Evaluating sampling sufficiency and the use of surrogates for assessing ant diversity in a Neotropical biodiversity hotspot

ARTICLE *in* ECOLOGICAL INDICATORS · JULY 2014

Impact Factor: 3.44 · DOI: 10.1016/j.ecolind.2014.06.036

8 AUTHORS, INCLUDING:

Tiago Luiz Massochini Frizzo

University of Brasília

9 PUBLICATIONS **29** CITATIONS

SEE PROFILE

Gabriela Procópio Camacho

Universidade Federal do Paraná

6 PUBLICATIONS **8** CITATIONS

SEE PROFILE

José Roberto Pujol Luz

University of Brasília

93 PUBLICATIONS **246** CITATIONS

SEE PROFILE

Ecological Indicators 46 (2014) 286–292

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Evaluating sampling sufficiency and the use of surrogates for assessing ant diversity in a Neotropical biodiversity hotspot

Heraldo L. Vasconcelos[a,*], Tiago L.M. Frizzo[b], Renata Pacheco[a], Jonas B. Maravalhas[a], Gabriela P. Camacho[c], Karine S. Carvalho[d], Elmo B.A. Koch[a], José R. Pujol-Luz[e]

[a] Instituto de Biologia, Universidade Federal de Uberlândia (UFU), Av. Pará 1720, 38400-902 Uberlândia, MG, Brazil
[b] Programa de Pós-Graduação em Ecologia, Universidade de Brasília (UnB), 70910-900 Brasília, DF, Brazil
[c] Programa de Pós-Graduação em Entomologia, Universidade Federal do Paraná (UFPR), 81531-980 Curitiba, PR, Brazil
[d] Departamento de Ciências Biológicas, Universidade Estadual do Sudoeste da Bahia (UESB), Jequié, BA, Brazil
[e] Departamento de Zoologia, Universidade de Brasília (UnB), 70910-900 Brasília, DF, Brazil

ARTICLE INFO

Article history:
Received 25 March 2014
Received in revised form 25 June 2014
Accepted 26 June 2014

Keywords:
Beta diversity
Indicator taxa
Formicidae
Higher-taxon surrogacy
Sampling protocols
Vertical stratification

ABSTRACT

Ecologists often seek sampling protocols that are both effective and relatively simple, i.e. those that provide a balance between the advantages obtained through sampling completeness and the costs involved in species sorting. Here we explored ways of simplifying a protocol devised for assessing geographic patterns of ant species richness and composition in the savanna-dominated *Cerrado* landscape of central Brazil. This protocol, which retrieved up to 88.5% of the expected number of species, was employed in five different sites located up to 1200 km apart from each other. In each site three transects (>1 km apart) were established, and within each transect 80 arboreal and 80 ground pitfall traps were installed. We then evaluated the degree of congruence in species richness and composition between data originated from the full data set and various subsets of the data, that either contained a reduced number of samples (in different spatial configurations) or potential surrogates of the entire ant fauna. Our main findings show that by sorting specimens from just a subset of the ant genera we retrieved most of the information existing in the full data set, even though the amount of work involved in ant sorting was reduced by about 50%. Reducing the number of samples taken in each site, as expected, also reduced the amount of work involved in ant sorting. However, in most of our analyses the level of congruence with the full data set was lower and/or became more erratic when sampling intensity was reduced to one-third or less of the original number of samples. More species were retrieved when samples were spread over multiple transects than when they were located in only one or two transects, indicating that even at relatively small scales (a few kilometers) there is a turnover of ant species within the same type of habitat. Based on these results we suggest that at least half of the samples from our original protocol are necessary to adequately describe large-scale patterns of ant species richness and beta diversity in Neotropical savannas. We also suggest that, whenever possible, sampling within a given site should be done in multiple locations in order to maximize the number of species collected.

1. Introduction

One of the main goals of ecology is to document and understand biodiversity patterns (Rosenzweig, 1995; Gaston, 2000). In addition to its relevance for the advance of ecological theory, studies of diversity patterns are essential for the formulation of proper conservation strategies. The savanna-dominated biome of central Brazil, known as *Cerrado*, is a biodiversity hotspot (Myers et al., 2000); yet, less than 3% of its 2 million square kilometers are strictly protected (Klink and Machado, 2005). Several areas of conservation priority for this biome have been identified based on information about the distribution patterns of various taxonomic groups (MMA, 1999). However, many other areas were considered to have insufficient information to evaluate their conservation importance (MMA, 1999), showing the urgent need to conduct field inventories in most of this region.

Insects are often neglected in conservation planning (Schuldt et al., 2009; McGeoch et al., 2011), and this is of concern given that insects are highly abundant and diverse in most terrestrial ecosystems (Mora et al., 2011), where they occupy a variety of niches

* Corresponding author. Tel.: +55 34 32182809.
E-mail address: heraldo@umuarama.ufu.br (H.L. Vasconcelos).

http://dx.doi.org/10.1016/j.ecolind.2014.06.036
1470-160X/

and play diverse ecological roles. Moreover, patterns of diversity of plants and vertebrates are not always coincident with those of insects (Landeiro et al., 2012; Jenkins et al., 2013), and the degree of congruence may depend on the region of study (Hess et al., 2006; Stork and Habel, 2013).

Ants are a key group of insects in many ecosystems, including tropical savannas (Hölldobler and Wilson, 1990). In addition, as Smith et al. (2005) have pointed out "ants often exhibit high rates of spatial turnover (replacement of species) and therefore provide the essential maps for assessing biodiversity at a scale at which conservation decisions are typically made on the ground". However, information about the geographic patterns of Cerrado ant diversity is still scant (but see Silva et al., 2004; Vasconcelos et al., 2008; Pacheco and Vasconcelos, 2012a). Therefore, new inventories over multiple sites are needed, and ideally such inventories should be as exhaustive as possible given that a low sampling effort can result in inflated pseudo-turnover rates (Chao et al., 2005; Cardoso et al., 2009). Nevertheless, a thorough ant inventory produces a massive number of specimens whose sorting is time and labor consuming. Thus, a balance between the advantages obtained through sampling completeness and the costs involved in species sorting must be obtained. In other words, one must seek a sampling protocol that is relatively simple but also effective for capturing spatial variation in community composition.

Recently, Vellend et al. (2008) have suggested that complete inventories are not necessarily required for a comparative, ecological characterization of biological communities. However, this raises the question: to what extent can sampling intensity be minimized, or the process of sorting samples simplified (e.g., Andersen and Majer, 2004), whilst maintaining robust comparative data? To answer this question we conducted an intensive inventory of ant assemblages in five different sites of the Cerrado biome. We then evaluated the degree of congruence in species richness and composition between data originated from the full data set and various subsets of the data, which contained either a reduced number of samples (in different spatial configurations) or potential surrogates of the entire ant fauna. Our aim was to find protocols that could both reduce the costs involved in ant sampling and sorting while retaining as much as possible of the original information.

2. Materials and methods

Ants were collected in five study sites which together covered an area spanning six degrees of latitude and 11 degrees of longitude. The study sites were located near the municipalities of Barreiras (12°08′S, 45°09′W), Mucugê (12°57′S, 41°28′W), Barra do Garças (15°51′S, 52°16′W), Bom Jesus (9°15′S, 44°47′W), and Brasília (15°57′S, 47°55′W). We only collected in relatively well-preserved areas of *cerrado sensu stricto*. The cerrado sensu stricto is the dominant savanna physiognomy of the Cerrado biome and is characterized by a mixture of plants of two distinct strata: a woody layer composed of trees and large shrubs 3–8 m tall, and a ground layer composed of grasses, herbs, and small shrubs (Oliveira-Filho and Ratter, 2002).

Our sampling protocol is similar to the one used in an intercontinental comparison of the savanna ant faunas of Brazil and Australia (Campos et al., 2011). In each site we established three transects distant at least 1 km from each other. Transects were linear, with a total length of approximately 380 m. Along each transect, the nearest tree (>3 m in height) was located at each 20 m interval, so that in each transect 20 trees were selected. Ants were collected using pitfall traps which are considered much more effective for sampling ants in savannas than other methods, including the Winkler method, baits, and subterranean traps (Parr and Chown, 2001; Lopes and Vasconcelos, 2008; Pacheco and Vasconcelos, 2012b). Four pitfall traps were buried in the ground around each selected tree (in a square grid of approximately 2.5 × 2.5 m) and four were fixed on the tree branches with masking tape, with each trap spaced at least 1 m from each other and placed on different branches whenever possible. We considered each set of four ground or four arboreal pitfall traps as a single sample. In this way we avoided recording individuals from the same colony collected in different pitfalls as belonging to different colonies (i.e. of overestimating the absolute frequency of a given species in a given location). A preliminary analysis of the contents of each pitfall trap in separate revealed that a set of four ground pitfalls collect, on average, 2.5 times more species than a single pitfall trap, whereas a set of four arboreal pitfalls collect 2.7 more species.

Pitfall traps consisted of a small plastic cup (250 ml, 8.5 cm high and 7.8 cm in diameter) partially filled with water and detergent. The arboreal traps were baited with human urine (diluted 1:2 in water). We used urine instead of sardines or honey as in the study of Campos et al. (2011) because pitfalls baited with urine were found to collect comparatively more species of arboreal ants than those baited with honey or sardines (with the latter capturing many ants but mostly from dominant species) (T. Frizzo, S. Powell and H. Vasconcelos, unpublished data). Ground pitfall traps were not baited since the disturbance made in the soil to install the traps (also known as the digging in effect) is itself attractive to ants (Greenslade, 1973). Ground and arboreal pitfall traps remained in operation for 48 h.

All ants collected were sorted to morphospecies and, whenever possible, identified to species using available taxonomic keys or through comparison with specimens previously identified by experts and deposited at the Zoological Collection of the Federal University of Uberlândia, in Brazil.

2.1. Data analysis

We evaluated the sampling completeness of our full sampling protocol by calculating the ratio between the number of species recorded (observed) in each site and the number of species expected to occur according to three different species richness estimators: Chao 2, ICE, and Jackknife 1 (Colwell, 2006).

To evaluate the effectiveness of the tree-dwelling and ground-dwelling assemblages as surrogates of the entire ant assemblage, two subsets of the full data set were created; one consisting of all arboreal samples and the other of all ground samples. We also evaluated the possibility of using a subset of the collected species (hereafter "indicator taxa") as surrogate of the entire ant fauna. For this we created a subset of the data consisting of species belonging to genera that we considered easier to sort to species or morphospecies level (Groc et al., 2010). Genera that did not satisfy this criterion included the speciose and taxonomically poorly resolved genera *Pheidole* and *Solenopsis*, as well as *Azteca*, *Brachymyrmex*, *Dorymyrmex*, *Forelius*, *Hypoponera*, *Linepithema*, *Nylanderia*, and most of the Attini (except *Apterostigma* and *Acromyrmex*). We also did not included as indicator taxa the army ants (Ecitoninae), since they have nomadic habits and thus are not properly sampled using conventional techniques. We evaluated the effectiveness of higher-taxon surrogacy by creating a data set based on identification of the specimens to genus level rather than to species or morphospecies level.

To evaluate how much of the original information would be lost by reducing the sampling intensity in each site, 34 subsets of the full data set were created (Fig. 1). Five sampling intensities, representing one-tenth, one-quarter, one-third, half, or two-thirds of the total number of samples taken in each site were evaluated. For each sampling intensity, two categories of data sets were created: one in which samples were spread over all the transects established in each site (hereafter widespread spatial configuration) and the

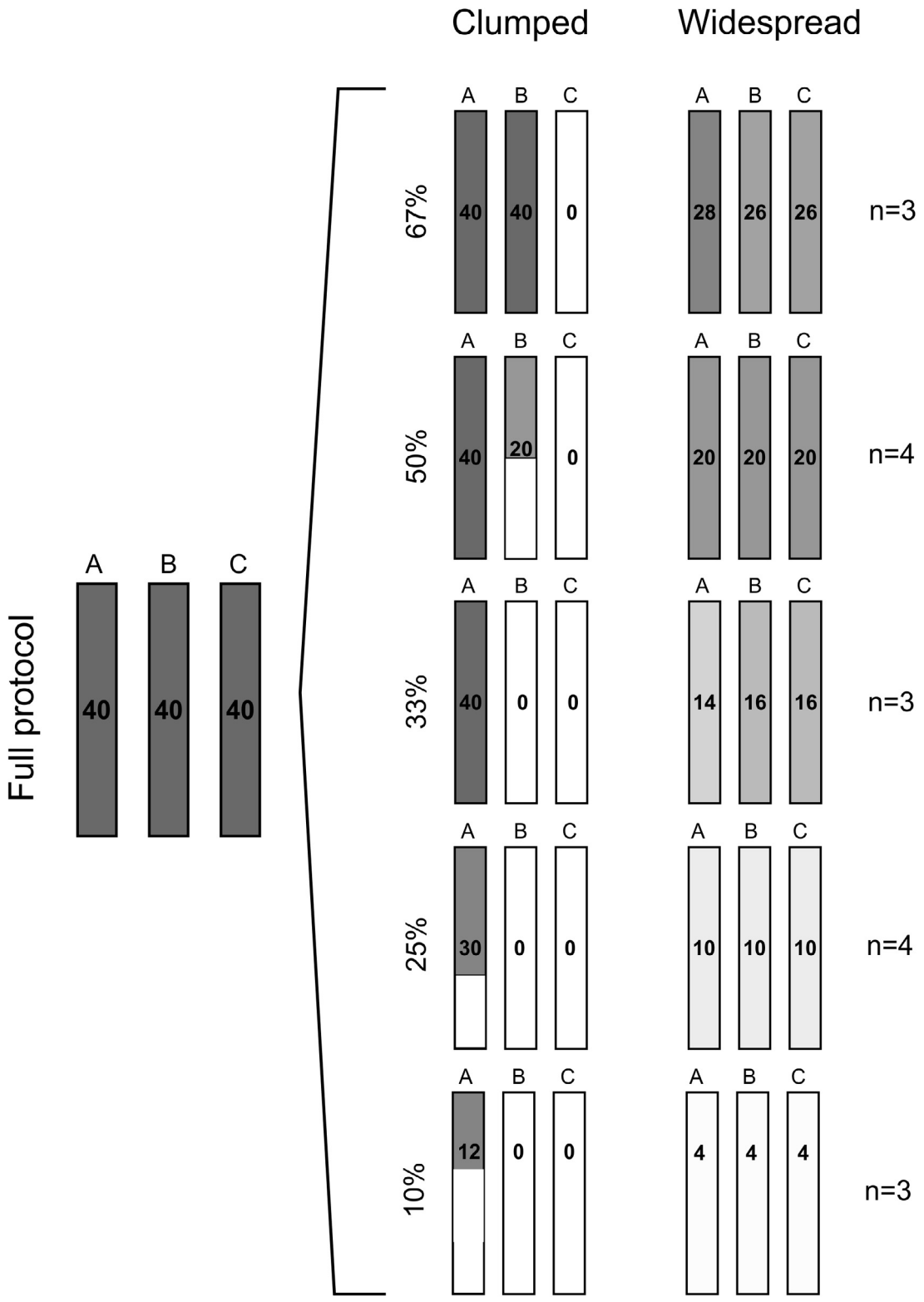

Fig. 1. Schematic view of the sub-sampling protocols. The full sampling protocol consisted of 120 samples (60 on ground and 60 in the arboreal vegetation) distributed in three transects (here denoted A, B, and C) located >1 km apart from each other. Sub-sampling protocols created to evaluate the influence of sampling intensity consisted of gathering data from two-thirds, half, one-third, one-quarter, or one-tenth of the samples. Each subset of the samples was either located in just one or two transects (clumped spatial configuration) or spread over all three transects established in each sampling site. Numbers inside bars represent the number of samples in each transect. In the widespread spatial configuration samples were spaced as evenly as possible along the entire transect, whereas in the clumped configuration samples were spatially aggregated along the transect. *N* represents the number of data sets generated for each sampling intensity and spatial configuration of the samples.

other in which all samples were located in the minimum possible number of transects (hereafter clumped spatial configuration; Fig. 1). Using information derived from samples located in all three transects (widespread spatial configuration) often resulted in a different number of possible combinations of the data than using information from fewer transects. For instance, when samples are located in the same transect there is only three possible data sets for a sampling intensity representing 33% of the total, whereas when

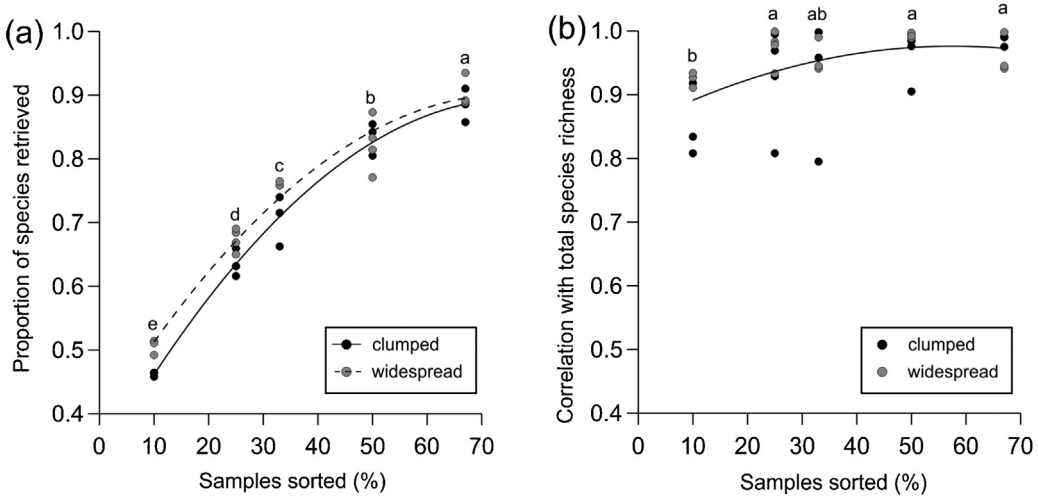

Fig. 2. (A) Proportion of the total number of species retrieved as a function of the sampling intensity (% of the total number of samples) and the spatial configuration of the samples. (B) Pearson correlation in site species richness between the full data set and subsets of the data representing 10%, 25%, 33%, 50% or 67% of the total number of samples, distributed either in all three transects (widespread) or in only one or two (clumped). Different letters above symbols indicate that there are significant differences in mean values among the different sampling intensities. Lines represent the fit of a quadratic function to the data.

samples are located in different transects this number is much higher. Thus, to avoid comparing the two spatial configurations using different number of replicates, we randomly chose three data sets from all possible data sets obtained using information from all transects.

To assess the level of congruence of the overall site species richness between the full data set and each subset of the data (38 subsets in total), simple linear Pearson correlations were calculated. To assess the level of congruence in the composition of the ant communities between the full data set and the different subsets of the data, we computed the "standardized Mantel statistic". The standardized Mantel statistics is a correlation coefficient that calculates the strength of the relationship between two matrices and thus indicates the effect size of the relationship (Peck, 2010, p. 114). Values close to 1 are indicative of a nearly perfect positive association with the full data set, whereas values close to 0 are indicative of no association. Therefore, the Mantel statistic and the Pearson correlation coefficient were used here as indexes of the amount of the original information retained by each subset of the original data set.

For each subset of the data two matrices were built: one containing data about species presence or absence in each of the five sampling sites and the other containing information about the frequency (number of samples in which the species was recorded) of each species in each site. Dissimilarity among sites was measured using the Bray–Curtis index (when matrices contained data on species frequencies) or the complement of the Sørensen index calculated as: $1-(2a/(S_1+S_2))$, where S_1 is the number of species in site "1", S_2 is the number of species in site "2", and "a" is the number of species that occurred in both sites. Prior to the analyses, the frequency data of each matrix was standardized by the column maximum values as to equate the differences between maximum frequencies possible for each sampling effort.

Analysis of Covariance (ANCOVA) was used to determine whether sampling intensity (the covariate, defined here as the proportion of the total number of samples included in the data subset) and the spatial distribution of the samples (clumped or widespread) had an influence on the above-described Pearson and Mantel correlation coefficients. To compare mean correlation values among the five different sampling intensities we used one-way ANOVA followed by a posteriori pair-wise comparison test (Fisher's LSD test) (Hsu, 1996, pp. 139–140). Because the data did not meet the assumptions of data normality, the ANCOVA and the ANOVA tests were performed using a randomization procedure (i.e. a permutation test). The permutation tests (number of randomizations = 5000) were run in the lmPerm package for R (R Core Development Team, 2014).

3. Results

In total we recorded 323 species or morphospecies from 63 ant genera. In the arboreal samples there was a total of 153 species, whereas the ground samples yielded 280 species. On average, we collected 4.5 species (range = 0–11 species) per sample in the arboreal vegetation, and 14.2 species (range = 1–30 species) per sample on the ground. Over half of the species (173 species or 54.5% of the total) were found in only one site, 76 in two sites, 35 in three sites, and 23 in four sites, whereas just 13 (4% of the total) were present in all five sampling sites. Barra do Garças was by far the most species-rich site with a total of 170 species, followed by Brasilia, Mucugê, Barreiras, and Bom Jesus (with 135, 100, 94 and 91 species, respectively). Sampling completeness, as inferred by different species richness estimators, was greater in Brasília, Barreira, and Bom Jesus (where, on average, 84.3% of the estimated number of species were collected, range = 80.5–88.5%), compared to Barra do Garças and Mucugê (mean = 71.9%, range = 65.4–75.9%).

3.1. Sampling intensity and spatial configuration of the samples

The number of species recorded in each site was affected both by variations in the sampling intensity and in the spatial configuration of the samples (Permutation ANCOVA $\eta^2 = 0.918$, sampling intensity $P < 0.001$, spatial configuration $P = 0.038$; Fig. 2a). More species were obtained when samples were spread over all three transects than when they were located in just one or two transects. The number of species recorded in each site decreased as sampling intensity decreased, with mean values differing significantly among the five sampling intensities analyzed (Permutation ANOVA $\eta^2 = 0.961$, $P < 0.001$; Fig. 2a).

Sampling intensity and the spatial configuration of the samples also significantly affected the level of congruence in site species richness with the full data set (Permutation ANCOVA $\eta^2 = 0.308$, sampling intensity $P = 0.004$, spatial configuration $P = 0.036$).

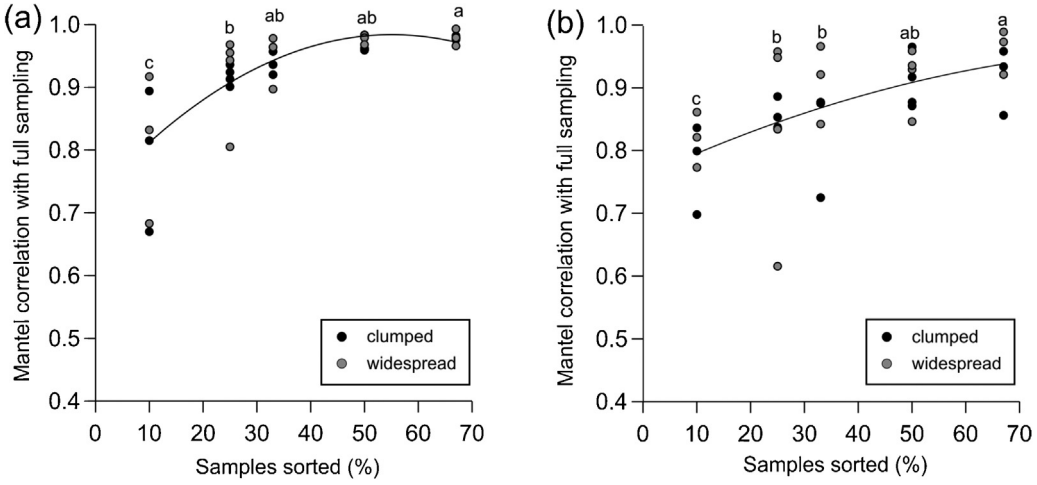

Fig. 3. Mantel correlation between the species matrix built using the full data set and matrices built using subsets of the data representing 10%, 25%, 33%, 50% or 67% of the total number of samples, distributed either in all three transects or in only one or two. (A) Matrices based on species frequency data. (B) Matrices based on species presence or absence data. Different letters above symbols indicate that there are significant differences in mean values among the different sampling intensities. Lines represent the fit of a quadratic function to the data.

Congruence was lower for a sampling intensity representing one-tenth of the total than for most of the remaining sampling intensities, and lower for the clumped than for the widespread spatial configuration (Fig. 2b).

The level of congruence in species composition with the full data set (measured by the standardized Mantel coefficient) was significantly affected by the sampling intensity (Permutation ANCOVA; frequency data: $P < 0.001$, $\eta^2 = 0.475$; presence or absence data: $P < 0.001$, $\eta^2 = 0.363$) but not by the spatial configuration of the samples (frequency data $P = 0.98$; presence or absence data $P = 0.22$). In general, congruence with the full data set was higher for a sampling intensity representing two-thirds or half of the total than for the remaining sampling intensities (Fig. 3). Furthermore, as sampling intensity decreased there was a trend toward finding higher variability in the resulting Mantel correlation coefficients (Fig. 3).

3.2. Surrogates

The "indicator taxa" data set was the best surrogate of species richness, followed by the ground-dwelling ant community, identification to genus level, and the tree-dwelling community (Table 1). "Indicator taxa" was also the best surrogate of the dissimilarities in species composition among our sampling sites when using species presence or absence data. When using data on species frequencies in each site, the best surrogate was the ground-dwelling fauna, followed by the indicator taxa, the arboreal fauna, and identification to genus level. Overall, Mantel correlation coefficients were much higher for the ground-dwelling fauna and the indicator taxa than for the remaining two surrogates (Table 1). The proportion of all species occurrences ($n = 5462$ ant species occurrences in the 600 samples), which can be considered as an index of the "cost" of sorting the ant samples, was much higher when using the ground than the tree-dwelling fauna as surrogates. Using the indicator taxa as surrogate reduced the amount of sorting by about 50% (Table 1). We did not calculate the amount of work involved in sorting ants to genus level rather than to species or morphospecies level, but it is certainly less work than any of the other surrogates analyzed here.

4. Discussion

4.1. Sampling intensity and spatial configuration of the samples

Species richness and composition are two essential measures for biodiversity assessment and monitoring. For speciose taxa (such as ants) the reliability of these measures depends, to a large extent, on the sampling effort and the level of complementarity of sampling methods employed (Souza et al., 2012). Undersampling can result in inadequate estimates of species richness (Longino et al., 2002) and/or inflate differences between communities because, simply by chance, many of the species that two communities have in common are not sampled in one or both communities (Cardoso et al., 2009). Nearly complete ant inventories of relatively small areas (e.g. La Selva Biological Station, with an area of 1500 ha) are possible, but they require both the use of multiple sampling methods and an enormous sampling effort (Longino et al., 2002). Although such complete inventories are desirable in any study, they are almost impossible to apply over multiple sites or sampling periods as commonly required in most ecological studies. It has been suggested that ecologists should consider the benefits of "surveying a subset of the total species pool and redirecting resources to increasing the study's sample size"(Vellend et al., 2008). However, there is little empirical evidence to support this contention. Here, we have employed an intensive sampling protocol in five different savanna areas and we were able to retrieve up to 88.5% of the species expected to occur in these sites. We then analyzed whether we

Table 1
Correlation (Pearson or Mantel) between the full data set and different subsets of the data ("surrogates"). The proportion of all species occurrences ($n = 5462$ species occurrences in a total of 600 samples) was used as an index of the "cost" of sorting the ant samples.

Surrogate	Species richness	Dissimilarity (presence/absence data)	Dissimilarity (frequency data)	"Cost" of sorting
Arboreal ants	0.878	0.617	0.645	0.233
Ground ants	0.945	0.876	0.959	0.767
Higher-taxon (Genus)	0.936	0.434	0.482	n.a.[a]
Indicator taxa	0.971	0.913	0.891	0.510

[a] Not applicable.

could reduce our sampling effort in each site without losing much of the original information (i.e. by measuring the amount of the original information retained by different subsets of the original data set). For species richness, the level of congruence with the full data set was similar for all levels of sampling intensity analyzed, except for the one-tenth level, for which the congruence was lower than for the remaining levels. When analyses involved species composition (i.e. dissimilarities among sites) rather than species richness, more marked differences among the different sampling intensities were detected. In particular, there was a higher congruence with the full data set for the sampling intensities representing two-thirds or half of the total than for the remaining sampling intensities. Furthermore, as one might expect, we found that the smaller the sampling effort the greater the variability in correspondence with the original data set. This indicates that a small sampling size increases the chances of obtaining results that did not reflect the actual differences among the sites being compared.

With regard to the spatial configuration of the samples we found that more species can be retrieved when samples are spread over multiple transects then when they are located in only one or two transects. This indicates that even at relatively small scales (few kilometers) there is a turnover of ant species within the same type of habitat, perhaps as consequence of the clumped spatial distribution of some species (e.g., Crist and Wiens, 1996), of minor variations in habitat structure (e.g., Bestelmeyer and Wiens, 2001), or both. Surprisingly, however, the spatial configuration of the samples had an effect on the level of correspondence with the full data set in just one of the three analyses we performed (the one involving species richness). Nevertheless, as one might expect, visual examination of the data indicates that the smaller the sampling effort the greater the difference in the number of species recorded according to the spatial configuration of the samples (Fig. 2). In particular, the difference in the number of species recorded between the two spatial configurations became more clear for sampling intensities representing one-third or less of the total, and this is probably because at these sampling intensities samples of the clumped configuration were all located in just one transect rather than in two as was the case for the other (and larger) sampling intensities. In other words, our test of the spatial configuration of the samples may have well been conservative since the contrast between the two configurations was relatively small for two of the five sampling intensities analyzed here (see Fig. 1).

One of the most widely used protocols to sample ants is the Ants of the Leaf-Litter (ALL) protocol (Agosti and Alonso, 2000), which was originally proposed by a group of ant specialists based on their experience on sampling sufficiency and complementarity of sampling methods. This protocol was specifically designed for sampling ground-dwelling forest ants and so far no other protocol has been proposed for sampling savanna ants. In this sense, it is interesting to note that the simplified protocol we are suggesting here (i.e. at least half of the samples from the original protocol) has some similarities to the ALL protocol. The latter recommends the use of two complementary sampling methods (ground pitfall-traps and the Winkler method) (Agosti and Alonso, 2000), while here we propose the use of pitfall traps on two different foraging/nesting strata. Savannas support relatively distinct ground and arboreal ant communities (Vasconcelos and Vilhena, 2006; Campos et al., 2011), and because of this arboreal and ground pitfall traps work as complementary methods. Forests obviously also support a distinct arboreal ant fauna (Yanoviak and Kaspari, 2000) but because accessing the forest canopy is much more difficult than accessing the crowns of savanna trees, the use of arboreal traps in forests is, consequently, also difficult. On the other hand, forests, most notably tropical ones, have a particularly rich fauna of cryptic species that live in the leaf-litter and that is why the use of the Winkler method is much more efficient is these habitats than in open savannas (Lopes and Vasconcelos, 2008; Parr and Chown, 2001, but see Souza et al., 2012). With regard to sampling effort, the ALL protocol recommends a minimum of 100 samples (50 from each method), but ideally more (Agosti and Alonso, 2000), while the analyses performed here indicate that at least 60 composite samples (of four pitfalls each) (30 in each stratum) are necessary. In other words, both protocols recommend a relatively high sampling effort. Several studies found in the ecological literature, however, employ a much lower sampling effort. This is not to say that these studies have any flaws, but it is important that myrmecologists should more often check the sufficiency of their sampling protocols.

4.2. Surrogates

As Leal et al. (2010) have pointed out "there is no accepted benchmark for assessing if a biodiversity surrogate provides a reliable prediction of total species richness for conservation planning purposes". These authors defined a surrogate as being 'reasonable' if it explains 60% of the variation in total species richness, 'good' if it explains 70%, and 'excellent' if it explains 80% or more (Leal et al., 2010). If we apply this same criterion here, then the arboreal-dwelling fauna can be regarded as a good surrogate of the species richness in our study sites, whereas the remaining surrogates are considered excellent. However, if we extend this criterion for the analyses involving species composition, than the higher-taxon surrogate and the arboreal fauna can be regarded as "below reasonable", whereas the ground-dwelling fauna and the indicator taxa are good or excellent (Table 1).

Sorting just the ground samples produced a much greater level of congruence with the full data set than sorting just the arboreal samples, and this is probably explained by the fact that ground-dwelling ants were both more abundant and diverse than the arboreal ants. While in the ground samples we recorded a total of 280 species (86.7% of the total), in the arboreal samples we recorded only 153. Thus, if one is only interested in estimating which species are present in a given site, it is better to use only ground pitfall traps than to use the same number of traps but with half of them being placed on ground and half in the arboreal vegetation. This is because most tree-dwelling species often also forage on the soil surface. However, these species are much less frequent on the ground than in the arboreal samples (Vasconcelos and Vilhena, 2006; Campos et al., 2011) and therefore if the aim of the study is to characterize the species in terms of abundance or frequency, than the sole use of ground pitfall traps will severely underestimate the abundance or frequency of arboreal species.

As was also observed in a previous study (Groc et al., 2010), by sorting samples from a subset of the ant genera (namely these less challenging for species level identification and/or separation into morphospecies) we retrieved most of the information existing in the full data set, even though the amount of work involved in mounting and sorting specimens was reduced by about 50%. This suggests that different ant genera show convergent levels of variation in species richness and composition across the savannas of Brazil, i.e. that the mechanisms affecting community assembly (species sorting and filtering) may be similar for species of different genera.

Although there is evidence that, within a given site or region, the relationship between species composition and genus composition can be strong (Andersen, 1995; Groc et al., 2010), here we found a poor level of congruence in species composition with the full data set by using higher-taxon surrogacy (identification to genus level). These contrasting results may be explained because our analyses involved sites located in different regions of the Brazilian Cerrado. Our results thus give additional support to the idea that, for ants,

the effectiveness of higher-taxon surrogacy is dependent on the spatial scale of the study (Andersen, 1995).

4.3. Concluding remarks and recommendations

Species richness and composition are two of the most widely used metrics for studies involving the assessment or monitoring of ant diversity, and the reliability of these metrics is dependent on how well the sites under analysis are sampled. However, because of the unfeasibility of making a thorough inventory over multiple sites, ant ecologists often seek a balance between the benefits of sampling completeness and the costs involved in species sorting. Here we show that by sorting a subset of the ant taxa we reduced the amount of work involved in specimen preparation and identification by half, and yet obtained a high level of congruence with the full data set. Reducing the number of samples taken in each sampling site, as expected, also reduced the amount of work involved in specimen preparation and identification. However, it is important to notice that the level of congruence with the full data set was lower and/or became more erratic when sampling intensity was reduced to one-third or less of the original number of samples. For this reason we suggest that at least half of the samples from our original protocol are necessary to adequately describe ant diversity patterns in Neotropical savannas. We also suggest that, whenever possible, sampling within a given site should be done in multiple locations as to maximize the number of species collected.

Acknowledgements

We are thankful for the Brazilian Council of Research and Scientific Development (CNPq grants 457407/2012-3 and 478107/2012-9) and the Foundation for Scientific Development of the Federal District (FAP-DF, Projeto PRONEX 563/2009) for funding this research. We also wish to thank Alan Andersen, Galen Priest, and two anonymous reviewers for reading and commenting on a previous version of the manuscript. Jorge Neves, Alessandra Bartimachi, Camila Bonizário, and Laura V.B. Silva were of great assistance during the fieldwork.

References

Agosti, D., Alonso, L.E., 2000. The ALL protocol: a standard protocol for the collection of ground-dwelling ants. In: Agosti, D., Majer, J.D., Alonso, L.E., Schultz, T.R. (Eds.), Standard Methods for Measuring and Monitoring Biodiversity. Smithsonian Institution Press, Washington and London, pp. 204–206.

Andersen, A.N., 1995. Measuring more of biodiversity: genus richness as a surrogate for species richness in Australian ant faunas. Biol. Conserv. 73, 39–43.

Andersen, A.N., Majer, J.D., 2004. Ants show the way down under: invertebrates as bioindicators in land management. Front. Ecol. Environ. 2, 291–298.

Bestelmeyer, B.T., Wiens, J.A., 2001. Ant biodiversity in semiarid landscape mosaics: the consequences of grazing vs. natural heterogeneity. Ecol. Appl. 11, 1123–1140.

Campos, R.I., Vasconcelos, H.L., Andersen, A.N., Frizzo, T.L.M., Spena, K.C., 2011. Multi-scale ant diversity in savanna woodlands: an intercontinental comparison. Aust. Ecol. 36, 983–992.

Cardoso, P., Borges, P.A.V., Veech, J.A., 2009. Testing the performance of beta diversity measures based on incidence data: the robustness to undersampling. Diver. Distrib. 15, 1081–1090.

Chao, A., Chazdon, R.L., Colwell, R.K., Shen, T.J., 2005. A new statistical approach for assessing similarity of species composition with incidence and abundance data. Ecol. Lett. 8, 148–159.

Colwell, R.K., 2006. EstimateS: statistical estimation of species richness and shared species from samples, version 8.2, http://viceroy.eeb.uconn.edu/estimates

Crist, T.O., Wiens, J.A., 1996. The distribution of ant colonies in a semiarid landscape: implications for community and ecosystem processes. Oikos 76, 301–311.

Gaston, K.J., 2000. Global patterns in biodiversity. Nature 405, 220–227.

Greenslade, P.J., 1973. Sampling ants with pitfall traps – digging-in effects. Insectes Sociaux 20, 343–353.

Groc, S., Delabie, J.H.C., Longino, J.T., Drivel, J., Majer, J.D., Vasconcelos, H.L., Dejean, A., 2010. A new method based on taxonomic sufficiency to simplify studies on Neotropical ant assemblages. Biol. Conserv. 143, 2832–2839.

Hess, G.R., Bartel, R.A., Leidner, A.K., Rosenfeld, K.M., Rubino, M.J., Snider, S.B., Ricketts, T.H., 2006. Effectiveness of biodiversity indicators varies with extent, grain, and region. Biol. Conserv. 132, 448–457.

Hölldobler, B., Wilson, E.O., 1990. The Ants. Belknap Press, Cambridge, Mass.

Hsu, J.C., 1996. Multiple Comparisons. Theory and Methods. Chapman & Hall, London.

Jenkins, C.N., Guenard, B., Diamond, S.E., Weiser, M.D., Dunn, R.R., 2013. Conservation implications of divergent global patterns of ant and vertebrate diversity. Diver. Distrib. 19, 1084–1092.

Klink, C.A., Machado, R.B., 2005. Conservation of the Brazilian Cerrado.Conserv. Biol. 19, 707–713.

Landeiro, V.L., Bini, L.M., Costa, F.R.C., Franklin, E., Nogueira, A., de 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. Ecol. Indicators 23, 366–373.

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. Biodiver. Conserv. 19, 3347–3360.

Longino, J.T., Coddington, J., Colwell, R.K., 2002. The ant fauna of a tropical rain forest: estimating richness three different ways. Ecology 83, 689–702.

Lopes, C.T., Vasconcelos, H.L., 2008. Evaluation of three methods for sampling ground-dwelling ants in the Brazilian cerrado. Neotrop. Entomol. 37, 399–405.

McGeoch, M.A., Sithole, H., Samways, M.J., Simaika, J.P., Pryke, J.S., Picker, M., Uys, C., Armstrong, A.J., Dippenaar-Schoeman, A.S., Engelbrecht, I.A., Braschler, B., Hamer, M., 2011. Conservation and monitoring of invertebrates in terrestrial protected areas. Koedoe 53, 13.

MMA, 1999. Áreas prioritárias para a conservação do Cerrado e Pantanal. Ministério do Meio Ambiente, Brasília, http://www.mma.gov.br/biomas/cerrado/areas-prioritarias

Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B., Worm, B., 2011. How many species are there on Earth and in the ocean? PLoS Biol. 9, e1001127.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858.

Oliveira-Filho, T., Ratter, J.A., 2002. Vegetation physiognomies and woody flora of the Cerrado Biome. In: Oliveira, P.S., Marquis, R.J. (Eds.), The Cerrados of Brazil – Ecology and Natural History of a Neotropical Savanna. Columbia University Press, New York, pp. 91–120.

Pacheco, R., Vasconcelos, H.L., 2012a. Habitat diversity enhances ant diversity in a naturally heterogeneous Brazilian landscape. Biodiver. Conserv. 21, 797–809.

Pacheco, R., Vasconcelos, H.L., 2012b. Subterranean pitfall traps: is it worth including them in your ant sampling protocol? Psyche (article ID 870794).

Parr, C.L., Chown, S.L., 2001. Inventory and bioindicator sampling: testing pitfall and Winkler methods with ants in a South African savanna. J. Insect Conserv. 5, 27–36.

Peck, J.E., 2010. Multivariate analysis for community ecologists: step-by-step using PC-ORD. MJM Software Design, Gleneden Beach, Oregon.

R Core Development Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, Retrieved from http://www.r-project.org/

Rosenzweig, M.L., 1995. Species Diversity in Space and Time. Cambridge University Press, Cambridge.

Schuldt, A., Wang, Z.H., Zhou, H.Z., Assmann, T., 2009. Integrating highly diverse invertebrates into broad-scale analyses of cross-taxon congruence across the Palaearctic. Ecography 32, 1019–1030.

Silva, R.R., Brandão, C.R.F., Silvestre, R., 2004. Similarity between Cerrado localities in central and southeastern Brazil based on the dry season bait visitors ant fauna. Stud. Neotrop. Fauna Environ. 39, 191–199.

Smith, A.M., Fisher, B.L., Hebert, P.D.N., 2005. DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philos. Trans.: Biol. Sci. 360, 1825–1834.

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. Appl. Soil Ecol. 56, 63–73.

Stork, N.E., Habel, J.C., 2013. Can biodiversity hotspots protect more than tropical forest plants and vertebrates? J. Biogeogra. 41, 421–428.

Vasconcelos, H.L., Araujo, B.B., Mayhe-Nunes, A.J., 2008. Patterns of diversity and abundance of fungus-growing ants (Formicidae: Attini) in areas of the Brazilian Cerrado. Revista Brasil. Zool. 25, 445–450.

Vasconcelos, H.L., Vilhena, J.M.S., 2006. Species turnover and vertical partitioning of ant assemblages in the Brazilian Amazon: a comparison of forests and savannas. Biotropica 38, 100–106.

Vellend, M., Lilley, P.L., Starzomski, B.M., 2008. Using subsets of species in biodiversity surveys. J. Appl. Ecol. 45, 161–169.

Yanoviak, S.P., Kaspari, M., 2000. Community structure and the habitat templet: ants in the tropical forest canopy and litter. Oikos 89, 259–266.