.1016/j.biocon.2005.09.005

Meyer C, Kreft H, Guralnick R, Jetz W (2015) Global priorities for an effective information basis of biodiversity distributions. Nat Commun 6:8221. https://doi.org/10.1038/ncomms9221

Mittermeier RA, Mittermeier CG, Gil PR, Pilgrim J, Fonseca G (2002) Wilderness: earth's last wild places. CEMEX, Mexico City

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

Offenberg J (2015) Ants as tools in sustainable agriculture. J Appl Ecol 52(5):1197–1205. https://doi.org/10.1111/1365-2664.12496

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

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR (2001) Terrestrial ecoregions of the world: a new map of life on earth. Bioscience 51(11):933–938. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Overbeck GE, Vélez-Martin E, Scarano FR, Lewinsohn TM, Fonseca CR, Meyer ST, Müller SC, Ceotto P, Dadalt L, Durigan G, Ganade G, Gossner MM, Guadagnin DL, Lorenzen K, Jacobi CM, Weisser WW, Pillar VD (2015) Conservation in Brazil needs to include non-forest ecosystems. Divers Distrib 21(12):1455–1460. https://doi.org/10.1111/ddi.12380

Prado DE (2003) As caatingas da américa do Sul. In: Leal IR, Tabarelli M, Silva JMC (ed) Ecologia e conservação da caatinga. Ed. Universitária da UFPE, Recife, pp 3–73

Quantum GIS Development Team (2015) QGIS GeographicInformation System. Open Source Geospatial Foundation Project. https://www.qgis.org Accessed 23 June 2017

R Development Core Team (2015) R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.r-project.org. Accessed 10 May 2017

Ram K (2014) AntWeb: programmatic interface to the AntWeb. R package version 0.6.7

Ribas CR, Sobrinho TG, Schoereder JH, Sperber CF, Lopes-Andrade C, Soares SM (2005) How large is large enough for insects? Forest fragmentation effects at three spatial scales. Acta Oecol 27(1):31–41. https://doi.org/10.1016/j.actao.2004.08.008

Ribas C, Campos R, Schmidt F, Solar R (2012) Review article ants as indicators in Brazil: a review with suggestions to improve the use of ants in environmental monitoring programs. Psyche: J Entomol. https://doi.org/10.1155/2012/636749

Santos J, Leal I, Almeida-Cortez J, Fernandes G, Tabarelli M (2011) Caatinga: the scientific negligence experienced by a dry tropical forest. Trop Conserv Sci 4:276–286. https://doi.org/10.1177/194008291100400306

Schultz TR, Brady SG (2008) Major evolutionary transitions in ant agriculture. Proc Natl Acad Sci 105(14):5435–5440. https://doi.org/10.1073/pnas.0711024105

Schultz TR (2000) In search of ant ancestors. Proc Natl Acad Sci 97(26):14028–14029. https://doi.org/10.1073/pnas.011513798

Silva RR, Brandão CRF (2010) Morphological patterns and community organization in leaf-litter ant assemblages. Ecol Monogr 80(1):107–124. https://doi.org/10.1890/08-1298.1

Silva RR, Brandão CRF (2014) Ecosystem-wide morphological structure of leaf-litter ant communities along a tropical latitudinal gradient. PLoS ONE 9:e93049. https://doi.org/10.1371/journal.pone.0093049

Silva RR, Feitosa RSM, Eberhardt F (2007) Reduced ant diversity along a habitat regeneration gradient in the southern Brazilian Atlantic forest. For Ecol Manage 240:61–69. https://doi.org/10.1016/j.foreco.2006.12.002

Smith J, Letten AD, Ke PJ, Anderson C et al (2018) A global test of ecoregions. Nat Ecol Evol 2:1889–1896. https://doi.org/10.1038/s41559-018-0709-x

Soberón JM, Llorente JB, Oñate L (2000) The use of specimen-label databases for conservation purposes: an example using Mexican papilionid and pierid butterflies. Biodivers Conserv 9:1441–1466. https://doi.org/10.1023/A:1008987010383

Soberón J, Jiménez R, Golubov J, Koleff P (2007) Assessing completeness of biodiversity databases at different spatial scales. Ecography 30(1):152–160. https://doi.org/10.1111/j.0906-7590.2007.04627.x

Sousa-Baena MS, Garcia LC, Peterson AT (2014) Completeness of digital accessible knowledge of the plants of Brazil and priorities for survey and inventory. Divers Distrib 20(4):369–381. https://doi.org/10.1111/ddi.12136

Underwood EC, Fisher BL (2006) The role of ants in conservation monitoring: if, when, and how. Biol Conserv 132(2):166–182. https://doi.org/10.1016/j.biocon.2006.03.022

Vale MM, Jenkins CN (2012) Across-taxa incongruence in patterns of collecting bias. J Biogeogr 39(9):1743–1748. https://doi.org/10.1111/j.1365-2699.2012.02750.x

Vasconcelos HL, Maravalhas JB, Feitosa RM, Pacheco R, Neves KC, Andersen AN (2018) Neotropical savanna ants show a reversed latitudinal gradient of species richness, with climatic drivers reflecting the forest origin of the fauna. J Biogeogr 45(1):248–258. https://doi.org/10.1111/jbi.13113

Vasconcelos HL, Vilhena JM, Facure KG, Albernaz AL (2010) Patterns of ant species diversity and turnover across 2000 km of Amazonian floodplain forest. J Biogeogr 37:432–440. https://doi.org/10.1111/j.1365-2699.2009.02230.x

Ward PS, Branstetter MG (2017) The acacia ants revisited: convergent evolution and biogeographic context in an iconic ant/plant mutualism. Proc R Soc B 284:20162569. https://doi.org/10.1098/rspb.2016.2569