

ECOGRAPHY

A large-scale assessment of ant diversity across the Brazilian Amazon Basin: integrating geographic, ecological and morphological drivers of sampling bias

Joudellys Andrade-Silva, Fabrício B. Baccaro, Lívia P. Prado, Benoit Guénard, Dan L. Warren, Jamie M. Kass, Evan P. Economo and Rogério R. Silva

J. Andrade-Silva (<https://orcid.org/0000-0002-5393-6502>) ✉ (joudellys@gmail.com), *L. P. Prado* (<https://orcid.org/0000-0003-1819-8767>) and *R. R. Silva* (<https://orcid.org/0000-0002-0599-2155>), Programa de Pós-Graduação em Zoologia, Univ. Federal do Pará/Museu Paraense Emílio Goeldi, Belém, Pará, Brazil and Museu Paraense Emílio Goeldi, Coordenação de Ciências da Terra e Ecologia, Laboratório de Morfologia e Ecologia Funcional de Formigas, Belém, Pará, Brazil. – *F. B. Baccaro* (<https://orcid.org/0000-0003-4747-1857>), Programa de Pós-Graduação em Zoologia, Univ. Federal do Amazonas, Amazonas, Brazil. LPP also at: Laboratório de Hymenoptera, Museu de Zoologia da Univ. de São Paulo, São Paulo, Brazil. – *B. Guénard* (<https://orcid.org/0000-0002-7144-1175>), School of Biological Sciences, Univ. of Hong Kong, Hong Kong, Hong Kong SAR. – *D. L. Warren* (<https://orcid.org/0000-0002-8747-2451>), *J. M. Kass* (<https://orcid.org/0000-0002-9432-895X>) and *E. P. Economo* (<https://orcid.org/0000-0002-9432-895X>), Biodiversity and Biocomplexity Unit, Okinawa Inst. of Science and Technology Graduate Univ., Okinawa, Japan. EPE also at: Radcliffe Inst. for Advanced Study, Harvard Univ., Cambridge, MA, USA.

Ecography

2022: e06295

doi: 10.1111/ecog.06295

Subject Editor: Tatsuya Amano
Editor-in-Chief: Miguel Araújo
Accepted 9 May 2022



Tropical ecosystems are often biodiversity hotspots, and invertebrates represent the main underrepresented component of diversity in large-scale analyses. This problem is partly related to the scarcity of data widely available to conduct these studies and the lack of systematic organization of knowledge about invertebrates' distributions in biodiversity hotspots. Here, we introduce and analyze a comprehensive data compilation of Amazonian ant diversity. Using records from 1817 to 2020 from both published and unpublished sources, we describe the diversity and distribution of ant species in the Brazilian Amazon Basin. Further, using high-definition images and data from taxonomic publications, we build a comprehensive database of morphological traits for the ant species that occur in the region. In total, we recorded 1067 nominal species in the Brazilian Amazon Basin, with sampling locations strongly biased by access routes, urban centers, research institutions and major infrastructure projects. Large areas where ant sampling is non-existent represent about 52% of the basin and are concentrated mainly in the northern, southeastern and western Brazilian Amazon. We found that distance to roads is the main driver of ant sampling in the Amazon. Contrary to our expectations, morphological traits had lower predictive power in predicting sampling bias than purely geographic variables. However, when geographic predictors were controlled, habitat stratum and traits contribute to explain the remaining variance. More species were recorded in better-sampled areas, but species richness estimation models suggest that areas in southern Amazonian edge forests are associated with especially high species richness. Our results represent the first trait-based, large-scale study for insects in Amazonian forests and a starting point for macroecological studies focusing on insect diversity in the Amazon Basin.

Keywords: Amazonian ants, ant morphology, neotropical biodiversity, sampling bias, vertical stratification



www.ecography.org

© 2022 The Authors. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Introduction

Tropical ecosystems are vital to maintaining the Earth's biodiversity. These environments host about three-quarters of all species, including 91% of terrestrial birds and over 75% of amphibians, terrestrial mammals, ants and flowering plants (Barlow et al. 2018). However, studies exploring patterns and mechanisms of species richness and geographic distributions have been primarily limited to vertebrate or plant taxa (Rull and Carnaval 2020, Teixido et al. 2020), suggesting that we should expect true species diversity to be considerably higher than what has so far been described.

Knowledge shortfalls in tropical environments are driven by biased sampling efforts around infrastructure projects, areas under strong anthropogenic impact (cities and managed areas) or areas that are easy to access (Oliveira et al. 2016, Barlow et al. 2018, Stropp et al. 2020). Furthermore, recent studies also show that biodiversity knowledge inside protected areas is scant, with more than 70% of these areas represented by fewer than 0.01 species records per km² (Oliveira et al. 2017). These shortfalls are present in many taxa, but some groups, such as invertebrates, may be more affected because they have a higher proportion of undescribed species (Cardoso et al. 2011) and are highly abundant, demanding more time and resources to be adequately sampled (Oliveira et al. 2016). For instance, it is estimated that only 20% of all insect species have been described, which represents a high barrier to understanding diversity (Stork et al. 2015). Further, this lack of taxonomic description (i.e., the Linnean shortfall) is followed by the absence of high-resolution distributional data (i.e., the Wallacean shortfall), which impairs our understanding of the mechanisms that lead a given species to establish and thrive in a given environment (Hutchinson 1957, Anderson and Martínez-Meyer 2004).

Taxonomic shortfalls typically cascade to other facets of biodiversity. For instance, the Raunkiæran shortfall represents the lack of species trait data (morphological, physiological or behavioral), which are essential to determine functional effects and responses of species in ecosystems (Hortal et al. 2015). Analyses on functional traits have been stressed as powerful approaches in community ecology (Fukami et al. 2005, Blonder et al. 2017), particularly for arthropods (Wong et al. 2019), expanding comparative analysis between ecosystems and regions. Interpreting species assemblages as a distribution of quantitative traits rather than a collection of taxonomic units may result in models with greater predictive power (McGill et al. 2006). The functional approach considers that traits or attributes are under selection from environmental filters and biotic interactions (Violle et al. 2007), linking morphological traits with ecological niches. Consequently, functional traits can also indicate other sampling biases. For example, casual collectors or particular sampling methods can more easily collect conspicuous or larger organisms (Lee and Guénard 2019), while hypogaeic organisms (usually smaller, with particular morphological traits, e.g., reduced eyes) require a targeted sampling effort (Kaspari and Weiser 1999, Schmidt and Solar 2010). However, to our knowledge,

Linnean, Wallacean and Raunkiæran shortfalls for tropical biodiversity have not been considered together so far.

Ants (Hymenoptera: Formicidae) are considered a highly diverse group and are numerically and ecologically dominant in tropical forests (Folgarait 1998, Lach et al. 2010). Furthermore, they are easily sampled and have a relatively well-established taxonomy (Hölldobler and Wilson 1990, Underwood and Fisher 2006), including standardized functional trait sampling protocols (Parr et al. 2017). Such characteristics make ants a prominent group for morphological studies compared to other invertebrates.

Morphological traits often reflect the range of habitats occupied by species in an ecosystem (Wong and Guénard 2017, Oliveira and Scheffers 2019). Classical hypotheses suggest that communities become more vertically stratified with increasing species richness, owing to reduced competition or finer niche subdivision (Hutchinson 1959, MacArthur and MacArthur 1961), leading to changes in the morphological diversity of assemblages. In ants, vertical stratification is typically well-defined, with distinct morphological traits from underground to canopy habitats (Fowler et al. 2000, Wong and Guénard 2017, Sosiak and Barden 2021). For example, subterranean ants have much shorter legs than surface-foraging ants, while arboreal ants have much larger eyes than subterranean, leaf-litter and ground-foraging ants (Weiser and Kaspari 2006, Sosiak and Barden 2021). These differences reflect the selective environmental pressures on species morphology.

To address shortfalls in biodiversity knowledge for the world's most expansive, contiguous tropical forest, here we present the first large-scale biodiversity assessment of a highly diverse group in the Brazilian Amazon Basin. On an unprecedented scale, we describe ant distribution, diversity and morphological traits, aiming to answer the following questions: 1) How many ant species have been recorded in the Brazilian Amazon? 2) How is sampling effort distributed across the basin and among habitat strata? 3) What are the main determinants of ant records based on geographic predictors (roads, rivers, cities and degree of urbanization) and ant characteristics (habitat strata and morphological traits)?

Material and methods

The occurrence database for ants in the Brazilian Amazon Basin

The Brazilian Amazon encompasses about 4 196 943 km² of the ~ 6 500 000 km² that make up the Amazon basin (IBGE 2004). We obtained all records available in the literature for the Brazilian Amazon (from 1817 to 2020) through the Global Ant Biodiversity Informatics (GABI – Guénard et al. 2017) project. Then, we compiled additional data on ant occurrences in the Brazilian Amazon from online databases and scientific repositories in Brazil. We also included checklists from non-published sources, mainly dissertations,

master's theses, field expeditions and environmental assessment reports, to compile the most comprehensive information on ant occurrences in the Brazilian Amazon (Supporting information). We obtained these checklists from Brazil's leading research centers on taxonomy, systematics and ant biology (see Acknowledgments). We only considered nominal ant taxa (species and subspecies); informal taxa (morphospecies) were not included.

From the specimen label or information found in the researchers' field notes, we used Gazetteer (<www.fallingrain.com>) and Google Earth (<<https://earth.google.com>>) software to search for the sampling location of each ant record that did not include coordinates, and georeferenced those records from the search results. We disregarded occurrences for first-order administrative divisions (i.e., states in Brazil), nonspecific localities (i.e., large rivers and roads) and information labeled 'Amazon' or 'Brazilian Amazon' (but see Analysis).

Species identifications were validated by examining the specimen (directly, sending to the specialist or by high-resolution images) and reviewing the identification tools used for species determination (for literature data). For questionable records and records of geographic expansions, experts in the taxon were consulted. Data that could not be validated using the tools described above were not considered. This taxonomic treatment was carried out by LPP and JAS and excluded 5.82% of the ant species previously surveyed for the Brazilian Amazon. Valid species names were based on the Online Catalog of the Ants of the World (Bolton 2022, checked in April 2022). The curated ant occurrence database for the Brazilian Amazon used here is available in the Supporting information and in the Dryad repository (<<https://doi.org/10.5061/dryad.ht76hdrj8>>).

Habitat stratification of the ant fauna

We classified the ant fauna in the Brazilian Amazon based on foraging strata. The information about foraging strata is the most widely available and accessible source of information about ant biology and was retrieved directly from the taxonomic literature and online datasets. However, given the absence of natural history information for many species, we consulted experts to record the foraging strata of those ant species without information in the taxonomic literature. Despite this, information on foraging stratum could not be obtained for some species (i.e., known species from a single record without collection details). In these cases (about 8% of total), the stratum predominantly occupied by the genus was assumed, except for *Pheidole* and *Camponotus*. Given the high species richness and behavioral variation within these particular genera, we instead used species groups and subgenera, respectively, to define strata. Due to differences in functions and morphological modifications between wingless and winged ant castes, ant species known only for winged individuals were excluded (about 3.5%).

We recognize four major foraging strata of ants in the Brazilian Amazon (Fig. 1): 1) arboreal: species that forage

primarily on vegetation (trees or shrubs); 2) epigaeic: species that forage mainly above the leaf litter; 3) epigaeic + arboreal: species that use both arboreal and epigaeic strata for their foraging; and 4) hypogaeic: species that forage underground or below the surface of the leaf litter. Due to the low number of ant species (39) that forage in both epigaeic and hypogaeic strata, we aggregated these species into the epigaeic stratum.

The morphological traits database

We constructed a database of morphological traits based on five continuous measurements (Silva and Brandão 2010, 2014, Del-Toro et al. 2015) (Table 1) for all ant species recorded in the Brazilian Amazon Basin (Supporting information). These traits were selected because they are classified as priority information in functional aspects of ant ecology (see the Global Ant Traits Database, Parr et al. 2017). Our database was based on more than 3000 high-definition images (available on Antweb or taken during visits to collections), including lateral, frontal and dorsal views. For species without high-definition images available, we obtained morphological traits from the taxonomic literature when possible, leading to data extracted from over 40 publications (Supporting information). We employed ImageJ software to record the measurements (<<http://imagej.nih.gov/ij>>).

At large scales, the extent of intraspecific variation depends strongly on the trait under consideration, but intraspecific variation contributes less to trait variation than the interspecific variability (Siefert et al. 2015). Studies on intraspecific trait variation in ant assemblages have suggested that intraspecific variation accounts for only 1–4% of total trait variation (Gaudard et al. 2019). Due to this, we measured one specimen for each species present in the Brazilian Amazon, and we always prioritized type specimens over non-type specimens. Morphological measurements from synonymized species were also avoided. However, unlike measurements made with specimens in hand – when it is possible to handle the specimen to standardize measurements of a structure in a single view – measurements made from images require adjustments. For example, when applying a standardized measurement protocol of a given structure was impossible, we used the image view that best represented the measured morphological trait. Although this can be the source of some degree of error (not quantified in the present study), this approach allows us to maximize the size of the morphological database.

Some morphological traits can only be measured at a specific view (e.g., Weber's length in a profile view image). Thus, measurements were not carried out when appropriate images were not available, and instead data were imputed (see Data preparation and morphological traits). Measurements performed on different specimens were never combined to fill in species data. Whenever possible, we used the minor workers to standardize the measurements, as is routinely done in studies of the morphological diversity of ants (Silva and Brandão 2010, Bishop et al. 2015, Schofield et al. 2016). However, when these were not available, we used major workers to obtain morphological measurements (3.58% of total).

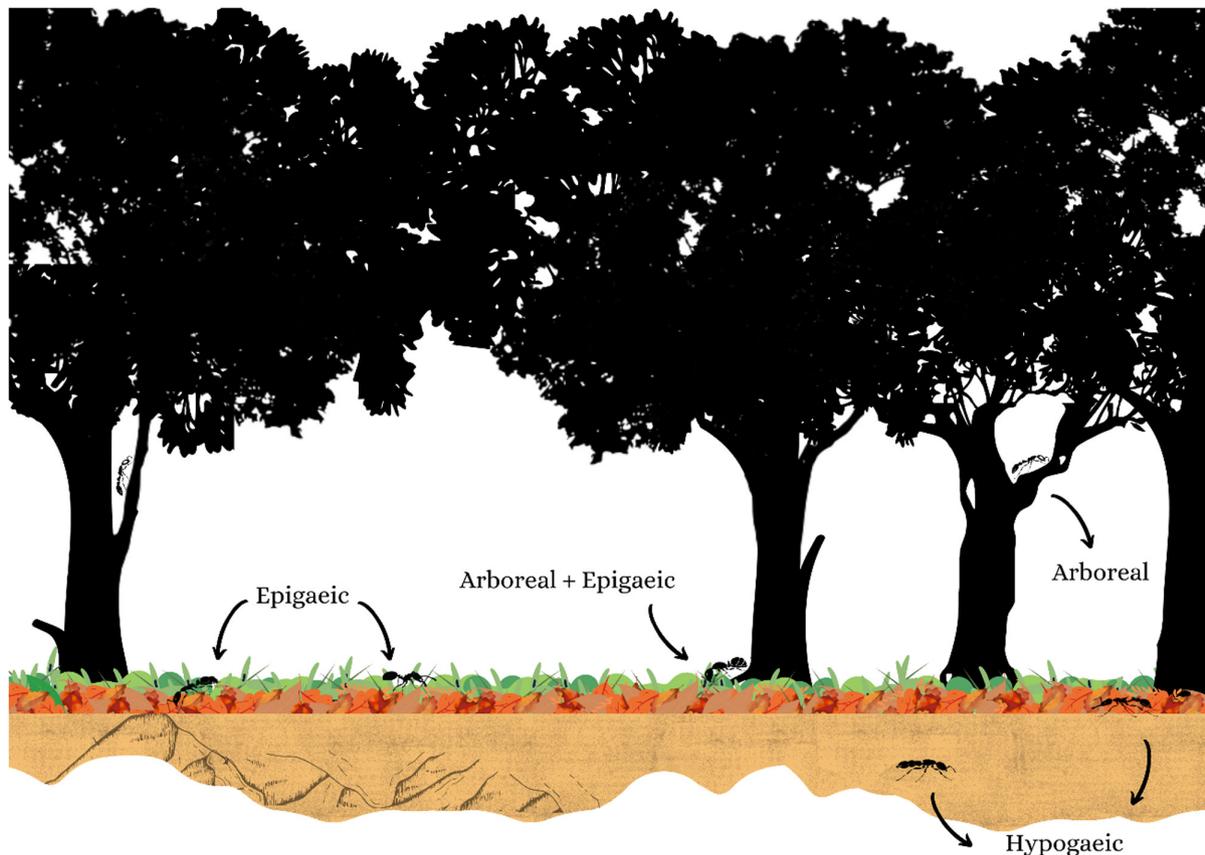


Figure 1. Illustration of the four ant foraging strata recognized for the species pool in the Brazilian Amazon Basin. Arboreal: on vegetation (trees or shrubs); Epigaeic: above the leaf litter or surface of the soil; epigaeic+arboreal: both arboreal and epigaeic strata; hypogaeic: underground or below the surface of the leaf litter.

Some ant species have vestigial or absent eyes, making it impossible to measure some morphological traits. We assigned the following rule for these species: when the species did not show eyes, we assigned a value equal to 0 (zero) for ‘maximum eye size’. The same procedure was adopted for morphological traits related to eyes, such as ‘interocular distance’. This protocol allows keeping such species in the analyses and maintains their unique morphological characteristics.

Data preparation and morphological traits

We combined morphological data extracted from images and taxonomic literature, but missing data were common

due to inadequate or unavailable specimen image views and damaged specimens (lacking the required morphological structures). Thus, we applied data imputation to fill 12.65% of the morphological matrix using multivariate imputation by chained equations (MICE) (Van Buuren and Oudshoorn 2000). The MICE approach preserves the observed data but explicitly provides a set of imputed values for missing data. We imputed 50 estimates for each morphological trait and used the median of these values to fill in missing data for each morphological trait. To run these imputations, we used the function *mice* in the R package *mice* (Van Buuren and Groothuis-Oudshoorn 2011, <www.r-project.org>).

Table 1. Morphological traits and biological significance associated with ant morphology used in the present study. Adapted from Silva and Brandão (2010).

Morphological traits	Biological significance
Mandible length	Determines the size of the captured food resource (Fowler et al. 1991, Weiser and Kaspari 2006).
Maximum eye size	Important character for food search (Weiser and Kaspari 2006).
Interocular distance	Eye position might influence the performance of visual predators (Fowler et al. 1991).
Weber’s length	Indicates body size and is related to several aspects of species’ life history, such as physiology and reproduction (Kaspari and Weiser 1999).
Hind femur length	A larger size favors agility, locomotion and finding resources during foraging but makes it challenging to enter cavities and is unfavorable in more complex environments (Gibb and Parr 2013). This trait is also important to avoid high surface temperatures, resulting in differences among strata (Sommer and Wehner 2012).

Analysis

Diversity and distribution of ant species in the Brazilian Amazon

Many occurrences retrieved from the literature (GABI) and research institutions were attributed with only the first-order administrative division. However, three of the nine states that make up the Brazilian Amazon do not have their whole territory covered by the basin. Therefore, we assembled two datasets: one to simply check the number of described species that occur in the Brazilian Amazon (including all occurrences, with or without geographic coordinate data), and another with the coordinate data, which were used in the analyses. Although we have included all the species records sampled in the whole territory of the three states partially covered by the Brazilian Amazon Basin, species with coordinate data outside the limits of the Brazilian Amazon were excluded in the following analysis.

We made species richness and sampling completeness estimates across the Brazilian Amazon using the rarefaction/extrapolation approach detailed by Chao et al. (2014). We calculated the observed and estimated ant richness, as well as sampling completeness, for each 10 arcmin grid cell (~20 km at the equator) based on the ant occurrences within a moving window, which had a width of 20 grid cells (total area ~160 000 km²). We experimented with different window widths (5 through 30 grid cells) and raster resolutions (2.5 through 10 arcmin) before deciding on this parameterization, which we think was the best balance between smoothness of the mapped predictions and appropriate spatial resolution to define communities for our dataset. Specifically, for each window we made a frequency table representing the number of occurrences per grid cell for each species (i.e., 'incidences'), then used the *iNEXT* function (with 'incidence_freq' setting for data type) in the R package *iNEXT* (Hsieh et al. 2016) to make diversity estimations based on these frequencies and the total number of cells as sampling units. This approach returns single estimates for estimated richness, observed richness and sampling completeness for the window, and this process was repeated for all grid cells within the study extent. To avoid unreliable estimates, we followed a similar methodology to Kusumoto et al. (2020) by omitting calculations for windows with too few grid cells containing occurrence points (in our case, < 2) or those with as many singletons as there were total occurrence points, though our approach differed as we applied no constraint to the number of species within a window.

Taxonomic and functional sampling biases in ants

Considering the high correlation between size measurements in ants (such as length or width of morphological structures), we standardized all measurements by Weber's length, removing the relationship between this trait and the other measurements. Therefore, we use ant size and relative measures in the following analyses.

We prepared a map based on kernel density estimation (KDE) with a bandwidth of 100 km to describe the ant sampling distribution in the Brazilian Amazon. We use KDE

to describe where the highest ant sampling densities are concentrated and which regions show deficits in sampling efforts in the Brazilian Amazon. Further, to describe whether sampling efforts were related to protected areas (conservation units and indigenous lands), we extracted the relative amount of sampling by counting points (geographic coordinates) inside and outside of protected areas using QGIS ver. 2.18.2 (QGIS Development Team 2019). Shapefiles for the Brazilian Amazon and protected areas were provided by Instituto Brasileiro de Geografia e Estatística (IBGE, <www.ibge.gov.br>) and Ministério do Meio Ambiente (<www.mma.gov.br>) in Brazil, respectively.

We used a generalized linear model (GLM) to estimate the determinants of the sampling bias in the ant occurrence records. Given that our goal was to measure the bias itself rather than to estimate the environmental associations of the species, we did not rarefy the occurrence data, as this would remove much of the signal that we were trying to measure. As predictor variables we used morphological traits, foraging stratum and geographic features, including highways, rivers, roads, urban areas and degree of urbanization. We randomly sampled background data over the Brazilian Amazon Basin with the same number of points as the occurrence data and performed a binomial regression that treats the background as 'pseudoabsences', similar to the use of background data in species distribution models (Peterson et al. 2011). Background point values for the geographic predictor variables were extracted by grid cell. For the other variables we randomly sampled the trait and foraging stratum matrices, representing the overall distribution of these variables, but without any correlation to the bias predictors. Therefore, these background data represent the expected trait distribution under the null hypothesis that species were equally common and occurred randomly in space. We built four GLMs: Model 1 is the full model containing the effects of foraging stratum, traits and geographic predictors. We further included an interaction term between stratum and other predictors to determine whether there is significantly different sampling based on habitat stratum. Model 2 also includes all variables, but no interaction terms. In this model, different strata may differ in overall sampling effort but the shape of the relationship between the geographic predictors or traits and the probability of detection will not be affected. Model 3 includes the spatial predictors only, modeling ant occurrence as a function of the geographic variables, with no trait data or habitat stratum in the model. Model 4 includes habitat strata and an interaction term between stratum and each of the geographic predictors. In this model, we removed predictors for ant morphology. We compared these four GLMs using Akaike information criterion (AIC, Burnham and Anderson 2002) values to find the model that best fits the data while controlling for overparameterization. Finally, spatial projections were made from the optimal model to predict sampling bias. Modeling and plots of results (variable importance in the models, predicted relative sampling intensity and variance explained by habitat stratum and

traits in the space) were implemented in the *ENMTools* R package (Warren et al. 2021) and modified as needed to match the data. Additionally, in order to test for the relative impact of different predictors on spatial sampling bias, we adopted the permutation approach implemented in the *vip* R package (Greenwell et al. 2020).

Because the morphological and foraging stratum variables cannot be projected onto a map directly, we also analyzed how much of the variance in spatial sampling in each area was explained by these variables. To do this, we chose five values for each morphological character representing the 0, 0.25, 0.5, 0.75 and 1.0 quantiles of the distribution of that character. For each of these values, we created raster layers for the study area where every grid cell was fixed at that value. We then projected the full GLM across every combination of stratum and quantile for each morphological variable, resulting in 100 model projections (4 strata \times 5 morphological variables \times 5 quantiles per variable). We then measured the variance in predictions between all levels of each variable, averaging across all values of the remaining variables.

Raster data were provided by IBGE (access routes) and Infraestrutura Nacional de Dados Espaciais (<<https://index.gov.br>>; populated place and degree of urbanization).

Results

Diversity of ants in the Brazilian Amazon Basin

We recorded 47 454 ant occurrences, comprising 1067 ant species in 106 genera and 12 subfamilies, from published and unpublished sources (Supporting information). However, 5088 ant occurrences had only first-order administrative division information (state) or nonspecific localities (i.e., large rivers and roads), and 70 ant species (1360 occurrences) did not contain any information about localities or precise geographic information. Therefore, the results presented here refer to a pool of 997 ant species retrieved from 41 006 records.

The most frequently sampled ant genera (> 2000 records) were *Camponotus* (4808), *Ectatomma* (4006), *Cephalotes* (3668), *Pseudomyrmex* (3667), *Dolichoderus* (2553), *Strumigenys* (2427), *Crematogaster* (2266) and *Pheidole* (2047). Together, these genera contain about 54% of all ant records in the Brazilian Amazon Basin. The most frequently collected species (> 1000 records) were *Ectatomma tuberculatum* (Olivier 1792) (1729), *Ectatomma brunneum* Smith 1858 (1612), *Camponotus atriceps* (Smith 1858) (1403), *Cephalotes atratus* (Linnaeus 1758) (1272) and *Strumigenys denticulata* Mayr, 1887 (1077). On the other hand, 315 ant species (about 29%) had fewer than five records.

Observed ant richness was highest in the central and eastern regions of the Brazilian Amazon and lowest in the northern, southern and western regions (except in areas of Acre and Rondônia States; Fig. 2A). Extrapolated richness estimates further indicate that regions to the south (in the State of Mato Grosso) and north (in the State of Roraima) are predicted to harbor high richness (Fig. 2B).

We found that arboreal (316 species) and epigaic (275 species) strata had the highest taxonomic diversity. The strata associated with hypogaic ($n=260$) and epigaic + arboreal ($n=103$) species had the lowest taxonomic diversity.

Distribution of ant records in the Brazilian Amazon Basin

Ant sampling was highly uneven throughout the Brazilian Amazon (Fig. 3). This deficit is concentrated mainly in the north, covering the whole east–west axis from the Amazon River to the border between the Brazilian Amazon and the international Amazon (Fig. 3). Further, we observed that many areas were never sampled, such as in the southeast and west of the Brazilian Amazon (Fig. 3) and inside protected areas (about 27% of records; Supporting information). Sampling completeness for the 20 km² cells suggested higher coverage along the main axis of the Amazon River to the eastern Amazon (Fig. 2C).

On the other hand, a few specific localities contained an unusually large number of records (Fig. 3). For example, about 39% of records occur in metropolitan areas in the Brazilian Amazon, such as Belém in the eastern Amazon (13%) and Manaus in the central Amazon (12.6%). These localities represent regions where the major scientific institutions in the Brazilian Amazon have historically performed biodiversity surveys and ecological studies (Museu Paraense Emílio Goeldi and Instituto Nacional de Pesquisas da Amazônia, respectively). We also detected a high number of records near major infrastructure projects such as the southeastern (the largest iron ore mine in the world in the Carajás Mountains) and southwestern (Jirau and Santo Antônio hydroelectric power plants) regions.

Taxonomic and functional ant sampling biases in the Brazilian Amazon

Of the 997 ant species recorded in the Brazilian Amazon, we could not locate any morphological information or high-quality images for 43 species. Therefore, the results involving trait data morphological diversity refer to a reduced pool of 954 ant species. Of these, 483 ant species had some trait value that had to be determined by data imputation (12.65% of the morphological dataset).

From the AIC comparison, the full model with trait data and interactions between habitat stratum and other variables (model 1) provided the best fit (Table 2). Variable importance tests showed that distance to roads and urban areas had a higher contribution than the distance to rivers, degree of urbanization or distance to highways (Fig. 4). Although the overall variance in predicted sampling intensity was only slightly affected by habitat stratum and traits (Fig. 5), we obtained better predictions of sampling intensity by incorporating these data. In this case, habitat stratum followed by body size and interocular distance were the main ant characteristics explaining spatial variation in ant sampling (i.e., when geographic predictors were fixed and did not

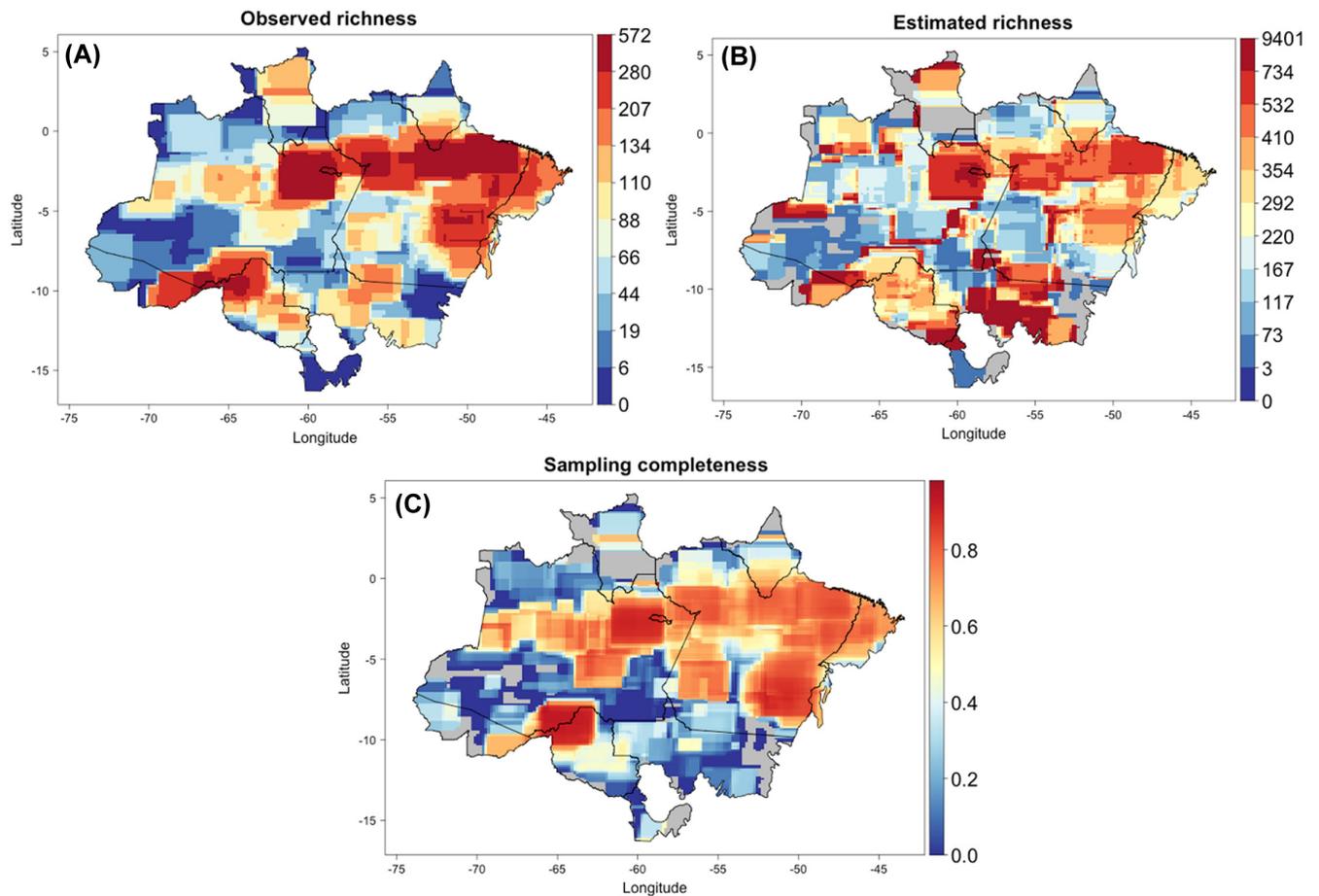


Figure 2. Estimates of observed richness (A), estimated richness (B) and sampling completeness (C) of ant species in the Brazilian Amazon Basin. These maps were generated using a moving window (400 km²) that calculated the rarefaction/extrapolation approach described in Chao et al. (2014) based on the ant occurrence data. Areas with no data (i.e. could not be calculated) are colored gray. An equal area (Eckert IV) projection was used in the analysis. Maps were projected to a geographic coordinate system for viewing purposes.

contribute to this variance). In general, the highest peaks of ant record density in the Brazilian Amazon are located < 1 km from the access routes and in grid cells with more than 75% urbanization (Supporting information).

Discussion

Although invertebrates are among the most abundant and diverse organisms in tropical forests, they are underrepresented in large-scale analyses (Stein et al. 2014, Teixido et al. 2020). For the first time, our study combines species occurrences and morphological traits over a large area of the world's largest tropical forest to describe species richness, distribution and sampling biases of a major invertebrate group. We found that ant sampling in the Brazilian Amazon Basin is highly uneven, mainly concentrated around roads and not near rivers, as is commonly recorded for diverse taxa (Vale and Jenkins 2012). Large areas where ant sampling is non-existent are abundant, while protected areas are sampled poorly or not at all. These sampling biases indicate substantial knowledge shortfalls for the ant fauna in the Brazilian Amazon (e.g.,

Wallacean, Linnean and Raunkiærán). Our effort is the first extensive assessment of two dimensions of ant diversity (distribution and morphology) and provides the foundation for future large-scale ecological studies on ants in the Amazon Basin, as well as indicates new areas for biodiversity surveys.

The known ant richness in the Brazilian Amazon is high, representing 61.8% of the diversity of ants currently recorded for Brazil (Guénard et al. 2017). Despite this, the true number of ant species in the Amazon Basin is underestimated due to spatial gaps (distribution and sampling methodology) and taxonomic impediments. Further, lower taxonomic resolution and many undescribed species may also hide some distributional patterns. For example, *Pheidole* is the most common, dominant and speciose genus of ants globally (Bolton 2022) and is often considered the most diverse in ecological studies in the Amazon (Fernandes and Souza 2018, Franco et al. 2019). However, it was only the eighth-most diverse genus in our study, due to the lack of species identifications in ecological surveys and limited geographic coverage in species descriptions. This disconnect is even more dramatic for *Solenopsis*, another ubiquitous and dominant genus that is not among the 10 most frequent genera in our database. In

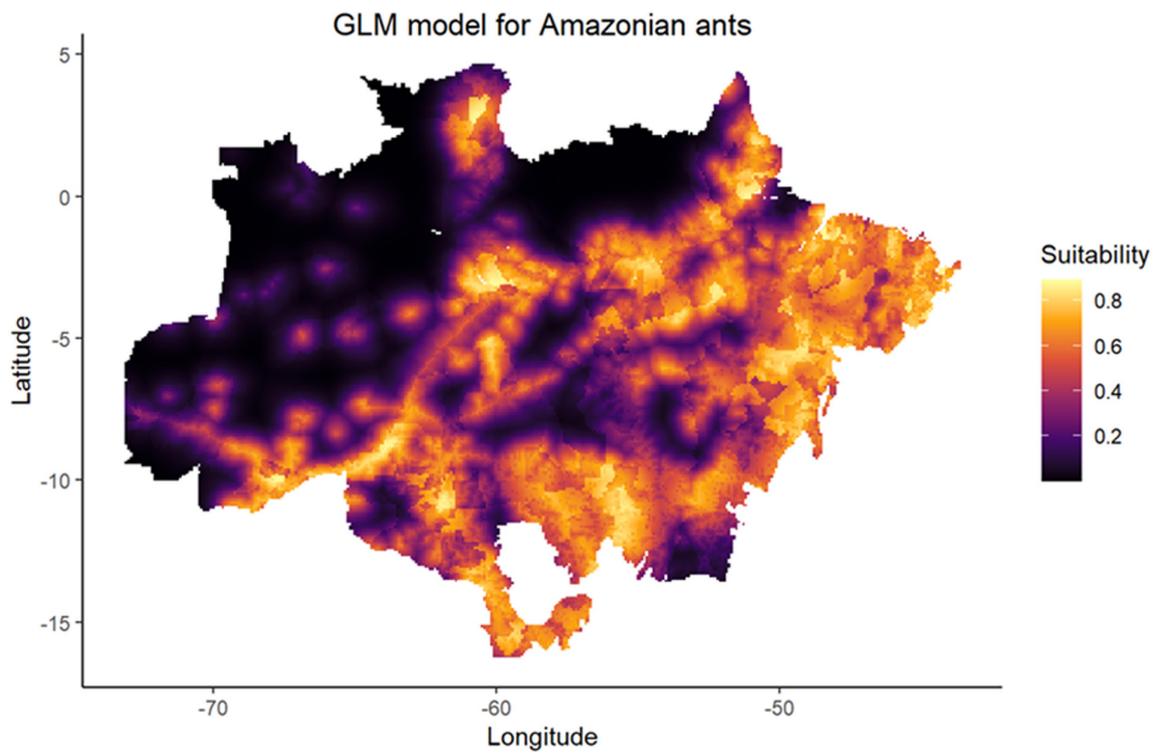
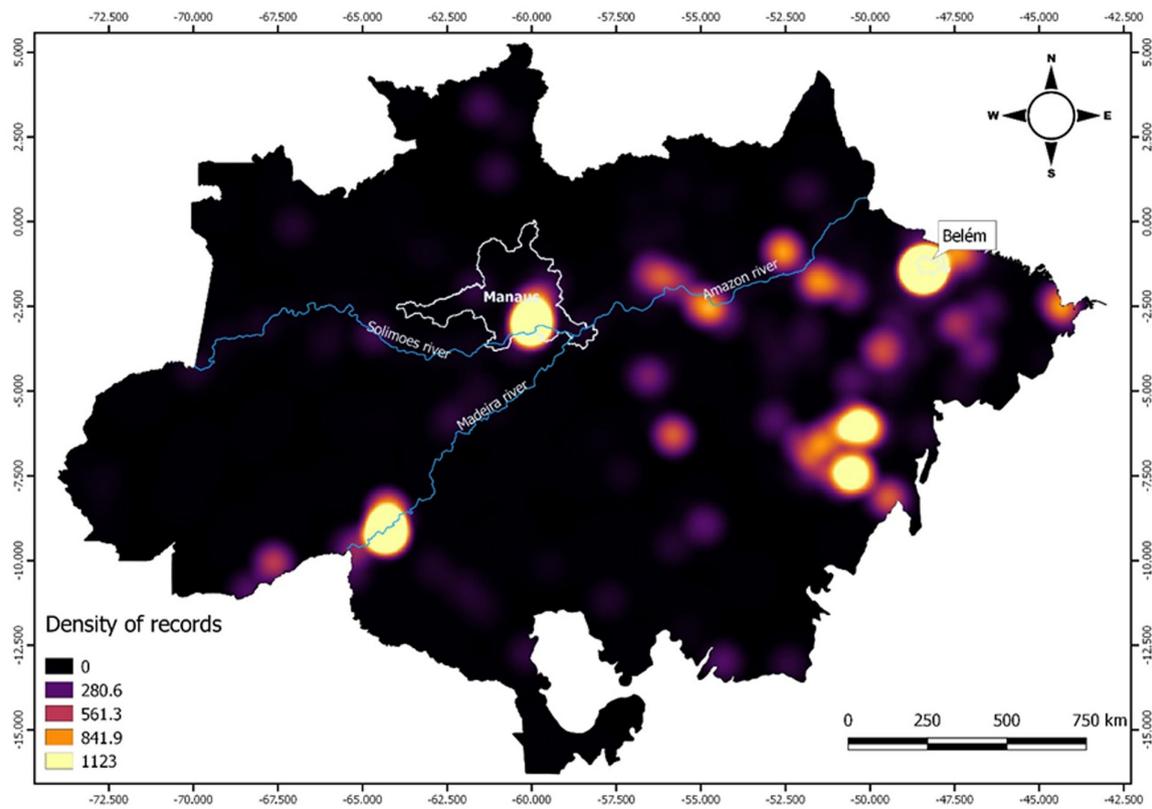


Figure 3. Ant sampling density (top) and predictions from a GLM model of sampling intensity (bottom) in the Brazilian Amazon Basin. The records (n = 41 006) are from published sources (1817–2020), online repositories, Brazilian research institutions and unpublished sources.

Table 2. Result of model selection based on the Akaike information criterion (AIC) for the determinants of ant records in the Brazilian Amazon Basin. See the analysis section for detailed description of models; numbers given after each model name in the table correspond to the description of that model in the text. K=number of parameters in the model; AICc=AIC score of the model with correction for small sample size; delta AICc=difference in AICc score between the best model and the model being compared; AICcWt=AICc weight, which is the proportion of the total amount of predictive power provided by the full set of models contained in the model being assessed; LnL: log-likelihood; Cum.Wt=sum of the AICc weights for all models less than or equal to the AICc of the given model.

Model names	K	AICc	Delta AICc	AICcWt	LnL	Cum.Wt
Full model (1)	44	79 920.21	0.00	0.999	-39 916.08	0.99
Full model without interactions (2)	14	80 661.84	741.62	< 0.001	-40 316.92	1.00
Spatial variables only (3)	6	80 663.44	743.23	< 0.001	-40 325.72	1.00
Spatial variables and stratum (4)	24	79 934.82	14.61	< 0.001	-39 943.40	1.00

addition to the taxonomic gaps, the difficulty in determining species boundaries based on their morphology makes identification at the species level especially challenging. Thus, the widespread use of morphospecies in ecological studies and large sampling gaps suggest that the true ant diversity in the Brazilian Amazon Basin may represent a larger portion of the Neotropical fauna than currently recognized.

The extent and degree of sampling bias we observed for collections of ants in the Brazilian Amazon is alarming. We found that the areas with highest sampling coverage are close to roads, urban centers and outside protected areas. Sampling biases are common across different tropical regions and biological groups (Gardner et al. 2009, Oliveira et al. 2016, 2017, Baldwin et al. 2018, Divieso et al. 2020). This pattern results from the cumulative sampling efforts that generally

occur around research institutions located primarily in large cities, resulting in many remote protected areas left unsampled in the Brazilian Amazon. In turn, this sampling bias has direct impacts on the observed characteristics of the sampled fauna (habitat stratum and trait data). Since deforestation generally begins close to the main roads (Amigo 2020), disturbance-tolerant species may be over-represented, while forest species less tolerant of disturbance may be underrepresented. Similar results were found for plants (Tobler et al. 2007, Daru et al. 2018), other arthropods and vertebrates (Oliveira et al. 2016). Together with recent findings, our results highlight an urgent need for coordinated actions and multi-taxon approaches in different and more isolated areas.

Historically, sampling in the Amazon was performed along the Amazon River, spreading along its tributaries and

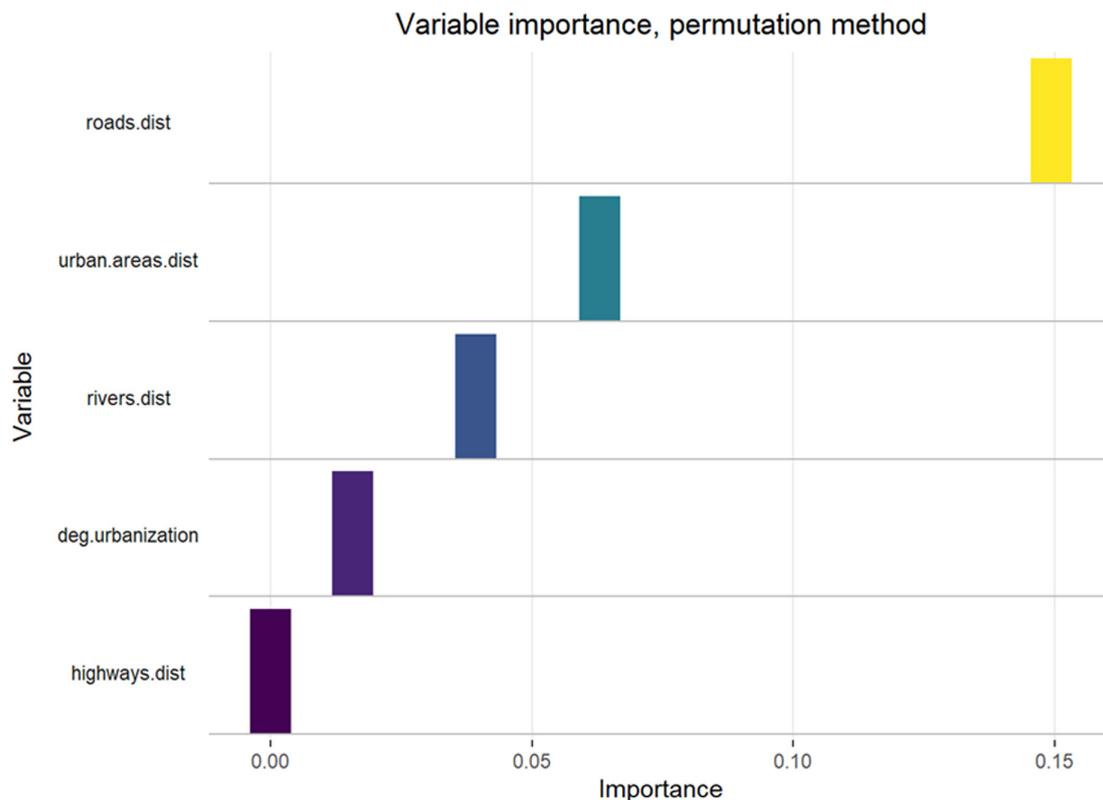


Figure 4. Variable importance in a generalized linear model (GLM) of spatial sampling bias. Importance is measured as the change in the area under the receiver-operating characteristic curve for the resulting model when the values for one predictor are randomized. Higher values indicate variables that more strongly affect quality of model predictions.

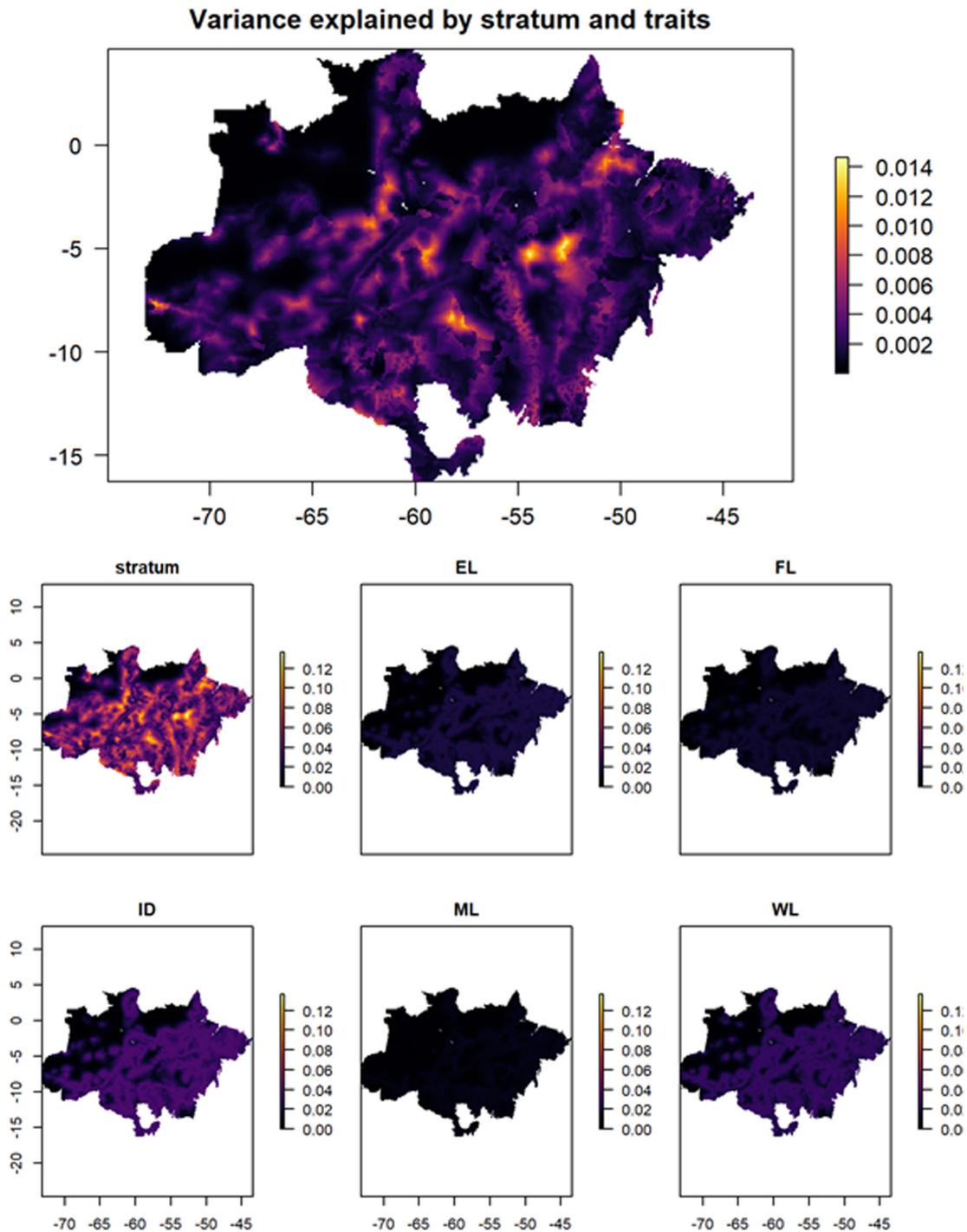


Figure 5. Amount of variation in each grid cell explained by stratum and morphological traits. Here we measure this as the standard deviation of model predictions in that grid cell across all quantile values of the focal trait, averaging over the predictions for the remaining traits. Although the overall variance explained by stratum and traits is small, it is not homogeneous; even within some well-sampled areas, there is an unequal sampling of strata and morphologies. EL=maximum eye size; FL=hind femur length; ID=interocular distance; ML=mandible length; WL=Weber's length.

around major towns (Santos et al. 2015). Although research has become more widespread in the last two decades, the trend toward sampling areas following rivers in the Amazon has persisted (Santos et al. 2015, Oliveira et al. 2016), mainly among major studied groups such as vertebrates and

plants (Hopkins 2007, Vale and Jenkins 2012). We found, however, that roads (not rivers) constitute the main driver of ant sampling bias in the Brazilian Amazon. This result may reflect limited financial investments (Costa and Magnusson 2010), as well as the scarcity of collaborative efforts to

sample regions unexplored in the Amazon (Santos et al. 2015, McMichael et al. 2017). Similar results have recently been documented for ants in the eastern Amazon (Albuquerque et al. 2021).

Estimated ant species richness patterns across the Brazilian Amazon suggest higher diversity in the best-sampled regions (e.g., central portions along the Amazon River and the eastern Amazon) but they also highlight that the often-cited ‘arc of deforestation’ region (Matricardi et al. 2020) along the east and south may harbor a high and undescribed ant diversity. These areas are at the interface with savannah vegetation and among the most threatened by the expanding deforestation frontier (Arruda et al. 2021).

Our comprehensive database clearly shows that knowledge shortfalls are enormous in the Amazon, mainly in the northern (from the Amazon River), southeastern and western regions and protected areas. Large knowledge shortfalls are problematic because they hamper a clear vision of the distribution of biodiversity and the possible taxonomic and functional loss caused, for example, by habitat conversion (Barlow et al. 2016, Divieso et al. 2020). A recent study pointed out that, by 2017, 12% of the Brazilian Amazon had been deforested without having a single tree specimen recorded (Stropp et al. 2020). Considering that Linnaean and Wallacean knowledge shortfalls affect invertebrate taxa more severely (Cardoso et al. 2011) and the current higher deforestation rates in the Amazon, we may never sufficiently document the species diversity and composition of these regions (Paiva et al. 2019), including multiple species unknown to science.

Protected areas are considered cornerstones for curbing the loss of biodiversity, but their effectiveness in the tropics has not been determined due to the lack of surveys (Cazalis et al. 2020). Protected areas cover about 51% of the Brazilian Basin (Salomão et al. 2019). Many of these areas are considered responsible for maintaining high levels of biological diversity with the potential to reveal many new species (Fearnside 2013). The scarcity of ant records within protected areas in the Brazilian Amazon seriously affects our understanding of the region’s biological diversity and compromises the conservation status of these environments (Hallmann et al. 2017, Cazalis et al. 2020).

Taxonomic and functional ant sampling biases

The lack of knowledge about ant morphological traits in the Brazilian Amazon is also pervasive. Morphological data offer a consistent approach to evaluate the processes that structure and organize ant communities, enabling predictions about species’ evolutionary and ecological aspects (Silva and Brandão 2010, Gibb et al. 2015). Despite this importance, morphological information for 43 ant species is entirely absent, and about 12% of the morphological matrix required data imputation for some trait value. The absence or constraints in available morphological data can affect our knowledge about the evolution, diversity and relationships among traits, and ultimately compromise efficient

conservation of species and ecosystems (Hortal et al. 2015, Freitas et al. 2021).

Another problem affecting trait data in our database is sampling methods (Lee and Guénard 2019). Particular sampling methods normally target one or a few strata, resulting in a biased view of overall ant diversity. Pitfall traps are widely used in monitoring programs in the Amazon for their low costs and highest complementarity with ant assemblages recorded by other techniques (Souza et al. 2012). However, in forest environments, the trait compositions of communities sampled by pitfall traps are dominated by traits relative to size (Lee and Guénard 2019). Thus, sampling methods may also be a source of biases in documenting morphological patterns of ants in the Brazilian Amazon.

Our results reveal that ant diversity in the Brazilian Amazon is high but, even with our comprehensive review, this diversity is largely unknown. The knowledge shortfalls are gaping, abundant and require coordinated action to reveal dimensions of ant biodiversity across different forest habitats. We strongly recommend carrying out ant surveys within protected areas, far from the main roads especially in the northern (from the Amazon River to the boundary of the international Amazon), western and southeastern regions. Sampling efforts should also prioritize undersampled strata (e.g., leaf-litter, canopy and underground). These coordinated actions, combined with increased investments in the training and qualification of researchers (e.g., taxonomists and ecologists) and infrastructure (laboratories and scientific collections), will maximize biodiversity discovery and thus reduce knowledge shortfalls for under-described taxa. Ultimately, these knowledge-based approaches are necessary to implement efficient conservation policies and management strategies to limit ongoing and future damage to the various ecological components of the globe’s largest tropical forest. Our study can be a baseline for the inclusion of insects in future macroecological analyses that seek to address biodiversity knowledge shortfalls.

Acknowledgements – We thank INPA, MZSP, CPDC, DZUP, MPEG (Kelly Sampaio, Diego Alves and William Overal), UFAC (Marília Costa, Amanda Silva and Fernando Schmidt), UFMT (Thiago Izzo), UNEMAT (Ricardo Vicente), UMC (Débora Souza-Campana) and UFMA (Ciro Santos and José Macário), for making available ant occurrences in the Brazilian Amazon Basin. To Gabriela Camacho (CAS) for confirming the occurrence of *Gnamptogenys* species. We are also grateful to researchers Jacques Delabie (CEPLAC), Thiago Ranzani (UFPR), Júlio Chaul (UFV), Emília Albuquerque (Smithsonian Institution), Otávio Silva and Emely Siqueira (MPEG), Marília Costa (UFPA), Benoit Jahyny (UNIVASF) and Itanna Fernandes (INPA) for their help in relation to ant stratification. To Rony Almeida (MPEG), for his help and critical analytic views. To Samara Costa (UFMA), for making the illustration presented in Fig. 1.

Funding – This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, to which JAS and LPP extend thanks for the scholarship received. JAS is also grateful for the financial support provided by Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (PCI-DB grant no. 317900/2021-8 and

grant no. 300981/2022-8). JMK was supported by the JSPS Postdoctoral Fellowships for Foreign Researchers Program. RRS thanks the Fundação Amazônia de Amparo a Estudos e Pesquisas (FAPESPA) for financing the project (ICAAF No. 012/2018 and process 2018/2999534).

Author contributions

Joudellys Andrade-Silva: Conceptualization (equal); Data curation (supporting); Formal analysis (lead); Methodology (equal); Validation (supporting); Writing – original draft (lead); Writing – review and editing (lead). **Fabrcio B. Baccaro:** Conceptualization (equal); Formal analysis (supporting); Methodology (equal); Writing – original draft (supporting); Writing – review and editing (equal). **Lívia P. Prado:** Data curation (supporting); Validation (lead); Writing – original draft (supporting); Writing – review and editing (supporting). **Benoit Guénard:** Data curation (lead); Writing – original draft (supporting); Writing – review and editing (equal). **Dan L. Warren:** Formal analysis (lead); Writing – review and editing (equal). **Evan P. Economo:** Data curation (lead); Writing – original draft (supporting); Writing – review and editing (supporting). **Jamie M. Kass:** Formal analysis (lead); Writing – review and editing (equal). **Rogério R. Silva:** Conceptualization (equal); Data curation (supporting); Formal analysis (supporting); Funding acquisition (lead); Methodology (equal); Project administration (lead); Supervision (lead); Writing – original draft (supporting); Writing – review and editing (equal).

Transparent peer review

The peer review history for this article is available at <<https://publons.com/publon/10.1111/ecog.06295>>.

Data availability statement

Data are available from the Dryad Digital Repository: <<https://doi.org/10.5061/dryad.ht76hdrj8>> (Andrade-Silva et al. 2022).

Supporting information

The Supporting information associated with this article is available with the online version.

References

- Albuquerque, E. Z. et al. 2021. Ants of the State of Pará, Brazil: a historical and comprehensive database of a key biodiversity hotspot in the Amazon Basin. – *Zootaxa* 5001: 1–83.
- Amigo, I. 2020. The Amazon's fragile future. – *Nature* 578: 505–507.
- Anderson, R. P. and Martínez-Meyer, E. 2004. Modeling species' geographic distributions for preliminary conservation assessments: an implementation with the spiny pocket mice (*Heteromys*) of Ecuador. – *Biol. Conserv.* 116: 167–179.
- Andrade-Silva, J. et al. 2022. Data from: A large-scale assessment of ant diversity across the Brazilian Amazon Basin: integrating geographic, ecological and morphological drivers of sampling bias. – Dryad Digital Repository, <<https://doi.org/10.5061/dryad.ht76hdrj8>>.
- Arruda, F. V. et al. 2021. Seasonal variation of ground and arboreal ants in forest fragments in the highly-threatened Cerrado-Amazon transition. – *J. Ins. Conserv.* 25: 897–904.
- Baldwin, C. C. et al. 2018. Below the mesophotic. – *Sci. Rep.* 8: 4920.
- Barlow, B. et al. 2018. The future of hyperdiverse tropical ecosystems. – *Nature* 559: 517–526.
- Barlow, J. et al. 2016. Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. – *Nature* 535: 144–147.
- Bishop, T. R. et al. 2015. Contrasting species and functional beta diversity in montane ant assemblages. – *J. Biogeogr.* 42: 1776–1786.
- Blonder, B. et al. 2017. New approaches for delineating n-dimensional hypervolumes. – *Methods Ecol. Evol.* 9: 305–319.
- Bolton, B. 2022. An online catalog of the ants of the world. – <<https://antcat.org>>, accessed 9 September 2020.
- Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference: a practical information-theoretic approach. – Springer.
- Cardoso, P. et al. 2011. The seven impediments in invertebrate conservation and how to overcome them. – *Biol. Conserv.* 144: 2647–2655.
- Cazalis, V. et al. 2020. Effectiveness of protected areas in conserving tropical forest birds. – *Nat. Commun.* 11: 4461.
- Chao, A. et al. 2014. Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. – *Ecol. Monogr.* 84: 45–67.
- Costa, F. R. C. and Magnusson, W. E. 2010. The need for large-scale, integrated studies of biodiversity – the experience of the Program for Biodiversity Research in Brazilian Amazonia. – *Nat. Conserv.* 8: 3–12.
- Daru, B. H. et al. 2018. Widespread sampling biases in herbaria revealed from large-scale digitization. – *New Phytol.* 217: 939–955.
- Del-Toro, I. et al. 2015. Predicted impacts of climatic change on ant functional diversity and distributions in eastern North American forests. – *Divers. Distrib.* 21: 781–791.
- Divieso, R. et al. 2020. How to prioritize areas for new ant surveys? Integrating historical data on species occurrence records and habitat loss. – *J. Insect Conserv.* 24: 901–911.
- Fearnside, P. M. 2013. Serviços ambientais provenientes de florestas intactas, degradadas e secundárias na Amazônia brasileira. – In: Peres, C. A. et al. (eds), *Conservação da biodiversidade em paisagens antropizadas do Brasil*. UFPR, pp. 26–57.
- Fernandes, I. O. and Souza, J. L. P. 2018. Dataset of long-term monitoring of ground-dwelling ants (Hymenoptera: Formicidae) in the influence areas of a hydroelectric power plant on the Madeira River in the Amazon Basin. – *Biodivers. Data J.* 6: e24375.
- Folgarait, P. J. 1998. Ant biodiversity and its relationship to ecosystem functioning: a review. – *Biodivers. Conserv.* 7: 1221–1244.
- Fowler, H. C. et al. 2000. Hypogaic and epigaic ant (Hymenoptera: Formicidae) assemblages of Atlantic coastal rainforest and dry mature and secondary Amazon forest in Brazil: continuums or communities. – *Trop. Ecol.* 41: 73–80.
- Fowler, H. G. et al. 1991. Ecologia nutricional de formigas. – In: Panizzi, A. R. and Parra, J. R. P. (eds), *Ecologia nutricional de insetos*. Manole, pp. 131–223.
- Franco, W. et al. 2019. First checklist of the ants (Hymenoptera: Formicidae) of French Guiana. – *Zootaxa* 4674: 509–543.
- Freitas, T. M. S. et al. 2021. Quantifying shortfalls in the knowledge on Neotropical Auchenipteridae fishes. – *Fish Fish.* 22: 87–104.

- Fukami, T. et al. 2005. Species divergence and trait convergence in experimental plant community assembly. – *Ecol. Lett.* 12: 1283–1290.
- Gardner, T. A. et al. 2009. Prospects for tropical forest biodiversity in a human-modified world. – *Ecol. Lett.* 12: 561–582.
- Gaudard, C. A. et al. 2019. Low level of intraspecific trait variation in a keystone invertebrate group. – *Oecologia* 190: 725–735.
- Gibb, H. and Parr, C. L. 2013. Does structural complexity determine the morphology of assemblages? An experimental test on three continents. – *PLoS One* 8: e64005.
- Gibb, H. et al. 2015. Does morphology predict trophic position and habitat use of ant species and assemblages? – *Oecologia* 177: 519–531.
- Greenwell, B. M. et al. 2020. Variable importance plots – an introduction to the vip package. – *R J.* 12: 343–366.
- Guénard, B. et al. 2017. The Global Ant Biodiversity Informatics (GABI) database: a synthesis of ant species geographic distributions. – *Myrmecol. News* 24: 83–89.
- Hallmann, C. A. et al. 2017. More than 75 percent decline over 27 years in total flying insect biomass in protected areas. – *PLoS One* 12: e0185809.
- Hölldobler, B. and Wilson, E. O. 1990. *The ants*. – Springer.
- Hopkins, M. J. 2007. Modelling the known and unknown plant biodiversity of the Amazon Basin. – *J. Biogeogr.* 34: 1400–1411.
- Hortal, J. et al. 2015. Seven shortfalls that beset large-scale knowledge of biodiversity. – *Annu. Rev. Ecol. Evol. Syst.* 46: 523–549.
- Hsieh, T. C. et al. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). – *Methods Ecol. Evol.* 7: 1451–1456.
- Hutchinson, G. E. 1957. Concluding remarks. – *Cold Spring Harb. Symp. Quant. Biol.* 22: 415–427.
- Hutchinson, G. E. 1959. Homage to Santa Rosalia, or why are there so many kinds of animals? – *Am. Nat.* 93: 145–159.
- IBGE (Inst. Brasileiro de Geografia e Estatística) 2004. Mapa de biomas do Brasil – primeira aproximação. Rio de Janeiro, Brasil, <www.ibge.gov.br>.
- Kaspari, M. and Weiser, M. 1999. The size-grain hypothesis and interspecific scaling in ants. – *Funct. Ecol.* 13: 530–538.
- Kusumoto, B. et al. 2020. Global distribution of coral diversity: biodiversity knowledge gradients related to spatial resolution. – *Ecol. Res.* 35: 315–326.
- Lach, L. et al. 2010. *Ant ecology*. – Oxford Univ. Press.
- Lee, R. H. and Guénard, B. 2019. Choices of sampling method bias functional components estimation and ability to discriminate assembly mechanisms. – *Methods Ecol. Evol.* 10: 867–878.
- MacArthur, R. H. and MacArthur, J. W. 1961. On bird species diversity. – *Ecology* 42: 594–598.
- Matricardi, E. A. T. et al. 2020. Long-term forest degradation surpasses deforestation in the Brazilian Amazon. – *Science* 369: 1378–1382.
- McGill, B. J. et al. 2006. Rebuilding ecology from functional traits. – *Trends Ecol. Evol.* 21: 178–185.
- McMichael, C. N. H. et al. 2017. Ancient human disturbances may be skewing our understanding of Amazonian forests. – *Proc. Natl Acad. Sci. USA* 114: 522–527.
- Oliveira, B. F. and Scheffers, B. R. 2019. Vertical stratification influences global patterns of biodiversity. – *Ecography* 42: 249–258.
- Oliveira, U. et al. 2016. The strong influence of collection bias on biodiversity knowledge shortfalls of Brazilian terrestrial biodiversity. – *Divers. Distrib.* 22: 1232–1244.
- Oliveira, U. et al. 2017. Biodiversity conservation gaps in the Brazilian protected areas. – *Sci. Rep.* 7: 9141.
- Paiva, P. F. P. R. et al. 2019. Deforestation in protect areas in the Amazon: a threat to biodiversity. – *Biodivers. Conserv.* 29: 19–38.
- Parr, C. L. et al. 2017. GlobalAnts: a new database on the geography of ant traits (Hymenoptera: Formicidae). – *Insect Conserv. Divers.* 10: 5–20.
- Peterson, A. T. et al. 2011. *Ecological niches and geographic distributions (MPB-49)*. – Princeton Univ. Press.
- QGIS Development Team 2019. *QGIS Geographic Information System*. – Open Source Geospatial Foundation Project.
- Rull, V. and Carnaval, A. C. 2020. Neotropical diversification: patterns and processes. – Springer.
- Salomão, R. et al. 2019. O estado das áreas protegidas: distribuição das áreas protegidas nos municípios da Amazônia Legal. – Imazon.
- Santos, J. G. et al. 2015. Geographic trends and information deficits in Amazonian conservation research. – *Biodivers. Conserv.* 24: 2853–2863.
- Schmidt, F. A. and Solar, R. 2010. Hypogaic pitfall traps: methodological advances and remarks to improve the sampling of a hidden ant fauna. – *Insectes Soc.* 57: 261–266.
- Schofield, S. F. et al. 2016. Morphological characteristics of ant assemblages (Hymenoptera: Formicidae) differ among contrasting biomes. – *Myrmecol. News* 23: 129–137.
- Siefert, A. et al. 2015. A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. – *Ecol. Lett.* 18: 1406–1419.
- Silva, R. R. and Brandão, C. R. F. 2010. Morphological patterns and community organization in leaf-litter ant assemblages. – *Ecol. Monogr.* 80: 107–124.
- Silva, R. R. and Brandão, C. R. F. 2014. Ecosystem-wide morphological structure of leaf-litter ant communities along a tropical latitudinal gradient. – *PLoS One* 9: e93049.
- Sommer, S. and Wehner, R. 2012. Leg allometry in ants: extreme long-leggedness in thermophilic species. – *Arthropod Struct. Dev.* 41: 71–77.
- Sosiak, C. E. and Barden, P. 2021. Multidimensional trait morphology predicts ecology across ant lineages. – *Funct. Ecol.* 35: 139–152.
- Souza, J. L. P. et al. 2012. Trade-offs between complementarity and redundancy in the use of different sampling techniques for ground-dwelling ant assemblages. – *Appl. Soil Ecol.* 56: 63–73.
- Stein, A. et al. 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. – *Ecol. Lett.* 17: 866–880.
- Stork, N. E. et al. 2015. New approaches narrow global species estimates for beetles, insects and terrestrial arthropods. – *Proc. Natl Acad. Sci. USA* 112: 7519–7523.
- Stropp, J. et al. 2020. The ghosts of forests past and future: deforestation and botanical sampling in the Brazilian Amazon. – *Ecography* 43: 979–989.
- Teixido, A. L. et al. 2020. Major biases and knowledge gaps on fragmentation research in Brazil: implications for conservation. – *Biol. Conserv.* 251: 108749.
- Tobler, M. et al. 2007. Implications of collection patterns of botanical specimens on their usefulness for conservation planning: an example of two Neotropical plant families (Moraceae and Myricaceae) in Peru. – *Biodivers. Conserv.* 16: 659–677.
- Underwood, E. C. and Fisher, B. L. 2006. The role of ants in conservation monitoring: if, when and how. – *Biol. Conserv.* 132: 166–182.

- Vale, M. M. and Jenkins, C. N. 2012. Across-taxa incongruence in patterns of collecting bias. – *J. Biogeogr.* 39: 1743–1748.
- Van Buuren, S. and Groothuis-Oudshoorn, K. 2011. Mice: multivariate imputation by chained equations in R. – *J. Stat. Softw.* 45: 1–67.
- Van Buuren, S. and Oudshoorn, K. 2000. Multivariate imputation by chained equations: MICE v1.0 User's Manual. – TNO Prevention and Health.
- Violle, C. et al. 2007. Let the concept of trait be functional! – *Oikos* 116: 882–892.
- Warren, D. L. et al. 2021. ENMTools 1.0: an R package for comparative ecological biogeography. – *Ecography* 44: 504–511.
- Weiser, M. D. and Kaspari, M. 2006. Ecological morphospace of New World ants. – *Ecol. Entomol.* 31: 131–142.
- Wong, M. K. L. and Guénard, B. 2017. Subterranean ants: summary and perspectives on field sampling methods, with notes on diversity and ecology (Hymenoptera: Formicidae). – *Myrmecol. News* 25: 1–16.
- Wong, M. K. L. et al. 2019. Trait-based ecology of terrestrial arthropods. – *Biol. Rev.* 94: 999–1022.