# Reliability and refinement of the higher taxa approach for bee richness and composition assessments

**Itai van Rijn, Thomas M. Neeson and Yael Mandelik**[#]

Department of Entomology
The Hebrew University of Jerusalem
P.O. Box 12
Rehovot 76100 Israel
[#]Corresponding author: Yael.Mandelik@mail.huji.ac.il
Tel: +972-8-9489224
Fax: +972-8-9466768

Current addresses:
I.V.R: Department of Zoology, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel.
T.M.N: Center for Limnology, University of Wisconsin, 680 North Park St., Madison WI 53706 USA

Running title: Higher taxa approach applied to bees

## Abstract

Limited resources and taxonomic expertise in biodiversity surveys often lead to the application of the Higher Taxa Approach (HTA), i.e. the identification of specimens to genus or higher taxonomic levels rather than to species. The reliability of the HTA varies significantly among studies, yet the factors underlying this variability have rarely been investigated. Bees are an ideal model taxon for testing the HTA because they are highly diverse, challenging to identify and there is widespread interest in their role as native pollinators, driving demand for efficient diversity assessment tools. Using extensive bee data sets collected across three biomes and various habitats, we assessed the performance of the HTA in reflecting bee species richness and composition patterns at local scales, factors affecting this performance, and ways to improve it. The performance of the HTA varied considerably among biomes, taxonomic levels (genus and sub-families) and diversity measures (species richness and composition); genus and sub-family richness accounted for 55-77% and 32-61% of the variation in species richness respectively; genus and sub-family composition accounted for 28-87% and 26-80% of the variation in species composition respectively. The number of species per higher taxon was a main factor influencing this performance (accounting for 63% of the variation), while the co-occurrence of taxonomically related species had no significant influence on the performance of the HTA. Further subdividing genera by body size contributed to the performance of the HTA and increased its accuracy in representation of compositional patterns by ca. 16%. Our results have several practical implications. The considerable variability found in the performance of the HTA in representing local-scale richness and composition patterns of bee species dictates caution in implementing this tool in bee surveys. When possible, an a priori evaluation of the expected performance of the HTA should be done, focusing on species distributions within higher taxonomic levels and the species/higher taxa ratio. Integrating morphological characteristics (such as body size) that consistently sub-divide genera will improve HTA's performance. Our results are likely applicable to the implementation of the HTA in other small-bodied and species-rich groups and contribute to the cost-effectiveness of biodiversity surveys.

**Key words:** Apoidea, bee communities, biodiversity assessment, biodiversity survey, conservation, species composition, taxonomic sufficiency

## Introduction

There is a growing need for quick and cost-effective methods for assessing biodiversity patterns for conservation planning and monitoring programs (Mandelik et al. 2010; Kessler et al. 2011; Tulloch et al. 2011). Since most conservation efforts are directed to the species level, i.e. protecting single species or an assemblage of species, biodiversity surveys often require the identification of specimens to species level. However, species identification is often limited by the low availability of taxonomic experts and requires significant investment of time and money (Williams and Gaston 1994; Gotelli 2004). This is especially true when surveying species-rich and small bodied taxa such as most insects groups (McGeogh 1998; Hodkinson and Jackson 2005). One of the promising surrogates for direct species identification is the 'Higher Taxa Approach' (HTA). In this method specimens are identified to higher taxonomic levels (e.g. genus, subfamily) and any analysis is performed using data on those higher levels (Gaston and Williams 1993). In most cases, the HTA has been used to predict patterns of species richness and community composition (Bevilacqua et al. 2011). The HTA has been widely applied in biodiversity monitoring programs (Dauvin et al. 2003; Muniz and Pires-Vanin 2005; Nakamura et al. 2007) and conservation planning (Cardoso et al. 2004; Gladstone and Alexander 2005). Despite the widespread use of higher taxon surrogates in conservation, empirical tests suggest that their performance, i.e. the correlation between species and higher level richness and composition patterns, may vary markedly with taxonomic resolution and biogeographic setting (Balmford et al. 1996; Baldi 2003; Cardoso et al. 2004; Gladstone and Alexander 2005; Guzman-Alvis and Carrasco 2005; Marshall et al. 2006; Lovell et al. 2007; Mandelik et al. 2007; Hirst 2008). This variability has led to confusion over the usefulness of these surrogates. Nevertheless, there is a shortage of studies on the mechanisms that underlie the HTA's performance and, ultimately, reliability (Bevilacqua et al. 2011).

Two main factors may influence the performance of the HTA. The first is the similarity in life history traits and ecological niche requirements among species belonging to the same higher taxonomic group. A high degree of similarity in ecological requirements may result either in high co-occurrence of taxonomically related species due to habitat filtering (Webb et al. 2002), or in low co-occurrence due to competition among ecologically similar species, i.e. niche separation (Lawton and Strong 1981), or in a mixture of both effects. In the case of co-occurrence of taxonomically related species, the HTA is expected to perform well in distinguishing species assemblages among different habitats or across environmental variation. Niche separation is expected to lower the HTA's performance in distinguishing species assemblages among different habitats, since species belonging to the same higher taxonomic group will be scattered across habitats.

A second important factor which may influence the performance of the HTA is the number of species per higher taxon and its variance. Bevilacqua et al. (2011) found that HTA's performance in marine invertebrates was unaffected by taxonomic relatedness (i.e. the relative closeness of species in the Linnaean taxonomic hierarchy), but strongly affected by species/higher taxa ratios. Neeson et al. (2013), in a theoretical model, found that the HTA is negatively affected by high variance in the distribution of individuals among species, as often found in ecological communities, including in bees (Williams et al. 2001). We are not aware of other empirical or theoretical studies investigating the mechanisms underlying the HTA's performance.

Here we assess factors affecting the reliability of the HTA in bees as a model terrestrial taxon. Bees (superfamily Apoidea) are a good system for testing ideas about the HTA's performance for several reasons. First, some taxonomically related species (i.e. species that are close in the Linnaean taxonomic hierarchy) exhibit similar nesting behavior, foraging behavior and body size (Michener 2007), which might at times lead to high co-occurrence of these related species. In other cases, taxonomically-related species differ in host breadth, sociality and habitat preferences, and are less likely to co-occur. Thus differences in the degree of species co-occurrence among taxa provide an opportunity to test hypotheses about the effects of species co-occurrence on the reliability of the HTA. Second, some bee genera contain tens of species in the same community and locality, while other genera are relatively species-poor. These disparities in the number of species per genera mean that bees can be used to explore whether the HTA is affected by variation in the species / higher taxon ratio. Third, bees are specifically suited to testing the performance of the HTA at local scales since they have limited living ranges (hundreds of meters to a few kilometers) (Greenleaf et al. 2007), thus sampled individuals are tightly linked to the site they were sampled in. Moreover, bees, as central place foragers, are highly dependent on local scale properties of sites such as suitable nesting substrate and foraging resources within flight distance of each other. These factors may create differences in bee communities found in adjacent habitats (Steffan-Dewenter and Tscharntke 1999; Potts et al. 2005; Gotlieb et al. 2011; Mandelik et al. 2012).

Bees are also of intrinsic interest due to their roles as crop and wild flower pollinators. About one third of the global human diet is dependent on crops pollinated by animals (Klein et al. 2007), mainly bees, and ca. 87% of all wild plants need animal pollination in order to reproduce (Ollerton et al. 2011). The proportion of animal-pollinated crops continues to increase globally (Aizen and Harder 2009). The decline in managed honey bee hives in parts of the world including U.S. and Europe (Potts et al. 2009; vanEngelsdorp et al. 2009) has raised interest in wild bee communities as providers of a pollination safety net (Winfree et al. 2007) and increased the demand for efficient methods for surveying wild bees around the world (Brown and Paxton 2009).

Finally, there are compelling practical reasons for assessing whether higher taxon surrogates might perform well for bees. Species-level identification of many wild bee samples requires special expertise and considerable investment of time (e.g. genitalia extraction, comparison to reference collections). Consequently, paid taxonomic identification can be costly (up to three USD/specimen based on our experience and on information from colleagues; an average bee survey or study can contain thousands of bees). For some groups and in some regions of the world expertise in wild bee taxonomy is very limited or nonexistent (Batley and Hogendoorn 2009; Eardley et al. 2009). Hence, specifically in the case of wild bees, there is an urgent need for a surrogate that will provide reliable richness and composition assessments. Genus level identification of bees can be conducted with relative ease after a short training (several days to a few weeks for identification of bee genera found in local scale surveys) compared to a much more extensive training required for species-level identification of bees (months to years, depending on the diversity of sites). Hence, the HTA could provide a significant shortcut in local bee surveys.

Here we investigate the ability of the HTA to serve as a reliable surrogate for direct species evaluation of bees and explore the factors that influence its performance. Our goal is to provide practical advice for using the HTA with greater confidence in future studies and surveys of bees and other small-bodied, species-rich taxa. Using extensive datasets from studies conducted in three different biomes and various habitats we aim to answer the following specific questions:

1. Can the HTA be used to reliably describe bee species richness and composition in local scale studies?
2. What are the main factors affecting the performance of the HTA? Specifically we investigate the effects of co-occurrence of taxonomically related species, the ratio between the number of species and higher taxonomic groups, and species distribution within higher taxa as possible drivers of variation in the HTA's performance.
3. Can the performance of the HTA be improved by subdivision of taxonomic groups according to body size?

## Materials and methods

### Data sets

We used bee data collected from three biomes. Accordingly, three datasets were formed: Desert ("Arid"), Mediterranean ("Med") and Temperate ("Temp"). Each data set contained two independent studies, conducted in different locations and years. Table 1 presents the number of individuals, species and higher taxonomic groups in each of the data sets.

*Arid data set*: These data were collected in the extreme arid ecosystem of the Arava, south of the Dead Sea, Israel. Annual average precipitation in this region is below 50 mm, wild plants are scattered, mainly in "wadis" (intermittent streams) and cultivated plants are abundant within gardens in settlements. Agriculture is one of the main land-use types and is conducted mainly in protected crop systems (net-houses, green-houses etc.). This data set contains bees collected in natural habitats (wadis), both adjacent and far from agriculture and in gardens inside settlements [for detailed description of methods, sampling effort and bee species sampled in the "gardens and natural habitats" study (Table 1) see Gotlieb et al. 2011, and for the "Agricultural habitats" study see Appendix A]. Specimens were identified following Michener (2007; see Appendix A for the list of taxonomists that identified the specimens). Specimens were housed in the National Collections of Natural History at the Tel Aviv University, and at the Hebrew University of Jerusalem.

*Mediterranean data set*: These data were collected in the Judean foothills, south-west of Jerusalem, central Israel, a region characterized by transient Mediterranean climate. Average annual precipitation in this region is 350-500 mm and main land-use categories are natural habitats, mainly Mediterranean scrub at various successional stages, planted forests (mainly pines) and agriculture (fields and orchards). This data set includes bees collected in natural habitats, planted pine forests and unmanaged areas adjacent to agriculture [for detailed description of methods, sampling effort and bee species sampled in the "Agriculture and natural habitats" study see Pisanty et al. 2014, and for the "Planted pine forests and natural habitats" study see Appendix A]. Specimens were identified and housed the same as specified for the Arid data set.

*Temperate data set*: These data were collected in central New Jersey and eastern Pennsylvania, USA. Average annual precipitation in this region is >1000 mm, and the main land-use categories are natural habitats, mainly temperate forests, abandoned agricultural fields (old fields) at various successional stages, agriculture and sub-urban areas. This data set includes bees collected in forests, agriculture areas and old fields of different ages [for detailed description of methods, sampling effort and bee species sampled in the "Forests and agricultural habitats" study see Mandelik et al. 2012, and for the "Abandoned agriculture fields" study see Appendix A].

Specimens were identified by Sam Droege and John Ascher. Specimens were housed at Rutgers University, USA.

Based on preliminary analysis of species accumulation rates in each dataset, we excluded sites with less than 30 specimens, as they likely fail to reflect bee assemblage at a site. The minimum threshold of 30 individuals/site reflects a balance between representative samples, based on moderate accumulation rates achieved under this sample size, and realistic sampling efforts, working in sites with very low bee abundance overall. Nevertheless, these samples are not exhaustive and should be regarded as a relative measure of richness.

## Statistical analysis

We conducted three sets of analyses aimed at:

### 1) Testing the performance of the HTA for bee richness and composition assessments

*Species richness*

For each data set, we calculated Pearson's correlation between species richness and richness of genera and subfamilies. Though the subfamily level is seldom used in bee classification, we used this level to explore the effect of coarser taxonomic levels on the HTA's performance. Since there are only seven bee families globally (Michener 2007), variation at this level, especially at the local scale, is too limited to allow meaningful analysis. The number of subfamilies in our datasets ranged between 2 and 12 per site.

*Species composition*

We used two complementary methods to explore the ability of the HTA to distinguish patterns of species composition. First, similarity matrices were drawn for each data set separately, for species, genus and subfamily data using the Bray-Curtis dissimilarity index (Beals 1984). Correlations between distance matrices of the species and higher taxonomic levels were calculated using Mantel tests (Dutilleul et al. 2000). A high correlation between species-level and genus- or subfamily-level similarity matrices would indicate that the HTA can be used to describe the similarity of species composition among sites.

A second analysis focused on the ability of the HTA to detect differences in species assemblages among habitats, a common goal of many surveys and monitoring programs (Harding et al. 1998; Bennett et al. 2006). In this analysis we included the two studies in which clear differences among habitats in species composition had previously been observed. The first is a study from the Arid data set conducted during 2008-2009 in which bees were sampled in four gardens inside settlements and in four natural areas in main "wadis" in the same region (Gotlieb et al. 2011). The second study is part of the 'Temperate' data set. This study was conducted in New Jersey in 2005-2006. Three habitats were sampled: secondary forests, agriculture fields of bee-attractive crops and old fields (Mandelik et al. 2012). In the other studies species compositions in the different habitats were overlapping (no significant habitat effect of species composition was found). The effect of habitats on community composition was visualized by plotting a two dimensional Non-metric Multidimensional Scaling (NMS) ordination based on Bray-Curtis dissimilarity index, for data on species, genera and subfamilies. In addition we conducted Multiresponse Permutation Procedure (MRPP) tests to quantify differences among habitats. The MRPP is a non-parametric procedure for testing the hypothesis of no difference among a priori determined groups. In this application, the MRPP tests whether the Bray-Curtis dissimilarity distance is smaller among sites with similar habitat than among sites grouped together at random (Zimmerman et al. 1985; Peltzer et al. 2000). If the degree of habitat clustering in both the NMS space and MRPP test is similar across taxonomic levels, we will infer that the HTA can be used to detect differences in species assemblages among habitats.

Mantel tests, NMS and MRPP test were conducted using PC-ORD statistical software (PC-ORD 5.31, MjM Software, Gleneden Beach, Oregon, USA)

### 2) Testing possible factors that affect the performance of the HTA in assessments of bee diversity patterns

*Effect of taxonomic relatedness*

We first examined if taxonomically related species overlapped in their distributions more than expected by chance. To do this, we performed an MRPP test where genera constituted the a priori grouping factor. We calculated a Bray-Curtis dissimilarity index among species according to the sites they occupied. Significant results from the MRPP test would indicate that species belonging to the same genus are more similar in the set of sites they occupy than species that are randomly grouped; i.e. that taxonomically related species co-occur.

Next we examined whether the HTA was affected by the degree of co-occurrence among taxonomically related species. Our hypothesis was that if such an effect exists, we would observe different outcomes when conducting the HTA on the current taxonomic division compared to a similar test where species were assigned randomly to higher taxonomic groups. We created genus-level Bray Curtis dissimilarity matrices by randomly assigning species to genera. For each data set (Arid, Med, Temp), the number of species assigned to a genus was set to match the number of species it contained in the original data. By doing this we aimed to avoid the possible effect of changing the numerical properties of species distribution within genera. We then made 10,000 runs of randomly created genus matrices. For each randomly created genus level matrix we calculated its correlation with the original species level Bray Curtis dissimilarity matrix, using a Mantel test. Then we plotted the distribution of Pearson's r values obtained by random runs and the Pearson's r value obtained when using the current taxonomic division. If the correlation between species-level and genus-level community matrices is significantly different when following the current taxonomic division compared with the randomly-generated divisions, we will conclude that patterns of co-occurrence or niche separation among related species affect the HTA's performance. In this test we used presence-absence data of species since we were interested in the effect of taxonomic relatedness regardless of differences in abundance. Code for randomization test was written in MATLAB software (MATLAB R2008A, The Mathworks Inc., Natick, Massachusetts, USA)

### *Effect of species distribution within higher taxonomic levels*

To explore how species' distributions varied among sites and among higher taxa, we calculated the average number of species in each genus and in each subfamily; the average number of species in each taxonomic group in each site; and the percentage of sites in which each taxonomic group was found. The HTA is predicted to be negatively affected by taxonomic groups which contain large numbers of species, especially if those species-rich groups are widespread across sites. Therefore, we also examined whether species-rich groups were widespread among sites by correlating the average number of species per site and the percentage of sites containing the genus using Spearman's rank correlation. Lastly we examined the effect of the higher taxa/species ratio (φ; Bevilacqua et al. 2011) on the HTA's performance using linear regression.

## 3) Exploring a possible way to improve the HTA's performance by adding body size measurements

Using the Mediterranean data set, we tested whether the reliability of the HTA could be improved by further subdividing genera using easily-measured morphologic data. This additional sub-division reduces the taxonomic resolution that is lost when using genera instead of species and may thus increase the HTA's performance. We explored the utility of body size, a prominent morphological characteristic commonly measured as the inter-tegular distance (the distance between the wing bases; ITD; Cane 1987), in increasing the HTA's performance at local scales. ITD has an a priori advantage of being consistently and objectively measured (Cane 1987), as opposed to other conspicuous morphological characteristics such as color patterns and hairiness, often used in identification keys. ITD is also appealing due to its functional significance through its effects on flying distance and on pollination efficiency (Greenleaf et al. 2007).

We measured ITD of up to five males and five females from each species (or fewer if less than five individuals/gender/species were available). Measurements were made under a stereoscope with an optic scale bar (Zeiss Stemi 2000-C). The average range of ITD within a species was 0.32 mm, corresponding to an average of 18.9% (SD: 11.6) of mean ITD of the different species. We then created seven discrete size groups: <1, 1-1.5, 1.5-2, 2-2.5, 2.5-3, 3-3.5, 3.5< (mm). Size ranges were set to 0.5mm since with this division the vast majority of species were assigned to no more than two groups, thus affording meaningful sub-division of genera. For species assigned to two size groups, specimens in the data set were randomly assigned to one of the groups, while keeping the ratio between the two groups according to the ratio found in the measured bees. Due to unavailability of some specimens in the "Med" data set (damaged, loaned etc.), we were only able to measure 83% of the total number of species. We tested whether division of genera into size categories strengthened the correlation to species data compared to the original genus data. We used the Bray-Curtis index and included only those species for which ITD measures were available. We compared the coefficients obtained for the species-genus and the species-genus-size correlations by a test of differences between two related correlation using 'Fisher's z' transformation (Xiao-Li et al. 1992) R programming language (R Core Team 2012)].

## Results

### 1) HTA's performance tests

*Species richness*

The correlation between species and generic richness was significant in all data sets and accounted for 55% (Arid) to 77% (Temp) of the variation in species richness (Arid: r=0.74 p<0.001, Med: r=0.84 p<0.001, Temp: r=0.88 p<0.001). Species and subfamilies richness showed significant yet weaker correlation, accounting for 32% (Arid) to 61% (Med) of the variation in species richness (Arid: r=0.57 p<0.002, Med: r=0.78 p<0.001, Temp: r=0.59 p<0.001).

*Species composition*

Generic level community composition similarity in the Arid and the Temp data sets explained 87% and 83%, respectively, of the observed species-level similarity (Arid: r=0.937 p<0.001, Temp: r=0.915 p<0.001). Subfamily data provided a somewhat lower correlation, accounting for 80% and 70% of the similarity observed at the species level in the Arid and Temp data sets respectively (Arid: r=0.899 p<0.001, Temp: r=0.836 p<0.001). In the Med data set both genus and subfamily provided a much weaker representation of the variation in species level composition, 28% and 26%, respectively (Genus: r=0.529 p<0.001, Subfamily: r=0.511 p<0.001).

In the NMS analysis, we found that differences in community composition among habitat types could be identified equally well when composition was described using species, genera, or subfamilies (Fig. 1). In the Temp study NMS plotting of all taxonomic levels showed a similar pattern where sites characterized by old field and agricultural habitat were more similar to each other than to forested sites (Fig. 1b). MRPP showed similarly significant differences across all taxonomic levels in both studies ($p<0.05$ for all tests; Fig. 1).

### 2) Factors affecting the HTA's performance

*Effect of taxonomic relatedness*

In all data sets, species belonging to the same genus co-occurred more often than species that were not congeneric (MRPP test results- Arid: $A=0.082$ $p<0.001$, Med: $A=0.045$ $p<0.001$, Temp: $A=0.055$ $p<0.001$). However, we found no evidence that these patterns of co-occurrence among related species had any effect on the performance of the HTA: in all three data sets the Mantel r values obtained for the current taxonomic division fell close to mean r values obtained from assigning species to genera at random (Arid: mean=0.69, Temp: mean=0.78, Med: mean=0.53) (Fig. 2).

*Effect of species distribution within higher taxonomic levels*

A hollow curve pattern was obtained for the distribution of species within genera in all three data sets while the distribution of species within subfamily showed greater variation among data sets (Fig. 3). In all three data sets there were genera with more than 10 species, but more than 40% of the genera contributed one species per site and more than 75% of the genera contributed up to two species on average per site. Nevertheless, the Med data set was distinguished from the other two datasets by the high occurrence of species-rich genera and subfamilies. The Med data set included the three richest genera in a single data set - *Andrena, Lasioglossum* and *Eucera,* with 88, 40 and 38 species, respectively. *Andrena* and *Lasioglossum* were represented by an average of 9.1 and 7.9 species respectively per site in the Med data set, while in the two other data sets no single genus contributed more than five species per site. In the Arid and the Temp datasets most subfamilies contained up to 10 species, whereas in the Med data set the majority of subfamilies contained more than 10 species, with more than a third containing more than 40 species. In all three data sets there was a significant positive correlation between the average number of species per site and the percentage of sites in which the genus was present (Arid: $\rho=0.74$ $p<0.001$; Med: $\rho=0.82$ $p<0.001$; Temp: $\rho=0.7$ $p<0.001$).

We found a marginally significant and positive effect of the genus/species ratio ($\varphi$) on the value of Mantel's r in community composition similarity ($r^2= 0.63$ $p=0.0592$). When there was a large number of species per genus, the similarity between the species-level community matrix and the genus-level community matrix was low; this similarity increased as the average number of species per genera decreased.

### 3) Improving the HTA's performance by adding body size measurements

The addition of body size classifications to the genus data increased the number of groups in the surrogate matrix from 37 to 83. This significantly improved the correlation to species level data ($t=-11.49$ $p<0.001$) and explained an additional 16% of the variation in species composition compared to using genus data only (genus+size: $r= 0.7$ $p<0.001$, genus only: $r=0.57$ $p<0.001$).

## Discussion

Our results show that in bees, a species-rich taxon, the reliability of genus-level data as a surrogate for species-level patterns may differ considerably among regions and community measures (richness vs. composition) and have limited reliability in some circumstances. Most studies of the HTA using genus-level data concluded that the HTA was an efficient surrogate for direct species identification in arthropods and in small-scale studies (Cardoso et al. 2004; Lovell et al. 2007; Mandelik et al. 2007). Similar results were observed in our study in both the desert and the temperate biomes, where genus-level data explained 84-87% of the variance in species composition. However, the use of genus-level data in the Mediterranean biome explained only

28% of this variance, indicating that even genus-level surrogates may have limited utility to describe patterns of community composition in this setting. Moreover, the HTA performed relatively well in reflecting variation among sites in species richness in the Mediterranean biome but performed only moderately well in reflecting this variance in the desert (representing 70% and 55% of the variation respectively).

As expected, the use of subfamily data reduced the performance of the HTA (relative to genera) in all three biomes (Cardoso et al. 2004; Lovell et al. 2007; Mandelik et al. 2007). However this reduction was relatively moderate for the composition measures and more variable and extensive for the richness measures; a 2-14% vs. 10-42% reduction in representation was found when using subfamily instead of genus for composition vs. richness measures respectively. This suggests that for bee composition assessment using subfamily instead of genus level data will cause only limited loss of information, while for richness assessment loss of information may be considerable and hard to predict.

We also tested the ability of the HTA to identify differences among bee communities in different habitats. In these tests both genus and subfamily level surrogates produced good results; we found that genus and subfamily data could be used to discriminate differences in community composition among habitats, and that they performed as well as species-level data in this regard. This is an encouraging result, because it suggests that higher taxon surrogates may be adequate to describe differences in community composition among different habitat types or land use categories. We emphasize however that both studies included in this analysis were part of the desert or the temperate data sets, in which the HTA performed well. A similar test in the Mediterranean data set might lead to a different result. Furthermore, the reliability of the HTA is known to be dependent on sample size (Neeson et al. 2013) and is thus limited by the completeness of samples.

We explored two factors that are hypothesized to affect the performance of the HTA. The first was the life history trait similarities of taxonomically related species. In all three data sets, species of the same genus showed a weak similarity in the composition of sites in which they were found. However, this similarity did not affect the performance of the HTA. These findings are similar to the findings of Bevilacqua et al. (2011) in marine invertebrates. The lack of influence of species relatedness on the HTA might be viewed as a shortcoming since it suggests that this method has limited ecological grounds. On the other hand it implies that taxonomic revisions and shifts of species among higher taxa, a common procedure in small bodied taxa including bees (Michener 2007), will not influence the HTA by altering the similarity in life history traits among related species.

The second factor tested for influencing the HTA is the numerical properties of the distribution of species among higher taxonomic levels. We found a marginally significant correlation between the $\varphi$ ratio (the genus/species ratio) and the strength of the correlation between species and genus composition patterns; the marginal values are possibly due to the limited number of studies used for establishing this correlation. The $\varphi$ ratio accounted for 63% of the variation in species-genus correlation in compositional patterns, pointing to an important role of the $\varphi$ ratio on the performance of the HTA. Similarly, the $\varphi$ ratio was found to be the main factor influencing the HTA in marine invertebrates (Bevilacqua et al. 2011) and in a mathematical model of the factors affecting surrogate reliability (Neeson et al. 2013). However, our results exhibit also some contrasting patterns to those expected by the proposed effect of $\varphi$ on the HTA. The ratio between species and subfamily richness found in both the desert and temperate data sets was lower than the ratio of species to genera found in the Mediterranean data set. Yet, the correlation coefficients obtained when using subfamily level in the first two data sets were still considerably higher than the genus-level result obtained for the Mediterranean data set. These results suggest that other properties of the species distribution within higher taxonomic levels may also be influencing the performance of the HTA.

In all data sets species-rich genera were spread across sites and not clustered in specific habitats/sites; species-rich genera contributed on average more species per site and were found in a higher percentage of sites compared to species-poor genera. This implies that indeed a few species-rich genera may have a strong influence on species composition patterns and thus on the performance of the HTA. The Mediterranean data set in particular, where a genus-based HTA performed poorly, is distinguished by containing the three most species-rich genera for a single data set and the two genera that contributed the highest average number of species per site for a single data set. These species-rich genera lowered the φ ratio in the Mediterranean data set. Furthermore, their presence might lead to poor performance of the HTA in this data set because they had a strong influence, compared to other genera, on shaping species composition among sites.

The use of morphospecies can reduce the need for identification by expert taxonomists (Oliver and Beattie 1996). In this paper we explored the utility of combining parataxonomic classification with the HTA, as a way to decrease the species/higher taxa ratio and thus improve the performance of the HTA. To reduce the subjectivity of parataxonomy classification (Krell 2004) we used body size (ITD) because of its a priori advantages of being easily, consistently and objectively measured and having functional significance (Greenleaf et al. 2007). We found that body size can efficiently subdivide local bee assemblages, as reflected by the relatively large number of discrete size categories that could be classified (up to seven size categories per genera), compared to other often used morphological characteristics such as color patterns and hairiness. Subdividing genera by body size significantly increased the φ ratio (i.e. decreased the number of species per group) and improved the representation of compositional patterns by 16%. However, there should be a greater refinement in going to genus-size classification for small genera than for hyperdiverse genera. Therefore the incorporation of size categories will likely cause a differential improvement in the HTA's performance depending on the distribution of species/genus (the configuration of the hollow curve) at the site. Every characteristic that can sub-divide genera in a consistent way and decrease the species/HT ratio, while transcending males and females, will likely increase the performance of the HTA. The choice of a specific characteristic to be integrated in the HTA should be based on its cost-efficiency i.e. the cost of acquiring this information vs. the expected decrease in species/HT ratio, based on the subdivisions that will be gained.

The ability of higher taxon surrogates to describe ecological drivers of community composition depends on the ecological similarity of taxonomically-related species. For example, when a single family or genus contains species that differ markedly in ecological characteristics such as body size and nesting and foraging habitat preferences, differences in the abundance of that family or genus across space or time cannot provide any information about the ecological drivers of community composition. In our study, we did not test whether differences in the relative abundances of genera and families were due to differences in life history traits of the constituent species.

We found high variability in the performance of the HTA in representing local-scale richness and composition patterns of bee species from different biomes. This variability dictates a cautious implementation of the HTA in surveys of bees and other species-rich taxa, and the need for an a priori evaluation of its expected performance in the designated region. The expected performance of the HTA in bees is affected mostly by numerical properties of the distribution of species among higher taxonomic levels. The ratio between species and higher taxa richness in particular is an important factor. In addition, the presence of a few species-rich groups can further weaken the ability of the HTA to represent patterns of species composition. Therefore, prior to applying the HTA, the characteristic species/higher taxa ratio and the presence of exceptionally rich groups in the surveyed region should be evaluated. This evaluation obviously relies on some prior knowledge of the bee fauna in the region; in cases where such prior knowledge is completely absent evaluating the ratio based on a subset of the collected material should be considered. When species/higher taxa ratio indicates possibly poor performance of the HTA, integration of morphological characteristics that can consistently subdivide genera and increase the ratio can also improve the performance of the HTA. Alternatively, identification of only a few species-rich genera to species level may improve the HTA dramatically without the cost of full species identification, although in bees, species-rich genera are typically the most difficult groups to identify to species.  The use of data collected from different biomes and from several types of habitats in this study enhances the reliability of our results and enabled us to identify both general trends and specific community structure features that affect the HTA at local scales. Our results may be applicable for the evaluation of the HTA in other small-bodied and species-rich groups and contribute to the cost-effectiveness of biodiversity surveys.

## Acknowledgments

We are grateful to the taxonomists who identified our bees (full list in appendix A); We thank two anonymous reviewers for helping to improve the paper. This study was funded by the Israel Science Foundation Research Grant No. 919/09**.** I.V.R was supported by the Ronald Appleby Fund and the Samuel and Lottie Rudin scholarship.

## Literature

Aizen, M.A. and Harder, L.D. (2009) The Global Stock of Domesticated Honey Bees Is Growing Slower Than Agricultural Demand for Pollination. *Current Biology,* **19,** 915-918.

Baldi, A. (2003) Using higher taxa as surrogates of species richness: a study based on 3700 Coleoptera, Diptera, and Acari species in Central-Hungarian reserves. *Basic and Applied Ecology,* **4,** 589-593.

Balmford, A., Jayasuriya, A.H.M. and Green, M.J.B. (1996) Using higher-taxon richness as a surrogate for species richness .2. Local applications. *Proceedings of the Royal Society of London Series B-Biological Sciences,* **263,** 1571-1575.

Batley, M. and Hogendoorn, K. (2009) Diversity and conservation status of native Australian bees. *Apidologie,* **40,** 347-354.

Beals, E., W. (1984) *Bray-curtis ordination: an effective strategy for analysis of multivariate ecological data*. Academic Press, London, ROYAUME-UNI.

Bennett, A.F., Radford, J.Q. and Haslem, A. (2006) Properties of land mosaics: Implications for nature conservation in agricultural environments. *Biological Conservation,* **133,** 250-264.

Bevilacqua, S., Terlizzi, A., Claudet, J., Fraschetti, S. and Boero, F. (2011) Taxonomic relatedness does not matter for species surrogacy in the assessment of community responses to environmental drivers. *Journal of Applied Ecology,* **49,** 357-366.

Brown, M. and Paxton, R. (2009) The conservation of bees: a global perspective. *Apidologie,* **40,** 410-416.

Cane, J.H. (1987) Estimation of bee size using intertegular span (Apoidea). *Journal of Kansas Entomological Society,* **60**, 145–147.

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

Dauvin, J.C., Gomez Gesteira, J.L. and Salvande Fraga, M. (2003) Taxonomic sufficiency: an overview of its use in the monitoring of sublittoral benthic communities after oil spills. *Marine Pollution Bulletin,* **46,** 552-555.

Dutilleul, P, Stockwell, J, D., Frigon, D and Legendre (2000) *The Mantel test versus Pearson's correlation analysis : Assessment of the differences for biological and environmental studies*. American Statistical Association, Washington, DC, ETATS-UNIS.

Eardley, C.D., Gikungu, M. and Schwarz, M.P. (2009) Bee conservation in Sub-Saharan Africa and Madagascar: diversity, status and threats. *Apidologie,* **40,** 355-366.

Gaston, K.J. and Williams, P.H. (1993) Mapping the World's Species-The Higher Taxon Approach. *Biodiversity Letters,* **1,** 2-8.

Gladstone, W. and Alexander, T. (2005) A test of the higher-taxon approach in the identification of candidate sites for marine reserves. *Biodiversity and Conservation,* **14,** 3151-3168.

Gotelli, N.J. (2004) A taxonomic wish-list for community ecology. *Philosophical transactions of the Royal society B Biological Sciences,* **359,** 585-597.

Gotlieb, A., Hollender, Y. and Mandelik, Y. (2011) Gardening in the desert changes bee communities and pollination network characteristics. *Basic and Applied Ecology,* **12,** 310-320.

Greenleaf, S., Williams, N., Winfree, R. and Kremen, C. (2007) Bee foraging ranges and their relationship to body size. *Oecologia,* **153,** 589-596.

Guzman-Alvis, A.I. and Carrasco, F. (2005) Taxonomic aggregation and redundancy in a tropical macroinfaunal assemblage of the southern Caribbean in the detection of temporal patterns. *Scientia Marina,* **69,** 133-141.

Harding, J.S., Benfield, E.F., Bolstad, P.V., Helfman, G.S. and Jones, E.B.D. (1998) Stream biodiversity: The ghost of land use past. *Proceedings of the National Academy of Sciences,* **95,** 14843-14847.

Hirst, A.J. (2008) Surrogate measures for assessing cryptic faunal biodiversity on macroalgal-dominated subtidal reefs. *Biological Conservation,* **141,** 211-220.

Hodkinson, I. and Jackson, J. (2005) Terrestrial and Aquatic Invertebrates as Bioindicators for Environmental Monitoring, with Particular Reference to Mountain Ecosystems. *Environmental Management,* **35,** 649-666.

Kessler, M., et al. (2011) Cost-effectiveness of plant and animal biodiversity indicators in tropical forest and agroforest habitats. *Journal of Applied Ecology*, **48**, 330-339.

Klein, A., Vaissiere, B., Cane, J., Steffan-Dewenter, I., Cunningham, S., Kremen, C. and Tscharntke, T. (2007) Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B-Biological Sciences,* **274,** 303-313.

Krell, F.-T. (2004) Parataxonomy vs. taxonomy in biodiversity studies–pitfalls and applicability of ‘morphospecies’ sorting. *Biodiversity and Conservation,* **13,** 795-812.

Lawton, J.H. and Strong, D.R., Jr. (1981) Community Patterns and Competition in Folivorous Insects. *The American Naturalist,* **118,** 317-338.

Lovell, S., Hamer, M., Slotow, R. and Herbert, D. (2007) Assessment of congruency across invertebrate taxa and taxonomic levels to identify potential surrogates. *Biological Conservation,* **139,** 113-125.

Mandelik, Y., Dayan, T., Chikatunov, V. and Kravchenko, V. (2007) Reliability of a higher-taxon approach to richness, rarity, and composition assessments at the local scale. *Conservation Biology,* **21,** 1506-1515.

Mandelik, Y., Roll, U. and Fleischer, A. (2010) Cost-efficiency of biodiversity indicators for Mediterranean ecosystems and the effects of socio-economic factors. *Journal of Applied Ecology,* **47,** 1179-1188.

Mandelik, Y., Winfree, R., Neeson, T. and Kremen, C. (2012) Complementary habitat use by wild bees in agro-natural landscapes. *Ecological Applications,* **22,** 1535-1546.

Marshall, J.C., Steward, A.L. and Harch, B.D. (2006) Taxonomic resolution and quantification of freshwater macroinvertebrate samples from an Australian dryland river: The benefits and costs of using species abundance data. *Hydrobiologia,* **572,** 171-194.

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

Michener, D.C. (2007) *The Bees of the World,* 2nd edn. Johns Hopkins University Press, Baltimore, Maryland, USA.

Muniz, P. and Pires-Vanin, A.M.S. (2005) More about taxonomic sufficiency: a case study using polychaete communities in a subtropical bay moderately affected by urban sewage. *Ocean Science Journal,* **40,** 127-143.

Nakamura, A., Catterall, C.P., House, A.P.N., Kitching, R.L. and Burwell, C.J. (2007) The use of ants and other soil and litter arthropods as bio-indicators of the impacts of rainforest clearing and subsequent land use. *Journal of Insect Conservation,* **11,** 177-186.

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

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

Ollerton, J., Winfree, R. and Tarrant, S. (2011) How many flowering plants are pollinated by animals? *Oikos,* **120,** 321-326.

Peltzer, D.A., Bast, M.L., Wilson, S.D. and Gerry, A.K. (2000) Plant diversity and tree responses following contrasting disturbances in boreal forest. *Forest Ecology and Management,* **127,** 191-203.

Pisanty, G., Klein, A.M. and Mandelik, Y. (2014) Do wild bees complement honeybee pollination of confection sunflowers in Israel? *Apidologie,* **45,** 235-247.

Potts, S.G., Roberts, S.P.M., Dean, R., Marris, G., Brown, M., Jones, R. and Settele, J. (2009) Declines of managed honey bees and beekeepers in Europe. *Journal of Apicultural Research,* **49,** 15-22.

Potts, S.G., Vulliamy, B., Roberts, S., O'Toole, C., Dafni, A., Ne'eman, G. and Willmer, P. (2005) Role of nesting resources in organising diverse bee communities in a Mediterranean landscape. *Ecological Entomology,* **30,** 78-85.

Prance, G.T. (1994) A Comparison of the Efficacy of Higher Taxa and Species Numbers in the Assessment of Biodiversity in the Neotropics. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* **345,** 89-99.

Steffan-Dewenter, I. and Tscharntke, T. (1999) Effects of habitat isolation on pollinator communities and seed set. *Oecologia,* **121,** 432-440.

Tulloch, A., Possingham, H.P. and Wilson, k. (2011) Wise selection of an indicator for monitoring the success of management actions. *Biological Conservation*, **144**, 141-154.

vanEngelsdorp, D., Evans, J.D., Saegerman, C., Mullin, C., Haubruge, E., Nguyen, B.K., Frazier, M., Frazier, J., Cox-Foster, D., Chen, Y., Underwood, R., Tarpy, D.R. and Pettis, J.S. (2009) Colony Collapse Disorder: A Descriptive Study. *PLoS ONE,* **4,** e6481.

Webb, C.O., Ackerly, D.D., McPeek, M.A. and Donoghue, M.J. (2002) Phylogenies and Community Ecology. *Annual Review of Ecology and Systematics,* **33,** 475-505.

Williams, N.M., Crone, E.E., Roulston, T.a.H., Minckley, R.L., Packer, L. and Potts, S.G. (2010) Ecological and life-history traits predict bee species responses to environmental disturbances. *Biological Conservation,* **143,** 2280-2291.

Williams, N.M., Minckley, R.L. and Silveira, F.A. (2001) Variation in native bee faunas and its implications for detecting community changes. *Conservation Ecology*, **5(1)**, article 7

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

Winfree, R., Williams, N.M., Dushoff, J. and Kremen, C. (2007) Native bees provide insurance against ongoing honey bee losses. *Ecology Letters,* **10,** 1105-1113.

Xiao-Li, M., Rosenthal, R. and Rubin, D.B. (1992) *Comparing correlated correlation coefficients*. American Psychological Association, Washington, DC, ETATS-UNIS.

Zimmerman, G.M., Goetz, H. and Mielke, P.W., Jr. (1985) Use of an Improved Statistical Method for Group Comparisons to Study Effects of Prairie Fire. *Ecology,* **66,** 606-611.

**Ecological Archives material**

Appendix A: Information on unpublished studies included in the analysis and a list of the taxonomists who identified the specimens

**Table 1**. Data sets used in this study

| Data set (number of sites) | Number of bee specimens in data set | Studies in each data set (number of bee specimens) | Species/Genera ratio | Number of sites included in data analysis |
|---|---|---|---|---|
| **Arid** (28) | 6099 | Gardens and natural habitats (2472) | 3.57 | 8 |
| | | Agriculture habitats (3627) | 3.33 | 20 |
| **Mediterranean** (46) | 6181 | Planted pine forests and natural habitats (1346) | 6.67 | 10 |
| | | Agriculture and natural habitats (4835) | 9.09 | 36 |
| **Temperate** (38) | 7306 | Forests and agriculture habitats (2045) | 3.85 | 18 |
| | | Abandoned agriculture fields (5261) | 3.85 | 20 |
| **Total** | 19,586 | | | 112 |

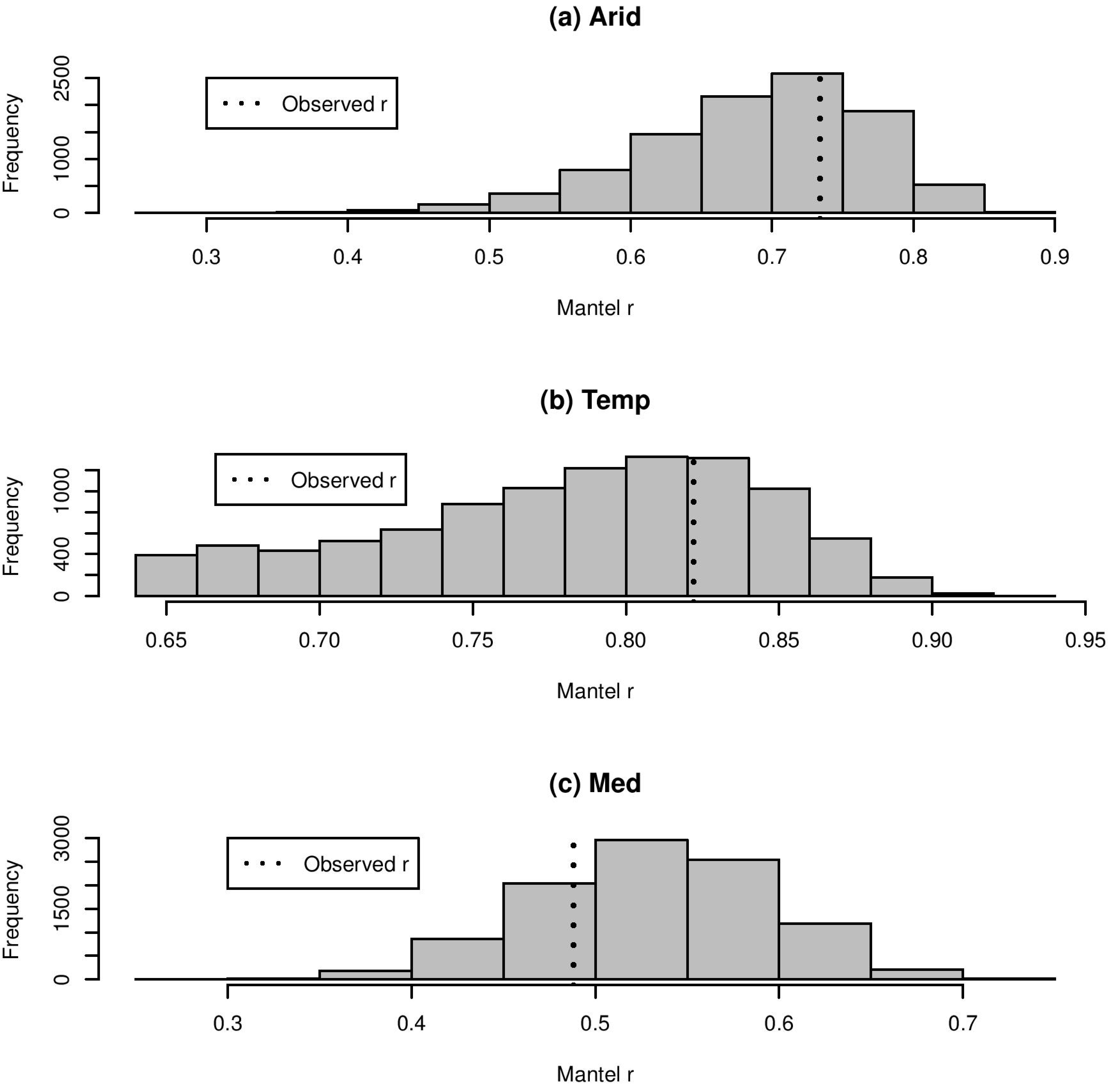
## Figure legends

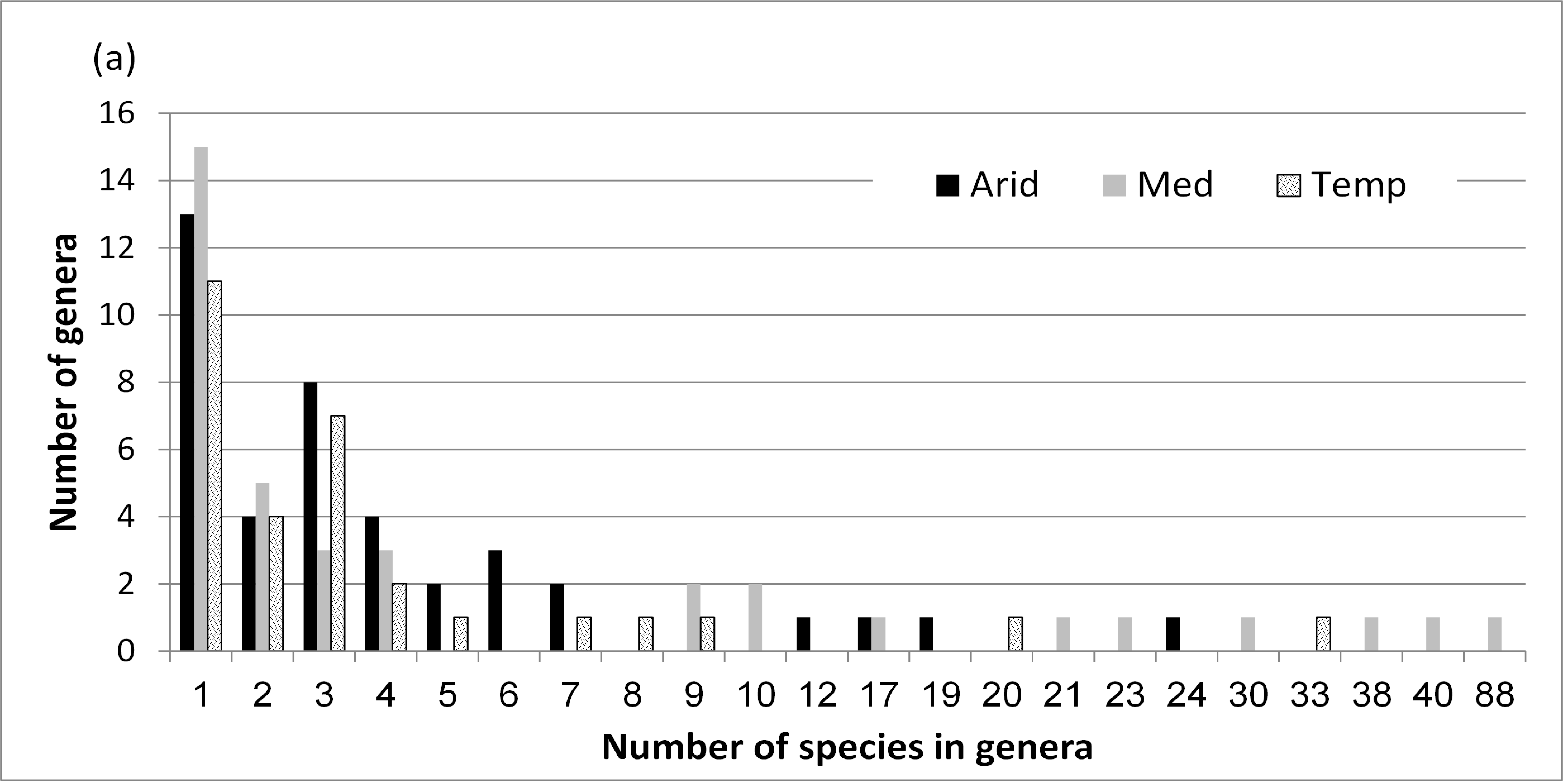
**Figure 1.** NMS results showing community composition similarity among habitats for species, genus and subfamilies in the (a) study form the Arid data set, comparing gardens and natural areas (b) study from the Temperate data set, comparing old fields, agriculture fields and forests. * $p<0.05$ in the MRPP.

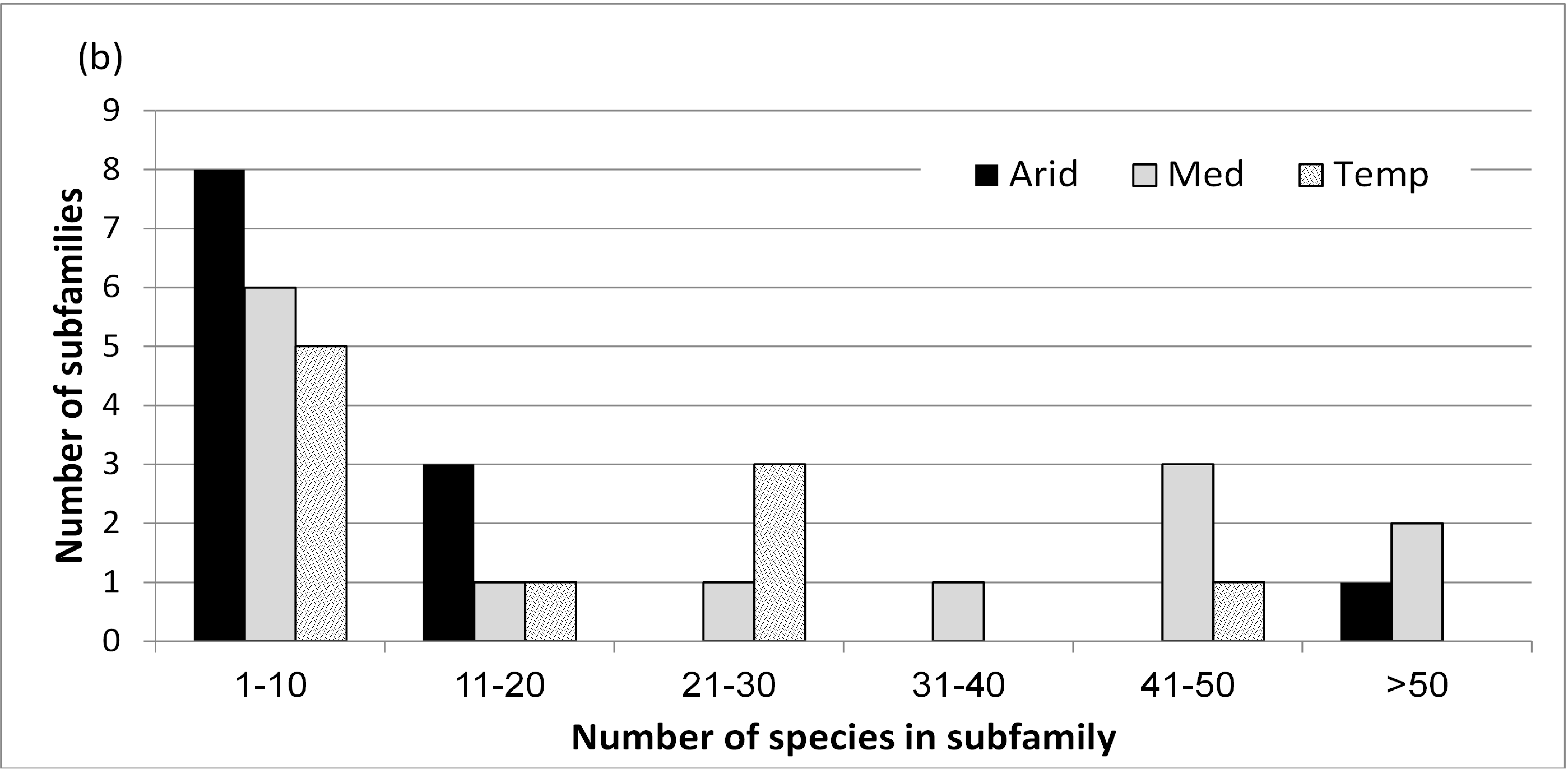
**Figure 2.** Distribution of Mantel r values from 10,000 runs where species were randomly assigned to genera. Dots indicate the observed mantel r value when following the current taxonomic division (a) Arid data set (b) Temperate data set and (c) Mediterranean data set.

**Figure 3.** Species distribution within (a) genera (b) subfamilies, in the three data sets.

(a)
Species
Genus
Subfamily
Garden
Natural
A=0.276*
A=0.350*
A=0.343*
(b)
Species
Genus
Subfamily
Old Field
Agriculture
Forest
A=0.194*
A=0.241*
A=0.285*


(a) Arid
Frequency
0
1000
2500
Observed r
0.3
0.4
0.5
0.6
0.7
0.8
0.9
Mantel r
(b) Temp
Frequency
0
400
1000
Observed r
0.65
0.70
0.75
0.80
0.85
0.90
0.95
Mantel r
(c) Med
Frequency
0
1500
3000
Observed r
0.3
0.4
0.5
0.6
0.7
Mantel r


(a)
Number of genera
16
14
12
10
8
6
4
2
0
Arid
Med
Temp
1
2
3
4
5
6
7
8
9
10
12
17
19
20
21
23
24
30
33
38
40
88
Number of species in genera


(b)
Number of subfamilies
9
8
7
6
5
4
3
2
1
0
Arid
Med
Temp
1-10
11-20
21-30
31-40
41-50
>50
Number of species in subfamily