

# Spatial assortment of soil organisms supports the size-plasticity hypothesis

Xin Sun (✉ [xsun@iue.ac.cn](mailto:xsun@iue.ac.cn))

Institute of Urban Environment, Chinese Academy of Sciences

Alain Isabwe

Haifeng Yao

Shixiu Zhang

<https://orcid.org/0000-0003-2192-7502>

Yuji Jiang

Institute of Soil Science <https://orcid.org/0000-0001-8290-3518>

Martin Breed

Flinders University <https://orcid.org/0000-0001-7810-9696>

---

## Article

**Keywords:** soil biota, community assembly, niche breadth, urbanization, DNA metabarcoding, conservation

**Posted Date:** August 30th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1896194/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

The size-plasticity hypothesis posits that larger size organisms are less plastic in their metabolic rates and therefore are more strongly environmental-filtered than smaller organisms. Many studies have supported this hypothesis by evaluating the relative roles of environmental filtration and dispersal for different taxonomic groups of soil organisms. Most observations are made at large spatial scales, which are assumed to have a wide array of varying habitats. However, since urbanization causes habitat fragmentation at smaller regional scales, testing the size-plasticity hypothesis at this scale would help better understand the spatial assortment of urban soil organisms which, in turn, would help to develop improved management and conservation strategies of urban soil health. Here, we used DNA metabarcoding on five groups of soil biota (bacteria, fungi, protists, nematodes, and soil invertebrates) to assess the relative importance of dispersal and environmental filters to examine the size-plasticity hypothesis at this spatial scale in an urban environment. We observed strong distance-decay of community similarities associated with higher levels of stochastic changes in bacteria, nematode and protist communities but not fungal or invertebrate communities. Bacterial communities occupied the widest niche followed by protists and nematodes, potentially because of their higher dispersal abilities compared to the larger soil organisms. Null deviation of communities varied with taxonomic groups where bacteria and nematodes were mainly driven by homogenizing dispersal, protists and fungi by drift and soil invertebrates by environmental selection. We further identified a small percentage of locally-adapted taxa (2.1%) that could be focal taxa for conservation and restoration efforts by, for example, restoring their habitats and enhancing their regional connectivity. These results support the size-plasticity hypothesis at the relatively unexplored regional scale in an urbanization context, and provide new information for improving urban soil health and sustainable city models and urban soil health.

## Introduction

Soil biota represents a large fraction of global biodiversity and is crucial for biochemical cycling, regulation of primary productivity, biomass decomposition and maintenance of soil health [1, 2]. Increasing evidence shows that soil biota is being directly and indirectly impacted by climate change, population growth, agricultural intensification and urbanization [1, 3–6]. These activities may lead to soil biota extinction, alteration of their trophic dynamics and increasing spatial fragmentation of their habitats [5, 6]. It is therefore crucial to evaluate the spatial patterns and ecological processes that shape soil biota to help enable improved and more targeted conservation and restoration approaches [7].

Global expansion of urban areas in recent years has induced substantial changes in soil physicochemical characteristics which in turn influence soil biota. Urbanization causes land-use transitions through which fragmentation of soil habitats is inevitable [5, 8]. This fragmentation, from large continuous natural areas to small habitat patches, has strong effects on soil biodiversity [9, 10]. Among the habitat patches in urban landscapes, green spaces serve as increasingly important reservoirs of soil biodiversity as the surrounding areas are developed [11, 12]. Given that spatial patterns of soil biota should be considered

when assessing the ecological conditions of urban soil biota [13], our knowledge about spatial assortment of different taxonomic groups of organisms in urban soils is crucial but poorly developed.

Soil biota encompasses both micro- and macro-organisms and is extremely variegated in terms of their morphology, biomass, diversity and lifestyles. Their spatial patterns are generally shaped by both deterministic and stochastic processes [14–16]. These two distinct but complementary ecological processes maintain species diversity at local, regional and global scales [17–19]. Deterministic processes are dominated by environmental filters which result from habitat preferences of species. Stochastic processes manifest mainly as random changes in species compositions [20]. Although these processes act simultaneously, their relative roles vary considerably across spatial scales [17, 18] and among different taxonomic groups [21]. Previous studies have revealed these processes for different taxonomic groups of soil biota in forests [21], farmlands [22], and in contaminated soils [23]. Evidence from large spatial scales (e.g., continental-scale) has supported the size-plasticity hypothesis which posits that smaller organisms are more plastic in their metabolic rates and therefore less impacted by environmental filters than larger organisms [24, 25]. According to the size-plasticity hypothesis, dispersal abilities should be reduced for the larger than the smaller taxonomic groups of soil biota due to the smaller organisms having greater metabolic plasticity and greater environmental tolerance. In other words, the strength of deterministic processes should increase with increasing sizes of organisms, in such a way that smaller organisms should occupy more diverse habitats than larger ones.

These expectations are also reflected in the niche breadth concept, which refers to the diversity of resources used or environments occupied by an individual, population, species, or clades [26]. Briefly, an organism with a broader niche should be more metabolically flexible at the community level. Both the niche concept and the size-plasticity hypothesis are central to estimating the risks of community collapse in a changing environment because smaller organisms have been assumed to be at greater risk of extinction as a result of higher stochastic fluctuations [27]. The two concepts, however, have rarely been integrated to detect taxa that are prone to dispersal or environmental filters. Moreover, studies investigating the relative importance of environmental filters and dispersal potential for different taxonomic groups of soil biota at regional scales are lacking. Testing this hypothesis at a regional scale is important as larger spatial scales allow for a greater diversity of organisms to coexist through ecological selection (e.g., latitudinal gradients of a taxonomic group might commonly vary by orders of magnitude between high and low latitudes), plus, from an applied perspective since conservation and restoration approaches are often implemented at regional (or smaller) scales [20, 28].

Recently, Zinger and colleagues showed that the size of soil biota was an important determinant of community assembly, and soil mesofauna were more stochastically-assembled than microorganisms in a tropical forest of French Guiana [21]. Hypothetically, the regional species pool should be reduced into taxa that generally occupy a particular site due to dispersal and environmental filters. Smaller soil organisms should occupy a broader niche than the larger organisms because they disperse more easily (Fig. 1). Only limited studies have jointly compared assembly processes in soil micro- and macroorganisms, due in part to technical limitations in characterizing soil biota. Soil biomonitoring has

greatly advanced recently via the advancement of DNA sequencing approaches, such as metabarcoding which has created an opportunity for rapid generation of soil biota data. This breakthrough has allowed detailed and comprehensive characterization of soil biota across space and time [29, 30].

In this study, we evaluated the strength of environmental filters and dispersal potential across different groups of soil biota along an urbanization gradient in a subtropical region. We hypothesized a reduced influence of dispersal and greater influence of environmental filters on the larger soil organisms compared with smaller organisms. Accordingly, we expected stochastic processes to dominate soil microorganisms. Specifically, our study aimed at answering three questions: 1) how do environmental filters act on different taxonomic groups of soil organisms along an urbanization gradient? 2) Do soil bacteria communities exhibit stronger dispersal abilities than soil macro-organisms? 3) Is the broader niche breadth being occupied by smaller than larger soil organisms? Our results support the size-plasticity hypothesis and, *in tandem* with the niche breadth concept, yielded important insights for urban soil biota conservation and management at a time when the world's population is more urbanized than ever before and continues to expand at an unprecedented rate [31].

## Materials And Methods

### Study area and sampling design

This study was conducted in Xiamen, Fujian, China. Xiamen has a subtropical monsoon climate with an average annual temperature of 20.7°C and an average annual rainfall of 1335.8 mm [32]. Under rapid urban development, Xiamen has experienced a decrease in farmland and an increase in built-up land by, respectively, 14.0% and 15.7% from 1980 to 2015 [33]. The sampling campaign was carried out in December 2020, the season during which optimum soil diversity is observed in the subtropics. Sampling plots were established based on the level of urbanization from fully urban areas in south east towards the mountain forests in the northwest. Briefly, four plots in four urban, four suburban parks, and three forest sites were surveyed. The sampled urban parks were Zhongshan Park (118.09E, 24.46N), Nanhu Park (118.11E, 24.48N), Zhonglun Park (118.15E, 24.49N) and Wuyuanwan Wetland (118.17E, 24.52N). The sampled suburban parks were Jiageng Park (118.11E, 24.57N), Jingxian Park (118.11E, 24.58N), Ningbao Park (118.06E, 24.57N) and Ridong Park (118.04E, 24.57N). The sampled forest plots were located in Huweishan (118.09 E, 24.49N), Xianyueshan (118.09E, 24.50N) and Tianzhusan (117.92E, 24.59N) forests. Each plot in the urban, suburban parks and forest was selected based on the dominant vegetation type including grassland, grassland and shrub, grassland, trees and shrub, or tree with grassland. Then three soil samples spanning a distance of ca. 100 cm from each other were collected by using a stainless soil sampling tube, thoroughly mixed to make a composite sample and then transported to the laboratory in clean polystyrene bags.

### Environmental variables

Immediately after sample collection, soil pH and moisture were measured by using a pH meter and gravimetric methods, respectively. The total organic carbon (TOC) was measured using the potassium

dichromate oxidation method with external heat. The concentration of total nitrogen (TN) and total carbon (TC) were quantified by using TOC/TN-VCPH elemental analyzer (Shimadzu, Kyoto, Japan). The resulting data were used to estimate the soil carbon-nitrogen ratio (C/N). Furthermore, the inductively coupled plasma mass spectrometry with ICP-MS Agilent 7500CX (Agilent Technologies, Santa Clara, CA, USA) was used to quantify the concentrations of total phosphorus (TP) and total sulphur (TS) through batch triacid digestion.

## DNA extraction and soil biota profiling

About 10 g of soil samples were thoroughly homogenized and 500 mg were sub-sampled for DNA extraction using the MP FastDNA spin kit (MP Biomedicals, Solon, OH, USA). The soil DNA was subjected to high-throughput sequencing on Illumina NovaSeq PE 250 (Illumina Inc., San Diego, CA, USA) by using a paired-end method. Four genes encoding for different soil organisms were targeted through specific primer pairs designed for PCR amplification (Table 1). The obtained paired-end sequence data were processed with QIIME2 as follows: the raw sequence data were processed using the DADA2 pipeline to filter out low quality sequences, denoising, and removing chimeras and obtain amplicon sequence variants (ASVs). Taxonomic assignments were performed by matching the representative sequence to specific annotation databases for each gene target (Table 1). To allow better comparison of distribution patterns among the studied soil organisms, the obtained ASV tables were rarefied at the same cut-off (*i.e.*, 20 000 sequences per sample) by using the *phyloseq* package in R.

Table 1  
Differentiating soil biota into five taxonomic groups based on the sequencing of specific gene targets and associated annotation database.

Taxonomy	Region	Primer pair [Ref.]	Annotation database [Ref.]
Bacteria	16S rRNA	515F/907R [34]	Greengenes [35]
Fungi	ITS	ITS1F/ITS2R [36]	Unite [37]
Protists	18S rRNA	TAReukFWD1F/TAReukREV3R [38]	PR2 [39]
Nematodes	18S rRNA	NF1/18Sr2b [40]	SILVA [41]
Invertebrates	COI	mICOI-intF/jgHCO219 [42]	COI database [43]

## Data analysis and visualization

Data analyses were carried out in R (v. 4.0) [44] and visualizations were aided by the packages *ggplot2* and *ggpubr* [45, 46]. Quantified environmental variables in the urban, suburban and forest plots were compared using the Kruskal-Wallis rank sum tests, and were visualized in a Principal Component Analysis (PCA) ordination. Prior to ordination, the environmental variables were log (x + 1) transformed, with exception of the pH. The function ‘autoplot’ of the package *ggfortify* was used to perform PCA. The Bray–Curtis similarity matrices were calculated using the *vegan* packages. These matrices were regressed against a matrix of pairwise geographical distances between the sites [47] to obtain distance-decay relationships as means to evaluate distribution patterns of the five taxonomic groups. The

relationships between spatial and community matrices were assessed using Mantel tests, and the significance level was estimated after 999 permutations as per *ecodist* package [48]. The community dissimilarity matrices measured with Jaccard index were further regressed against a pairwise Euclidean distance of the sampled environmental variables to detect community-environment relationships. Then the Spearman's correlations between the matrices and each of the measured environmental variables were calculated and visualized using *ggcor* package.

## Neutral community model

The neutral community model was used to test the importance of stochastic processes for each taxonomic group as means to test the size-plasticity hypothesis. The neutral community model assumes that stochastic processes are associated with a significant relationship between the abundance and occurrence frequency of taxa collected across different local sites in a region [49]. In practice, the neutral community model returns a parameter  $Nm$  which is an estimate of dispersal between communities with  $N$  being the metacommunity size (*i.e.*, total number of sequences) and  $m$  is the immigration rate. The model fit was evaluated using the proportion of local communities in which each ASV is detected, and their abundance in the metacommunity which represents regional species pool. This was done by using non-linear least-squares method with the *minpack* package to generate the best fit distribution curve. ASVs above and below 95% confidence intervals of the neutral community model fit are assumed not to be stochastically distributed [50, 51].

## Null deviation

To improve our observations and to confirm the size-plasticity hypothesis, the neutral community model was complemented with null deviation of communities. This model uses modified Raup–Crick dissimilarity metric that is independent of sample size and allows to calculate the proportion of observed dissimilarities across samples that are higher than those estimated from the null model after 1000 randomization [52]. The output results range from  $-1$  to  $1$ , with values approaching  $-1$  (*i.e.*,  $-0.95$  to  $-1$ ) indicating homogenizing dispersal, values near  $1$  ( $0.95$  to  $1$ ) indicating environmental selection, and values  $-0.95$  to  $0.95$  indicating drift.

## Niche breadth

Through the niche breadth concept, we expected that a taxonomic group with a wider niche breadth to be more metabolically flexible [53]. The niche breadth was calculated by using Levins' niche breadth index ( $B$ ):

$$B_j = 1 / \sum_{j=1}^N P_{ij}^2$$

where  $B_j$  is the habitat niche breadth of an ASV  $j$  in the whole species pool;  $N$  is the total number of communities of each metacommunity;  $P_{ij}$  is the proportion of an ASV  $j$  in community  $i$  [53]. In this model, high  $B$  indicates that an ASV occurs widely and evenly along a wide range of sampling sites resulting in the wider niche breadth. The average  $B$ -values from all taxa in a single community as an indicator of

habitat niche breadth at the community level. Whereas habitat specialists should have low  $B$ -values, habitat generalists should have higher  $B$ -values and be more evenly distributed along a wider range of habitats compared with habitat specialists. We ran this analysis using functions available in the *EcoUtil* package [54].

## Identifying locally-adapted taxa

In this study, locally-adapted taxa were defined as taxa that emerge from the regional species pool to become suitable to only a habitat type (Fig. 1). We assumed that each of the studied soil organisms (ASV level) underwent two major ecological filters before being observed at a given ecosystem type. First, some taxa are dispersal-filtered. These taxa are expected to be above or below the neutral community model as they may not fit the model distribution due to low dispersal abilities. Second, among taxa that are found above and below the model predictions, some taxa may be environmental-filtered. Hypothetically, these taxa may reach a given locality but are unable to adapt to local environmental conditions (e.g., under selective pressure of soil moisture, pH etc.). Through the niche breadth, these taxa are the specialists because of their abilities to cope with habitat-specific environmental condition. They have significantly low  $B$ -values compared to expected patterns under 999 randomizations. To improve the accuracy of these taxa, we selected the locally-adapted taxa at the intersection between the specialist taxa and taxa above or below the neutral community model predictions. The responses of the locally-adapted taxa to environmental changes were assessed by means of redundancy analysis (RDA).

## Results

### Environmental overview

The concentrations of total sulphur and total phosphorus were significantly higher ( $p < 0.05$ ) in the urban than suburban parks and forest sites (Fig. S1). Soil pH was significantly lower ( $p < 0.05$ ) in forests than the urban and suburban parks, and there was no clear separation between the suburban and urban sampling sites when all samples were pooled in a PCA ordination (Fig. S2). Other environmental variables such as the moisture, TN, TOC and TC did not vary significantly among the three ecosystem types ( $p > 0.05$  in each case).

### Stronger distance-decay of community similarity for soil bacteria than other soil biota

The rarefied communities (Fig. S3) were used for further analyses including analysis of spatial change of communities through distance-decay relationships supplemented with Mantel tests. This analysis revealed clear patterns of different taxonomic groups of soil biota (Fig. 2). Distance-decay relationships were stronger for bacteria ( $r = -0.51$ ;  $p < 0.001$ ) than for fungi ( $r = -0.33$ ,  $p < 0.001$ ). Similar level of soil community distance decay was observed for nematodes ( $r = -0.32$ ,  $p < 0.001$ ) and protists ( $r = -0.31$ ,  $p < 0.05$ ). Compared with other groups, soil invertebrates exhibited an opposite trend of distance-decay relationships ( $r = 0.33$ ,  $p < 0.01$ ). The comparison of community similarity matrices of the five taxonomic

groups showed that the similarities were significantly different among the groups ( $p < 0.001$ ) except between protists and nematodes (Fig. S4). Despite the low number of amplicon sequence variants (ASVs), both protists and nematodes had higher community similarity than fungi and invertebrates. The lowest community similarity was observed for invertebrates within the studied region that spans ca. 27 km between the most distant sites.

## High levels of stochastic change in bacteria and nematodes but not fungi, protists or invertebrates

The neutral community model was used to assess the effect of stochastic processes on the assembly of each group of soil biota. Stochastic processes accounted for 69% of the variation in bacterial community assembly and 50% for nematodes. Like protists (28%), the role of stochastic processes in structuring fungi (< 1%) and invertebrates (< 1%) was low and not significant (Fig. 3A). The parameter  $Nm$  determines the relationship between the occurrence frequency, the size of the regional species pool and the level of immigration rate was greatest for bacteria ( $Nm = 556$ ), followed by protists ( $Nm = 82$ ), fungi ( $Nm = 32$ ), and nematodes ( $Nm = 3$ ) (Fig. 3A). Only small percentages of taxa (6%-10%) were observed above and below the neutral model prediction (95% confidence intervals) as compared to the whole metacommunity size for all taxonomic groups (Fig. 3B). Specifically, there was a greater percentage of taxa that occurred more or less frequently than predicted by neutral community model for fungi (10%), followed by invertebrates (9.7%), bacteria (8.6%), protists (7.4%) and nematodes (6.3%; Fig. 3B). At the lower taxonomic levels, the neutral community model fit generally decreased in order of taxonomic classes of bacteria, nematodes, protists, fungi and invertebrates (Fig. 3C).

## Environmental filtration acts more strongly on soil macro-organisms than microorganisms

By partitioning the effects of selection, dispersal and drift using null deviation of communities for different taxonomic groups, we found that the average deviation was  $-0.39$  and  $-0.14$  for bacteria and nematodes, respectively (Fig. 4A). While fungal and protist communities had almost similar null deviation of  $-0.06$  and  $-0.04$ , respectively, invertebrates had a positive mean null deviation of  $0.35$ . The relationships between the Jaccard dissimilarity matrices of the studied soil biota with the overall environmental gradient was significant for fungi, protists and nematodes and not for bacteria and invertebrates (Fig. 4B). The correlation analysis of soil biota beta diversity with the sampled environmental variables indicated strong positive associations between bacterial turnover and pH plus bacteria and total sulphur. Spatial turnover in both fungi and nematodes was also strongly correlated with pH. Whereas protists were significantly correlated with total phosphorus, spatial turnover in invertebrate communities showed no significant association with environmental variables (Fig. 4C).

## Soil bacteria occupied a wider niche than other soil biota



The niche occupancy (*Bcom*) of soil biota was significantly different among the studied taxonomic groups, except for fungi and nematodes (Fig. 5). The bacterial community occupied the widest niche, followed by protists, nematodes and fungi. Although invertebrates occupied the narrowest niche, the largest number of specialist taxa (6.3%) was found for fungi (Fig. 5).

## Locally-adapted taxa reflect habitat-specific environmental conditions

Taxa that were adapted to a given ecosystem type (*i.e.*, urban park, sub-urban park and forest) were identified from the whole regional species pool by determining the taxa at the intersection between specialist taxa, and the taxa above and below the neutral community model. We observed a strong and significant relationship between the number of taxa identified as specialists and taxa that were above and below the neutral model prediction ( $R = 0.99$ ,  $p < 0.001$ ; Fig. 6A) when considering the absolute number of ASVs. However, this relationship was not significant when using the percentage of taxa for each group ( $R = 0.59$ ,  $p > 0.05$ ; Fig. 6B). The intersection of these taxa returned only 2.1% of taxa that were locally adapted (Fig. 6C). Of these taxa, a total of 534 bacterial ASVs and 146 fungal ASVs were locally-adapted microbial taxa in the study area. There were 151 protist, 8 nematode and 10 invertebrate locally-adapted ASVs (Fig. 6C). Although bacterial taxa were highly variable at the family level (*i.e.*, more ASVs classified as 'others') the families *Rhodospirillaceae* and *Hypomicrobiaceae* dominated the urban parks and forest sites. Whereas the family *Streptosporangiaceae* prevailed in sub-urban parks, the family *Koribacteraceae* was the main locally-adapted taxon in forest sites. Among the dominant fungal families was the family *Archaeorhizomycetaceae* in both urban and sub-urban parks. The protist families *Gregarinomorpha* and *Colpodellidea* primarily dominated the urban and suburban sites (Fig. 6D). The family *Tubulinea* tended to be locally-adapted in forest sites. Among the eight locally adapted nematode ASVs, *Xiphinema taylori* was specific to forest sites while *Tylenchulus semipenetrans* and *Dolylaimellus virginianus* were dominant in sub-urban and urban parks, respectively (Fig. 6D). Finally, the locally-adapted taxa in the invertebrate group were composed of the order *Lepidoptera* and *Aranea* which were also dominant in forest sites. In contrast, the *Enchytraeida* appeared more dominant in the sub-urban and urban park sites. The later ecosystem type was uniquely characterized by ASVs belonging to the order *Haplotaxida* (Fig. 6D). Redundancy analysis showed environmental variables associated with the locally-adapted taxa (Fig. S5). Most bacterial families tended to follow an opposite direction with the environmental variables in the RDA ordination. *Streptosporangiaceae* was associated with pH, *Koribacteraceae* with moisture and *Rhodospirillaceae* with total sulphur. The fungal family *Archaeorhizomycetaceae*, the protistan *Colpodellidea*, the nematode *Dolylaimellus virginianus* and the invertebrate *Aranae* were negatively associated with increase in total carbon, total organic carbon and total nitrogen (Fig. S5).

## Discussion

We evaluated the spatial patterns of soil organisms from five taxonomic groups along an urbanization gradient in Xiamen, a subtropical region of China. We show that: (i) the soil bacterial communities exhibited stronger dispersal patterns than other soil organisms; (ii) there was a stronger environmental filtration effect on the larger (e.g., invertebrates) than the smaller soil organisms (e.g., bacteria); and (iii) the bacteria occupied a wider niche than other soil organisms. Overall, microorganisms and small soil organisms (i.e., bacteria, fungi, nematodes and protists) were generally more affected by dispersal-driven processes than large organisms (e.g., invertebrates).

### **Weaker effects of environmental filters on smaller than larger soil organisms: supporting evidence of the size-plasticity hypothesis**

Our results support the size-plasticity hypothesis where smaller organisms are filtered less by their environment than larger ones. With the size-plasticity hypothesis, environmental filters generally dictate how species without certain physiological traits may not occur in local communities [22, 55]. In soil systems, the size-plasticity has been supported in microbial systems targeting bacteria, archaea and fungi [22]. Both smaller and larger soil organisms should exhibit widely different degree of metabolic flexibility, dispersal environmental tolerance, and evolutionary adaptation [21, 56]. These features may enable some taxa to occupy wider niches, resulting in greater coexistence of species within metacommunities [56, 57].

In order to identify the relative roles of dispersal and environmental filtration-related processes that generated the observed distribution patterns, we integrated the neutral community model, the null deviation and the niche breadth. We show that stochastic processes were dominant for bacteria and nematodes as a significant percentage of taxa (69% and 50%) could be explained by the neutral community model. Fungi, protists and soil invertebrates were potentially structured by non-random deterministic processes. Deviation from the null model expectations in community turnover of the studied soil biota returned mean values that were significantly lower than expected by the null model for bacteria, nematodes, protists, and fungi but greater than expected for invertebrates (Fig. 4A). These results indicate that bacteria and nematodes exhibited, on average, a negative null deviation that tended to be closer to -0.95 suggesting that homogenizing dispersal has a major role in structuring both bacteria and nematodes [52]. In conjunction with their large percentage of community variation for both bacteria and nematodes, these findings explained the observed strong distance-decay patterns (Fig. 2). The communities of fungi and protists exhibited similar patterns and underlying processes, as the null deviations were closer to -0.95 to 0.95 – a range that indicates a major role of drift [52]. Invertebrates were more driven by environmental selection because the null deviation tended to be between 0.95 to 1 (Fig. 4A). From these observations, the strength of environmental selection generally decreases with soil organisms sizes, from microbial communities to invertebrates. High dispersal rates of microbial communities could partially explain steeper distance-decay relationships observed for both bacteria and nematodes [58]. Similar patterns and processes between bacteria and nematodes could also result from previously confirmed roles of nematodes in stimulating bacterial dynamics in a spatially-dependent manner [59].

Both fungi and protists underwent similar assembly processes as the neutral community model did not explain a significant proportion of taxa and null deviation, indicating the major role of drift. According to both the neutral community model observations and null deviation, niche-based deterministic processes were the main force in shaping invertebrates. This observation could be explained by the fact that, on one hand, soil bacteria have higher environmental tolerance and rapid growth rates as compared to large organisms such as invertebrates. On the other hand, both low sinking rates and high population sizes may allow them to undergo strong passive dispersal. The contrasting patterns observed for fungi, may be explained by the fact that most of their body comprises of vegetative spreading mycelium in soil, reaching up to 100 m in some cases [60]. Therefore, in certain fungal groups, our sampling might have captured the same individual multiple times, which less likely occurs in other taxonomic groups. Through the niche breadth concept it is possible to identify how the diversity of resources used or environments is being tolerated by a set of taxa [26]. However, most previous studies evaluated the niche breadth across large spatial scale (e.g. many degrees of latitude), yet at these scales, climatic gradients are likely to exert an influence on resource availability [61]. Whereas the niche breadth can be defined along any number of ecological axes [62], limited number of studies have applied this concept at a regional scale influenced by human activities such as urbanization. In our study, the order of niche breadth occupancy, from the wider to narrow niche was as follows: bacteria > protists > fungi > nematodes > invertebrates. Dispersal is a major concept that explains distribution patterns, in which species distributions are explained by differences in arrival [56, 57, 63].

## Distribution patterns of soil organisms

Distance–decay relationships are widely used to test whether communities that exist closer in space are generally more compositionally similar than those further apart [47, 64]. Generally, communities with lower dispersal potential should have steeper trends of community similarity along a distance as compared with communities with high dispersal potentials [64]. By using Distance-decay relationships, many previous studies have showed distribution patterns of different taxonomic groups by gauging their spatial community turnover [e.g., 65, 66, 67], but rarely has a wide range of taxonomic groups in soils been jointly compared, as we did in our study. Fungi exhibited stronger Distance-decay relationships than protists, bacteria and soil animals in the alpine grasslands on the Tibetan Plateau [68] and in paddy soils along a transect across East Asia [18]. In previous studies, fungal communities exhibited stronger Distance-decay relationships than bacteria in far-distant termite mounds of Australia [69], in dryland habits of northwest China [67], across Eastern China forests [66] and in north China agricultural soils [65]. These observations were made at large spatial scales yet biological processes act at different spatial and temporal scales [70, 71], including those that are escalated by increasing urbanization worldwide [5, 6]. In contrast to the above studies, we observed stronger distance-decay relationships for bacteria than fungi and protists at a regional scale, pointing out a non-ubiquitous distribution of soil microorganisms. The results also suggested that major taxonomic groups, within which body size traits are conserved, is a key factors influencing community assembly even at a regional scale [18, 72]. We expected to find similar distribution patterns between bacteria and fungi as microorganisms, but, while distance-decay

relationships of fungi were negligible, bacteria followed similar distance-decay relationships patterns as nematodes and protists. This indicates that distribution patterns observable through distance-decay relationships may result from the presence of major dispersal barriers, dispersal limitation or a decreasing similarity in environmental features [64]. Thus, we could go one step further evaluating the ecological mechanisms underlying the observed distribution patterns (see below).

## Locally-adapted taxa across ecosystem types

By integrating the neutral community model and niche breadth analysis, we identified a small percentage of taxa (2.1%) that appear to be locally-adapted to each of the three studied ecosystem types (i.e., urban, suburban and forest) along our urbanization gradient. These taxa could reflect environmental changes and could be of great importance in designing appropriate conservation and restoration practices in the face of increasing urban-induced spatial fragmentation.

Locally-adapted taxa from each of the five studied taxonomic groups were identified and their potential to locally colonize each of the three ecosystem types including urban, suburban parks and forest sites (Fig. 6). For bacteria, members of the family *Koribacteraceae* were predominantly unique to forest sites. This family plays a pivotal role in plant-biomass degradation [73] and was previously found to dominate in soils after a native forest was turned into teak plantation [74]. Although they occur naturally in forest soils [75], their dominance in managed forests such as teak could be linked to their opportunistic behaviour to high levels of available organic matter from forest debris. The purple photosynthetic bacteria (*Rhodospirillaceae* and *Hypomicrobiaceae*) that are often found in stagnant waters and mud were among the main locally-adapted bacteria in urban parks. These taxa deserve further attention due to their extremely varying morphological and physiological features [76]. The urban and suburban parks were also enriched for locally adapted putative saprophyte fungi belonging to the family *Archaeorhizomycetaceae* that were previously found dominant in areas with greater water levels such as baldcypress-dominated swamps in USA [77] and balsam fir- dominated parks in Quebec, Canada [78]. The families *Gregarinomorpha* and *Colpodellidea* which consist mainly of the apicomplexan parasites that use invertebrates as hosts [79], co-occurred with invertebrate worms of the family *Enchytraeidae* and the family *Haplotaxida* in the urban and suburban parks. Remarkably, some members of the family *Gregarinomorpha* were previously found surviving the gut of *Enchytraeidae* [80] indicating that further research on co-occurrence of invertebrates and parasitic protists in urban soils is needed. Whereas the plant parasite nematode *Tylenchulus semipenetrans* appeared adapted to suburban park soils, the omnivorous nematode *Dorylaimellus virginianus* was a potential locally-adapted taxon in urban park soils, which suggests that urbanization alters the functional composition of the soil nematode community [81]. While the former tends to follow a gradient of total phosphorus increase, the later tends to follow a negative trend of total phosphorus increase (Fig. S5). Like in this study in which families *Araneae* and *Lepidoptera* were locally-adapted in forests sites (Fig. 6), they comprised the major proportion of the total abundance (78%) of all invertebrates found in forests of the Azores archipelago [82]. However, both taxa differed in the way they correlated with soil moisture (Fig. S4). Whereas *Araneae* were previously negatively correlated with soil moisture [83], *Lepidoptera* showed the opposite trend [84].

Further field and experimental studies could be useful to confirm the observed concordant negative responses of *Aranae*, *Archaeorhizomycetaceae*, *Colpodellidea* and *Dolyllaimellus virginianus* taxa to organic matter availability.

## Implications for soil biota conservation in urban landscapes

Urbanization causes habitat fragmentation, resulting in urban green spaces being important soil biodiversity reservoirs. Due to the rapid rate of global urbanization, it is imperative for biodiversity conservation to better understand how ecological processes diverge in these urban systems [19]. In turn, understanding ecological processes that shape distribution patterns of soil organisms along environmental gradients such as those induced by urbanization is an important step towards conservation of soil biodiversity [17, 18, 85]. Green spaces (e.g., urban parks, suburban parks and natural forests) may be useful to examine the effect of urbanization on ecological processes of soil biodiversity. Communities whose distribution patterns are more strongly influenced by environmental filtering processes are assumed to be unique to a given locality [86]. In this case, to better inform management actions, restoration efforts could focus on improving regional patterns in the communities concerned. Because the movement of soil taxa is ubiquitous and facilitates important ecological mechanisms that drive local community and regional species pool composition, and hence biodiversity [87], restoring the local conditions or enhancing regional connectivity could be pathways for soil biota conservation in the face of increasing urbanization. With increasing anthropogenic threats to soil biodiversity, the potential of community assembly inferences to promote soil biota conservation should be further explored.

## Conclusion

We aimed to determine whether micro-organisms exhibit stronger dispersal abilities than large soil biota along an urbanization gradient. The relevance of the size plasticity hypothesis which is generally portrayed under the above assumption was supported by our findings. One potential implication of our study is the possibility that conservation efforts could use higher taxonomic groups and inferences from community assembly to determine taxa that are locally-adapted in urban soil systems, which could be useful for targeted conservation efforts. Further research is needed to determine the degree to which the size-plasticity hypothesis stands in multiple urban soils with differing magnitudes of environmental gradient, and incorporate more taxonomic groups (e.g., viruses and archaea) by taking advantage of soil DNA metabarcoding potential.

## Declarations

### DATA AVAILABILITY

Raw sequence data are available at .... R scripts for data analyses are provided in the supplementary materials.

### ACKNOWLEDGMENTS

This study was supported by the National Natural Science Foundation of China (No. 42021005, 41571052), the Science and Technology Development Project of Jilin Province of China (No. 20200201003JC), the National Science and Technology Fundamental Resources Investigation Program of China (No. 2018FY100303). We thank Prof. Yong-Guan Zhu for providing the platform to implement this study, Xiaobo Liu and Zihan Guo for their contributions to the fieldwork, and Zhihong Qiao for her assistance with laboratory work.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## References

1. Geisen S, Wall DH, van der Putten WH (2019) Challenges and opportunities for soil biodiversity in the anthropocene. *Curr Biol* 29(19):R1036–R1044
2. Fierer N (2017) Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat Rev Microbiol* 15(10):579–590
3. Gossner MM et al (2016) Land-use intensification causes multitrophic homogenization of grassland communities. *Nature* 540(7632):266–269
4. Leff JW et al (2015) *Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe*. *Proceedings of the National Academy of Sciences*, **112**(35): p. 10967–10972
5. Alberti M et al (2017) *Global urban signatures of phenotypic change in animal and plant populations*. *Proceedings of the National Academy of Sciences*, **114**(34): p. 8951–8956
6. El-Sabaawi R (2018) Trophic structure in a rapidly urbanizing planet. *Funct Ecol* 32(7):1718–1728
7. Yu S et al (2022) *Inconsistent patterns of soil fauna biodiversity and soil physicochemical characteristic along an urbanization gradient*. *Frontiers in Ecology and Evolution*, : p.1029
8. Zambrano L, Aronson MFJ, Fernandez T (2019) *The Consequences of Landscape Fragmentation on Socio-Ecological Patterns in a Rapidly Developing Urban Area: A Case Study of the National Autonomous University of Mexico*. *Frontiers in Environmental Science*, **7**
9. Wilson MC et al (2016) Habitat fragmentation and biodiversity conservation: key findings and future challenges. *Landscape Ecol* 31(2):219–227
10. Guillard C et al (2018) Biodiversity of urban soils for sustainable cities. *Environ Chem Lett* 16(4):1267–1282
11. Dou Y, Kuang W (2020) A comparative analysis of urban impervious surface and green space and their dynamics among 318 different size cities in China in the past 25 years. *Sci Total Environ* 706:135828
12. Francini G et al (2018) *Soil biota in boreal urban greenspace: Responses to plant type and age*. **118**: p. 145–155

13. Corline NJ et al (2021) *Understanding community assembly rules in managed floodplain food webs*. *Ecosphere*, **12**(2)
14. Tripathi BM et al (2018) Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. *ISME J* 12(4):1072–1083
15. Liu W et al (2021) Balanced stochastic versus deterministic assembly processes benefit diverse yet uneven ecosystem functions in representative agroecosystems. *Environ Microbiol* 23(1):391–404
16. Thakur MP et al (2020) Towards an integrative understanding of soil biodiversity. *Biol Rev* 95(2):350–364
17. Bahram M et al (2016) Stochastic distribution of small soil eukaryotes resulting from high dispersal and drift in a local environment. *ISME J* 10(4):885–896
18. Luan L et al (2020) Organism body size structures the soil microbial and nematode community assembly at a continental and global scale. *Nat Commun* 11(1):1–11
19. Isabwe A et al *Riverine bacterioplankton and phytoplankton assembly along an environmental gradient induced by urbanization*. *Limnology and Oceanography*. n/a(n/a)
20. Nemergut DR et al (2013) Patterns and processes of microbial community assembly. *Microbiol Mol biology reviews: MMBR* 77(3):342–356
21. Zinger L et al (2019) Body size determines soil community assembly in a tropical forest. *Mol Ecol* 28(3):528–543
22. Jiao S et al (2020) Balance between community assembly processes mediates species coexistence in agricultural soil microbiomes across eastern China. *ISME J* 14(1):202–216
23. Jiao S, Chen W, Wei G (2017) Biogeography and ecological diversity patterns of rare and abundant bacteria in oil-contaminated soils. *Mol Ecol* 26(19):5305–5317
24. Wu W et al (2018) Contrasting the relative importance of species sorting and dispersal limitation in shaping marine bacterial versus protist communities. *ISME J* 12(2):485–494
25. Farjalla VF et al (2012) Ecological determinism increases with organism size. *Ecology* 93(7):1752–1759
26. Carscadden KA et al (2020) Niche breadth: causes and consequences for ecology, evolution, and conservation. *Q Rev Biol* 95(3):179–214
27. Beissinger SR (2000) Ecological mechanisms of extinction. *Proc Natl Acad Sci USA* 97(22):11688–11689
28. Poiani KA et al (2000) Biodiversity Conservation at Multiple Scales: Functional Sites, Landscapes, and Networks. *Bioscience* 50(2):133–146
29. Yang et al (2021) Recent advances in environmental DNA-based biodiversity assessment and conservation. *Divers Distrib* 27(10):1876–1879
30. Breed MF et al (2019) The potential of genomics for restoring ecosystems and biodiversity. *Nat Rev Genet* 20(10):615–628
31. DESA U (2018) 2018 Revision of World Urbanization Prospects. United Nations

32. Isabwe A et al (2019) Community assembly mechanisms underlying the core and random bacterioplankton and microeukaryotes in a river–reservoir system. *Water* 11(6):1127
33. Shrestha S et al (2021) Impact of Land Use Change Due to Urbanisation on Surface Runoff Using GIS-Based SCS–CN Method: A Case Study of Xiamen City, China. *Land* 10(8):839
34. Stubner S (2002) Enumeration of 16S rDNA of *Desulfotomaculum* lineage 1 in rice field soil by real-time PCR with SybrGreen™ detection. *J Microbiol Methods* 50(2):155–164
35. DeSantis TZ et al (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72(7):5069–5072
36. Toju H et al (2012) High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. *PLoS ONE* 7(7):e40863
37. Abarenkov K et al (2010) The UNITE database for molecular identification of fungi–recent updates and future perspectives. *New Phytol* 186(2):281–285
38. Stoeck T et al (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* 19:21–31
39. Guillou L et al (2012) The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res* 41(D1):D597–D604
40. Porazinska DL et al (2009) Evaluating high-throughput sequencing as a method for metagenomic analysis of nematode diversity. *Mol Ecol Resour* 9(6):1439–1450
41. Quast C et al (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41(D1):D590–D596
42. Leray M et al (2013) A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Front Zool* 10(1):1–14
43. Porter TM, Hajibabaei M (2018) Over 2.5 million COI sequences in GenBank and growing. *PLoS ONE* 13(9):e0200177
44. Team RC (2020) *R: A language and environment for statistical computing (v. 4.0. 2)[Computer software]. R Foundation for Statistical Computing.*
45. Wickham H, Chang W, Wickham MH (2016) Package ‘ggplot2’. Create Elegant Data Visualisations Using the Grammar of Graphics Version 2(1):1–189
46. Kassambara A, Kassambara MA (2020) *Package ‘ggpubr’.*
47. Morlon H et al (2008) A general framework for the distance–decay of similarity in ecological communities. *Ecol Lett* 11(9):904–917
48. Goslee S, Urban D, Goslee MS (2020) *Package ‘ecodist’.*
49. Ofițeru ID et al (2010) *Combined niche and neutral effects in a microbial wastewater treatment community.* *Proceedings of the National Academy of Sciences*, **107**(35): p. 15345–15350



50. Burns AR et al (2016) Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME J* 10(3):655–664
51. Chen W et al (2019) Stochastic processes shape microeukaryotic community assembly in a subtropical river across wet and dry seasons. *Microbiome* 7(1):138
52. Chase JM et al (2011) Using null models to disentangle variation in community dissimilarity from variation in  $\alpha$ -diversity. *Ecosphere* 2(2):1–11
53. Pandit SN, Kolasa J, Cottenie K (2009) Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. *Ecology* 90(8):2253–2262
54. Salazar G (2019) *EcolUtils: Utilities for community ecology analysis*. R Package Version 0.1. Available online: <https://github.com/GuillemSalazar/EcolUtils> (accessed on 22 2019)
55. Kraft NJB et al (2015) Community assembly, coexistence and the environmental filtering metaphor. *Funct Ecol* 29(5):592–599
56. Cadotte MW, Tucker CM (2017) Should environmental filtering be abandoned? *Trends Ecol Evol* 32(6):429–437
57. Kraft NJ et al (2015) Community assembly, coexistence and the environmental filtering metaphor. *Funct Ecol* 29(5):592–599
58. Evans S, Martiny JB, Allison SD (2017) Effects of dispersal and selection on stochastic assembly in microbial communities. *ISME J* 11(1):176–185
59. Jiang Y et al (2017) Nematode grazing promotes bacterial community dynamics in soil at the aggregate level. *ISME J* 11(12):2705–2717
60. Douhan GW et al (2011) Population genetics of ectomycorrhizal fungi: from current knowledge to emerging directions. *Fungal Biology* 115(7):569–597
61. Granot I, Belmaker J (2020) Niche breadth and species richness: Correlation strength, scale and mechanisms. *Glob Ecol Biogeogr* 29(1):159–170
62. Sexton JP et al (2017) Evolution of ecological niche breadth. *Annu Rev Ecol Evol Syst* 48:183–206
63. Fraaije RGA et al (2015) Dispersal versus environmental filtering in a dynamic system: drivers of vegetation patterns and diversity along stream riparian gradients. *J Ecol* 103(6):1634–1646
64. Soininen J, McDonald R, Hillebrand H (2007) The distance decay of similarity in ecological communities. *Ecography* 30(1):3–12
65. Zhang K et al (2020) Space is more important than season when shaping soil microbial communities at a large spatial scale. *Msystems* 5(3):e00783–e00719
66. Ma B et al (2017) Distinct biogeographic patterns for archaea, bacteria, and fungi along the vegetation gradient at the continental scale in Eastern China. *Msystems* 2(1):e00174–e00116
67. Wang J et al (2017) *The Patterns and Drivers of Bacterial and Fungal  $\beta$ -Diversity in a Typical Dryland Ecosystem of Northwest China*. *Frontiers in Microbiology*, 8(2126)
68. Kang L et al (2021) *Stochastic processes regulate belowground community assembly in alpine grasslands on the Tibetan Plateau*. *Environmental Microbiology*, n/a(n/a)

69. Chen Q-L et al (2021) Deterministic selection dominates microbial community assembly in termite mounds. *Soil Biol Biochem* 152:108073
70. Huang S et al (2021) Scale-dependent effects of niche specialisation: The disconnect between individual and species ranges. *Ecol Lett* 24(7):1408–1419
71. Rapacciolo G, Blois JL (2019) Understanding ecological change across large spatial, temporal and taxonomic scales: integrating data and methods in light of theory. *Ecography* 42(7):1247–1266
72. van der Gast CJ (2015) Microbial biogeography: the end of the ubiquitous dispersal hypothesis? *Environ Microbiol* 17(3):544–546
73. Levy-Booth DJ et al (2019) Seasonal and ecohydrological regulation of active microbial populations involved in DOC, CO<sub>2</sub>, and CH<sub>4</sub> fluxes in temperate rainforest soil. *ISME J* 13(4):950–963
74. De Gannes V et al (2016) *Microbial community structure and function of soil following ecosystem conversion from native forests to teak plantation forests*. *Frontiers in microbiology*, 7: p. 1976
75. Männistö M et al (2018) Bacterial and fungal communities in boreal forest soil are insensitive to changes in snow cover conditions. *FEMS Microbiol Ecol* 94(9):fiy123
76. Sakarika M et al (2020) Purple non-sulphur bacteria and plant production: benefits for fertilization, stress resistance and the environment. *Microb Biotechnol* 13(5):1336–1365
77. Kernaghan G, Patriquin G (2015) Diversity and host preference of fungi co-inhabiting *Cenococcum mycorrhizae*. *Fungal Ecol* 17:84–95
78. Lumibao CY et al (2020) Divergent biotic and abiotic filtering of root endosphere and rhizosphere soil fungal communities along ecological gradients. *FEMS Microbiol Ecol* 96(7):fiaa124
79. Rueckert S, Betts EL, Tsaousis AD (2019) The symbiotic spectrum: where do the gregarines fit? *Trends Parasitol* 35(9):687–694
80. Butaeva F, Paskerova G, Entzeroth R (2006) *Ditrypanocystis* sp.(Apicomplexa, Gregarina, Selenidiidae): the mode of survival in the gut of *Enchytraeus albidus* (Annelida, Oligochaeta, Enchytraeidae) is close to that of the coccidian genus *Cryptosporidium*. *Tsitologiya* 48(8):695–704
81. Pavao-Zuckerman MA, Coleman DC (2007) Urbanization alters the functional composition, but not taxonomic diversity, of the soil nematode community. *Appl Soil Ecol* 35(2):329–339
82. Gaspar C, Borges PA, Gaston KJ (2008) Diversity and distribution of arthropods in native forests of the Azores archipelago.
83. Suter RB, Doyle G, Shane CM (1987) *Oviposition site selection by Frontinella pyramitela (Araneae, Linyphiidae)*. *Journal of Arachnology*, : p.349–354
84. Tian T et al (2022) Too dry or too wet soils have a negative impact on larval pupation of fall armyworm. *J Appl Entomol* 146(1–2):196–202
85. Barnett SE, Youngblut ND, Buckley DH (2019) *Soil characteristics and land-use drive bacterial community assembly patterns*. *FEMS Microbiology Ecology*, 96(1)
86. Corline NJ et al (2021) Understanding community assembly rules in managed floodplain food webs. *Ecosphere* 12(2):e03330

87. Schlägel UE et al (2020) Movement-mediated community assembly and coexistence. Biol Rev Camb Philos Soc 95(4):1073–1096

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryinformation.docx](#)