ACCESS MICROBIOLOGY

RESEARCH ARTICLE

Mills *et al.*, *Access Microbiology* 2022;4:000320

DOI 10.1099/acmi.0.000320

MICROBIOLOGY SOCIETY

OPEN ACCESS

Rare genera differentiate urban green space soil bacterial communities in three cities across the world

Jacob G. Mills[1,*], Caitlin A. Selway[1], Laura S. Weyrich[1,2], Chris Skelly[3,4], Philip Weinstein[1,5,6], Torsten Thomas[7], Jennifer M. Young[1,8], Emma Marczylo[9], Sudesh Yadav[10], Vijay Yadav[10]†, Andrew J. Lowe[1,5] and Martin F. Breed[1,5,8]

Abstract

Vegetation complexity is potentially important for urban green space designs aimed at fostering microbial biodiversity to benefit human health. Exposure to urban microbial biodiversity may influence human health outcomes via immune training and regulation. In this context, improving human exposure to microbiota via biodiversity-centric urban green space designs is an underused opportunity. There is currently little knowledge on the association between vegetation complexity (i.e. diversity and structure) and soil microbiota of urban green spaces. Here, we investigated the association between vegetation complexity and soil bacteria in urban green spaces in Bournemouth, UK; Haikou, China; and the City of Playford, Australia by sequencing the 16S rRNA V4 gene region of soil samples and assessing bacterial diversity. We characterized these green spaces as having 'low' or 'high' vegetation complexity and explored whether these two broad categories contained similar bacterial community compositions and diversity around the world. Within cities, we observed significantly different alpha and beta diversities between vegetation complexities; however, these results varied between cities. Rare genera (<1% relative abundance individually, on average 35% relative abundance when pooled) were most likely to be significantly different in sequence abundance between vegetation complexities and therefore explained much of the differences in microbial communities observed. Overall, general associations exist between soil bacterial communities and vegetation complexity, although these are not consistent between cities. Therefore, more in-depth work is required to be done locally to derive practical actions to assist the conservation and restoration of microbial communities in urban areas.

INTRODUCTION

Micro-organisms are important to every major biogeochemical process on Earth. They fix nitrogen, draw carbon-dioxide down from the atmosphere, weather rocks, decompose organic material, and, among many other things, form the base of the food web [1, 2]. Furthermore, micro-organisms form symbiotic relationships with many plants and animals where they often have important roles in regulating host health [3–5]. However, these ecosystem functions and services are being degraded by anthropogenic global change leading to climate, biodiversity and health crises. Urbanization in particular is linked to a public health crisis of rapidly rising non-communicable disease rates that are linked to losses of human exposure to microbial biodiversity [4, 6]. Indeed, there have been repeated calls to conserve and restore microbial biodiversity [3, 7–9] due to the impact of human activities on ecosystem and human health [6, 10].

Received 01 July 2021; Accepted 09 December 2021; Published 27 January 2022

Author affiliations: [1]School of Biological Sciences, The University of Adelaide, Adelaide, Australia; [2]Department of Anthropology and Huck Institutes of the Life Sciences, Pennsylvania State University, Pennsylvania, USA; [3]Healthy Urban Microbiome Initiative; [4]Research & Intelligence, Public Health Dorset, Dorset County Council, Dorset, UK; [5]Environment Institute, The University of Adelaide, Adelaide, Australia; [6]School of Public Health, The University of Adelaide, Adelaide, Australia; [7]Centre for Marine Science and Innovation, School of Biological, Environmental and Earth Sciences, University of New South Wales, Sydney, Australia; [8]College of Science and Engineering, Flinders University, Bedford Park, South Australia; [9]Toxicology Department, Centre for Radiation, Chemical and Environmental Hazards, Public Health England, Chilton, Oxfordshire, UK; [10]School of Environmental Sciences, Jawaharlal Nehru University, New Delhi, India.

***Correspondence:** Jacob G. Mills, jacob.mills@adelaide.edu.au

Keywords: microbial conservation; urban green space; vegetation complexity; soil bacteria; urban design.

Abbreviation: PD, phylogenetic diversity.

†Deceased 24th November 2017

Two supplementary tables and three supplementary figures are available with the online version of this article.

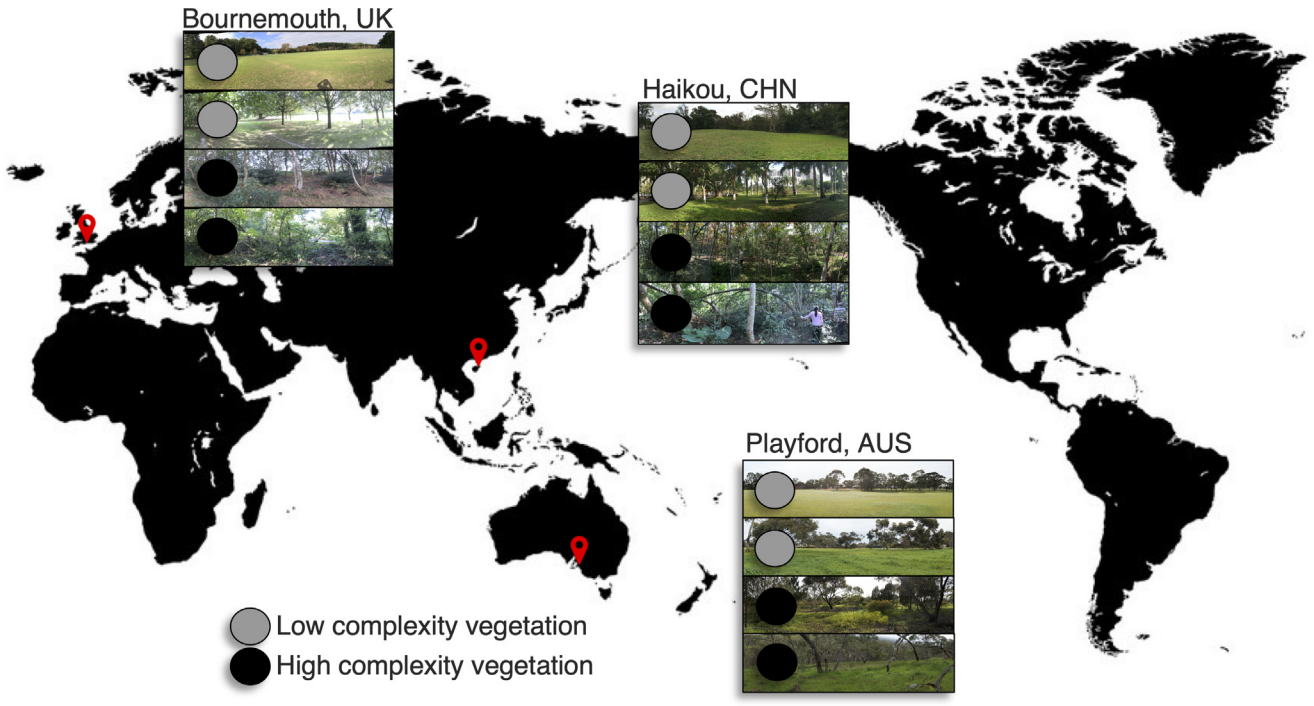

Fig. 1. 'Low' and 'high' complexity vegetation urban green spaces were sampled in Bournemouth, UK; Haikou, CHN; and Playford, AUS.

One potential area where microbial communities could be conserved and restored is urban green spaces, and these areas are already used to help mitigate many issues that urbanization has on public health in general [9, 11]. Certain urban green space designs can reduce air pollution [12, 13] and heat island effects [14], while potentially restoring microbial biodiversity to benefit ecosystem services [15, 16]. Indeed, restoring the urban microbiota by planting native vegetation could improve the exposure to microbes that humans need for immune training and regulation, thus contributing to reducing the immune disease prevalence found in cities [4, 6, 17]. Further, there is growing evidence that environmental microbiota can transfer readily to humans through inoculated play-ground media [18] or by simply using green spaces [19], and that vegetation type or diversity near the home is associated with human microbial diversity [20].

Community characteristics of vegetation, such as species richness and functional diversity, are closely linked to microbial communities, including urban soils [21–23]. Soil in revegetated urban areas have microbial communities more representative of remnant areas compared to typical Victorian-era green spaces, such as lawns [23, 24]. These associations are likely driven by plant-microbe-soil chemistry feedback loops [25, 26]. However, this evidence for the relationship between vegetation complexity of urban green space and their associated soil microbiota remains limited. As such, here we build on our earlier work in a single city [23] to focus on the association of vegetation complexity and soil bacterial communities both within and between three cities across different regions of the world.

METHODS

Study sites

We focused our study on urban green spaces that represented 'low' or 'high' complexity vegetation in the cities of Bournemouth, UK; Haikou, China; and the City of Playford (hereafter known as Playford), Australia (Fig. 1). Bournemouth has a 'marine' climate with short dry summers and heavy precipitation during mild winters, Haikou has a 'humid subtropical' climate, and Playford has a 'Mediterranean' climate. The green spaces were categorized as 'high' (i.e. remnant woodlands, revegetated woodlands or regenerated woodlands) or 'low' (i.e. lawns, vacant lots or parklands) complexity vegetation based on the diversity and structure of their vegetation (Fig. 1). These two categories were based on our previous quantification of vegetation diversity (i.e. plant species richness) and structure (i.e. layers of plant growth-forms creating 3D structure) in the urban green space sites of Playford (quantified in [23]). In each city, we selected six sites of 'low' and six sites of 'high' complexity vegetation (example photos in Fig. 1) by using local knowledge of existing urban green space vegetation types. Within each site, a 25×25 m quadrat in a NSEW orientation was sampled, with geo-references and photos taken at the SW corner.

Soil sampling

Soils were sampled for DNA extraction according to the Biomes of Australian Soil Environments (BASE) project protocol [27] in September and October 2016. In brief, 100–200 g of soil from nine points within the quadrat were randomly sampled, pooled and homogenized. From this pooled sample, 50 g were stored at −20 °C until microbial analysis. The Bournemouth and Haikou samples have not been analysed previously. The Playford samples are a subset of those reported in Mills *et al.* [23], including all samples except those from 'Parkland' sites.

Microbial community analysis

Soil DNA was extracted (one extraction from 0.2 g of each 50 g sample) using the DNeasy Powerlyser soil kit (QIAGEN) in the country of sampling, as per the manufacturer's instructions. Extraction blank controls were not used; however, high biomass samples, such as soil, are less susceptible to contamination compared to those of low biomass and are therefore unlikely to produce results heavily swayed by contaminants [28, 29]. Extracted DNA was then shipped to the University of Adelaide for downstream analysis as per Selway *et al.* [19]. Briefly, the bacterial 16S rRNA V4 gene region was amplified using primers 515F and barcoded 806R (Table S1, available in the online version of this article) [30, 31], and PCR components and cycling conditions were followed as previously described in Selway *et al.* [19]. PCR products were pooled into groups of approximately 30 samples at equimolar concentrations. Pools were cleaned (AxyPrep Mag Clean-up kit; Axygen Scientific), quantified and pooled together into a final sequencing pool before sequencing the DNA (with primers in Table S2) at the Australian Genome Research Facility using a 2×150 bp kit on an Illumina MiSeq.

In QIIME2 (v 2018.8), DNA sequences were merged, trimmed to 150 bp, and quality filtered (>Q4), resulting sequences were then denoised with deblur [32] to create amplicon sequence variants (ASVs), as previously described in Selway *et al.* [19]. Representative ASVs were assigned to the SILVA database (version 132). To remove laboratory contaminant sequences, ASVs were identified from PCR negative controls using the prevalence method within the *decontam* package (v 1.8.0 [33]; in R (v 4.0.0 [34]; and with a threshold probability of 0.5. Any identified contaminants were removed from all biological samples before downstream analysis. Additionally, ASVs assigned to mitochondria, chloroplast, Archaea or 'unknown' kingdom were also removed, and ASVs with fewer than ten reads across all samples in the dataset were excluded. Post-filtering, there were at least five 'low' and five 'high' complexity vegetation replicates for each city (see sample metadata via links in '*Data access*').

Statistical analyses

All statistics were done in R (v 4.0.0 [34]. ASVs were agglomerated to genus level for statistical analysis using the 'tax_glom' function of the *phyloseq* package (v 1.32.1 [35]. During the genus agglomeration, all unresolved taxa at genus level (i.e. 'NA' or 'blank') were removed.

Before alpha diversity was calculated, the agglomerated genus level data was rarefied to 2396 reads with the 'rarefy_even_depth' function of the *phyloseq* package. Alpha diversity was calculated as observed genus richness and Shannon's diversity with the 'estimate_richness' function in *phyloseq*, and Faith's phylogenetic diversity was calculated with the 'pd' function of the *picante* package (v 1.8.1 [36]. We used generalized linear mixed models (GLMMs) to test for difference in alpha diversity by crossing the fixed factors of 'city' and 'vegetation complexity' and nesting the random factor of 'site' within 'city'. GLMMs were done with the 'glmer' function of the *lme4* package v 1.1–25 (v 1.1–25) [37]. Distributions for the GLMMs were Poisson for observed genus richness (count data), and Gamma for Faith's phylogenetic diversity and Shannon's diversity (positive, non-integer, non-parametric data). The Poisson GLMM was tested for over-dispersion (result: ratio=0.46). Main effects of the GLMMs were tested by Type II Wald Chi^2 tests with the 'Anova' function of the *car* package (v 3.0–10 [38]. Pairwise comparisons of 'city' and 'vegetation complexity' combinations were tested by z-tests with Holm–Bonferroni P-adjustment with the 'glht' function of the *multcomp* package v 1.4–15 (v 1.4–15) [39].

Ordinations of beta diversity were done with the 'ordinate' function in *phyloseq*. Ordinations were based on unrarefied data in principal coordinates analysis (PCoA) with Bray–Curtis and Jaccard distance matrices. We used PERMANOVA, with 999 iterations, with the 'adonis' function of the *vegan* package (v 2.5–6) [40]; to test the model of 'vegetation complexity' nested within 'city'. Pairwise comparisons between nested vegetation complexities (e.g. Bournemouth Low vs. Bournemouth High) were tested by PERMANOVA with 999 iterations with the 'pairwise.adonis2' function of the *pairwise.adonis* package (v 0.0.1) [41].

We created a relative abundance stack plot by converting the rarefied genus abundances to percentages. All genera with total rarefied sequences across all samples being less than 1% of total rarefied sequences were pooled into a single group named '<1% abund.'. The less than 1% cut-off was determined by a rank-abundance curve of percentage abundance across all samples (Fig. S1). We tested for differentially abundant bacterial genera between 'low' and 'high' complexity vegetation sites within each city. Log-2 fold-change measurement of bacterial genera was done using the 'DESeq' function of the *DESeq2* package in *phyloseq* (v 1.28.1) [42]. *DESeq2* does an internal normalization, where the count of each genus within a sample is divided by the mean of that genus across samples. Differentially abundant genera (alpha=0.05) were plotted into heatmaps using

the 'pheatmap' function of the *pheatmap* package (v 1.0.12) [43]. The differential abundance heatmap scale represents the mean abundance of each genus across samples as 0 with ±3 standard deviations. The differential abundance heatmap rows and columns were clustered based on Manhattan distance to most efficiently arrange the grid. The heatmap trees represent how closely related a row or column are, not taxa, based on the scale in each cell. Unclassified genera were not included in the heatmap.

RESULTS AND DISCUSSION

Each city had distinct soil microbial communities

We compared soil bacterial genera between cities and found that the communities were quite distinct from each other, regardless of vegetation complexity, both in terms of alpha diversity [observed genus richness, Chi^2=28.67, Pr(>Chi^2)<0.001; Faith's phylogenetic diversity, Chi^2=21.02, Pr(>Chi^2)<0.001; Shannon's diversity, Chi^2=21.80, Pr(>Chi^2)<0.001; Fig. 2a] and beta diversity distances [Bray–Curtis, F=19.61, Pr(>F)=0.001; Jaccard, F=10.68, Pr(>F)=0.001; Fig. 2b]. Further, beta diversity at the ASV-level had similar patterns to the genus-level results; however, the data were over-dispersed (i.e. significantly more variable than predicted for the model) and therefore not used further (Fig. S2). These differences between cities were expected given their differences in geography and climate, where for example, temperature, aridity and distance from the equator vary and each are strong predictors of soil microbial diversity [44]. Moreover, strong biogeographic zoning and distance–decay relationships have previously been observed for urban soil bacterial communities across ten cities within China [45].

'Low' and 'high' complexity vegetation soils have similar diversity

We next compared diversity of sites with 'low' versus 'high' vegetation complexity within all three cities. In Bournemouth, 'high' complexity vegetation green spaces were significantly more diverse than 'low' complexity spaces for their bacterial genera (observed genus richness, z=3.17, *P*=0.014; Faith's PD of genera, z=−2.93, *P*=0.034), (Table 1, Fig. 2). In contrast, alpha diversity of bacterial genera in soil from Playford and Haikou for all three measures were non-significantly different between the 'low' and 'high' complexity vegetation (Table 1). However, there was a significant interaction between 'city' and 'vegetation complexity' for all three alpha diversity measures (Fig. 2a). This interaction was caused by the Playford soils being lower in diversity in the 'high' complexity vegetation soils relative to the 'low' complexity soils, whereas diversity was higher in these 'high' sites in Bournemouth and Haikou.

The difference in diversity between 'low' and 'high' complexity vegetation soils in Playford compared to Bournemouth and Haikou may be due to Playford's relatively drier climate and the tendencies of native vegetation in this part of Australia to prefer relatively arid conditions. Such conditions are less conducive to supporting high microbial biodiversity [44]. In these drier environments, areas of lower vegetation complexity, such as urban lawns, are often heavily watered and fertilized. This practice can lead to higher nutrient loads relative to higher vegetation complexity native soils, potentially increasing microbial diversity independent of vegetation complexity. However, we note that our previous work with the Playford samples [23] indicated a consistent pattern in alpha diversity as found here in Bournemouth and Haikou (i.e. more vegetation complexity associated with greater bacterial alpha diversity). Although, our earlier study reported data from the V1-3 region of the 16S rRNA gene, rather than the V4 region reported here. As such, future work should further explore the effect of marker gene choice on vegetation-bacterial diversity associations.

Differences in bacterial composition between vegetation complexities vary between cities

We next tested relationships between soil bacterial composition at the genus-level and the vegetation complexity of urban green spaces. The composition of bacterial communities was significantly different between 'low' and 'high' complexity vegetation in both Haikou [Bray–Curtis, F=4.05, Pr(>F)<0.05; Jaccard, F=3.19, Pr(>F)<0.01] and Playford [Bray–Curtis, F=4.42, Pr(>F)<0.05; Jaccard, F=3.22, Pr(>F)<0.05) (Table 2, Fig. 2]. However, Bournemouth had no significant difference between the vegetation complexities for both Bray–Curtis and Jaccard distances (Fig. 2b).

Vegetation type (e.g. lawn, remnant woodland) is a known driver of microbial diversity and composition in urban soil [21, 23]. However, there is little consistency between soil microbial communities in what seem to be broadly similar ecological settings, as in our study, due to a complexity of multiple driving factors. Such factors include plant species turnover and soil properties that vary on broad spatial scales, such as temperature [32] and, at finer scales, pH and salinity [46]. Certainly, pH and salinity have previously been found to strongly associate with urban soil bacterial community composition [16, 23]. While we did not measure soil physicochemical properties here, they may, in some instances, override any effect of the vegetation community on the soil community and potentially lead to results as we saw in Bournemouth.

Rare genera contribute to differences in community structure

We performed differential abundance testing to investigate which genera may have been driving the differences between the 'low' and 'high' complexity vegetation soils. Rare genera (i.e. <1% relative abundance) dominated the significantly differentially

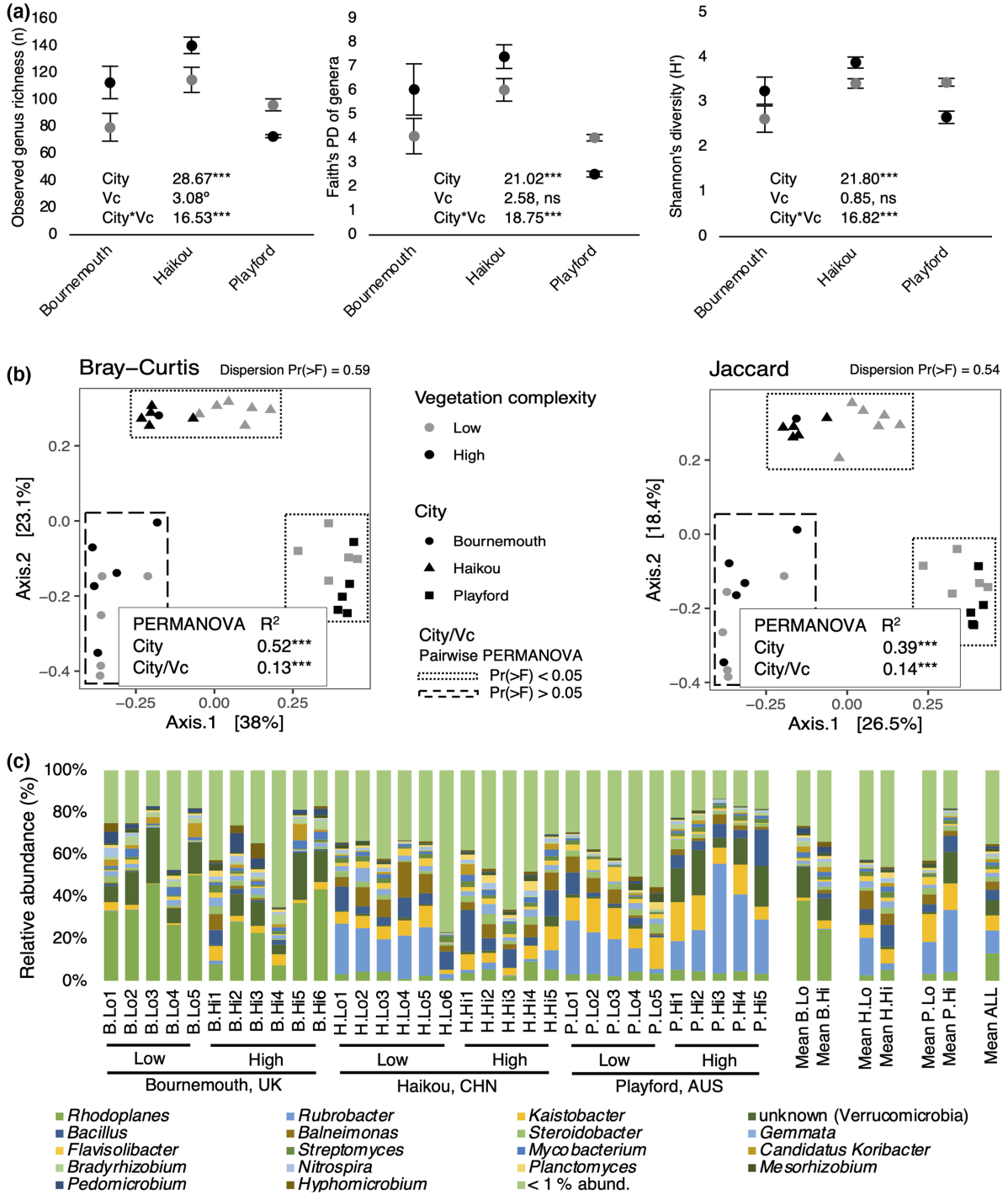

Fig. 2. (a)'City' by 'Vegetation complexity' (Vc) for the alpha diversity GLMMs on observed genus richness, Faith's phylogenetic diversity (PD), and Shannon's diversity. Results are Chi^2 values from Type II Wald Chi^2 tests on the GLMMs followed by significance codes for Pr(>Chi^2). See Table 1 for pairwise results. Significance codes: 'ns' not significant; '°' P<0.10; '*' P<0.05; '**' P<0.01; '***' P<0.001. (b) PCoAs of soil bacterial genus communities in urban green spaces by Bray–Curtis and Jaccard distance. Main PERMANOVA test with 999 iterations of 'Vegetation complexity' (Vc) nested within 'City' (distance ~City/Vc); R^2 and P-value significance codes ['***', Pr(>F)<0.001]. Within city 'Vegetation complexity' differences were tested with pairwise PERMANOVA. Cities surrounded by dotted boxes were significantly different between their 'low' and 'high' vegetation complexity green spaces. Cities surrounded by dashed boxes were not significantly different between their 'low' or 'high' vegetation complexity green spaces. For detailed main and pairwise PERMANOVA results see Table 2. (c) Relative abundance (%) of soil bacterial genera across all sites. Genera read left to right by rows in the legend and correspond to bottom to top in the stack plot.

Table 1. Pairwise alpha diversity – observed genus richness, Faith's phylogenetic diversity (PD), and Shannon's diversity – of soil bacterial genera under the GLMM interaction of 'City' by 'Vegetation complexity'. Significance codes: 'ns' not significant; '°' P<0.10; '*' P<0.05; '**' P<0.01; '***' P<0.001

City * vegetation complexity	Observed genus richness			Faith's PD of genera			Shannon's diversity		
	z-value	Pr(>|z|)	Sig.	z-value	Pr(>|z|)	Sig.	z-value	Pr(>|z|)	Sig.
Bournemouth High - Haikou High	−2.22	0.183	ns	2.98	0.032	*	2.78	0.058	°
Bournemouth High - Playford High	3.81	0.002	**	−2.04	0.209	ns	−1.21	0.790	ns
Bournemouth High - Bournemouth Low	3.17	0.014	*	−2.93	0.034	*	−2.79	0.058	°
Bournemouth High - Haikou Low	−0.24	1.000	ns	0.64	1.000	ns	1.38	0.790	ns
Bournemouth High - Playford Low	1.36	0.521	ns	0.75	1.000	ns	1.70	0.619	ns
Haikou High - Playford High	5.76	0.000	***	−4.80	0.000	***	−3.82	0.002	**
Haikou High - Bournemouth Low	5.15	0.000	***	−5.66	0.000	***	−5.33	0.000	***
Haikou High - Haikou Low	2.00	0.276	ns	−2.37	0.124	ns	−1.46	0.790	ns
Haikou High - Playford Low	3.43	0.007	**	−2.13	0.198	ns	−1.03	0.790	ns
Playford High - Bournemouth Low	−0.62	1.000	ns	−0.86	1.000	ns	−1.51	0.790	ns
Playford High - Haikou Low	−4.04	0.001	***	2.65	0.068	°	2.53	0.092	°
Playford High - Playford Low	−2.36	0.145	ns	2.67	0.068	°	2.79	0.058	°
Bournemouth Low - Haikou Low	−3.40	0.007	**	3.54	0.005	**	4.10	0.001	***
Bournemouth Low - Playford Low	−1.74	0.405	ns	3.53	0.005	**	4.30	0.000	***
Haikou Low - Playford Low	1.59	0.447	ns	0.14	1.000	ns	0.39	0.790	ns

abundant bacteria between 'low' and 'high' complexity vegetation soils. For example, in Bournemouth, *Bacillus* (characteristic of 'high' complexity vegetation soils) was the only genus out of seven differentially abundant genera (P<0.05, Fig. 3) that was also dominant in relative abundance (>1% relative abundance, Fig. 2c) between the vegetation complexity levels – the other six genera were less than 1% in relative abundance. In Haikou, *Rubrobacter* (characteristic of 'low' complexity vegetation soils) was the only genus out of four differentially abundant genera to also be greater than 1% in relative abundance (P<0.05, Fig. 3), and in Playford, *Flavisolibacter* and *Gemmata* (both characteristic of 'low' complexity vegetation soils) were the only differentially abundant genera out of 23 that were also dominant (P<0.05, Fig. 3). Overall, differential abundance tests showed that there are soil bacteria characteristics of either 'low' or 'high' complexity vegetation within each city and that rare taxa are important in defining these communities. This result is consistent with other findings that rare bacteria biogeographically distinguish forensic soil samples [47, 48].

The rare genera (<1% relative abundance) that dominated the differential abundance testing between the vegetation complexities (P<0.05, Fig. 3) were potentially functionally important to these locations as has been found in greenhouse soils [49]. For example, *Luteibacter* was the only genus to significantly represent 'high' vegetation complexity across all cities in this study and species of this genus are known to live both in soil and on humans [50, 51]. The 'high' complexity vegetation soils of Bournemouth had significantly more *Sinorhizobium* (nitrogen-fixers) [52]; and *Kaistia* (methanotrophs) [53]; of the order Rhizobiales [54] than 'low' complexity soils. However, most differentially abundant genera in Bournemouth were higher in only one site relative to others; therefore, they are not characteristic of either vegetation complexity studied here. In Haikou, *Haliangium* (producer of fungicidal haliangicins) [55], *Rubrobacter* and *Acetobacter* were significantly more abundant in the 'low' than in the 'high' vegetation diversity soils, whereas *Erwinia* (genus of many plant pathogen species) [56, 57]; was more abundant in the 'high' vegetation diversity soils. In Playford, *Rhizobium* (nitrogen-fixers) [54]; were significantly more abundant in the 'high' complexity vegetation.

The differential abundance findings that imply rare genera are driving the community differences are further supported by the similarity between Bray–Curtis and Jaccard ordinations (Fig. 2b). Further, ordinations of only the rare genera (those <1% relative abundance) were similar to ordinations using the whole community (Fig. S3), therefore implying that rare genera are driving these patterns; however, these data were over-dispersed. Additionally, the rank-abundance curve showed there were 300 of 318 genera with less than 1% relative abundance across all sites (Fig. S1), and, when pooled, had an average relative abundance of 35% across all sites (Fig. 2c). These findings indicate that rare genera may be quite valuable to urban soils and that they should not be overlooked when planning soil microbial conservation. To that end, rare micro-organisms have been identified to play key functional roles, from biogeochemical cycles to holobiont health [58]. Further exploration of the functional contribution of rare bacteria in urban green spaces would provide deeper understanding of their value in conservation and restoration efforts.

Table 2. Main and pairwise PERMANOVA on soil bacterial genus communities for vegetation complexity nested within cities. Significance codes Pr(>F): 'ns' not significant; '°' $P<0.10$; '*' $P<0.05$; '**' $P<0.01$; '***' $P<0.001$

Formula=distance~city/vegetation complexity							
Main PERMANOVA		**Bray–Curtis**			**Jaccard**		
		R^2	**F**	**Pr(>F)**	**R^2**	**F**	**Pr(>F)**
City	**$df_{2,26}$**	**0.52**	**19.61**	***	**0.39**	**10.68**	***
City/Vegetation complexity	**$df_{3,26}$**	**0.13**	**3.25**	***	**0.14**	**2.56**	***
Pairwise PERMANOVA		Bray-Curtis			Jaccard		
		R^2	F	Pr(>F)	R^2	F	Pr(>F)
Bournemouth High - Haikou High	$df_{1,9}$	0.30	3.82	**	0.23	2.70	*
Bournemouth High - Playford High	$df_{1,9}$	0.60	13.36	**	0.45	7.35	**
Bournemouth High - Bournemouth Low	$df_{1,9}$	0.16	1.78	ns	0.14	1.49	ns
Bournemouth High - Haikou Low	$df_{1,10}$	0.42	7.19	***	0.32	4.75	**
Bournemouth High - Playford Low	$df_{1,9}$	0.54	10.54	**	0.40	5.95	**
Haikou High - Playford High	$df_{1,8}$	0.69	17.99	**	0.53	8.94	**
Haikou High - Bournemouth Low	$df_{1,8}$	0.56	10.30	**	0.43	6.13	*
Haikou High - Haikou Low	$df_{1,9}$	0.31	4.05	*	0.26	3.19	**
Haikou High - Playford Low	$df_{1,8}$	0.61	12.58	**	0.46	6.76	**
Playford High - Bournemouth Low	$df_{1.8}$	0.72	20.44	**	0.56	10.35	*
Playford High - Haikou Low	$df_{1,9}$	0.58	11.81	**	0.45	7.34	**
Playford High - Playford Low	$df_{1,8}$	0.36	4.42	*	0.29	3.22	*
Bournemouth Low - Haikou Low	$df_{1,9}$	0.61	13.86	**	0.48	8.25	**
Bournemouth Low - Playford Low	$df_{1,8}$	0.67	16.22	*	0.52	8.52	*
Haikou Low - Playford Low	$df_{1,9}$	0.49	8.74	**	0.39	5.71	**

CONCLUSIONS

Our study suggests that a global comparison of cities in terms of vegetation factors driving microbial diversity may be limited due to the overall strength of their differences driven by geographic or climatic factors. However, investigating trends related to vegetation complexity within cities may produce general recommendations about fostering microbial biodiversity. Certainly, rare taxa should not be overlooked when considering the conservation of microbial biodiversity. Urban green space design for conservation of microbial biodiversity, biogeochemical cycling, public health outcomes and public usability will likely require complementary proportions of both 'low' and 'high' complexity vegetation green spaces. However, what those proportions are will need to be investigated on a city-by-city, or region-by-region basis.

Restoration of biodiversity in urban green spaces has the potential to build native microbial communities. Such endeavours will require local adaptive management within urban green space landscapes that will allow practitioners to understand the knowledge gaps pertaining to their city and properly investigate the outcomes of their efforts [59]. Local knowledge gaps may include understanding functional microbial biodiversity, examining how differently designed green spaces influence environmental and human microbiota (e.g. 'low' and 'high' complexity, native and novel species mixtures), and determining if the use of remnant inoculations accelerates the recovery of native microbial phylogenetic and functional diversity. Further, there is currently a strong call to 'decolonize' public spaces in colonial and imperial countries [60, 61]. Therefore, it would be interesting to track whether such cultural modifications to urban designs influences environmental and human microbiota given that 'low' and 'high' complexity green spaces are somewhat representative of these cultural differences. More work is needed to describe the functional contributions of rare bacteria in urban soils and to determine the best ways to conserve and restore microbial biodiversity to provide the breadth of ecosystem services that they could provide to the urban landscape.

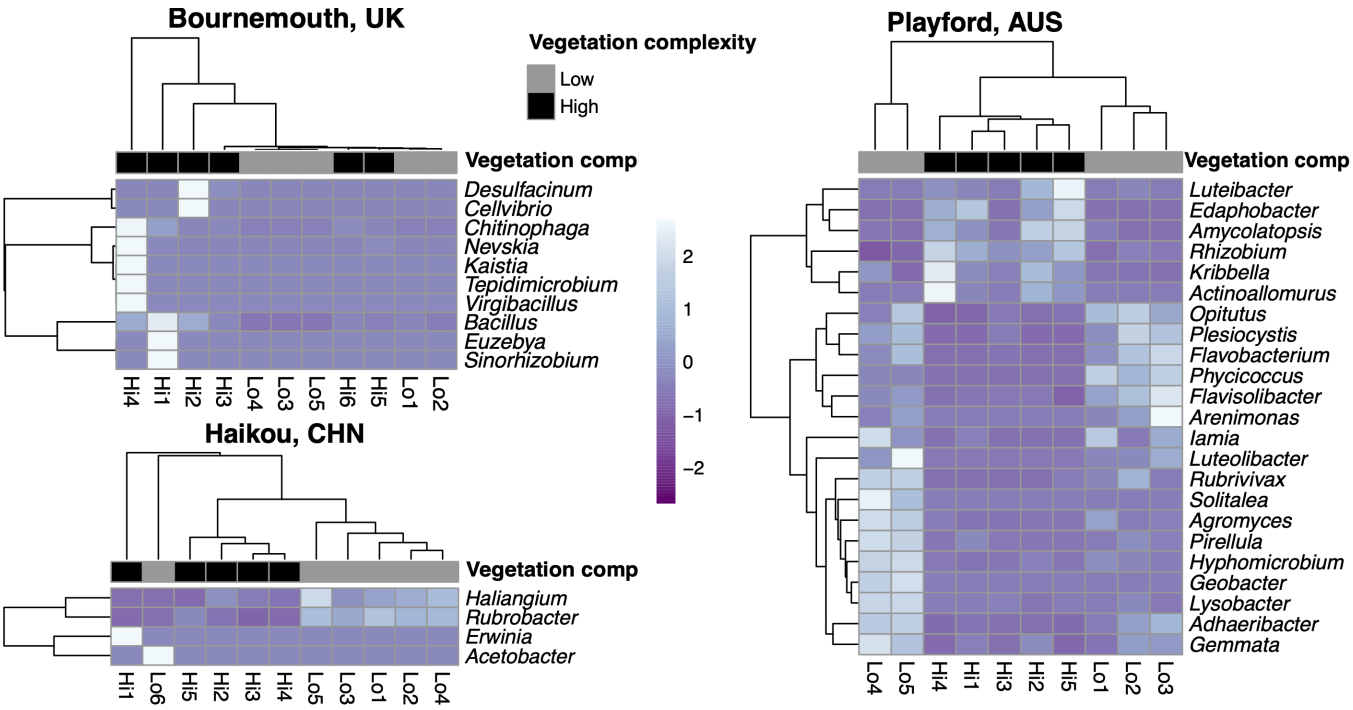

Fig. 3. Differentially abundant bacterial genera from Bournemouth, Haikou and Playford measured by log-2 fold-change with *P*-value<0.05. Extreme ends of the heat colour scale represent three standard deviations from the mean rarefied abundance for each genus across samples. Hierarchical clustering of genera (rows) and vegetation replicates (columns) are both by Manhattan distance.

Data access

Raw sequence reads are available on QIITA Gonzalez *et al.* [62] with the following details: Playford, AUS: Study ID 13064, EBI accession ERP124074 (https://www.ebi.ac.uk/ena/browser/view/PRJEB40432); Bournemouth, UK and Haikou CHN: Study ID 13559, EBI accession ERP126847 (https://www.ebi.ac.uk/ena/browser/view/PRJEB42921). Sample metadata, ASV, reference sequence and taxonomy tables and scripts used for analysis are available at .

Funding information
This research was funded by the Environment Institute (The University of Adelaide). P.W., A.L., M.B. were all Environment Institute researchers during the period of this research. The directors of the Environment Institute did not contribute to the study design, collection, analysis and interpretation of data, writing of the manuscript, nor the decision to submit the article for publication.

Acknowledgements
Michael Rowland (Bournemouth, Christchurch and Poole Council) and Soumya Prasad (Jawaharlal Nehru University, New Delhi) for assistance in selecting sites in Bournemouth and New Delhi (even though this dataset was not used), respectively. Jo Park, Ken Daniel and the City of Playford for assistance with site selection and access. Qiyong Liu and Keke Liu from the Chinese CDC, and Peng Bi (University of Adelaide) for coordinating and providing samples from China.

Author contributions
J.M., C.A.S., L.W., P.W., A.L., M.B.: conceptualization. J.M., C.A.S., C.S., L.W., J.Y., E.M., S.Y., V.Y., M.B.: investigation. J.M., C.A.S.: data curation. J.M., C.A.S., L.W., J.Y., T.T.: formal analysis. P.W., M.B.: funding acquisition. J.M., C.A.S., L.W., J.Y., T.T., A.L., M.B.: methodology. J.M., C.A.S.: software. J.M., C.A.S., L.W., C.S., P.W., A.L., M.B.: project administration. J.M., L.W., E.M., C.S., S.Y., V.Y., A.L.: resources. L.W., C.S., T.T., P.W., A.L., M.B.: supervision. All: validation. J.M.: visualization. J.M.: writing – original draft. All: writing – review and editing.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References

1. Cockell CS, Jones HL. Advancing the case for microbial conservation. *Oryx* 2009;43:520.
2. Rousk J, Bengtson P. Microbial regulation of global biogeochemical cycles. *Front Microbiol* 2014;5:103.
3. Carthey AJR, Blumstein DT, Gallagher RV, Tetu SG, Gillings MR, *et al*. Conserving the holobiont. *Funct Ecol* 2020;34:764–776.
4. Rook GAW, Martinelli R, Brunet LR. Innate immune responses to mycobacteria and the downregulation of atopic responses. *Curr Opin Allergy Clin Immunol* 2003;3:337–342.

5. Rosado BHP, Almeida LC, Alves LF, Lambais MR, Oliveira RS. The importance of phyllosphere on plant functional ecology: a phyllo trait manifesto. *New Phytol* 2018;219:1145–1149.
6. von Hertzen L, Hanski I, Haahtela T. Natural immunity. Biodiversity loss and inflammatory diseases are two global megatrends that might be related. *EMBO Rep* 2011;12:1089–1093.
7. Bello MGD, Knight R, Gilbert JA, Blaser MJ. Preserving microbial diversity. *Science* 2018;362:33–34.
8. Blanco H, Lal R. *Principles of Soil Conservation and Management*. New York, USA: Springer; 2008.
9. Mills JG, Brookes JD, Gellie NJC, Liddicoat C, Lowe AJ, *et al*. Relating Urban Biodiversity to Human Health With the "Holobiont" Concept. *Front Microbiol* 2019;10:550.
10. Zuo T, Kamm MA, Colombel J-F, Ng SC. Urbanization and the gut microbiota in health and inflammatory bowel disease. *Nat Rev Gastroenterol Hepatol* 2018;15:440–452.
11. Kabisch N, van den Bosch M, Lafortezza R. The health benefits of nature-based solutions to urbanization challenges for children and the elderly - A systematic review. *Environ Res* 2017;159:362–373.
12. Ferkol T, Schraufnagel D. The global burden of respiratory disease. *Ann Am Thorac Soc* 2014;11:404–406.
13. Xing Y, Brimblecombe P. Traffic-derived noise, air pollution and urban park design. *Journal of Urban Design* 2020;25:590–606.
14. Aflaki A, Mirnezhad M, Ghaffarianhoseini A, Ghaffarianhoseini A, Omrany H, *et al*. Urban heat island mitigation strategies: A state-of-the-art review on Kuala Lumpur, Singapore and Hong Kong. *Cities* 2017;62:131–145.
15. Hoch JMK, Rhodes ME, Shek KL, Dinwiddie D, Hiebert TC, *et al*. Soil Microbial Assemblages Are Linked to Plant Community Composition and Contribute to Ecosystem Services on Urban Green Roofs. *Front Ecol Evol* 2019;7.
16. Joyner JL, Kerwin J, Deeb M, Lozefski G, Prithiviraj B, *et al*. Green infrastructure design influences communities of urban soil bacteria. *Front Microbiol* 2019;10:982.
17. Mills JG, Weinstein P, Gellie NJC, Weyrich LS, Lowe AJ, *et al*. Urban habitat restoration provides a human health benefit through microbiome rewilding: the Microbiome Rewilding Hypothesis. *Restor Ecol* 2017;25:866–872.
18. Hui N, Grönroos M, Roslund MI, Parajuli A, Vari HK, *et al*. Diverse environmental microbiota as a tool to augment biodiversity in urban landscaping materials. *Front Microbiol* 2019;10:536.
19. Selway CA, Mills JG, Weinstein P, Skelly C, Yadav S, *et al*. Transfer of environmental microbes to the skin and respiratory tract of humans after urban green space exposure. *Environ Int* 2020;145:106084.
20. Pearson AL, Pechal J, Lin Z, Benbow ME, Schmidt C, *et al*. Associations detected between measures of neighborhood environmental conditions and human microbiome diversity. *Sci Total Environ* 2020;745:141029.
21. Hui N, Jumpponen A, Francini G, Kotze DJ, Liu X, *et al*. Soil microbial communities are shaped by vegetation type and park age in cities under cold climate. *Environ Microbiol* 2017;19:1281–1295.
22. Laforest-Lapointe I, Paquette A, Messier C, Kembel SW. Leaf bacterial diversity mediates plant diversity and ecosystem function relationships. *Nature* 2017;546:145–147.
23. Mills JG, Bissett A, Gellie NJC, Lowe AJ, Selway CA, *et al*. Revegetation of urban green space rewilds soil microbiotas with implications for human health and urban design. *Restor Ecol* 2020;28:S322–S334.
24. Baruch Z, Liddicoat C, Laws M, Kiri Marker L, Morelli H, *et al*. Characterising the soil fungal microbiome in metropolitan green spaces across a vegetation biodiversity gradient. *Fungal Ecology* 2020;47:100939.
25. Berendsen RL, Pieterse CMJ, Bakker P. The rhizosphere microbiome and plant health. *Trends Plant Sci* 2012;17:478–486.
26. Fierer N. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat Rev Microbiol* 2017;15:579–590.
27. Bissett A, Fitzgerald A, Court L, Meintjes T, Mele PM, *et al*. Introducing BASE: the Biomes of Australian Soil Environments soil microbial diversity database. *Gigascience* 2016;5:21.
28. Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, *et al*. Contamination in low microbial biomass microbiome studies: issues and recommendations. *Trends Microbiol* 2019;27:105–117.
29. Velásquez-Mejía EP, de la Cuesta-Zuluaga J, Escobar JS. Impact of DNA extraction, sample dilution, and reagent contamination on 16S rRNA gene sequencing of human feces. *Appl Microbiol Biotechnol* 2018;102:403–411.
30. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, *et al*. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* 2011;108 Suppl 1:4516–4522.
31. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, *et al*. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 2012;6:1621–1624.
32. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, *et al*. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 2017;551:457–463.
33. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 2018;6:226.
34. R Core Team. *R: A language and environment for statistical computing. p. ^pp*. Vienna, Austria: R Foundation for Statistical Computing; 2019.
35. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;8:e61217.
36. Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, *et al*. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 2010;26:1463–1464.
37. Bates D, Sarkar D, Bates MD, Matrix L. The lme4 package. *R package version* 2007;2:74.
38. Fox J, Weisberg S, Adler D, Bates D, Baud-Bovy G, *et al*. *Package Car*. Vienna: R Foundation for Statistical Computing; 2012.
39. Hothorn T, Bretz F, Westfall P, Heiberger RM, Schuetzenmeister A, *et al*. *Multcomp: Simultaneous Inference in General Parametric Models. R Package Version 1.3-2*. 2014.
40. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, *et al*. vegan: Community Ecology Package. R package version 2.4-3. p. ^pp. Github, 2017. https://CRAN.R-project.org/package=vegan
41. Arbizu PM. Github, 2017. https://github.com/pmartinezarbizu/pairwiseAdonis
42. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
43. Kolde R. 2015. https://CRAN.R-project.org/package= pheatmap
44. Delgado-Baquerizo M, Reith F, Dennis PG, Hamonts K, Powell JR, *et al*. Ecological drivers of soil microbial diversity and soil biological networks in the Southern Hemisphere. *Ecology* 2018;99:583–596.
45. Yang T, Shi Y, Zhu J, Zhao C, Wang J, *et al*. The spatial variation of soil bacterial community assembly processes affects the accuracy of source tracking in ten major Chinese cities. *Sci China Life Sci* 2021;64:1546–1559.
46. Fierer N, Jackson RB. The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A* 2006;103:626–631.
47. Damaso N, Mendel J, Mendoza M, von Wettberg EJ, Narasimhan G, *et al*. Bioinformatics approach to assess the biogeographical patterns of soil communities: the utility for soil provenance. *J Forensic Sci* 2018;63:1033–1042.
48. Habtom H, Pasternak Z, Matan O, Azulay C, Gafny R, *et al*. Applying microbial biogeography in soil forensics. *Forensic Sci Int Genet* 2019;38:195–203.
49. Xue M, Guo Z, Gu X, Gao H, Weng S, *et al*. Rare rather than abundant microbial communities drive the effects of long-term greenhouse

cultivation on ecosystem functions in subtropical agricultural soils. *Sci Total Environ* 2020;706:136004.

50. Akter S, Huq MA. *Luteibacter pinisoli* sp. nov., a casein degrading bacterium isolated from rhizospheric soil of *Pinus koraiensis*. *Arch Microbiol* 2018;200:1017–1023.
51. Kämpfer P, Lodders N, Falsen E. *Luteibacter anthropi* sp. nov., isolated from human blood, and reclassification of *Dyella yeojuensis* Kim *et al.* 2006 as *Luteibacter yeojuensis* comb. nov. *Int J Syst Evol Microbiol* 2009;59:2884–2887.
52. Mitsui H, Sato T, Sato Y, Ito N, Minamisawa K. *Sinorhizobium meliloti* RpoH1 is required for effective nitrogen-fixing symbiosis with alfalfa. *Mol Genet Genomics* 2004;271:416–425.
53. Im W-T, Yokota A, Kim M-K, Lee S-T. *Kaistia adipata* gen. nov., sp. nov., a novel alpha-proteobacterium. *J Gen Appl Microbiol* 2004;50:249–254.
54. Garrido-Oter R, Nakano RT, Dombrowski N, Ma K-W, AgBiome Team, *et al*. Modular traits of the rhizobiales root microbiota and their evolutionary relationship with symbiotic rhizobia. *Cell Host Microbe* 2018;24:155-167..
55. Fudou R, Iizuka T, Sato S, Ando T, Shimba N, *et al*. Haliangicin, a novel antifungal metabolite produced by a marine myxobacterium. 2. isolation and structural elucidation. *J Antibiot* 2001;54:153–156.
56. Barras F, van Gijsegem F, Chatterjee AK. Extracellular enzymes and pathogenesis of soft-rot erwinia. *Annu Rev Phytopathol* 1994;32:201–234.
57. Vanneste JL. *Fire blight: the disease and its causative agent, Erwinia amylovora*. Wallingford: CABI; 2000.
58. Jousset A, Bienhold C, Chatzinotas A, Gallien L, Gobet A, *et al*. Where less may be more: how the rare biosphere pulls ecosystems strings. *ISME J* 2017;11:853–862.
59. Gellie NJ, Breed MF, Mortimer PE, Harrison RD, Xu J, *et al*. Networked and embedded scientific experiments will improve restoration outcomes. *Front Ecol Environ* 2018;16:288–294.
60. Giblin J, Ramos I, Grout N. Dismantling the master's house: thoughts on representing empire and decolonising museums and public spaces in practice an introduction. *Third Text* 2019;33:471–486.
61. Parker LA. The ecologist. *Decolonising our minds and public spaces*. 2018
62. Gonzalez A, Navas-Molina JA, Kosciolek T, McDonald D, Vázquez-Baeza Y, *et al*. Qiita: rapid, web-enabled microbiome meta-analysis. *Nat Methods* 2018;15:796–798.