Diversity and Distributions, (Diversity Distrib.) (2016) **22**, 111–122

Diversity and Distributions
A Journal of Conservation Biogeography

BIODIVERSITY
RESEARCH

Taxonomic sufficiency and indicator taxa reduce sampling costs and increase monitoring effectiveness for ants

Jorge Luiz Pereira Souza[1]*, Fabricio Beggiato Baccaro[2], Victor Lemes Landeiro[3], Elizabeth Franklin[4], William Ernest Magnusson[4], Pedro Aurélio Costa Lima Pequeno[5] and Itanna Oliveira Fernandes[1]

[1]*Programa de Pós-Graduação em Entomologia, Instituto Nacional de Pesquisas da Amazônia-INPA, Caixa Postal 2223, CEP 69080–971, Manaus, AM, Brazil,* [2]*Departamento de Biologia, Universidade Federal do Amazonas-UFAM, Av. General Rodrigo Octávio, 6200, Coroado I, CEP 69077–000, Manaus, AM, Brazil,* [3]*Departamento de Botânica e Ecologia, Instituto de Biociências, Universidade Federal de Mato Grosso-UFMT, CEP 78060–900, Cuiabá, MT, Brazil,* [4]*Instituto Nacional de Pesquisas da Amazônia-INPA, Coordenação de Biodiversidade, Caixa Postal 2223, CEP 69080–971, Manaus, AM, Brazil,* [5]*Programa de Pós-Graduação em Ecologia, Instituto Nacional de Pesquisas da Amazônia-INPA, Caixa Postal 2223, CEP 69080–971, Manaus, AM, Brazil*

*Correspondence: Jorge Luiz Pereira Souza, Programa de Pós-Graduação em Entomologia, Instituto Nacional de Pesquisas da Amazônia-INPA, Caixa Postal 2223, CEP 69080–971, Manaus, AM, Brazil.
E-mail: souza.jorge@gmail.com

ABSTRACT

Aim Despite the accelerating loss of biodiversity and the increased number of methods for conservation planning, the availability of information about the spatial distribution of biodiversity remains limited. One way to overcome this problem is to focus on surrogate resolutions that are able to represent species-level data and can be efficiently measured. Surrogates are only useful if the ecological patterns detected at the species-level still hold when based on coarser taxonomic identification, and if these responses are consistent across regions. We present a comprehensive analysis using data from a large-scale evaluation of ground-dwelling ants, to evaluate the use of surrogates.

Location Amazon basin.

Methods The sampling design covered 13 sites in eight phytophysiognomies, which in conjunction with other environmental characteristics (altitude, soil granulometry and slope) were used to validate the ecological patterns (ability of the surrogates to reproduce the ecological responses identified for species) of coarser surrogate taxa (indicator taxa, mixed-level approach, genus and subfamily). The surrogates were evaluated for their capacity to predict variation in total species richness and composition. We also estimated the monetary and time costs, in order to evaluate the cost-effectiveness of using different surrogate levels.

Results Genus was the most cost-effective surrogate: it predicted 81% of site variation in species richness, was highly correlated ($r^2 = 0.76$) with species composition, very highly correlated ($r^2 = 0.97$) with ecological patterns detected at species level and saved ~40% of total project costs. The mixed-level approach, indicator taxa and subfamily were not effective in representing the species-level data.

Main conclusions Genus can be used as a surrogate for species, due to its high predictive value, independent of environmental heterogeneity. Genus may be useful as a surrogate for species in other megadiverse regions, especially where savings in project costs can be applied to increase sampling effort.

Keywords
alpha diversity, amazonia, beta diversity, latitudinal gradient, monitoring, sampling standardized protocol, tropical forest

INTRODUCTION

Biodiversity loss is accelerating rapidly in response to increasing human influence on the Earth's natural ecosystems (Laurance *et al.*, 2012). Despite the increased availability of methods for conservation planning, adequate information about the spatial distribution of biodiversity in large regions, such as the Amazon basin, remains sparse for most biological groups (Margules *et al.*, 2002). One way to overcome this problem is to focus on fewer taxonomic groups or surrogate

DOI: 10.1111/ddi.12371

taxonomic levels that are able to represent other taxa and provide satisfactory answers within a short time (Andersen, 1995; Landeiro *et al.*, 2012). Surrogates could be a higher taxonomic unit (i.e. Genus or Family) to bridge a taxonomic gap for species identification in the same group. Organisms can often be easily and quickly identified to coarser-resolution taxonomic categories (Olsgard *et al.*, 2003), but these levels of identification are only useful if patterns identified from coarser-resolution sorting reflect the patterns that would be found by identifying the specimens to species level.

The use of coarser taxonomic units as surrogates is more common for highly diverse taxa, such as terrestrial or aquatic invertebrates (Gaspar *et al.*, 2010; Gallego *et al.*, 2012). For example, ant genera have been proposed as taxonomic surrogates for ant species (Andersen, 1995; Oliver & Beattie, 1996a). Ants comprise a highly diverse group, can be sampled quickly and easily (Agosti *et al.*, 2000; Souza *et al.*, 2012), are associated with important ecological functions (Hoffmann & Andersen, 2003), and the colonies are relatively stationary, hence facilitating long-term monitoring (Kaspari & Weiser, 2000). Ants normally represent a high proportion (~30%) of terrestrial-animal biomass, especially in forested areas (Fittkau & Klinge, 1973; King *et al.*, 2013). Unfortunately, comprehensive ant surveys normally result in thousands of individuals and hundreds of species to be sorted and identified (e.g. Lopes & Vasconcelos, 2008; Souza *et al.*, 2012) and require much more time than vertebrate taxa, increasing the project costs (Pik *et al.*, 1999; Moreno *et al.*, 2008).

Three main recommendations have been made to shorten the identification process for ants. The first is to use morphospecies as a surrogate for species (Oliver & Beattie, 1993, 1996b). This recommendation usually generates a large number of morphotypes, which are generally difficult to align with morphotypes erected by other researchers. To reduce this problem, several recent studies in the Amazon Basin have established standardized morphotypes based on morphological characters for investigations of different areas (e.g. Souza *et al.*, 2012). The second shortcut is to use species rich genera to represent ant diversity, which may be more informative than genus richness (Andersen, 1995). Higher taxonomic levels, such as genus and family, can also be used as surrogates for species (Pik *et al.*, 1999; Gardner *et al.*, 2008). In Australia, Andersen (1995) and Pik *et al.* (1999) found a strong correlation between the number of species and the number of genera in mesic areas and eucalyptus plantations respectively. The third shortcut is based on the concept of taxonomic sufficiency (Ellis, 1985), in which the identification is carried to species level, but focuses on specific 'indicator taxa', which are sometimes referred as 'biological indicators' (*sensu* McGeoch, 1998). Groc *et al.* (2010) suggested the use of a mixed-level approach that combines the species listed as indicator taxa with identifications only to genus for those genera for which specifies-level identification is difficult.

Some studies have used ecological responses, based on environmental variables that determine the distribution of organisms, to evaluate the decision to reduce the number of sampling techniques (Souza *et al.*, 2012) or sample volume (Santos *et al.*, 2008), or the selection of surrogate species (Franklin *et al.*, 2013). Investigating the ecological relationships between surrogate taxa and environmental variables is of more practical importance than simply showing correlations between numbers of taxonomic entities. In many tropical areas, topography and soil texture are correlated with water availability, with better-drained clayey soils on the plateaus and relatively poorly drained sandy soils in the valleys (Hodnett *et al.*, 1997). This landscape feature generates microhabitat variability that can affect spatial patterns of ground-dwelling ant assemblages at local scales (Mezger & Pfeiffer, 2011). Ant diversity and activity seem to be correlated with moisture availability and are higher in wetter seasons and localities (Kaspari & Weiser, 2000). Therefore, useful surrogate taxa should retrieve trends related to topographic and soil texture variation at the local scale, similar to those found when using species-level identification.

In this study, we evaluated the use of coarser-resolution taxonomic categories as surrogates to describe patterns in ground-dwelling ant assemblages distributed over a latitudinal range covering 1800 km in the Brazilian Amazon. First, we documented differences between species composition and four potential surrogate resolutions (subfamily, genus, indicator taxa and mixed-level approach) at 13 sites covering eight vegetation associations and different topographical and soil characteristics across the region. Then, we measured the degree of congruence between the species-level data and the four potential surrogate resolutions. The aim was to find the coarsest level of information that could be used as a substitute for the species-level identifications and to test whether the ecological patterns observed with the species-level dataset could be retrieved from the data on higher taxonomic categories. We also estimated the monetary and time costs to obtain data at different taxonomic resolutions, in order to evaluate the cost-effectiveness of using the different surrogates in monitoring programs.

METHODS

Study sites

The study was conducted at 13 sites associated with the Brazilian Biodiversity Research Program (PPBio). The sites cover a latitudinal gradient of 1800 km in the Amazon Basin (Fig. S1), the coordinates, vegetation types, elevation range, rainfall and spatial sampling design of the study sites are summarized in Table 1.

Sampling design

We sampled ants using the RAPELD sampling design, which is based on a system of trails and permanent plots where a diverse range of taxa can be sampled (Magnusson *et al.*, 2005). The permanent plots are 250 m long and positioned to follow

Table 1 Vegetation, rainfall, elevation range, number of samples and sampling techniques for the 13 sites in the Brazilian Amazon.

Sites	Coordinates	Vegetation type (number of plots)	Elevation range (m.a.s.l.)	Mean rainfall (mm)	Sampling area (km^2)	Number of plots	Number of samples per plot	Technique used	Total of samples	Sampling date
Maracá	3°22′56.73″N 61°27′52.31″W	Open ombrophilous forest (14), Deciduous forest (7), Semi-deciduous forest (8), Campinarana seasonal forest (1)	54–85	1718	25	30	10	Bait, Pitfall & Winkler	900	February/2007
Cauamé	2°52′12.00″N 60°38′24.00″W	Open savannas (12)	67–85	1650	5	12	10	Bait & Pitfall	240	February/2011
Viruá	1°27′49.28″N 61°1′30.59″W	Open ombrophilous forest (12), Campinarana seasonal forest (8), Seasonal campinarana (6), Seasonal shrubby campinarana (4)	43–130	1682	25	30	10	Bait, Pitfall & Winkler	900	February/2007
UFAM	2°38′26.51″S 60°5′44.55″W	Dense ombrophilous forest (21)	42–130	2362	24	21	10	Pitfall	210	September/2011
Ducke	2°57′51.69″S 59°56′27.26″W	Dense ombrophilous forest (30)	46–110	2507	25	30	10	Bait, Pitfall & Winkler	850	September/2006
Manaquiri	3°41′31.25″S 60°14′51.60″W	Open ombrophilous forest (10)	30–36	2200	5	10	10	Pitfall & Winkler	200	November/2009
Orquestra	4°59′2.39″S 61°34′30.00″W	Dense ombrophilous forest (5)	36–61	2200	5	5	10	Pitfall	50	November/2010
Campanã	5°36′36.00″S 62°12′0.00″W	Dense ombrophilous forest (5)	70–72	2200	5	5	10	Pitfall	50	November/2010
Jari	5°58′11.99″S 62°29′24.00″W	Dense ombrophilous forest (5)	70–72	2200	5	5	10	Pitfall	50	November/2010
Teotônio	8°50′28.50″S 64°3′43.92″W	Dense ombrophilous forest (6)	69–112	2246	5	6	5	Winkler	120*	September/2011 to June
Ilha Búfalos	9° 9′6.56″S 64°30′6.97″W	Dense ombrophilous forest (6)	82–115	2246	5	6	5	Winkler	120*	September/2011 to June
Ilha Pedras	9°10′36.22″S 64°36′38.83″W	Dense ombrophilous forest (6)	76–113	2246	5	6	5	Winkler	120*	September/2011 to June
Jaci-Paraná	9°27′44.43″S 64°23′32.97″W	Dense ombrophilous forest (6)	103–134	2246	5	6	5	Winkler	120*	September/2011 to June

*Amounts related to 30 samples taken four times.

terrain contours, to minimize the effects of topographical variation within plots. At each site, the plots were 1 km distant from each other and all plots were sampled in a standardized spatial design. We used two arrangements, with five or ten samples per sampling technique in each permanent plot. The number of plots per site ranged from 5 to 30 and the number of samples per site ranged from 50 to 900 (number of samples X number of plots X number of techniques used). Collections were made four times at the Teotônio, Ilha Búfalos, Ilha Pedras and Jaci-Paraná sites, during September 2011 and June 2012. In total, we sampled 369 plots and took 3930 samples (Table 1). The differences in the number of sampling units per sampling site did not affect the conclusions, because the study was focused on comparisons between different taxonomic resolutions within sites, not on comparisons of assemblages among sites.

Ant sampling

Ground-dwelling ants were collected using pitfall traps, sardine baits and litter samples obtained with a Winkler extractor. One to three sampling techniques were used per site (Table 1). At each sampling station, the ants were extracted from 1 m^2 of sifted litter in a Winkler extractor through a 1 cm^2 mesh sieve and placed in Winkler sacks. The sieved litter was placed in a mesh bag suspended inside a cotton bag for 48 h. Pitfall traps and sardine baits were placed at the same stations after the litter was collected. The pitfall traps were 8 cm deep and 9.5 cm in diameter and were buried with the rim at ground level and left open for 48 h. After the pitfall traps were removed, approximately 5 g of canned sardine was placed on a plastic card on the litter surface. All ants on the plastic card were collected after 45 min and preserved in 90% ethanol (see Souza *et al.*, 2012 for further details).

All ants were sorted to morphospecies, and whenever possible were identified to species, using available taxonomic keys or by comparison with specimens in collections previously identified by experts. Vouchers are deposited in the INPA Entomological Collection. The raw data used here (except for those from Santo Antônio) are available from the PPBio web site (https://ppbio.inpa.gov.br). The species list per study site is provided in Table S1.

Surrogate resolution

We tested four levels of potential surrogates for ant species. First, we reduced the taxonomic resolution by pooling species into subfamily and genus (higher-taxon surrogacy; Bolton *et al.*, 2005). Second, we selected a subset of genera to be used as indicator taxa, based on the concept described by Andersen (1995), where widespread genera and with higher species richness can be good predictors of all species. To be used as indicator taxa, genera were chosen that fit the following criteria: (1) had at least one species in 10 of 13 sampling sites, (2) had at least 10 species in all pooled site data, (3) were taxonomically well known and (4) contained species that can be easily identified from morphological characters. The genera that satisfied these criteria were *Camponotus*, *Crematogaster, Dolichoderus, Gnamptogenys* and *Pheidole*. Third, we used the mixed-level group proposed by Groc *et al.* (2010). This approach combines two identification levels, genera and species. The matrices of the mixed-level approach are composed by the five genera selected above (indicator taxa) identified to species level, plus the other ants identified taxonomically to genus level. Fourth, as the indicator data matrix comprises genera with high species richness (speciose), we also tested each individual genus as a predictor for the level of species. In general, individual species rich genera were less effective for predicting patterns of species compared with the results of the indicator taxa matrix, and therefore, these data are presented only in the supplemental material.

All plot-dissimilarity matrices based on surrogates were compared with the species-level matrix at each study site (see data analysis).

Data analysis

All analyses described below were undertaken with site as the unit of analysis. Data for each identification level (species, mixed-level, indicator taxa, genus and subfamily) were standardized (mean = 0; SD = 1). All analyses were run in the R environment for statistical computing (R Development Core Team, 2014). We used the reference values proposed by Leal *et al.* (2010) for predictions of species richness by surrogates and expanded its use for surrogates in predicting species composition, recovery of ecological patterns and efficiency in maintaining ecological patterns. A surrogate was defined as ‘reasonable’ if it predicted $\geq$ 60% and < 70% of the variation found with species-level data, ‘good’ if it predicted $\geq$ 70% and < 80%, and ‘excellent’ if it predicted $\geq$ 80%. Surrogates were considered useful when they could predict species richness, species composition, information about ecological patterns of species and the ability to maintain the relative order of importance of the predictor variables in > 60% of cases.

Surrogate resolution and richness

We individually regressed an index of ant species richness (the number of species found per site) against indices of richness of all surrogate levels (subfamily, genus, indicator taxa and mixed level), to estimate whether surrogate richness can predict species richness. The significance level (alpha) for multiple tests of the same hypothesis was adjusted with the Bonferroni correction.

Surrogate resolution and composition

We used relative-frequency data (i.e. number of traps in a plot in which a species was collected) to avoid giving more weight to species that have larger nests when constructing the taxa-composition matrix (Gotelli *et al.*, 2011). We used the Mantel test with the Bray–Curtis distance to test the correlation between site-dissimilarity matrices calculated with species resolution, and all other resolutions (Mantel, 1967). Mantel correlations were used to assess whether coarser taxonomic resolution (subfamily, genus, indicator taxa and mixed level) would change the structure within the dissimilarity matrices (i.e. pairwise similarity). The statistical significance of the Mantel tests was estimated based on 1000 permutations. Differences in how much each surrogate matrix predicted the species matrix across sites, as measured by Mantel correlations, were tested for by analysis of variance (ANOVA), followed by the Tukey multiple-range test.

Ecological analysis

We evaluated if datasets with coarser taxonomic resolutions showed similar responses to environmental variables, to those of species-resolution datasets. The main focus of the analysis with environmental variables was not to detect how species and potential surrogates are distributed along environmental gradients. Therefore, we were not interested in higher or significant values, but rather in whether potential surrogates were able to reproduce the responses identified at the species level. Data on percentage of clay content, slope of terrain and altitude were used as predictor variables in the analyses. The environmental datasets are available in the PPBio web site (http://ppbio.in pa.gov.br/), as are details of the sampling protocols (metadata) for each variable. Altitude, slope and soil clay contents were selected because they were sampled using the same protocol at all 13 sites, and previous studies have shown them to be associated with variations in ant-assemblage composition (Vasconcelos *et al.*, 2003; Oliveira *et al.*, 2009).

We used redundancy analysis (RDA) to estimate how much of the variation in the response data (species-level or surrogate matrices) is explained by the environmental variables. RDA is a direct extension of multiple regression analysis to model multivariate response data (Borcard *et al.*, 2011). Species- and surrogate-frequency matrices were transformed using the Hellinger transformation, as recommended for this kind of analysis (Legendre & Gallagher, 2001). The statistical significance of RDA models was based on 1000 permutations. The variance explained by the environmental variables was estimated by the adjusted R^2 (adjR^2) of the RDA models. To evaluate the capacity of the surrogate candidates in predicting the ecological patterns detected with species, we regressed the adjR^2 values obtained for each surrogate resolution, within each sampling site, on the adjR^2 obtained with species-level identification for the respective sampling site.

We also calculated the adjR^2 of each environmental predictor to assess the relationship between the variance explained by each predictor when using species data or surrogate data. To evaluate the capacity of the candidate surrogates to predict the ecological patterns detected with species, we compared the congruence between the responses of the surrogate matrices (significance of explained variance of the partial RDA) and those obtained with species identification.

As the significance of predictions of ecological patterns may change between surrogate levels, we evaluated the capacity of the surrogates to identify ecological patterns revealed by the species-level analyses. By 'capacity to identify patterns', we mean that an environmental effect, whether statistically significant or not, detected in one analysis with species level, can still be detected in a subsequent analysis based on a coarser taxonomic resolution. Similarly, the individual contribution of each environmental variable to the model may change between the surrogate levels and reduce the capacity of the surrogates to reproduce the conclusions based on species. The ability to identify ecological patterns was determined by comparing the change in order of importance of the predictor variables in the RDA models (surrogate and species matrixes). The percentage maintenance of conclusions about ecological patterns (relationships with clay and slope) was estimated from the number of times that each surrogate level captured the pattern identified with species-level data.

Time and monetary cost

Ant sampling performs particularly well, according to the criteria of monetary cost, time spent and ease of surveying (e.g. Majer *et al.*, 2007; Souza *et al.*, 2012). Therefore, the time and monetary costs for the four surrogates were considered in relation to the finest identification level (species), and the fractions of these costs were calculated for each surrogate resolution. The costs were based on expenses for field and laboratory material, field sampling and staff costs (i.e. salary and scholarships). The laboratory costs included those related to species sorting, mounting, identifying and chemicals for conservation of voucher species. The staff costs were estimated as the time taken during field and laboratory activities. All costs were measured for each site and salary, and scholarships were based on current Brazilian federal government payments. The mean value of currency conversion rates for the Brazilian currency (Real = R$) used to estimate costs in US dollars was 1US$ = R$ 1.94 (average of annual means from 2006 to 2013; source: Brazilian Central Bank). Costs of laboratory equipment and accommodation buildings for field staff were not included, as there is no qualitative change in total cost (Gardner *et al.*, 2008).

Time and monetary costs were estimated to evaluate how much could be saved if a surrogate resolution is used instead of species-level identifications. The total effort spent on identification to species level was determined. This effort represented 100% of the costs of time and money invested. For each surrogate resolution, the costs were calculated for each site, and we calculated the proportion of the total effort used for species processing and identification.

RESULTS

A total of 515 species/morphospecies distributed in 11 subfamilies and 79 genera were identified. The number of taxonomic units identified in mixed-level and indicator taxa was 261 and 187, respectively (Table S2).

Richness estimates for all surrogate resolutions were able to predict site variation in species richness, with the r^2 ranging from 0.45 to 0.81 (Fig. 1). Genus was rated as an 'excellent' surrogate, and the mixed-level approach was rated as a 'good' surrogate for species richness. The lowest values were detected for indicator taxa and subfamily, which did not reach the minimum reasonable values of species richness prediction.

The Mantel correlation coefficients were statistically different (ANOVA, $P < 0.001$). In the pairwise test, differences were detected between mixed level and subfamily (Tukey test, $P < 0.001$), indicator taxa and mixed -level (Tukey test, $P < 0.05$), and genus and subfamily (Tukey test, $P < 0.05$). The Mantel correlation coefficients were significant in almost all comparisons between the potential surrogate resolutions and the species-level matrix. The number of non-significant values of Mantel correlation coefficients increases inversely with the quality of prediction of potential surrogates (Table 2). The highest correlations with species-level determinations were for the mixed-level approach (mean = 0.89), which was rated as an 'excellent' surrogate for species composition. Genus (mean = 0.76) was rated as a 'good' surrogate for species composition, followed by indicator taxa (mean = 0.69), which was rated as a 'reasonable' surrogate for species composition. The lowest correlation was for subfamily (mean = 0.55), which did not reach the minimum reasonable value for a species-composition prediction. In nine of the 13 Mantel correlations, the coefficients obtained for genus were higher than those found for indicator taxa. In two of the 13 Mantel correlations, the coefficients obtained for genus were higher than those found for the mixed-level approach. Subfamily had higher coefficients than indicator taxa in five of the 13 Mantel correlations (Table 2).

Redundancy analysis indicated that the ant-assemblage composition was significantly related to environmental variables within sites, and all potential surrogate resolutions were able to predict the ecological pattern demonstrated by species reasonably well, with r^2s ranging from 0.97 to 0.84 being rated as 'excellent' surrogates (Fig. 2). In general, significant RDA values were restricted to sites with a larger sample area and number of plots; usually at these sites, environmental variables tend to have a wider range. Conversely, non-significant RDA values were restricted to the sites with a small sample area, which had less environmental variation (Table S3).

The ecological patterns captured with species-level data were also captured for genus in all of the comparisons. The capacity of this surrogate resolution to identify the ecological pattern captured with species level was high (100%), but dropped to 77% for the mixed-level approach and to 54% for indicator taxa and subfamily (Table S3).

The efficacy of surrogates in capturing the ecological patterns detected with species was moderate to high, with the r^2s ranging from 0.35 to 0.99 (Fig. 3). The best model fit obtained using altitude as the predictor was found for the mixed-level approach (0.99), followed by genus (0.89), subfamily (0.86) and indicator taxa (0.85), that is all four ranked as 'excellent' surrogates. Using clay content as the predictor, the highest value was detected for the mixed-level

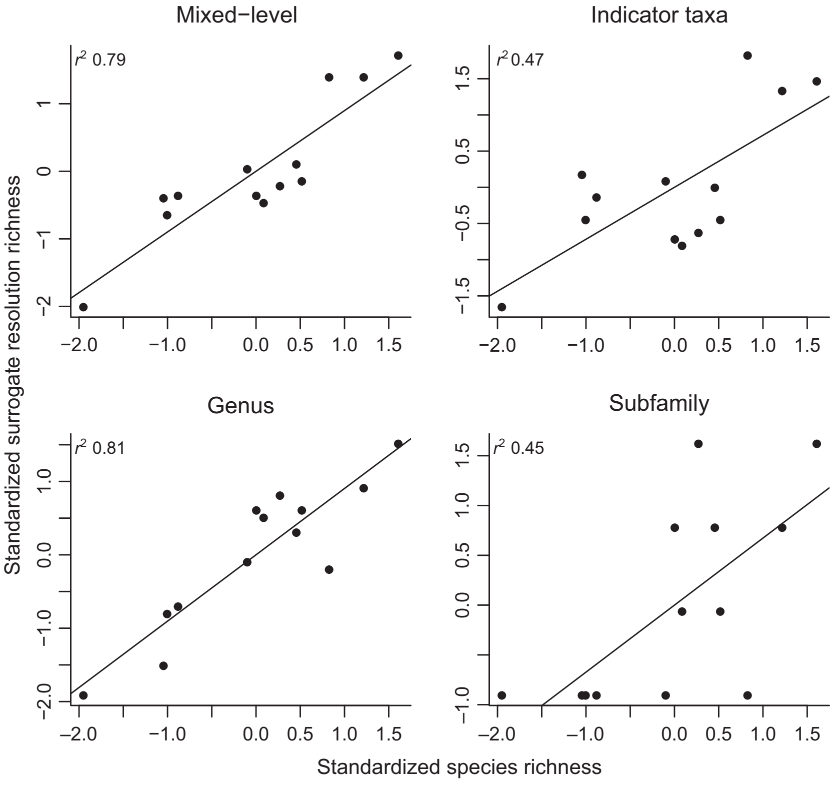

Figure 1 Regression between overall species richness and surrogate richness for the four levels of potential taxonomic surrogate resolutions used for 13 study sites. The statistical significance of models was tested using 1000 permutations. All regressions were significant at $P < 0.001$.

Table 2 Values of Mantel correlation coefficients for comparisons between matrices of species and all potential surrogate levels, for 13 sites.

Sites	Sample Techniques	Mixed-level	Indicator taxa	Genus	Subfamily
Maracá	Pitfall, Winkler and Bait	0.9342***	0.7240***	0.8779***	0.7793***
Cauamé	Pitfall and Bait	0.9173***	0.7354***	0.8195***	0.7975***
Viruá	Pitfall, Winkler and Bait	0.9662***	0.8635***	0.9014***	0.8211***
UFAM	Pitfall	0.9431***	0.8397***	0.7133***	0.4748***
Ducke	Pitfall, Winkler and Bait	0.9182***	0.7819***	0.6351***	0.3132***
Manaquiri	Pitfall and Winkler	0.8689***	0.6795***	0.6637***	0.1946
Orquestra	Pitfall	0.9183**	0.4384	0.6706**	0.0980
Campanã	Pitfall	0.5809	0.6676*	0.7576*	0.5491
Jari	Pitfall	0.9133*	0.6548*	0.3938	0.3382
Teotônio	Winkler	0.8985**	0.5483	0.7926**	0.4653
Ilha Búfalos	Winkler	0.9610**	0.4959	0.9094***	0.6612**
Ilha Pedras	Winkler	0.9905***	0.9585**	0.9902**	0.9842**
Jaci-Paraná	Winkler	0.8045***	0.5569*	0.6633*	0.6332**
Mean of Mantel coefficients		0.8935	0.7604	0.7558	0.5548

Significance levels: $^*P \leq 0.05$; $^{**}P \leq 0.01$; $^{***}P \leq 0.001$.

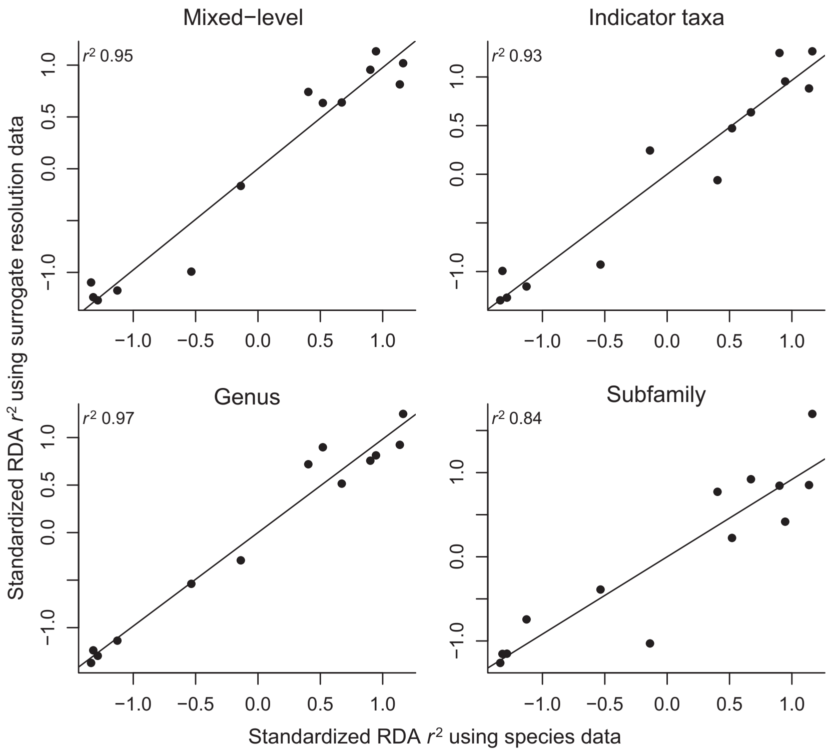

Figure 2 Regression between standardized r^2 of RDA of species and all potential taxonomic surrogate resolutions for 13 study sites. The statistical significance of models was tested using 1000 permutations. All regressions were significant at $P < 0.001$.

approach (0.97), followed by genus (0.89), ranking both as 'excellent' surrogates. The lowest values were for indicator taxa (0.52) and subfamily (0.35), which did not reach the minimum reasonable values. Using slope of terrain as the predictor, high values were detected for the mixed-level approach (0.98), genus (0.95) and indicator taxa (0.84), ranking the three as 'excellent' surrogates. The lowest value was for subfamily (0.65), ranking this surrogate as 'reasonable'. Similarly to the RDA values, the values of significant partial RDAs were also restricted to sites with larger areas and numbers of plots sampled, and consequently the non-significant values were mostly for sites with smaller areas and fewer plots sampled (Table S4).

Estimates of maintenance of the magnitudes of the effects of altitude, clay content and slope varied among sites (Table S4). Only genus met the eligibility criteria, being reasonably effective (~62%) in maintaining the order of importance of the variables and had 14 inversions in the importance in 39 comparisons (i.e. the variable that contributed most to the model of species was in the second or

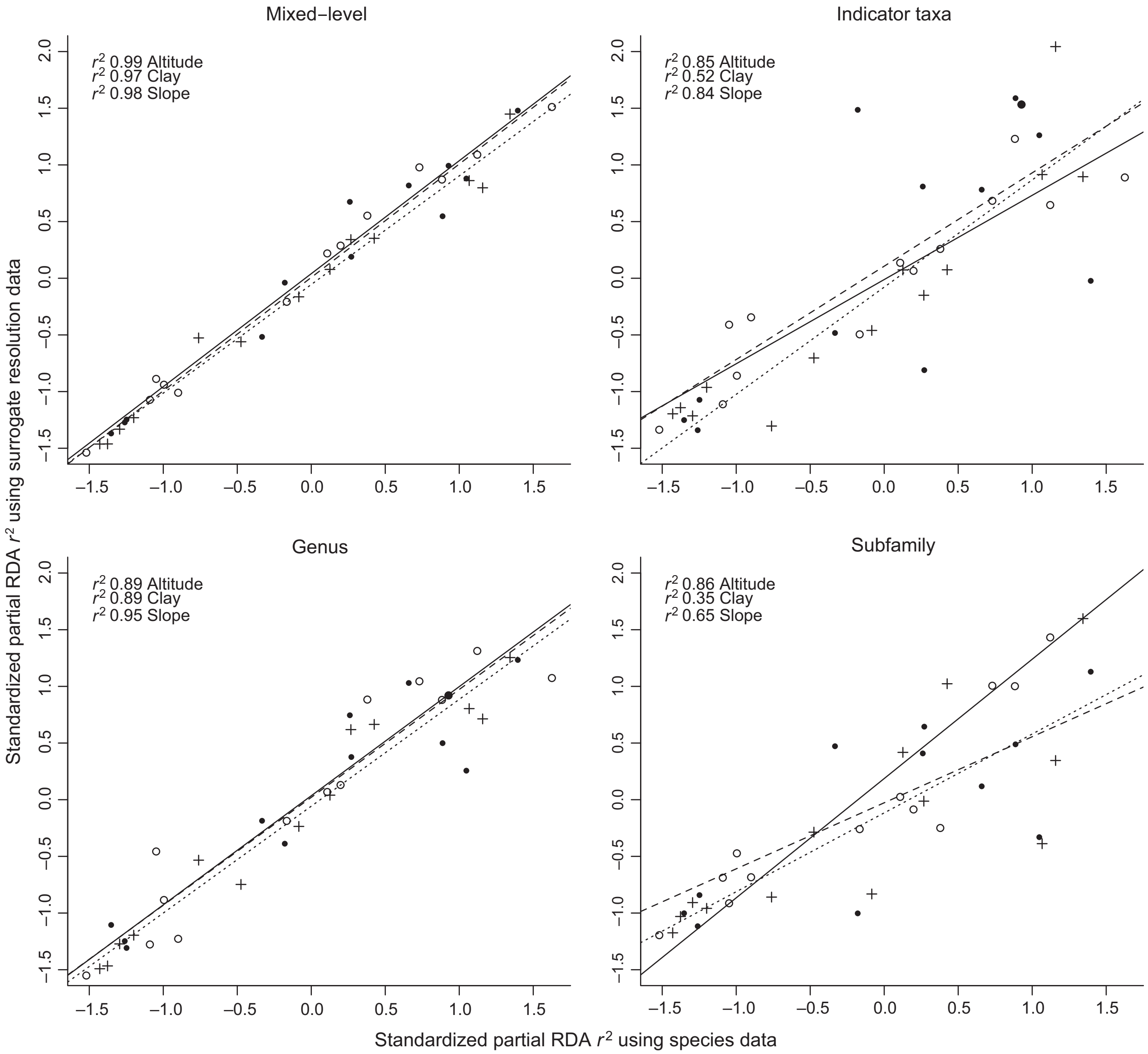

Figure 3 Regressions of standardized partial r^2 of RDA of species and all potential taxonomic surrogate resolutions for 13 study sites. The statistical significance of models was tested using 1000 permutations. All regressions were significant at $P < 0.001$. Open symbols and continuous trend line refer to the altitude, closed symbols and dashed trend line refer to soil clay content, and cross-symbols and pointed trend line refer to slope of terrain.

third position in the model for mixed level). Nineteen inversions were recorded for subfamily, reducing the capacity to maintain the order of importance of the variables to 51.3%. Twenty-one inversions in the importance of the variables with the mixed-level approach and indicator taxa were detected, resulting in retention of the order of importance of the variables in 46% of the tests for these surrogates (Table S4). Among the environmental variables used, the clay content most influenced the distribution of ant species. It was the first variable in order of importance in the model tested in six of 13 locations. The second most important variable for the distribution of species of ground-dwelling ants was the altitude, followed by the slope of the terrain. Similarly, using genus level, the variable that best determined the distribution of ants was the clay content, followed by the altitude. The slope of the terrain was not the most important variable for any of the 13 sites (Table S4).

The surrogate resolutions differed in their relative costs (Table 3, Table S5). The mixed-level approach was the surrogate combination that required the most effort, on average accounting for ~82% of the cost for species-level sorting for the 13 sites. This was followed by the indicator taxa and genus level, which required averages of ~74% and ~62% of the total costs, respectively. Sorting ants to subfamily was cheaper, reducing the total costs to ~28% of that of species-level. Similarly to the monetary cost, the potential surrogate that required the most time to obtain the final identification was mixed-level (~79%), followed by indicator taxa (~72%),

Table 3 Summary of the relative effort (cost and time) required for each surrogate resolution for 13 sites in the Brazilian Amazon.

	Mixed-level		Indicator taxa		Genus		Subfamily	
Sites	Cost (%)	Time (%)	Cost (%)	Time (%)	Cost (%)	Time (%)	Cost (%)	Time (%)
Maracá	84	81	80	76	63	54	31	13
Cauamé	85	82	79	74	67	59	33	17
Viruá	84	81	80	76	63	54	31	13
UFAM	80	78	72	69	62	57	28	18
Ducke	84	81	80	76	63	54	31	13
Manaquiri	81	78	72	69	61	57	27	18
Orquestra	81	78	72	69	61	57	27	18
Campanã	81	79	72	69	61	57	27	18
Jari	81	78	72	69	61	57	27	18
Teotônio	80	78	72	71	62	58	28	20
Ilha Búfalos	80	78	72	71	62	58	28	20
Ilha Pedras	82	78	71	71	62	58	28	20
Jaci-Paraná	81	78	72	71	62	58	28	20
Mean	81.8	79.1	74.3	71.6	62.3	56.8	28.8	17.4

genus (~57%) and subfamily. Subfamily was the fastest, reducing the time to ~17% of that needed for species-level sorting.

DISCUSSION

In the present study, four surrogate resolutions for ant species recovered much of the information on assemblage composition (taxonomic approach) and ecological patterns (ecological approach) across the Amazon basin. Indicator taxa and genus had the highest predictive values for species-level data and were the most cost-effective surrogates. Variation in the composition of ground-dwelling-ant assemblages among sampling techniques and sites has been documented in the Amazon previously (Souza *et al.*, 2012). Nevertheless, the relationship between species and surrogates remained strong regardless of the geographic location of the sampling site or the type of vegetation. The use of a coarser taxonomic resolution for terrestrial species generally reduces time and monetary costs (Williams & Gaston, 1994; Grimbacher *et al.*, 2008), especially if non-specialists are involved.

The number of studies on the effectiveness of surrogates has grown in recent decades, for diverse environments and scales (Andersen, 1995; Ricketts *et al.*, 2002; Sætersdal *et al.*, 2005; Bhusal *et al.*, 2014), and these studies have produced varied and often inconsistent results (Lawler & White, 2008; Neeson *et al.*, 2013). Andersen (1995) indicated that idiosyncrasies in results might be due to differences in the number of species within the genera. As studies differ in spatial scale, in the methods used to test surrogate levels, and in the groups tested, it is difficult to generate uniform conclusions or derive general trends (Lawler & White, 2008). To reduce these problems, we can standardize sampling methods to better describe diversity distributions at large scales (Andersen, 1999; Reyers *et al.*, 2000). However, there is no consensus on the surrogate group or taxonomic resolution to be used, because studies at large scales are relatively few or were carried out in small numbers of vegetation associations. Our sampling design covered a latitudinal gradient of approximately 14 degrees in the Amazon basin, and gave consistent results. Regardless of sampling area, geographic site or type of vegetation, we found a high correlation between surrogate matrices and the matrix of species (taxonomic approach). Furthermore, our results demonstrated high congruency in RDA, which reflects ecological patterns shared among groups of potential surrogates and data based on sorting to species level (ecological approach). The significant results of the environmental analyses were restricted to sites with larger areas covered and more plots sampled, whereas non-significant values were associated with sites with smaller areas and fewer plots. Souza *et al.* (2009) obtained similar results in a study of ants, using both genus- and species-level taxonomic identification in the Amazon.

Following the criteria suggested by Leal *et al.* (2010), we concluded that the mixed-level approach and genus were well able to predict (i.e. to predict > 70% of the variation in species data) the richness, composition, ecological patterns and quality of the ecological patterns detected with species. Indicator taxa and subfamily did not reach the minimum reasonable values of species prediction (i.e. < 60% of the variation in species data) for the taxonomic or ecological approaches. The cost to identify ants to genus level was lower than the cost to use the mixed-level approach. The mixed-level approach proposed by Groc *et al.* (2010) also had a total cost very similar to species-level identification, reducing the cost by only 18%. The lowest costs were achieved using subfamily, which represented only 29% of the cost for species, but this surrogate did not perform well (≥ 60% congruence) in predicting species associations.

The best surrogate resolution to maintain the ecological patterns observed with the species-level data was genus, which maintained the ecological pattern in all cases. Genus was the best surrogate in maintaining the quality of the ecological pattern detected with species in 24 of 39 comparisons. Therefore, genus was the most efficient and reliable of the four surrogate resolutions tested in this study, and met the criteria suggested by Leal *et al.* (2010), predicting at least 76% of the patterns detected with the species, reducing costs by about 40%, maintaining the ecological patterns at all sites, and retaining the ranking of the ecological variables in 62% of the cases tested.

Previous studies have found inconsistent results between regions (Andersen, 1997; Heino, 2014), and as sampling intensity and spatial extent increase, the predictive power of surrogates may weaken (Cardoso *et al.*, 2004). However, our results were surprisingly consistent, considering the wide latitudinal and environmental coverage (from open savannas to dense forests). We did not detect changes or weakening of the predictive power of the surrogate resolutions along the latitudinal gradient studied, indicating that these relationships between surrogate resolutions and species persist among various vegetation types in the Amazon basin. Our results echo the findings of other studies that genus level was the most effective surrogate resolution to capture the patterns detected with species-level data (Pik *et al.*, 1999; Heino & Soininen, 2007; Gallego *et al.*, 2012).

The use of genera as taxonomic surrogates for species is helpful to answer many questions, especially in situations where the available resources are limited, such as in biodiversity monitoring (Kallimanis *et al.*, 2012). One of the main benefits of the approach proposed here is the saving of time and money, since the genus has a high predictive value for species-level data. This allows inclusion of new sampling sites or additional replicates in existing sampling sites, increasing the number of species encountered, and the statistical power and the generality of the results (Costa & Magnusson, 2010).

Identification to genus will be sufficient to predict species-level relationships in studies of ant community ecology and environmental monitoring programs. Also, specimens deposited in collections will have a much higher chance of eventually being identified to species level by experts if collection lots are already sorted to genus level. In biodiversity monitoring, the aim is to survey a given study area several times over many years, which is much more expensive than a fine-scale taxonomic inventory. This is an example of the practical limitations and restrictions of time and money in biodiversity assessment, which can be overcome using a surrogate approach. The cost (time and money) of environmental monitoring is a crucial factor for its viability (Margules & Pressey, 2000), and any reliable alternative should be used to increase the area covered or the length of the time series. In 2010, approximately US$ 0.01 per hectare was invested in monitoring actions in the Brazilian Amazon, and increases in this investment are not expected in the near future (Magnusson *et al.*, 2013). Therefore, we propose that the use of higher taxonomic surrogates can be applied to large-scale surveys of ants in the Amazon, despite the wide environmental heterogeneity, and that this taxonomic resolution should be tested in other megadiverse regions.

ACKNOWLEDGEMENTS

We thank Everaldo Pereira, Juliana Araújo, Pollyana Cavalcante, Camila Gomes, Marcos Torres, Carlos Nogueira, Márcia Cidade and Claudio Santos-Neto for their help in sampling ants. Fernando Fernández, Jacques Delabie, John Longino, José Vilhena and Rodrigo Feitosa confirmed the species identifications for this study. The concessionaire responsible for building and operating the Santo Antônio Hydroelectric Plant – SAE provided financial and logistical support. Financial support was provided by PIPT/FAPEAM 1750/08; PNPD/CAPES 03017/19-05; FAPEAM 062.01325/2014; CNPq via PRONEX 16/2006, the Program for Biodiversity Research (PPBio) 558318/2009-6, 457545/2012-7; the Long Term Ecological Research (PELD) 403764/2012-2 and the Center for Integrated Studies of the Amazonian Biodiversity (CENBAM). J.L.P.S. was supported by CNPq and FAPEAM post-doctoral scholarship, I.O.F. is supported by CNPq doctoral scholarship and P.A.C.L.P. was supported by CAPES doctoral scholarship. The authors are grateful to Alan Andersen, Jonathan Majer and all anonymous referees for their comments on the earlier versions of the manuscript.

REFERENCES

Agosti, D., Majer, J.D., Alonso, L. & Schultz, T.R. (2000) *Ants: standard methods for measuring and monitoring biodiversity*. Smithsonian Institution Press, Washington DC.

Andersen, A.N. (1995) Measuring more of biodiversity: Genus richness as a surrogate for species in Australian ant faunas. *Biological Conservation*, **73**, 39–43.

Andersen, A.N. (1997) Using ants as bioindicators: multiscale issues in ant community ecology. *Ecology and Society*, **1**, 1–8.

Andersen, A.N. (1999) My bioindicator or yours? Making the selection. *Journal of Insect Conservation*, **3**, 61–64.

Bhusal, D.R., Kallimanis, A.S., Tsiafouli, M.A. & Sgardelis, S.P. (2014) Higher taxa vs. functional guilds vs. trophic groups as indicators of soil nematode diversity and community structure. *Ecological Indicators*, **41**, 25–29.

Bolton, B., Alpert, G., Ward, P.S. & Naskrecki, P. (2005) *Bolton's catalogue of ants of the world: 1758–2005*. Harvard University Press, Cambridge, MA.

Borcard, D., Gillet, F., Legendre, P., Gentleman, R., Hornik, K. & Parmigiani, G.G. (2011) *Numerical ecology with R*. Springer, New York, New York, NY.

Cardoso, P., Silva, I., Oliveira, N.G. & Serrano, A.R.M. (2004) Higher taxa surrogates of spider (Araneae) diversity and their efficiency in conservation. *Biological Conservation*, **117**, 453–459.

Costa, F.R.C. & Magnusson, W.E. (2010) The need for large-scale, integrated studies of biodiversity – The experience of the program for biodiversity research in Brazilian Amazonia. *Natureza & Conservação*, **08**, 3–12.

Ellis, D. (1985) Taxonomic sufficiency in pollution assessment. *Marine Pollution Bulletin*, **16**, 459.

Fittkau, E.J. & Klinge, H. (1973) On biomass and trophic structure of the central Amazonian rain forest ecosystem. *Biotropica*, **5**, 2–14.

Franklin, E., de Moraes, J., Landeiro, V.L., Souza, J.L.P., de Pequeno, P.A.C.L., Magnusson, W.E. & de Morais, J.W. (2013) Geographic position of sample grid and removal of uncommon species affect multivariate analyses of diverse assemblages: the case of oribatid mites (Acari: Oribatida). *Ecological Indicators*, **34**, 172–180.

Gallego, I., Davidson, T.A., Jeppesen, E., Pérez-Martínez, C., Sánchez-Castillo, P., Juan, M., Fuentes-Rodríguez, F., León, D., Peñalver, P., Toja, J. & Casas, J.J. (2012) Taxonomic or ecological approaches? Searching for phytoplankton surrogates in the determination of richness and assemblage composition in ponds. *Ecological Indicators*, **18**, 575–585.

Gardner, T.A., Barlow, J., Araujo, I.S. *et al.* (2008) The cost-effectiveness of biodiversity surveys in tropical forests. *Ecology Letters*, **11**, 139–150.

Gaspar, C., Gaston, K.J. & Borges, P.A.V. (2010) Arthropods as surrogates of diversity at different spatial scales. *Biological Conservation*, **143**, 1287–1294.

Gotelli, N.J.G., Ellison, A.M., Dunn, R.R. & Sanders, N. (2011) Counting ants (Hymenoptera: Formicidae): biodiversity sampling and statistical analysis for myrmecologists. *Myrmecological News*, **15**, 13–19.

Grimbacher, P.S., Catterall, C.P. & Kitching, R.L. (2008) Detecting the effects of environmental change above the species level with beetles in a fragmented tropical rainforest landscape. *Ecological Entomology*, **33**, 66–79.

Groc, S., Delabie, J.H.C., Longino, J.T., Orivel, J., Majer, J.D., Vasconcelos, H.L. & Dejean, A. (2010) A new method based on taxonomic sufficiency to simplify studies on Neotropical ant assemblages. *Biological Conservation*, **143**, 2832–2839.

Heino, J. (2014) Taxonomic surrogacy, numerical resolution and responses of stream macroinvertebrate communities to ecological gradients: are the inferences transferable among regions? *Ecological Indicators*, **36**, 186–194.

Heino, J. & Soininen, J. (2007) Are higher taxa adequate surrogates for species-level assemblage patterns and species richness in stream organisms? *Biological Conservation*, **137**, 78–89.

Hodnett, M.G., Vendrame, I., Marques Filho, A.O., Oyama, M.D. & Tomasella, J. (1997) Soil water storage and groundwater behaviour in a catenary sequence beneath forest in central Amazonia. II. Floodplain water table behaviour and implications for streamflow generation. *Hydrology and Earth System Sciences*, **1**, 279–290.

Hoffmann, B.D. & Andersen, A.N. (2003) Responses of ants to disturbance in Australia, with particular reference to functional groups. *Austral Ecology*, **28**, 444–464.

Kallimanis, A.S., Mazaris, A.D., Tsakanikas, D., Dimopoulos, P., Pantis, J.D. & Sgardelis, S.P. (2012) Efficient biodiversity monitoring: which taxonomic level to study? *Ecological Indicators*, **15**, 100–104.

Kaspari, M. & Weiser, M.D. (2000) Ant activity along moisture gradients in a Neotropical forest. *Biotropica*, **32**, 703–711.

King, J.R., Warren, R.J. & Bradford, M.A. (2013) Social insects dominate eastern US temperate hardwood forest macroinvertebrate communities in warmer regions. *PLoS ONE*, **8**, e75843.

Landeiro, V.L., Bini, L.M., Costa, F.R.C., Franklin, E., Nogueira, A., Souza, J.L.P., Moraes, J. & Magnusson, W.E. (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. *Ecological Indicators*, **23**, 366–373.

Laurance, W.F., Useche, D.C., Rendeiro, J. *et al.* (2012) Averting biodiversity collapse in tropical forest protected areas. *Nature*, **489**, 290–294.

Lawler, J.J. & White, D. (2008) Assessing the mechanisms behind successful surrogates for biodiversity in conservation planning. *Animal Conservation*, **11**, 270–280.

Leal, I.R., Bieber, A.G.D., Tabarelli, M. & Andersen, A.N. (2010) Biodiversity surrogacy: indicator taxa as predictors of total species richness in Brazilian Atlantic forest and Caatinga. *Biodiversity and Conservation*, **19**, 3347–3360.

Legendre, P. & Gallagher, E. (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia*, **129**, 271–280.

Lopes, C.T. & Vasconcelos, H.L. (2008) Evaluation of three methods for sampling ground-dwelling ants in the Brazilian cerrado. *Neotropical Entomology*, **37**, 399–405.

Magnusson, W.E., Lima, A.P., Luizão, R., Luizão, F., Costa, F.R.C., Castilho, C.V. & Kinupp, V.F. (2005) Rapeld: a modification of the Gentry method for biodiversity surveys in long-term ecological research sites. *Biota Neotropica*, **5**, 1–6.

Magnusson, W., Braga-Neto, R., Pezzini, F. *et al.* (2013) *Biodiversidade e monitoramento ambiental integrado Biodiversity and integrated environmental monitoring*. Áttema Editorial, Manaus.

Majer, J.D., Orabi, G. & Bisevac, L. (2007) Ants (Hymenoptera: Formicidae) pass the bioindicator scorecard. *Myrmecological News*, **10**, 69–76.

Mantel, M. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

Margules, C.R. & Pressey, R.L. (2000) Systematic conservation planning. *Nature*, **405**, 243–253.

Margules, C.R., Pressey, R.L. & Williams, P.H. (2002) Representing biodiversity: data and procedures for identifying priority areas for conservation. *Journal of Biosciences*, **27**, 309–326.

McGeoch, M.A. (1998) The selection, testing and application of terrestrial insects as bioindicators. *Biological Reviews*, **73**, 181–201.

Mezger, D. & Pfeiffer, M. (2011) Partitioning the impact of abiotic factors and spatial patterns on species richness and community structure of ground ant assemblages in four Bornean rainforests. *Ecography*, **34**, 39–48.

Moreno, C.E., Guevara, R., Sánchez-Rojas, G., Téllez, D. & Verdú, J.R. (2008) Community level patterns in diverse systems: a case study of litter fauna in a Mexican pine-oak forest using higher taxa surrogates and re-sampling methods. *Acta Oecologica*, **33**, 73–84.

Neeson, T.M., Van Rijn, I. & Mandelik, Y. (2013) How taxonomic diversity, community structure, and sample size determine the reliability of higher taxon surrogates. *Ecological Applications*, **23**, 1216–1225.

Oliveira, P.Y., Souza, J.L.P., Baccaro, F.B. & Franklin, E. (2009) Ant species distribution along a topographic gradient in a "terrafirme" forest reserve in central Amazonia. *Pesquisa Agropecuária Brasileira*, **44**, 852–860.

Oliver, I. & Beattie, A.J. (1993) A possible method for the rapid assessment of biodiversity. *Conservation Biology*, **7**, 562–568.

Oliver, I. & Beattie, A.J. (1996a) Designing a cost-effective invertebrate survey: a test of methods for rapid assessment of biodiversity. *Ecological Applications*, **6**, 594–607.

Oliver, I. & Beattie, A.J. (1996b) Invertebrate morphospecies as surrogates for species: a case study. *Conservation Biology*, **10**, 99–109.

Olsgard, F., Brattegard, T. & Holthe, T. (2003) Polychaetes as surrogates for marine biodiversity: lower taxonomic resolution and indicator groups. *Biodiversity and Conservation*, **12**, 1033–1049.

Pik, A.J., Oliver, I. & Beattie, A.J. (1999) Taxonomic sufficiency in ecological studies of terrestrial invertebrates. *Australian Journal of Ecology*, **24**, 555–562.

R Development Core Team. (2014) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Reyers, B., Van Jaarsveld, A.S. & Krüger, M. (2000) Complementarity as a biodiversity indicator strategy. *Proceedings of the Royal Society B: Biological Sciences*, **267**, 505–513.

Ricketts, T.H., Daily, G.C. & Ehrlich, P.R. (2002) Does butterfly diversity predict moth diversity? Testing a popular indicator taxon at local scales. *Biological Conservation*, **103**, 361–370.

Sætersdal, M., Gjerde, I. & Blom, H.H. (2005) Indicator species and the problem of spatial inconsistency in nestedness patterns. *Biological Conservation*, **122**, 305–316.

Santos, E.M., Franklin, E. & Magnusson, W.E. (2008) Cost-efficiency of subsampling protocols to evaluate oribatid-mite communities in an Amazonian savanna. *Biotropica*, **40**, 728–735.

Souza, J.L.P., Baccaro, F.B., Landeiro, V.L., Franklin, E. & Magnusson, W.E. (2012) Trade-offs between complementarity and redundancy in the use of different sampling techniques for ground-dwelling ant assemblages. *Applied Soil Ecology*, **56**, 63–73.

Souza, J.L.P., Moura, C.A.R. & Franklin, E. (2009) Efficiency in inventories of ants in a forest reserve in Central Amazonia. *Revista Agropecuária Brasileira*, **44**, 940–948.

Vasconcelos, H.L., Macedo, A.C.C. & Vilhena, J.M.S. (2003) Influence of topography on the distribution of ground-dwelling ants in an Amazonian forest. *Studies on Neotropical Fauna and Environment*, **38**, 115–124.

Williams, P.H. & Gaston, K.H.J. (1994) Measuring more of diversity: can higher-taxon richness predict wholesale species richness? *Biological Conservation*, **67**, 211–217.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1 Site locations on Amazon basin.

Table S1 Ground-dwelling ant species sampled at 13 sites across the Brazilian Amazon.

Table S2 Richness values of species and surrogate resolution in 13 sampling sites.

Table S3 Proportion of variance in the ant-assemblage composition jointly explained by the environmental variables in Redundancy Analysis (RDA) for models within each site sampled in the Brazilian Amazon.

Table S4 Proportion of variance explained by each environmental variable in Redundancy Analysis (RDA) models within each site in the Brazilian Amazon.

Table S5 Absolute and relative values of time and monetary cost for species and surrogate levels at each site.

BIOSKETCH

Jorge Luiz Pereira Souza is PhD in Entomology by National Institute for Amazonian Research-INPA, Manaus, Brazil. Interested in: Community ecology, monitoring of biodiversity, spatial and temporal analysis with ants. Use of surrogates for species and bio-design for product and patent with focus on ants. He is postdoctoral fellow, member of Program for Biodiversity Research (PPBio) and the Center for Integrated Studies of the Amazonian Biodiversity (CENBAM).

Authors contribution: JLPS, FBB, VLL and PACLP performed statistical analyses; JLPS, FBB and IOF collected data; JLPS, EF and WEM were responsible for the sampling design; JLPS wrote the first draft of the manuscript; and all authors contributed substantially to revisions.

Editor: Alan Andersen