ORIGINAL PAPER

Robustness of biodiversity surrogates to temporal variation and imperfect detection

Ana Cristina da Silva Utta[1] · Pedro Aurélio Costa Lima Pequeno[2] · William Ernest Magnusson[1,3] ·
Jorge Luiz Pereira Souza[4,5]

Abstract

Introduction Common species can predict overall species richness and composition, serving as a shortcut to biodiversity and supporting well-informed management of areas. However, the impacts of temporal variation in underlying species occurrence and imperfect detection on surrogacy effectiveness are unknown. We hypothesized that common species will represent total species composition both across sites and along environmental gradients and that this surrogacy effectiveness is maintained over time and robust to imperfect detection.

Materials and methods We sampled ground-dwelling ants using pitfall traps over 25 km^2 on Amazon. We measured the congruence between a complete dataset and subsets containing a smaller number of species based on their commonness and compared the results among three sampling years. Further, we simulated virtual assemblages to determine the effect of species detection probability on surrogacy effectiveness and compared the costs associated with using all species vs. common species only. The common species composition responded to environmental gradients similarly to the overall species composition, regardless of year. The simulated data showed similar trends, although including imperfect detection increased the minimum number of common species required to recover overall patterns reasonably.

Conclusion - Implications for insect conservation This study represents the first report on the effectiveness of common species in predicting ant assemblage patterns over time. We propose that imperfect detection should be considered in biodiversity surrogacy studies, especially regarding its impact on surrogate effectiveness over time.

Keywords Amazon · Commonness · Environmental variables · Ground-dwelling ants · Tropical rainforest

Introduction

Processes that structure ecological communities are intrinsically dynamic and knowing how environmental changes act on species over time and helps predict how these changes influence them (Báez et al. 2016; Vellend 2016). There are studies, for example, that demonstrate how organisms can be affected by climate change through phenotypic plasticity and also by genetic changes (Root et al. 2003). Therefore, it would be useful if biomonitoring programs could provide information on environmental changes (e.g. global warming, human disturbance, invasive species, El Niño Southern Oscillation) that affect organisms over time (Magurran et al. 2010), aiding in the adaptive management of conservation projects but this does not occur in all studies (Buchart et al. 2010; Paknia and Pfeiffer 2011). However, measuring diversity patterns over time has several constraints, such as, for example, the presence of rare species, which makes the study

✉ Ana Cristina da Silva Utta
uttaacsu18@gmail.com

1 Programa de Pós-Graduação Em Ecologia, Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus, Amazonas CEP 69067-375, Brazil

2 Programa de Pós-Graduação Em Recursos Naturais, Universidade Federal de Roraima, Boa Vista, Roraima CEP 69317-810, Brazil

3 Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus, Amazonas CEP 69067-375, Brazil

4 Instituto Nacional da Mata Atlântica (INMA), Santa Teresa, Espirito Santo CEP 29650-000, Brazil

5 Programa de Pós-Graduação Em Zoologia, Universidade Federal Do Amazonas (UFAM), Manaus, Amazonas, Brazil

more expensive and needs more time to be developed and the lack of institutional and government funding over a long period (Margules et al. 2002; Whittaker et al. 2005). These constraints can be even more pronounced in regions of high biodiversity, such as tropical forests (Costa and Magnusson 2010; Magnusson et al. 2013), thus emphasizing the need for reliable, cost-effective proxies for biodiversity.

Under limited data, time, or financial resources for diversity surveys or even target groups with unstable taxonomy and often associated with reduced body size, researchers may consider a surrogate approach as an alternative to reduce costs (McGeoch 1998; Margules et al. 2002; Hodkinson & Jackson 2005; Thomson et al. 2018; Souza et al. 2022). Surrogates can be umbrella species (Caro 2010), rare or endemic species (Lamoreux et al. 2006), coarse taxonomic levels (Andersen 1995; Souza et al. 2016), or common species (Vellend et al. 2008) which reasonably inform about diversity patterns of the whole assemblage. Ideally, the selected species or selected group should represent a large part of the variation in species richness or composition in the broader community in the focal area and respond to the same ecological processes that generate and maintain total diversity (Khan 2006; Heino and Soininen 2007; Siqueira et al. 2012). A surrogate approach or the use of a subset of species would facilitate the identification process, and the responses obtained through these approaches should ideally be able to be extrapolated to the whole target group (Williams and Gaston 1994; Olsgard et al. 2003). Although the effectiveness of surrogate approach is not unanimously addressed by researchers, there are numerous examples where surrogates have been demonstrated to predict major diversity patterns of broader communities (Williams and Gaston 1994; de Oliveira et al. 2020), on wide spatial and temporal scales (Souza and Fernandes 2021, 2022, 2023; Souza et al. 2016), although there are some situations in which they fail (Gaspar et al. 2010).

A study examining the effectiveness of using taxa as surrogates for insect species richness and diversity was conducted with data from Coleoptera and Formicidae collected during biodiversity surveys in five tropical countries and two temperate countries (Rosser and Eggleton 2012). The objective was to determine whether this approach was suitable for ants. The study found that while genera were a suitable proxy for species richness for Coleoptera, they were not adequate for Formicidae. This highlights the limitations of the surrogate approach for ants and the need for alternative methods to accurately measure their species richness and diversity (Rosser and Eggleton 2012).

One surrogate approach suggested by some researchers is the use of a subgroup of species, usually the most common or abundant ones, which can predict species richness and/or composition patterns and assist in the management of protected areas, and this has been observed in invertebrates,

for example, for both mites (Franklin et al. 2013) and butterflies (Graça et al. 2017). Accordingly, many studies have shown that common species contribute disproportionately to diversity patterns of more inclusive groups of species, especially in highly diverse taxa, such as ants (Vázquez and Gaston 2004; Šiding et al. 2009; Mazaris et al. 2010; Keizer-Vlek et al. 2012; Franklin et al. 2013). Moreover, if properly applied, common species could help protect most of the biodiversity (Warman et al. 2004; Gaston 2010).

Ants are a megadiverse and ubiquitous group of social insects, with a total biomass of 12 megatons of dry carbon, which exceeds the combined biomass of wild birds and mammals and amounts to 20% of human biomass (Schultheiss et al. 2022). In addition to well-resolved taxonomy, ants play important ecological roles (Hölldobler and Wilson 1990; Kaspari et al. 2000; Elizalde et al. 2020). Furthermore, ants are easily sampled using standardized methods (Agosti et al. 2000; Souza et al. 2012), which makes them well suited for biomonitoring (Delabie et al. 2021).

One potential limitation of the common species approach in predicting the total diversity of an assemblage which has received little attention to date is imperfect detection, especially at large spatial and temporal scales, for both moving and sessile organisms (Chen et al. 2013; Guillera-Arroita 2017). Imperfect detection in biodiversity studies refers to the failure to observe or detect all individuals of a species present in a given area during a study, resulting in underestimation or misrepresentation of species occurrence and abundance. Various factors contribute to imperfect detection, including methodological limitations, behavioral traits of species, spatial and temporal variability, population density, and observer effects (Guillera-Arroita 2017). Disregarding imperfect detection can lead to negative impacts on data quality and interpretation, for example, may lead to an underestimation of true species richness, biased estimates of species abundance, and a misrepresentation of community composition (Benoit et al. 2018). In datasets containing species with low population densities, imperfect detection leads to false absences that can reduce the accuracy of distribution models and estimates of diversity (Dorazio et al. 2011). This is of particular concern when considering the importance of these estimates, which are often used to support conservation and management decisions (Yoccoz et al. 2001).

Accordingly, we investigated the efficiency of the use of common species as indicators of patterns of change in ant species composition along a marked topo-edaphic gradient over 25 km^2 of Amazonian rain forest. (1) We investigated the degree of congruence of spatial patterns of variation in species composition (both among sites and along environmental gradients) using the Procrustes dissimilarity test. We used a complete dataset (with all species) and subsets containing gradually smaller number of species based on their commonness (i.e. number of sites in which they occurred)

to assess how similar the responses of the subsets were in relation to full data. (2) In parallel, we simulated virtual assemblages based on by the ant data to determine the effect of species detection probability on the minimum number of common species necessary to recover the known species composition patterns of the "true" ant assemblage. We hypothesized that a selected group of common species could adequately represent the spatial variation in species composition among sites and along the environmental gradient and that this pattern would be maintained over time. However, we expected the minimum number of species required to reveal whole-assemblage patterns to increase as detection probability decreased, due to increased noise in the data. (3) To compare the cost-efficiency of using these two types of data in biodiversity-monitoring programs we estimated the monetary and time costs associated with the use of common species and total ant species composition data.

Methods

Study area

Fieldwork was carried out at the Ducke Reserve (Fig. S1), a protected area, located at km 26 on the Manaus-Itacoatiara Road (02º 55′ S, 59º 53′ W, and 03º01′ S, 59º 5′ W). The climate of the reserve is characterized by a rainy season from November to May and a dry season during the remaining months of the year (Marques-Filho et al. 1982), with a mean annual rainfall of 3017 mm in the period from April 1965 to November 1980. Ducke Reserve is covered by dense ombrophilous forest with topographic and edaphic variations that are typical of several areas in central Amazon. The reserve has a total area of 10,000 ha, with a nutrient-poor soils that range from Xanthic Hapludox on the plateaus, to sandy soils—Typic Epiaquods and Typic Endoaquods in the valleys (Chauvel et al. 1987).

Sampling design

The study was carried out within a 25 km² system of trails used by the Brazilian Biodiversity Research Program (PPBio). The sampling design was based on the RAPELD system (Costa and Magnusson 2010). The RAPELD system can be applied at a variety of spatial scales. Smaller modules (combinations of trails and plots) are used for rapid cross-site assessments. The acronym RAP was coined by Conservation International for rapid biological surveys. Larger grids are used to investigate the ecological processes in more accessible sites. These are often long-term ecological research sites (LTER or PELD in Portuguese (Magnusson et al. 2005). The grid system in Ducke Reserve consists of North–South and East–West trails giving access to 30

plots of 250 m in length spaced every 1 km. For ants, these are usually 1 m wide (Figure S1). Plots follow topographic contours to reduce within-plot variation (Costa et al. 2005; Magnusson et al. 2013).

Ant sampling

Pitfall traps of 500 ml (8 cm height and 9.5 cm diameter) containing approximately 100 ml of alcohol to preserve the ants were buried at ground level following the methodology described in Souza et al. (2012). In each of the 30 plots, 10 pitfalls were installed (one every 25 m), corresponding to 300 pitfalls in the grid during each collection event. To prevent leaves and rainwater falling into the pitfall traps, a plastic dish was placed above the trap. The pitfall traps remained in the field for 48 h (Bestelmeyer et al. 2000). Ants were sampled in 2006, 2012 and 2014 during the period of lowest rainfall. The ants were initially identified to genus level using the taxonomic keys of Baccaro et al. (2015). After this, ants were classified in morphospecies and, when feasible, to species level using specific keys, in addition to comparison with collection material identified by taxonomists. Voucher specimens were deposited in the National Institute for Amazonian Research Invertebrate Collection. The species list is available in Table S1.

Environmental variables

We investigated whether common species composition responded to environmental gradients similarly to overall species composition. The variables used in this study were (1) slope of terrain: measured in April and September 2001, with the aid of a clinometer, from six points, spaced every 50 m along the plot (2) soil clay content: determined from physical–chemical analysis of soil samples with sodium pyrophosphate, was measured with an clinometer, collected in March and April 2008 from six points spaced every 50 m along the plot; (3) litter volume calculated by collecting litter from 1 m² at each point using a Winkler extractor, and then measuring the volume using a graduated plastic bucket (10 l). Litter volume was measured in the two periods studied (2006 and 2012). For more details and metadata, see PPBio website https://ppbio.inpa.gov.br/manuais (Table S2).

Database

We created one community matrix containing species occurrences (presence or absence) as columns and plots as rows (n = 30) for each of the three study years. We used occurrences rather than some measure of abundance to avoid giving too much weight to species with larger colonies (Hölldobler and Wilson 1990; Gotelli et al. 2011). The number of occurrences was measured within each plot. Thus, it can

range from 0 (no occurrence of the species in the plot) to 10 (the species occurred in all pitfall traps in the plot). All analyses were run in the R environment for statistical computing (R Core Team 2020, version 4.0.0), using the packages vegan 2.5–6 (Oksanen et al. 2019) and CommEcol 1.7.0 (Melo et al. 2019).

Selecting a potential set of common species

Although some studies dealing with common species have used a priori criteria for their selection (Borcard and Legendre 1994; Lavoie et al. 2009; Minor 2011), we followed the method proposed by Franklin et al. (2013), and determined species commonness by directly assessing the number of plots in which each species occurred. Species were removed from the complete data matrix (hereafter: complete matrix) based on their rarity in a sequential procedure (first species that occurred in a single plot, then species that occurred in two plots, and so on), thus generating reduced data matrices (hereafter: reduced matrices). We excluded species until we reached a reduced matrix containing exclusively those species that occurred in all 30 plots.

Congruence between reduced and complete matrices

For a surrogate to be considered effective, it must adequately predict or represent the spatiotemporal as well as ecological patterns of the entire assemblage of species (Williams and Gaston 1994; Olsgard et al. 2003; Heino and Soininen 2007; Siqueira et al. 2012; Souza et al. 2016). Thus, the congruence of spatial variation in composition between each reduced matrix and the complete matrix was quantified by applying a Procrustes correlation,which ranges between 0 and 1, (a Procrustes correlation value of 1 indicates a complete similarity between the matrices, while a value of 0 it does not have similarity, with 999 Monte Carlo permutations to test for statistical significance, to evaluate the efficiency of the use of common species as indicators of patterns of change in ant species composition (Peres-Neto and Jackson 2001). In the Procrustes analysis, we measured the dissimilarities among ant assemblages, comparing each reduced matrices and the complete species matrix using a rotational-fit algorithm that minimizes the sum of squared residuals between the matrices. One of the main advantages of using Procrustes is to perform dissimilarity comparisons with the raw data of the matrices, instead of using distance measures as in the Mantel test, for example (Peres-Neto and Jackson 2001). The m^2 statistic (the goodness-of-fit statistic that measures the level of congruence between two ordination configurations) was transformed to the Procrustes correlation coefficient (r) using the following equation: $r = \sqrt{(1 - m2)}$ (Landeiro et.al. 2012; Oksanen et al. 2019).

Removed species were replaced by columns with zeros to match the dimensionality of the complete matrix and allow their comparison (Peres-Neto and Jackson 2001). To evaluate the reliability of common species as indicators, we followed the values used by Leal et al. (2010), dos Santos Ribas and Padial (2015) and Souza et al. (2016, 2018), which define a "reasonable" prediction, if reduced matrices predicted ≥ 60% and < 70% of the variation found with complete matrix data, 'good' if it predicted ≥ 70% and < 80%, and 'excellent' if it predicted ≥ 80%.

Relationships with environmental gradients

We used terrain slope, soil clay content and litter volume as environmental predictors. We selected these variables because they were sampled in a standardized way at the study site and their effects on the composition of the ground-dwelling ant assemblages in the Brazilian Amazon were known from earlier studies (Vasconcelos et al. 2003; Gomes et al. 2018). We used distance-based redundancy analysis (db-RDA) to test the effect of environmental gradients on the complete species composition matrix, and also on the reduced species composition matrices. db-RDA is a type of direct ordination that combines multiple regression and PCoA (Borcard et al. 2011). We assessed how the statistical signal of the effects of the environmental gradients changed as species were removed from the data matrix. The objective was not to seek for significant effects per se, but to detect with what number of common species in the reduced matrix there is a change in the signal of the environmental effects, from significant to non-significant P values or vice versa.

Impact of imperfect detection on effectiveness of common species as surrogates

In order to assess how well common species can capture species composition patterns in an entire assemblage, we conducted simulations using virtual assemblages with predetermined patterns of species composition and detection probabilities. This allowed us to estimate the impact of species detection probability on the ability of common species to retrieve the overall species composition. Assemblages were simulated using the program COMPAS (Minchin 1987), implemented in R with function "compass" of the "CommEcol" package (Melo et al. 2019). This simulator has been used extensively in assessments of the effectiveness of various methods in recovering known patterns in ecological data (e. g. Austin et al. 2006; Nekola and Mcgill 2014; Hui et al. 2017). The simulator mimics the response of species abundances to an environmental gradient ranging from 0 to 100, whose values are sampled from a uniform distribution. Species expected responses are given by a Beta function, which creates unimodal, potentially asymmetric, response

curves whose optima are randomly distributed along the gradient. The main parameters to vary are the number of species, the average of the optimal species abundance along the gradient, and the average width of species responses along the gradient.

Based on data previously collected at the same study site (Oliveira et al. 2020; Souza et al. 2012), we simulated the abundance of 200 species, with an abundance average of 1.64 at their optima and an average width of 30 along the gradient. This created a species turnover pattern similar to that observed in the empirical datasets (Fig. 2), but whose structure is known rather than inferred. Next, we added detection error: we simulated the detection probability of each species as a random draw from a normal distribution with a given mean (in logit scale) and standard deviation of one (Yamaura et al. 2019).

Using this setting, we simulated 100 datasets, varying the mean detection probability randomly across simulations from 0.05 to 1.0, thus encompassing the entire possible range. For each simulated data set, we repeated the two previous analyses described for the empirical datasets (i.e. Procrustes correlations and db-RDA excluding increasing numbers of uncommon species), and retrieved the minimum number of species that still reasonably recovered the simulated structure. "Good" was defined as a Procrustes correlation above 0.7 and a statistical significance (significant or non-significant) equivalent that found with the whole dataset in the case of db-RDA. To determine how imperfect detection affected the surrogate efficacy of common species, we regressed the minimum number of species obtained in simulations to the mean probability of detection, separately for obtaining the two recovery criteria (n = 100 in each case).

Estimating time and financial resources

We estimated the absolute and relative (%) time and financial costs for species composition and common species identification. We measure the time spent according to the number of days it takes to collect, sort, and identify the material. We considered the costs as expenses associated with field and laboratory materials, field sampling, and the salary of team members. To calculate the salaries expense and the time consumed, we include the duration of activities in the field and the laboratory. We measured all costs (time and financial) for the complete matrix and reduced matrices, and the salary and scholarships were based on current Brazilian Federal Government payments. To estimate the costs in US dollars, we used the average value of the currency conversion rates for Brazilian currency (Real = R$), which was US $ 1 = R$ 2.61 (average annual resources from 2006 to 2020; source: Central Bank of Brazil). Costs of laboratory equipment and accommodation buildings for field staff were not included, as there was no qualitative difference in total cost (Gardner et al. 2008).

Results

For three sampling events (2006, 2012 and 2014), 16,090 ants distributed in 11 subfamilies, 57 genera and 292 species were collected, 161 were formally described and 131 were morphospecies (Table S1). In all sampled years, *Pheidole* was the most species-rich genus, with registers of 50, 55 and 26 species in 2006, 2012 and 2014 respectively. Sixty species (20.55%) were found in all years, 76 species were collected exclusively in 2006, 71 species only in 2012 and 30 only in 2014 (Table S1). The number of singletons was 63 for 2006, 68 for 2012 and 40 for 2014, while the number of doubleton*s* was 27 for 2006, 26 for 2012 and 22 for 2014 (Table S1).

The species with the highest number of occurrences on 300 pitfalls installed in 2006 was *Crematogaster tenuicula* Forel, 1904 (96.7%), followed by *Trachymyrmex* sp. 5 (76.7%) and *Ectatomma lugens* Emery, 1894 (70%). In 2012, the species with the highest occurrences were *C. tenuicula* (93.3%), *Strumigenys denticulata* Mayr, 1887 (86.7%) and *Pheidole fracticeps* Wilson, 2003 (83.3%). In 2014, *Pheidole biconstricta* Mayr, 1870 (86.7%), *Pheidole cataractae* Wheeler, 1916 (80%) and *Solenopsis clytemnestra* Emery, 1896 (76.7%) showed the highest occurrences (Table S1).

In all sampling events, the reduced matrices retaining a subset of common species were able to predict the spatial pattern of composition observed (Fig. 1). The minimum number of common species capable of recovering a statistically significant amount of (albeit low) information for all species was three in 2006 (r = 0.379, P = 0.019, Fig. 1a, Table S3), seven in 2012 (r = 0.523, P = 0.001, Fig. 1b, Table S3) and three in 2014 (r = 0.471, P = 0.001, Fig. 1c, Table S3); however, low congruence values were associated with the complete matrix for these subsamples. The minimum number of species required to achieve a 'good' congruence rating (Procrustes correlation > 0.7), was 20 species in 2006, and 2012 and nine species in 2014 (Table S3).

Overall, we observed a significant association between common species composition and environmental gradients, although this relationship varied across different years. Specifically, during the first two collection events in 2006 and 2012, a reduction to fewer than 10 common species resulted in non-significant associations (see Fig. 2a, b). Similarly, in 2014, the threshold for non-significant relationships occurred when the number of common species dropped below 25. (Fig. 2c). Twenty-five common species are enough to have significant results of ant associations with environmental gradients in the three periods studied
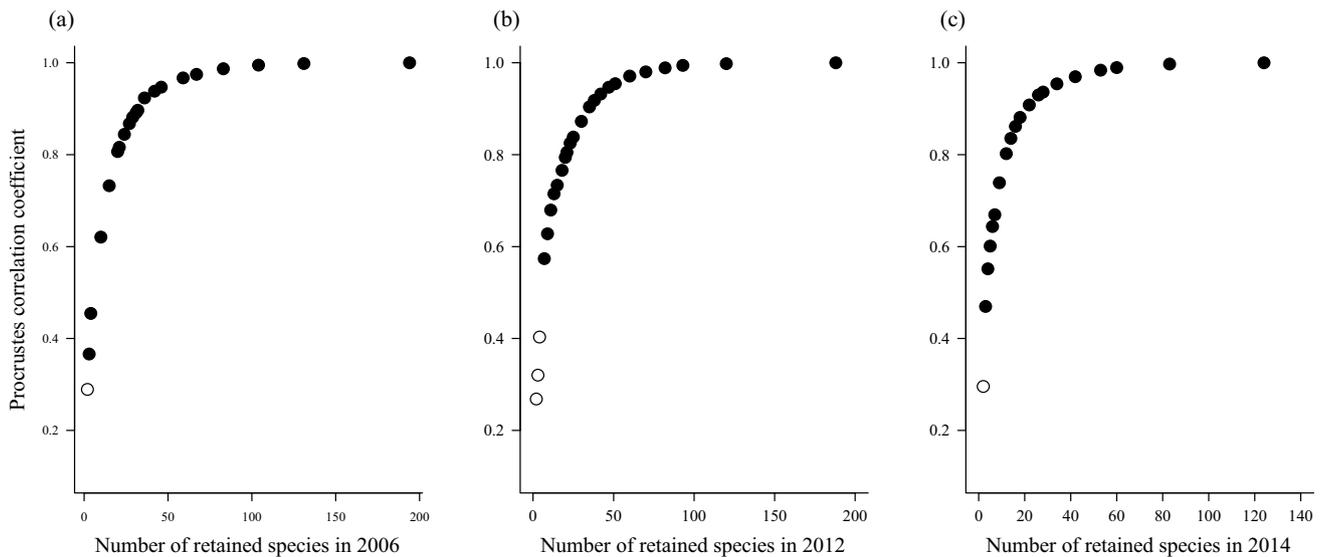
Fig. 1 Procrustes correlation coefficients between the complete matrix and the reduced matrices of ant species composition in the years 2006 (**a**), 2012 (**b**) and 2014 (**c**). Black symbols indicate significant values for an alpha ≤ 0.05; white symbols indicate non-significant values
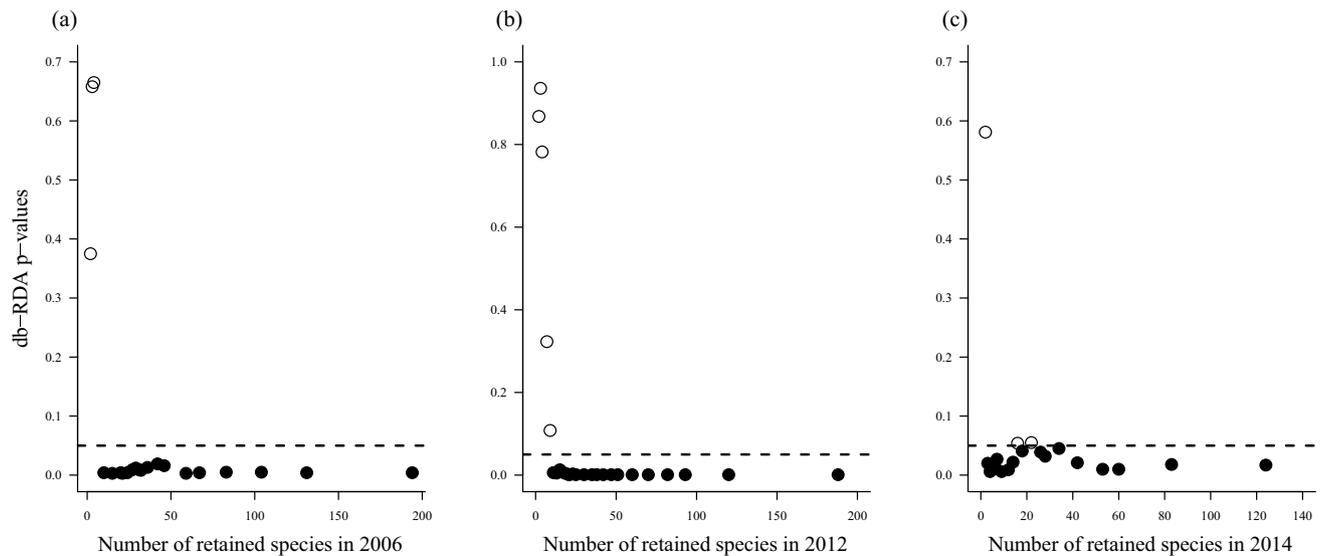

Fig. 2 P values of the partial RDA for the effect of environmental variables on the number of retained species in 2006 (**a**), 2012 (**b**), and 2014 (**c**). The dashed line indicates the significance limit of the P value, defined a priori (0.05). Black symbols indicate significant values for an alpha ≤ 0.05; white symbols indicate non-significant values

(Table S3), achieving a 'good' or 'excellent' congruence rating (Procrustes correlation > 0.7).

Common species in the simulated data had a similar result to that observed in the empirical data. However, the minimum number of species required to maintain a reasonable congruence between the complete and reduced matrices (in addition to detecting environmental effects) was 45 (Fig. 3a, b). Further, with increasing mean probability of detection, fewer species were needed to reach the established

criterion of a Procrustes correlation equal to or large than 0.7 ($R^2 = 0.372$, P = 0.001, Fig. 4a). By contrast, as the mean probability of detection increased, more species were needed to detect the influences of environmental variables ($R^2 = 0.433$, P = 0.001, Fig. 4b), although statistical significance also became variable (Fig. 4b).

Using the observed data, the minimum number of common species required to significantly recover the overall spatial pattern of species composition and its
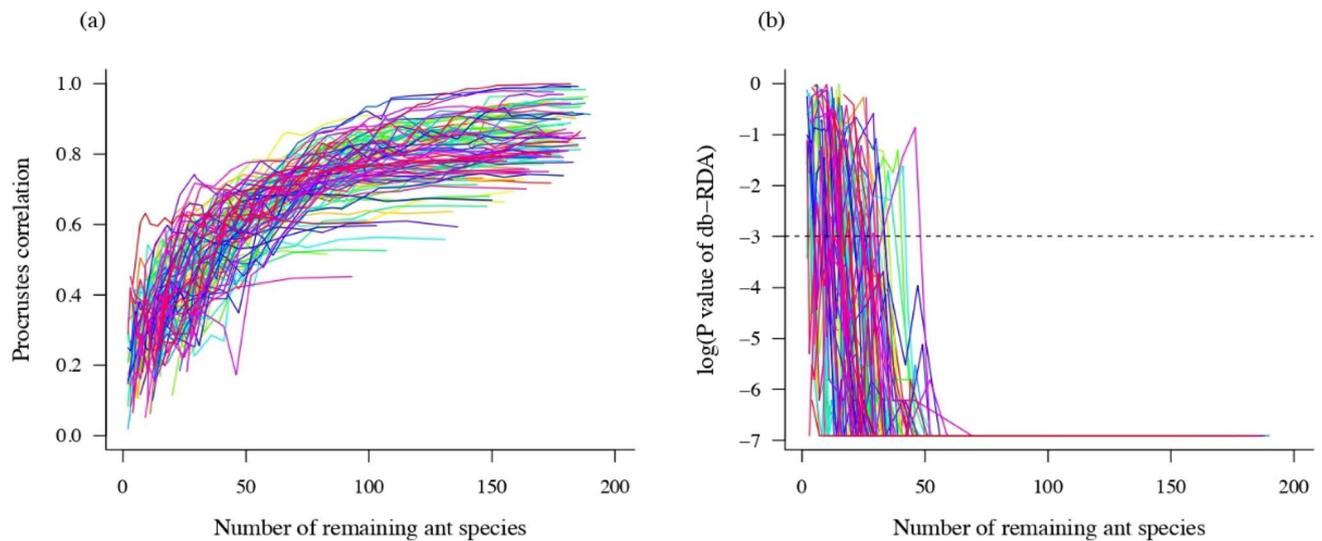
Fig. 3 Procrustes correlation coefficients between the complete matrix and the reduced matrices of ant species composition using simulated data (**a**). P values of the partial RDA for the effect of environmental variables on the complete species matrix and the reduced matrices of ant species composition using simulated data (**b**). For the best P values visualization, the y-axis was transformed into a log scale
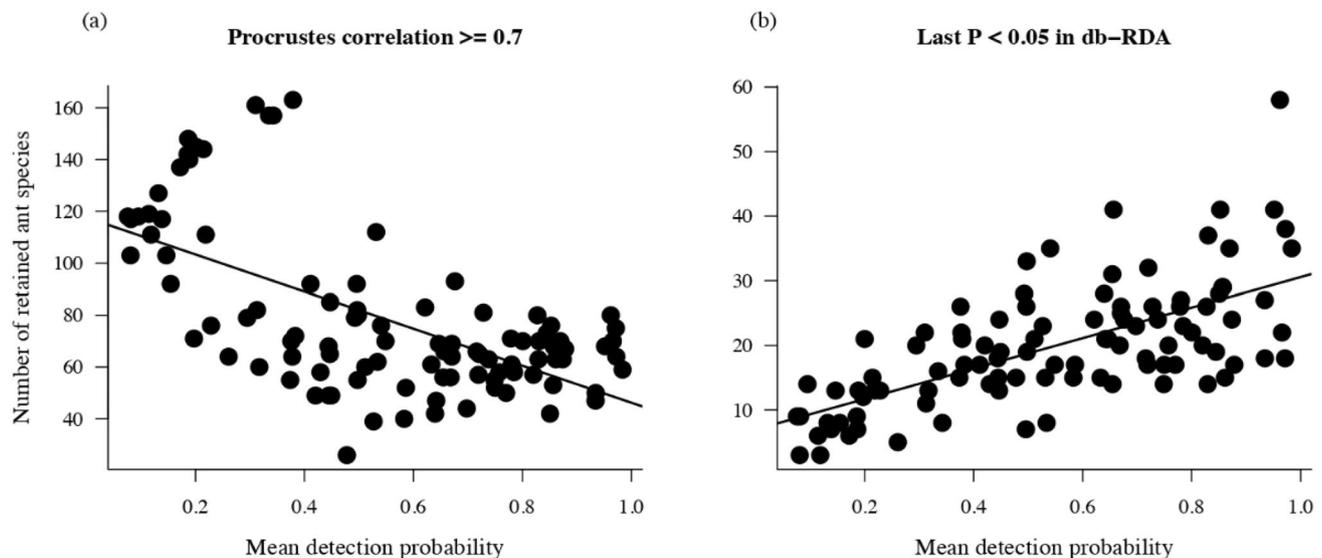

Fig. 4 Mean detection probability of species retained in associations classified as good (Procrustes > 0.7) between common species and the total species composition (**a**). Mean detection probability of species retained in associations with environmental gradient (dbRDA P value ≤ 0.05) (**b**)

environmental response was 25, contributing to estimated savings in time and monetary cost of 20–40% (Table 1; S4). Based on the simulated data, the minimum number of species should be on average 57 (Table S3), which would lead to savings up to 50% (Table 1 and S4). The savings in time in the field and laboratory and the total monetary cost were similar in the years 2006 and 2012. The savings in 2014 were lower compared to previous years when the limit of 57 common species was applied (Table 1 and S4). Using 57 common species as a minimum threshold, in addition to saving time and money, we also maintain associations between common species and total species, in addition to recovering significant environmental responses, achieving an 'excellent' congruence rating (Table S3).

Table 1 Summary of the relative effort (cost and time) required for each group of common ant species sampled in 30 plots of the Ducke Reserve in the years 2006, 2012, and 2014 in addition to the simulated data

2006			2012			2014			Simulated data		
Species retained	Time (%)	Cost (%)	Species retained	Time (%)	Cost (%)	Species retained	Time (%)	Cost (%)	Species retained (average)	Time (%)	Cost (%)
194	100	100	188	100	100	124	100	100	200	100	100
131	96.91	100	120	96.91	100	83	96.91	100	169	96.91	100
104	93.83	100	93	93.83	100	60	93.83	100	157	93.83	100
83	90.74	100	82	90.74	100	53	90.74	100	147	90.74	100
67	87.65	100	70	87.65	100	42	87.65	100	137	87.65	100
59	83.33	96.12	60	83.33	96.12	34	83.33	96.12	128	83.33	96.12
46	80.25	82.15	51	80.25	82.15	28	80.25	82.15	119	80.25	82.15
42	77.16	82.15	47	77.16	82.15	26	77.16	82.15	111	77.16	82.15
36	74.07	82.15	42	74.07	82.15	22	74.07	82.15	103	74.07	82.15
32	70.99	82.15	38	70.99	82.15	18	70.99	82.15	96	70.99	82.15
31	66.67	78.28	35	66.67	78.28	16	66.67	78.28	88	66.67	78.28
29	63.58	78.28	30	63.58	78.28	14	63.58	78.28	82	63.58	78.28
27	60.49	64.31	25	60.49	64.31	12	60.49	64.31	76	60.49	64.31
24	57.41	64.31	23	57.41	64.31	9	57.41	64.31	70	57.41	64.31
21	54.32	64.31	21	54.32	64.31	7	54.32	64.31	64	54.32	64.31
20	50	60.44	20	50	60.44	6	50	60.44	59	50	60.44
15	46.91	60.44	18	46.91	60.44	5	46.91	60.44	54	46.91	60.44
10	43.83	60.44	15	43.83	60.44	4	40.74	46.47	51	43.83	60.44
4	40.74	46.47	13	40.74	46.47	3	30.25	46.47	47	40.74	46.47
3	37.65	46.47	11	37.65	46.47	2	24.07	43.74	42	37.65	46.47
2	30.25	43.74	9	33.33	43.74				39	33.33	43.74
			7	30.25	43.74				32	27.16	43.74
			4	27.16	43.74				29	24.07	43.74
			3	20.99	29.78				25	20.99	29.78
			2	17.29	19.33				22	17.28	25.67
									18	14.20	25.67
									15	11.11	19.33
									12	8.02	19.33

Discussion

This study reports for the first time the use of common ground-dwelling ant species as predictors of species composition patterns of the total ant assemblage over time. It also indicates that the common species of ants have a similar ecological response to those detected using the complete species composition, regardless of the studied year. Additionally, the primary utility of employing common species may lie in longitudinal monitoring of the same location, proving valuable amid swift global environmental changes for understanding real-time community dynamics at minimal cost. The simulated data revealed similar patterns, although they were more conservative regarding the minimum number of common species required to recover the full pattern. Importantly, imperfect detection eroded the information contained by common species data on spatial variation in species composition and its response to environmental factors, but surprisingly decreased the number of species required to detect ecological relationships, suggesting a trade-off between these abilities. These results are consistent with previous findings that dominant species (i.e. most abundant species) can be reliably used to infer patterns of compositional dissimilarity (Roden et al. 2018), but highlight how imperfect detection can modify this ability. These results are of great importance since any reduction in effort must be accompanied by maintenance of the quality of inferences on taxonomic and ecological responses so that the protocol can be considered efficient (Magnusson et al. 2013; Souza et al. 2016; Graça et al. 2017).

The congruence detected between the complete and the reduced matrices indicates that common ground-dwelling

ant species can adequately predict variation in ant species composition over time (2006–2014). A similar result occurred in the analyses of ecological predictors, with the responses of common species being close to that detected using the complete matrix. This was true even when using a small number of common species in all years studied. If we used at least 25 common ant species, we could reproduce all the results found with the complete matrix of 292 species in all studied periods, saving up to 50% of the time and monetary costs of the initial sampling protocol. From a biomonitoring perspective, this is useful as most of these species have not yet been described (Lewinsohn and Prado 2005). Although our study focused on ants, common species may be suitable surrogates in other species-rich groups (e.g. megadiverse groups like Coleoptera, Diptera, Lepidoptera, and other social Hymenoptera), especially those with a restricted group of frequent species and a large group of rare species.

In tropical environments, a large number of rare species is a typical pattern (Hubbell 2001). A possible cause of a large number of singletons is that these may be passerby species or specialists with a low population density (Novotný and Basset 2000), or rare species have different niches relative to common species (Brasil et al. 2020), highlighting the importance of rare species in ecological processes (Gaston 2012; Leitão et al. 2016). Alternatively, they may be simply poorly detected (see below).Surveys of invertebrates as well as vertebrate communities are most often biased by imperfect detection (Yoccoz et al. 2001; Yamaura et al. 2019). There is likely an excess of singleton species in empirical datasets due to under sampling bias, especially for species-rich groups such as tropical invertebrates (Coddington et al. 2009; Chao et al. 2017). For instance, species density can be expected to be a strong predictor of detection probability regardless of environmental conditions, as it is easier find species with more individuals. Furthermore, environmental conditions can also affect detection, which may confound environmental responses. This creates uncertainty about the extent to which rare species provide meaningful information on species composition patterns: seemingly rare species may be truly rare and thus provide reliable information on community patterns, or they may be artifacts of imperfect detection and provide no relevant information (statistical noise) or even bias community patterns, e.g. if such species respond differently to environmental conditions compared to those which are easier to detect but this pattern cannot be detected due to imperfect detection.

Our simulation revealed that imperfect detection increased the number of more common species necessary to adequately represent broader patterns of spatial variation in species composition, but decreased the number of species required to adequately represent environmental relationships. For instance, with imperfect detection, it would take 50 common species to reproduce all the results detected with the complete assemblage, saving between 10 and 20% in the analyzed period vs. 25 species and 50% saving with perfect detection. This result was consistent with our hypothesis that imperfect detection should erode information of species composition patterns. In particular, the number of required species was roughly constant as mean detection probability decreased to ca. 0.4 and then became highly variable, suggesting relative robustness to imperfect detection when detection probability is > 0.4. However, the second result was unexpected, as it suggested that imperfect detection actually facilitated detecting environmental relationships. We suggest this occurred because, for a given detection probability, rarer species should be naturally less observed. Hence, their occurrence patterns are more prone to sampling bias and noise than those of common species. This way, excluding rarer species from the analysis increases the amount of information about the true patten, allowing for the exclusion of most species.

In this light, there would be a trade-off between detecting spatial patterns in species composition and detecting community environmental responses from common species distributions, even though mean detection probability had a stronger effect on the latter than on the former. It should also be noted that P values became increasingly variable (i.e. more different values at each simulation) as the community dataset was reduced. Thus, whereas fewer species were required to detect environmental responses on average, the uncertainty about this inference increased. Overall, our simulation suggested that, under constant imperfect detection, ignoring the species expected to be rarest in a given sampling protocol (e.g. singletons and doubletons) had little impact on species composition inferences, and known patterns can be depicted similarly well if only a relatively small fraction of the species in a community is considered. However, further reductions could have undesired effects, such as imprecise estimates of community environmental responses (Roden et al. 2018).

Any way of saving time or funding in monitoring biodiversity is desirable. This is especially important in regions with mega diversity (Magnusson et al. 2013), or in situations where science has been suffering several funding cuts (Escobar 2015; Angelo 2017; Magnusson et al. 2018; Kowaltowski 2021), or even where the environment has undergone anthropic pressure over the years, resulting in the suppression of local biodiversity (Costa et al. 2020), where time for a complete inventory, even with low diversity, can be a crucial constraint, as is the case of the Brazilian Amazon. Our results using common species suggest time and money savings of 20 to 50% compared to monitoring using the total ground-dwelling ant assemblage. These savings could be used in future collections or in expanding the sampling area (Costa and Magnusson 2010; Souza et al. 2016). Yet, although these results here are promising,

generalizations require caution, as imperfect detection has been shown decrease the ability to infer compositional dissimilarity patterns while increase that ability to detect ecological relationships under the conditions simulate here. This suggests that the cost of ignoring imperfect detection depends on the research goal. Nonetheless, if researchers are more interested in detecting ecological responses rather than in accurately describing compositional dissimilarity on its own and need to reduce costs or reallocate sampling effort, using common species as surrogates is a reasonable option.

More generally, the utility of surrogates in ecology and conservation biology is known to depend on some factors, as indicated by meta-analyses. One the one hand, using one taxon to predict species distribution patterns of another taxon (cross-taxon surrogates) is generally of little or no use (Rodrigues and Brooks 2007; Westgate et al. 2014). On the other hand, using higher-level taxa to predict species-level distribution patterns is generally informative, especially as higher-taxa-to-species ratio increases and evenness of species abundances decreases (Neeson et al. 2013, de Oliveira Jr. et al. 2020). The utility of common or dominant species has received less attention compared to those other approaches. However, simulations and data also suggest that relatively few dominant species are useful in recovering compositional dissimilarity patterns of entire communities of a given taxon, again with increasing utility as species abundance evenness decreases (Roden et al. 2018). In this light, our results based on common or more widespread species support the latter findings, as species with higher local abundance also tend to occur across more sites (Borregaard and Rahbek 2010). Importantly, we showed that the usefulness of common ant species as surrogates for the total assemblage is maintained over time, and this is possibly one of the greatest applicability of this type of approach. Furthermore, this feature is suited to long-term ecological research or environmental monitoring programs. We currently experience frequent environmental changes, which undoubtedly lead to the suppression of biodiversity (Ripple et al. 2021; Caro et al. 2022), and the common species can be a relevant strategy to get some idea of how communities are changing in "real-time" at as little cost as possible. The alternative of not using surrogates for biodiversity (or some equivalent approach), not having enough information (or not being fast enough), and therefore being thoroughly in the dark about what is going on does not seem to be the best choice for the environment (Bergallo et al. 2023; Souza 2023).

Conclusions

In this study, common ant species (approximately 1/4 of the total assemblage) were efficient in predicting overall species composition in the studied Amazonian ground-dwelling ants, as well as in detecting the effects of environmental variables on overall species composition. These results held in three different years spanning a decade, suggesting consistency over time. Simulations indicated that these findings were relatively robust to imperfect detection, although very low detection probability limited the utility of using only the most common species, by reducing the precision of estimates of community environmental responses. Our simulation was limited to a relatively simple scenario, i.e. constant imperfect detection (Yamaura et al. 2019). Further study is warranted in more complex circumstances, such as counting data under environment-dependent imperfect detection. The common species as indicators of species composition patterns of the total assemblage may also be profitable for other megadiverse biological groups (i.e., high species richness and abundance) in localities with vegetation types similar to this study and when the objective is to infer community environmental responses. Nonetheless, our results reiterate the importance of creating and establishing programs that support studies of this scope to develop reliable and efficient shortcuts for assessing biodiversity, allowing a reduction of expenses in future long-term ecological research or environmental monitoring programs and also promoting management efforts.

Data availability The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request. For more details and metadata, please visit PPBio website https://ppbio.inpa.gov.br/manuais

Declarations

References

Agosti D, Majer JD, Alonso LE, Schultz TR (2000) Agosti et al 2000.pdf

Andersen AN (1995) Measuring more of Biodiversity: Genus richness as a surrogate for species richness in Australian. Biol Conserv 73:39–43

Angelo C (2017) Brazilian scientists reeling as federal funds slashed by nearly half. Nature. https://doi.org/10.1038/nature.2017.21766

Austin MP, Belbin L, Meyers JA et al (2006) Evaluation of statistical models used for predicting plant species distributions: role of artificial data and theory. Ecol Modell 199:197–216. https://doi.org/10.1016/j.ecolmodel.2006.05.023

Baccaro FB, Feitosa RM, Fernandez F, et al (2015) Guia para os gêneros de formigas do Brasil

Báez S, Jaramillo L, Cuesta F, Donoso DA (2016) Effects of climate change on Andean biodiversity: a synthesis of studies published until 2015. Neotrop Biodivers 2:181–194. https://doi.org/10.1080/23766808.2016.1248710

Benoit D, Jackson DA, Ridgway MS (2018) Assessing the impacts of imperfect detection on estimates of diversity and community structure through multispecies occupancy modeling. Ecol Evol 8:4676–4684. https://doi.org/10.1002/ece3.4023

Bergallo HG, Rosa C, Ochoa AC et al (2023) Long-term ecological research: chasing fashions or being prepared for fashion changes? An Acad Bras Cienc 95:e20230051. https://doi.org/10.1590/0001-3765202320230051

Bestelmeyer BT, Agosti D, Leeanne F, Alonso T, Brandão CRF, Brown WL, Delabie JHC, Silvestre R (2000) Field techniques for the study of ground-living ants An Overview, description, and evaluation. In: Agostel D, Majer JD, Alonso L, Schulizz TD (eds) Standard methods for measuring and monitoring biodivrsity. Springer, Charm, pp 122–144

Borcard D, Legendre P (1994) Borcard_et_legendre_1994_Environmental control and spatial structure in ecological communities. pdf. 61:37–53

Borcard D, Gillet F, Legendre P (2011) Numerical ecology with R. New York: springer 2:688

Borregaard MK, Rahbek C (2010) Causality of the relationship between geographic distribution and species abundance. Q Rew Biol 85:3–25. https://doi.org/10.1086/650265

Brasil LS, Vieira TB, Andrade AFA et al (2020) The importance of common and the irrelevance of rare species for partition the variation of community matrix: implications for sampling and conservation. Sci Rep 10:1–8. https://doi.org/10.1038/s41598-020-76833-5

Butchart SH, Walpole M, Collen B et al (2010) Global biodiversity: indicators of recent declines. Science 328(5982):1164–1168. https://doi.org/10.1126/science.1187512

Caro T (2010) Conservation by Proxy: Indicator, Umbrella, Keystone, Flagship, and Other Surrogate Species. Island Press, Washington

Caro T, Rowe Z, Berger J et al (2022) An inconvenient misconception: climate change is not the principal driver of biodiversity loss. Conserv Lett. https://doi.org/10.1111/conl.12868

Chao A, Colwell RK, Chiu CH, Townsend D (2017) Seen once or more than once: applying Good-Turing theory to estimate species richness using only unique observations and a species list. Methods Ecol Evol 8:1221–1232. https://doi.org/10.1111/2041-210X.12768

Chauvel A, Lucas Y, Boulet R (1987) On the genesis of the soil mantle of the region of Manaus, Central Amazonia, Brazil. Experientia 43:234–241. https://doi.org/10.1007/BF01945546

Chen G, Kéry M, Plattner M et al (2013) Imperfect detection is the rule rather than the exception in plant distribution studies. J Ecol 101:183–191. https://doi.org/10.1111/1365-2745.12021

Coddington JA, Agnarsson I, Miller JA et al (2009) Undersampling bias: the null hypothesis for singleton species in tropical arthropod surveys. J Anim Ecol. https://doi.org/10.1111/j.1365-2656.2009.01525.x

Costa FRC, Magnusson WE (2010) The need for large-scale, integrated studies of biodiversity—the experience of the program for biodiversity research in Brazilian Amazonia. Nat Conserv 08:3–12. https://doi.org/10.4322/natcon.00801001

Costa FRC, Magnusson WE, Luizao RC (2005) Mesoscale distribution patterns of Amazonian understorey herbs in relation to topography, soil and watersheds. J Ecol. https://doi.org/10.1111/j.1365-2745.2005.01020.x

Costa FRC, Zuanon JAS, Baccaro FB et al (2020) Effects of climate change on central amazonian forests: a two decades synthesis of monitoring tropical biodiversity. Oecologia Aust 24:317–335. https://doi.org/10.4257/oeco.2020.2402.07

de Oliveira Jr SS, Ortega JC, dos Santos Ribas LG, Lopes VG, Bini LM (2020) Higher taxa are sufficient to represent biodiversity patterns. Ecol Indic 111:105994. https://doi.org/10.1016/j.ecolind.2019.105994

de Souza JLP, Baccaro FB, Landeiro VL 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. https://doi.org/10.1016/j.apsoil.2012.01.004

Delabie JHC, Koch E, Dodonov P et al (2021) Sampling and analysis methods for ant diversity assessment. In: Santos JC, Fernandes GW (eds) Measuring arthropod biodiversity—a handbook of sampling methods. Springer, Dordrecht, pp 13–54

Dorazio RM, Gotelli NJ, Ellison AM (2011) Modern methods of estimating biodiversity from presence-absence surveys. In: Biodiversity Loss in a Changing Planet. InTech, pp 277–302

dos Santos Ribas LG, Padial AA (2015) The use of coarser data is an effective strategy for biological assessments. Hydrobiologia 747:83–95. https://doi.org/10.1007/s10750-014-2128-6

Elizalde L, Arbetman M, Arnan X et al (2020) The ecosystem services provided by social insects: traits, management tools and knowledge gaps. Biol Rev 95:1418–1441. https://doi.org/10.1111/brv.12616

Escobar H (2015) Fiscal crisis has Brazilian scientists scrambling. Science 349:909–910

Franklin E, De Moraes J, Landeiro VL et al (2013) Geographic position of sample grid and removal of uncommon species affect multivariate analyses of diverse assemblages: the case of oribatid mites (Acari: Oribatida). Ecol Indic 34:172–180. https://doi.org/10.1016/j.ecolind.2013.04.024

Gardner TA, Barlow J, Araujo IS et al (2008) The cost-effectiveness of biodiversity surveys in tropical forests. Ecol Lett 11:139–150. https://doi.org/10.1111/j.1461-0248.2007.01133.x

Gaspar C, Gaston KJ, Borges PAV (2010) Arthropods as surrogates of diversity at different spatial scales. Biol Conserv 143:1287–1294. https://doi.org/10.1016/j.biocon.2010.03.007

Gaston KJ (2010) Valuing common species. Science 327:154–155. https://doi.org/10.1126/science.1182818

Gaston KJ (2012) The importance of being rare. Nature 487:46–47. https://doi.org/10.1038/487046a

Gomes CB, Souza JLP, Franklin E (2018) A comparison between time of exposure, number of pitfall traps and the sampling cost to capture ground-dwelling poneromorph ants (Hymenoptera: Formicidae). Sociobiology 65:138–148. https://doi.org/10.13102/sociobiology.v65i2.1207

Gotelli NJG, Ellison AM, Dunn RR, Sanders N (2011) Counting ants (Hymenoptera: Formicidae): biodiversity sampling and statistical analysis for myrmecologists. Myrmecological News 15:13–19

Graça MB, Souza JLP, Franklin E et al (2017) Sampling effort and common species: optimizing surveys of understorey fruit-feeding butterflies in the Central Amazon. Ecol Indic 73:181–188. https://doi.org/10.1016/j.ecolind.2016.09.040

Guillera-Arroita G (2017) Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. Ecography (cop) 40:281–295. https://doi.org/10.1111/ecog.02445

Heino J, Soininen J (2007) Are higher taxa adequate surrogates for species-level assemblage patterns and species richness in stream organisms? Biol Conserv 137:78–89. https://doi.org/10.1016/j.biocon.2007.01.017

Hodkinson ID, Jackson JK (2005) Terrestrial and aquatic invertebrates as bioindicators for environmental monitoring, with particular reference to mountain ecosystems. Environ Manage 35:649–666. https://doi.org/10.1007/s00267-004-0211-x

Hölldobler B, Wilson EO (1990) The ants. Harvard University Press, Cambridge

Hubbell SP (2001) The unified neutral theory of biodiversity and biogeography. Princeton University Press

Hui FKC, Warton DI, Ormerod JT et al (2017) Variational approximations for generalized linear latent variable models. J Comput Graph Stat 26:35–43. https://doi.org/10.1080/10618600.2016.1164708

Kaspari M, 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/10.1086/303313

Keizer-Vlek HE, Verdonschot PFM, Verdonschot RCM, Goedhart PW (2012) Quantifying spatial and temporal variability of macroinvertebrate metrics. Ecol Indic 23:384–393. https://doi.org/10.1016/j.ecolind.2012.04.025

Khan SA (2006) Is species level identification essential for environmental impact studies? JSTOR 91:29–34

Kowaltowski AJ (2021) Brazil's scientists face 90% budget cut. Nature 598:566. https://doi.org/10.1038/d41586-021-02882-z

Lamoreux JF, Morrison JC, Ricketts TH et al (2006) Global tests of biodiversity concordance and the importance of endemism. Nature 440:212–214. https://doi.org/10.1038/nature04291

Landeiro VL, Bini LM, Costa FRC et al (2012) How far can we go in simplifying biomonitoring assessments? An integrated analysis of taxonomic surrogacy, taxonomic sufficiency and numerical resolution in a megadiverse region. Ecol Indic 23:366–373

Lavoie I, Dillon PJ, Campeau S (2009) The effect of excluding diatom taxa and reducing taxonomic resolution on multivariate analyses and stream bioassessment. Ecol Indic 9:213–225. https://doi.org/10.1016/j.ecolind.2008.04.003

Leal IR, Bieber AGD, Tabarelli M, Andersen AN (2010) Biodiversity surrogacy: indicator taxa as predictors of total species richness in Brazilian Atlantic forest and Caatinga. Biodivers Conserv. https://doi.org/10.1007/s10531-010-9896-8

Leitão RP, Zuanon J, Williams SE et al (2016) Rare species contribute disproportionately to the functional structure of species assemblages. Proc R Soc B: Biol Sci. https://doi.org/10.1098/rspb.2016.0084

Lewinsohn TM, Prado PI (2005) How many species are there in Brazil? Conserv Biol 19:619–624. https://doi.org/10.1111/j.1523-1739.2005.00680.x

Magnusson WE, Lima AP, Luizão R et al (2005) RAPELD: a modification of the Gentry method for biodiversity surveys in long-term ecological research sites. Biota Neotrop 5:19–24. https://doi.org/10.1590/S1676-06032005000300002

Magnusson WE, Grelle CEV, Marques MCM et al (2018) Effects of Brazil's political crisis on the science needed for biodiversity conservation. Front Ecol Evol 6:1–5. https://doi.org/10.3389/fevo.2018.00163

Magnusson W, Braga-Neto R, Baccaro FPF, et al (2013) Biodiversidade e Monitoramento Ambiental Integrado/Biodiversity and Integrated Environmental Monitoring

Magurran AE, Baillie SR, Buckland ST et al (2010) Long-term datasets in biodiversity research and monitoring: Assessing change in ecological communities through time. Trends Ecol Evol 25:574–582. https://doi.org/10.1016/j.tree.2010.06.016

Margules CR, Pressey RL, Williams PH (2002) Representing biodiversity: data and procedures for identifying priority areas for conservation. J Biosci 27:309–326. https://doi.org/10.1007/BF02704962

Marques-Filho AO, Ribeiro MNG, Santos HM, Santos JM (1982) Estudos climatológicos da Reserva Florestal Ducke. Acta Amaz 11:759–768

Mazaris AD, Kallimanis AS, Tzanopoulos J et al (2010) Can we predict the number of plant species from the richness of a few common genera, families or orders? J Appl Ecol 47:662–670. https://doi.org/10.1111/j.1365-2664.2010.01814.x

McGeoch MA (1998) The selection, testing and application of terrestrial insects as bioindicators. Biol Rev 73:181–201. https://doi.org/10.1111/j.1469-185X.1997.tb00029.x

Melo DHA, Filgueiras BK, Iserhard CA et al (2019) Effect of habitat loss and fragmentation on fruit-feeding butterflies in the Brazilian Atlantic Forest. Canad J Zool 97(7):588–596. https://doi.org/10.1139/cjz-2018-0202

Minchin PR (1987) Simulation of multidimensional community patterns: towards a comprehensive model. Vegetatio 71:145–156. https://doi.org/10.1007/BF00039167

Minor MA (2011) Spatial patterns and local diversity in soil oribatid mites (Acari: Oribatida) in three pine plantation forests. Eur J Soil Biol 47:122–128. https://doi.org/10.1016/j.ejsobi.2011.01.003

Neeson TM, Rijn IV, Mandelik Y (2013) How taxonomic diversity, community structure, and sample size determine the reliability of higher taxon surrogates. Ecol Appl 23:1216–1225. https://doi.org/10.1890/12-1167.1

Nekola JC, Mcgill BJ (2014) Scale dependency in the functional form of the distance decay relationship. Ecography (cop) 37:309–320. https://doi.org/10.1111/j.1600-0587.2013.00407.x

Novotný V, Basset Y (2000) Rare species in communities of tropical insect herbivores: pondering the mystery of singletons. Oikos 89:564–572. https://doi.org/10.1034/j.1600-0706.2000.890316.x

Oksanen J, Blanchet F, Friendly M, et al (2019) vegan: Community Ecology Package. R package version 2.5–6

Olsgard F, Brattegard T, Holthe T (2003) Polychaetes as surrogates for marine biodiversity: lower taxonomic resolution and indicator groups. Biodivers Conserv 12:1033–1049. https://doi.org/10.1023/A:1022800405253

Paknia O, Pfeiffer M (2011) Hierarchical partitioning of ant diversity: implications for conservation of biogeographical diversity in arid and semi-arid areas. Divers Distrib 17:122–131. https://doi.org/10.1111/j.1472-4642.2010.00719.x

Peres-Neto PR, Jackson DA (2001) How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. Oecologia 129:169–178. https://doi.org/10.1007/s004420100720

R Development Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Ripple WJ, Wolf C, Newsome TM et al (2021) World scientists' warning of a climate emergency 2021. Bioscience 71:894–898. https://doi.org/10.1093/biosci/biab079

Roden VJ, Kocsis ÁT, Zuschin M, Kiessling W (2018) Reliable estimates of beta diversity with incomplete sampling. Ecology 99:1051–1062. https://doi.org/10.1002/ecy.2201

Rodrigues ASL, Brooks TM (2007) Shortcuts for biodiversity conservation planning: the effectiveness of surrogates. Annu Rev Ecol Evol Syst 38(1):713–737. https://doi.org/10.1146/annurev.ecolsys.38.091206.095737

Root TL, Price JT, Hall KR et al (2003) Fingerprints of global warming on wild animals and plants. Nature 421(6918):57–60. https://doi.org/10.1038/nature01333

Rosser N, Eggleton P (2012) Can higher taxa be used as a surrogate for species-level data in biodiversity surveys of litter/soil insects? J Insect Conserv 16:87–92. https://doi.org/10.1007/s10841-011-9395-6

Schultheiss P, Nooten SS, Wang R, Wong MKL, Brassard F, Guénard B (2022) The abundance, biomass, and distribution of ants on Earth. Proc Nat Acad Sci 119:e2201550119. https://doi.org/10.1073/pnas.2201550119

Šiding AL, Šizlingová E, Storch D et al (2009) Rarity, commonness, and the contribution of individual species to species richness patterns. Am Nat 174:82–93. https://doi.org/10.1086/599305

Siqueira T, Bini LM, Roque FO et al (2012) Common and rare species respond to similar niche processes in macroinvertebrate metacommunities. Ecography (cop) 35:183–192. https://doi.org/10.1111/j.1600-0587.2011.06875.x

Souz JLP, Fernandes IO, Agosti D, Johnson NF, Baccaro FB (2022) Assessing the efficacy of higher-taxon approach for ant species surveys to improve biodiversity inventories. Anim Conserv 25:370–381. https://doi.org/10.1111/acv.12758

Souza JLP (2023) Effectiveness of higher-taxon approach on ants and sample size effect: an assessment in Brazilian biomes and states. Biodivers Conserv 32:635–652. https://doi.org/10.1007/s10531-022-02516-4

Souza JLP, Fernandes IO (2021) Do spatial and temporal scales affect the efficiency of surrogates in ant monitoring on the hydroelectric power-plant area in Brazilian Amazon? Ecol Indic 121:107158

Souza JLP, Baccaro FB, Landeiro VL et al (2016) Taxonomic sufficiency and indicator taxa reduce sampling costs and increase monitoring effectiveness for ants. Divers Distrib 22:111–122. https://doi.org/10.1111/ddi.12371

Souza JLP, Baccaro FB, Pequeno PACL et al (2018) Effectiveness of genera as a higher-taxon substitute for species in ant biodiversity analyses is not affected by sampling technique. Biodivers Conserv 27:3425–3445. https://doi.org/10.1007/s10531-018-1607-x

Thomson SA et al (2018) Taxonomy based on science is necessary for global conservation. PLoS Biol 16:1–12. https://doi.org/10.1371/journal.pbio.2005075

Vasconcelos HL, Macedo ACC, Vilhena JMS (2003) Influence of topography on the distribution of ground-dwelling ants in an Amazonian forest. Stud Neotrop Fauna Environ 38:115–124. https://doi.org/10.1076/snfe.38.2.115.15923

Vázquez LB, Gaston KJ (2004) Rarity, commonness, and patterns of species richness: the mammals of Mexico. Glob Ecol Biogeogr 13:535–542. https://doi.org/10.1111/j.1466-822X.2004.00126.x

Vellend M (2016) The theory of ecological communities (MPB-57). Princeton University Press, Princeton. https://doi.org/10.1515/9781400883790

Vellend M, Lilley PL, Starzomski BM (2008) Using subsets of species in biodiversity surveys. J Appl Ecol 45:161–169. https://doi.org/10.1111/j.1365-2664.2007.01413.x

Warman LD, Forsyth DM, Sinclair ARE et al (2004) Species distributions, surrogacy, and important conservation regions in Canada. Ecol Lett 7:374–379. https://doi.org/10.1111/j.1461-0248.2004.00590.x

Westgate M, Barton P, Lane P et al (2014) Global meta-analysis reveals low consistency of biodiversity congruence relationships. Nat Commun 5:3899. https://doi.org/10.1038/ncomms4899

Whittaker RJ, Araújo MB, Jepson P et al (2005) Conservation biogeography: assessment and prospect. Divers Distrib 11:3–23. https://doi.org/10.1111/j.1366-9516.2005.00143.x

Williams PH, Gaston KJ (1994) Measuring more of biodiversity: can higher-taxon richness predict wholesale species richness? Biol Conserv 67:211–217. https://doi.org/10.1016/0006-3207(94)90612-2

Yamaura Y, Blanchet FG, Higa M (2019) Analyzing community structure subject to incomplete sampling: hierarchical community model vs. canonical ordinations. Ecology. https://doi.org/10.1002/ecy.2759

Yoccoz NG, Nichols JD, Boulinier T (2001) Monitoring of biological diversity in space and time. Trends Ecol Evol 16:446–453. https://doi.org/10.1016/S0169-5347(01)02205-4