Urban soil microbiomes exhibit taxonomic and functional potential for enhanced contaminant cycling

Izabel L. Stohel[1], Young C. Song[1], Anna B. Turetcaia[2], Andrew J. Wilson[2], Dietrich Epp Schmidt[3], Stephanie A. Yarwood[3], Andrew Townsend[1], Emily B. Graham[2*]

[1]Pacific Northwest National Laboratory, Environmental Molecular Sciences Laboratory, Richland, WA 99354, United States.

[2]Pacific Northwest National Laboratory, Environmental and Biological Sciences Directorate, Laboratory, Richland, WA 99354, United States.

[3]University of Maryland, College Park, MD 20742, USA

*Corresponding author: emily.graham@pnnl.gov

Abstract.

Urbanization is a leading cause of global biodiversity loss, but its impact on soil microorganisms and biogeochemistry remain uncertain. Although all urban soil microbiomes are influenced by common anthropogenic processes, we lack generalizable patterns in how urbanization impacts the composition and functional potential of the global soil microbiome. These patterns are vital for grasping cities' roles in biogeochemical processes and managing urban land expansion. We re-analyze soil metagenomic sequences from the Global Urban Soil Environment Ecology Network (GLUSEEN) to identify coordinated changes in microbial taxonomy and functionality and to pinpoint core taxa and metabolisms across five global cities spanning ecoregions and management regimes. Our findings reveal greater spatial variability in microbial taxonomy compared to functional potential, indicating that common anthropogenic processes might shape the genetic potential of diverse taxa. We identify 111 gene annotations that represent core urban soil functions, present in all cities but rare (<5%) in reference soils. Across low-to-high urban land use gradients; nitrogen, trace/heavy metal, glycerol, and fructose/fructan metabolic processes are linked with highly urbanized areas. Core urban microbial taxa – defined by overrepresentation in urban land uses compared to reference soils – included diverse soil bacteria and a more constrained set of methane- and nitrogen-cycling archaea. Overall, our work aids in deriving testable hypotheses and in modeling the feedbacks between growing metropolitan areas and the processes impacting global environmental change.

Introduction

Despite urban areas covering only 1% of global land, 70% of the world's population is projected to live in cities by 2050 (United Nations 2018; Ritchie, Samborska and Roser 2024). Human activity in these urban centers is reshaping the way organisms evolve and interact with each other (Pickett *et al.* 2011). Research on urbanization's effect on ecosystems is becoming more common, but knowledge about urban soil microbiomes is still sparse, even though they play a crucial role in global biogeochemistry (Dubey *et al.* 2019; Naylor *et al.* 2020). Therefore, a comprehensive understanding of soil microbial community structure and functional potential is essential for predicting the ecological roles of microbes in urban ecosystems.

Rapidly increasing urbanization and associated changes in land-use are responsible for biodiversity loss of macroorganisms (i.e., animal and plant communities) and for increased pollution, both of which may subsequently impact urban biogeochemistry and soil microbiome function (Pauchard *et al.* 2006; Nagendra *et al.* 2014). Land clearing for agriculture and urban expansion, for instance, have resulted in fragmentation of forest biomes across the world (Reinmann *et al.* 2020). This practice also leads to increased pollutants, like nitrate and ammonium, which in turn, change soil biogeochemical properties and microbiome function (Wang *et al.* 2017). Additionally, industrial activities can lead to heavy metal and organic contamination in urban soils, in particular polycyclic aromatic hydrocarbons (PAHs), that are toxic to humans and are added to soil through processes like burning fossil fuels and waste incineration (Fernández-Luqueño *et al.* 2011; Delegan *et al.* 2022). Finally, urban land use and management practices like fertilization, irrigation, and the installation of impervious surfaces increase nutrient inputs into soils via runoff (Fenn *et al.* 2003). Nitrogen runoff in particular can enhance microbial denitrification processes that release nitrous oxide – a potent greenhouse gas (Kettering *et al.* 2013; Toor *et al.* 2017).

In parallel, soil microorganisms (i.e., bacteria, archaea, and fungi) are major conductors of global carbon and nutrient cycles. Soil biogeochemical properties that are impacted by urbanization like moisture, temperature, pH, and carbon and nutrient concentrations have a strong effect on microbial communities (Rousk *et al.* 2010; Brockett, Prescott and Grayston 2012; Karhu *et al.* 2014; Graham *et al.* 2016; Widdig *et al.* 2020). Microbial diversity is also dependent on interspecies and plant-microbe interactions (Wilpiszeski *et al.* 2019; Sharma *et al.* 2020; McClure *et al.* 2022), which have been shown to increase the relative abundance of genes involved in greenhouse gas emissions in urban soils (Delgado-Baquerizo, Eldridge and Yu-Rong 2021). More generally, urban soil microbiomes have been shown to exhibit homogenization of bacterial communities compared to surrounding natural ecosystems (Gill *et al.* 2020; Delgado-Baquerizo, Eldridge and Yu-Rong 2021).

The extent to which the urban environment drives soil microbiomes towards a common state (i.e., microbiome composition and function; urban convergence hypothesis) versus supporting a robust, diverse microbial community (i.e., the urban mosaic) remains debated (Pouyat *et al.* 2015). Within the dataset re-analyzed here, for example, it was reported that soil biogeochemical properties like organic carbon and total nitrogen content converged due to anthropogenic factors (Pouyat *et al.* 2015). Other studies have also seen the effect of urban convergence on soil properties, such as soil particle size and soil carbon content (Herrmann, Schifman and Shuster 2020). However, due to the plethora of land covers and uses contained within different metropolitan areas, it remains to be explored whether or not urbanization has a uniform effect that will drive microbial communities towards greater biodiversity loss or greater taxonomic and functional potential. Further, understanding the extent to which urbanization acts upon microbial taxa versus the functional traits they contain is important for representing the impacts of population growth on global biogeochemistry. Standardized spatial observations distributed within and across cities are key to resolving these questions.

Determining a core soil urban microbiome is essential to study patterns associated with anthropogenic factors. Broadly defined, a core microbiome is a set of shared taxa and/or functional potential across all samples in a dataset or a group within a dataset (Turnbaugh *et al.* 2007; Hamady and Knight 2009; Shade and Handelsman 2012; Neu, Allen and Roy 2024). Microorganisms consistently found in all assemblages within a specific habitat have been proposed as essential for the community's functioning (Turnbaugh and Gordon 2009; Vael and Desager 2009; Benson *et al.* 2010; Sommer and Bäckhed 2013). Therefore, pinpointing specific taxa and functional potential in a core microbiome is the initial step in identifying a 'healthy' community and forecasting how the community will react to disturbances. Methods to determine a core microbiome can vary with experimental factors such as spatial or temporal scale and can leverage occurrence-based data, relative abundance data, or a combination of relative abundance and occurrence data (Shade and Handelsman 2012; Neu, Allen and Roy 2024). Here, we define the core urban soil microbiome as organisms and functional genes that are distributed across global urban land uses and are also scarce or absent from more pristine reference soils.

We use metagenomic sequences from soils collected by the Global Urban Soil Environment Ecology Network (GLUSEEN; Pouyat et al., 2017) to investigate the taxonomic and functional diversity of the urban soil microbiome. We aim to examine if soil taxonomic and functional diversity are conserved across worldwide urban cities, if changes in taxonomy and functional potential are coupled, and if there are ubiquitous (“core”) taxa and functions that are characteristic of urban soils regardless of climate and geography. Previous work with this dataset, by Epp Schmidt, et al. (2019, 2017) has shown some consistency in archaeal and bacterial communities, with archaeal convergence being driven by ammonia- and methane-oxidizing organisms and genes and bacterial convergence being driven by nitrite-oxidizing organisms. However, Epp Schmidt, et al. (2019, 2017) did not explore aspects of urban

microbiomes related to contaminants and other carbon and nitrogen cycling processes. We further extend this work by reporting concomitant changes in taxonomy and functional diversity within and across cities; extracting the core urban microbiome across all cities; and uncovering a newly reported set of widespread soil genomic potential for the tolerance and/or cycling of urban pollutants.

Methods

Dataset Description

We used publicly available metagenomic sequencing data from the Global Urban Soil Environment Ecology Network (GLUSEEN) to investigate the urban soil microbiome from 66 samples across five global cities: USA: Baltimore (n=11), Finland: Helsinki (n=12), Finland: Lahti (n=14), Hungary: Budapest (n=14), and South Africa: Potchefstroom (n =15) (Figure 1). Soils spanned three levels of urbanization that encompass a gradient of common anthropogenic impacts: (1) remnant: within city limits, (2) turf: managed land, and (3) ruderal: high degree of physical disturbance; in addition to reference soils with low urban impacts near each city. Further details on land use classifications, soil collection, and metagenomic sequencing are described by Pouyat, et al. (2015) and Epp Schmidt, et al. (2017). Briefly, soils were collected prior to 2017 from 0-10 cm depth, and shotgun metagenomic sequencing was performed on Illumina HiSeq 3000 using standardized methods. Samples used for this publication are available at MG-RAST (Yarwood and Epp Schmidt). (https://www.mg-rast.org/mgmain.html?mgpage=project&project=mgp16346).

Taxonomic and Functional Inference of Metagenomes

Metagenomic sequences from each soil were retrieved from Metagenomic Rapid Annotations using Subsystems Technology (MG-RAST) and processed using the Joint Genome

Institute (JGI) Integrated Microbial Genomes and Microbiomes (IMG/M) standard workflows (Clum *et al.* 2021). Briefly, reads meeting the QC standards were assembled into contigs using BBDuk version 38.79, bbcms version 38.44 and metaSPADES version 3.13.0 (Nurk *et al.* 2017; Bushnell 2023). Read-based taxonomy was assigned to assembled contigs using the genome-based taxonomy database, and assembled contigs were further annotated against the Kyoto Encyclopedia of Genes and Genomes (i.e., KEGG Orthology, KO; (Kanehisa et al., 2019, 2000, 2023)) as part of standard JGI workflows (Parks *et al.* 2018; Chaumeil *et al.* 2020). While read-based taxonomy is a useful tool for determining microbial clades via metagenomics, future work that assesses taxonomy by sequencing selected portions of conserved marker genes may be a useful complement to the results presented here (Wood and Salzberg 2014; Wood, Lu and Langmead 2019; Clum *et al.* 2021). After sequence processing, we summarized taxa and KO for each soil sample into matrices by counting sequences in each soil assigned to every unique taxon (order-level due inconsistencies in taxonomic annotations extending beyond microbial orders) or KO. Counts in each matrix were normalized using a relative log expression (RLE; Pereira et al., 2018) implemented in the R package 'DESeq2' (Love, Huber and Anders 2014). KO were mapped to pathways using the 'KEGGREST' package in R (version 4.3.1) (Tenenbaum and Maintainer 2023). Code is available on Figshare (https://figshare.com/s/d788b28a869481c25689).

Statistical Analysis

All statistical analyses were implemented in R (version 4.3.1) with base statistical packages and with the 'vegan' and 'usedist' packages (Bittinger 2020; Oksanen *et al.* 2022). We examined how taxonomic diversity varied across land uses and cities by calculating alpha- (richness and Shannon Diversity Index) and beta- diversity (Bray-curtis dissimilarity), followed by ANOVA and Tukey's honestly significant difference (HSD) post hoc test or by permutational multivariate analysis of variance (PERMANOVA) to test for statistically significant differences in

156 alpha- and beta-diversity, respectively. Diversity metrics were independently calculated based
157 on the distribution of microbial orders (for microbial taxonomy) and KO (for microbial functional
158 potential). Then, we individually compared how taxonomic and functional composition changed
159 with geographic distance, using the Haversine dissimilarity matrix within and across cities, using
160 Mantel tests. Relative changes in taxonomic and functional composition were assessed using
161 Mantel test of Bray Curtis dissimilarity for microbial taxonomy vs. functional potential.

162 Additionally, we visualized data in the following ways. For visualizations of read-based
163 taxonomy, we first generated phylogenetic trees using phyloT v2 (phyloT : a phylogenetic tree
164 generator), an online tree generator based on the Genome Taxonomy Database. Then, we
165 visualized the tree in R using the packages, 'ggtree' (Yu *et al.* 2017), 'treeio' (Wang *et al.* 2020),
166 and 'ggnewscale' (Campitelli 2020). We generated heatmaps using the 'pheatmap' package
167 (Kolde 2019). Distributions of shared and unique taxa and KO across cities were visualized
168 using the 'ComplexUpset' package in R (Krassowski 2020; Lex *et al.* 2020). All other
169 visualizations were generated with 'ggplot2' (Wickham 2016).

170 ***Results***

171 *Microbial Diversity of Urban Soils*

172 When comparing diversity metrics across cities, the Baltimore soil microbiome appeared
173 to be distinct. The taxonomic diversity of soil microbiomes in Baltimore was both higher (alpha)
174 and distinct (beta) compared to other cities (Figure 2). Excluding reference soils, taxonomic
175 richness was nearly two-fold higher in Baltimore soils than all other cities (261.46, vs. 120.57-
176 172.20, $P < 0.001$) and Shannon diversity followed a similar trend (**H'**=3.77, vs. **H'**=0.80-1.87, P
177 < 0.001, Figure 2). Baltimore soils also had lower functional richness compared to Budapest
178 and Potchefstroom (mean = 2,189.636 vs. 3440.714 and 3325.067 unique KO per sample, P=
179 0.021 and 0.040, respectively) and the lowest Shannon diversity index when compared to all

180 other metropolitan areas (**H'**= 7.110, **H'** range= 7.669-7.816, P<0, Figure 2). Taxonomic and
181 functional composition in the Baltimore soil microbiome was also distinct when compared with
182 other cities (PERMANOVA, both with and without reference soils P < 0.001, Figure 2).

183 Highly urbanized land uses in most cities contained distinct soil microbiomes relative to
184 reference soils. Though there was no significant change in taxonomic or functional alpha
185 diversity metrics within urban land uses (ruderal, remnant, turf) and only slight differences in
186 alpha diversity between urban and reference land uses (Figure 2); soil microbiomes associated
187 with the most highly urbanized land use (ruderal) were distinct from reference soils with respect
188 to taxonomy and/or functional potential in Baltimore, Potchefstroom, and Helsinki
189 (PERMANOVA, P = 0.02, 0.02, and 0.09 for taxonomy and P = 0.02, 0.18, and 0.01 for
190 functional potential, respectively).

191 We then tested relationships between microbial diversity (taxonomic and functional) and
192 geographic position, and between taxonomic and functional dissimilarity, to assess the extent to
193 which taxonomic and functional attributes of urban microbial communities were coordinated
194 through space (Figure 3). When examining soils from all cities collectively, both taxonomic and
195 functional dissimilarity increased with geographic distance (P= 0.01), an effect that was present
196 both with and without the inclusion of reference soils. For each city individually, neither
197 taxonomic nor functional diversity varied significantly with geographic distance within urban land
198 uses in any city (P > 0.10). We obtained similar results when including reference soils, where
199 the only significant correlations with distance were found in Baltimore (P = 0.04, 0.05 for
200 taxonomy and functional potential, respectively). However, taxonomic and functional
201 composition were partially coupled across all cities collectively (all P = 0.01, with and without
202 reference soils) and within Budapest, Helsinki, and Potchefstroom (all P=0.01, with and without
203 reference soils).

Core Urban Microbial Taxa and Metabolic Potential

Two hundred fifty-seven microbial orders were common to soil microbiomes across all cities (hereafter, 'core' organisms, Figure 4). Although most of these orders were also contained within reference soils, they exhibited clear increases in abundance from reference to more urbanized soils (Figure 5, Figure S1). Among core urban microbial orders, eight bacterial and two archaeal orders had greater than 50% of their total abundance contained within ruderal soils (the most urban land use type, unclassified Candidatus Saccharibacteria, unclassified Crenarchaeota, Caldisericales, Clostridiales, Desulfurellales, Petrotogales, Micrococcales, Archaeoglobales, Chlamydiales, and unclassified Firmicutes), with low abundance in reference soils (0-15% for eight orders and 17-29% for the remaining two orders; Table S1). Eleven more orders (incl. seven archaeal orders) had more than 50% of their total abundance contained within turf soils, with low corresponding abundance in reference soils (<12% of total abundance contained within reference soils, Aeromonadales. Caulobacterales, Burkholderiales, Candidatus Nitrosocaldales, Methanopyrales, unclassified Thermoplasmata, Nitrosopumilales, unclassified Aigarchaeota, unclassified Calescamantes, Thermoplasmatales, unclassified Thaumarchaeota). Overall (regardless of presence in reference soils), the most abundant shared orders of bacteria included Enterobacterales, Burkholderiales, Lactobacillales, Bacillales, Rhizobiales, Micrococcales, Caulobacterales, Acidobacteriales, Clostridiales, and Pseudomonadales.

Core urban archaeal orders were generally more evenly distributed across reference and urban land uses, however, several taxa still appeared to be more prevalent in urbanized soils. No core orders were found only in urban land uses, but 7 of 18 core archaeal orders were at least twice as abundant in ruderal vs. reference soils: Methanopyrales, Nitrosopumilales, Nitrososphaerales, Thermoproteales, Thermoplasmatales, Archaeoglobales, and Methanocellales. Overall (regardless of presence in reference soils), Nitrososphaerales was by far the most abundant shared archaeal order, followed by other nitrogen- and methane-cycling

229 (Nitrosopumilales, Methanosarcinales, Methanomicrobiales, Methanobacteriales) and salt-
230 tolerant archaea (Halobacteriales, Haloferacales, Natrialbales). Interestingly, after excluding
231 reference soils, Baltimore soils displayed five times as many auxiliary (i.e., not in core
232 microbiome) microbial orders as any other city (25 vs. five or less in all other cities; Figure 4).

233 Further, we found 3,667 KO shared across urban land uses in all cities (Figure 4).
234 Regardless of their presence in reference soils, the top 10 most abundant core KO encoded
235 proteins that are associated with generalized cell functions (K07347 outer membrane usher
236 protein, K01992, K02004, and K01990 ABC-2 type transport system permease and ATP-binding
237 proteins, K02529 LacI family transcriptional regulator, galactose operon repressor, K03088
238 polymerase sigma-70 factor ECF subfamily, K01652 acetolactate synthase I/II/III large subunit
239 [EC:2.2.1.6], K07497 putative transposase), notably including two metal-dependent functions:
240 K02014 (iron complex outer membrane receptor protein) and K02035 (peptide/nickel transport
241 system substrate-binding protein).

242 When narrowing these potential functions to those not widely contained within reference
243 soils, our results point to a core set of 111 gene annotations that were found in urban land uses
244 in all five cities and in less than 5% of reference soils. In particular, 29 KO were common to
245 urbanized land in all cities but were not found in any reference soil. Of these, 12 were found in
246 at least 25% of soils from the most urbanized land use type (ruderal). This included 4 genes
247 associated with nitrogen cycling (K02591 *nifK*, nitrogenase molybdenum-iron protein beta chain,
248 K02586 *nifD*, nitrogenase molybdenum-iron protein alpha chain, K02588 *nifH*, nitrogenase iron
249 protein NifH, K12256 *spuC* putrescine---pyruvate transaminase [EC:2.6.1.113]). In addition to
250 the two Mo-dependent genes related to N cycling (*nifK* and *nifD*), five other genes in this group
251 are closely associated with heavy metals, contaminants, and/or microbial stress responses
252 (K05297, rubredoxin---NAD+ reductase; K18601 aldehyde dehydrogenase [EC:1.2.1.-]; K07231
253 putative iron-regulated protein, K03332, *fruA*, fructan beta-fructosidase [EC:3.2.1.80], K01960,

254 *pycB* pyruvate carboxylase subunit B). The three remaining genes were associated with
255 common cellular functions (K19123 *casA*, CRISPR system Cascade subunit CasA, K03041,
256 *rpoA1*, DNA-directed RNA polymerase subunit A' [EC:2.7.7.6], K10726, replicative DNA
257 helicase Mcm [EC:5.6.2.3]).

258 If we further extend this analysis to urban soil KO found in less than 5% of reference
259 samples, we generate an additional 82 KO for consideration, of which 34 are found in more than
260 25% of soils from the most urbanized land use type (Table S1). Notable KO include: K08352
261 (thiosulfate reductase/polysulfide reductase chain A [EC:1.8.5.5], in 74% of ruderal soils);
262 K13794 (LysR family transcriptional regulator, regulatory protein for *tcuABC*, in 70% of ruderal
263 soils); K00869 and K00938 (mevalonate kinase [EC:2.7.1.36] and phosphomevalonate kinase
264 [EC:2.7.4.2], in 48% and 39%); K02587 and K02229 (*nifE*, nitrogenase molybdenum-cofactor
265 synthesis protein NifE, and *cobG*, precorrin-3B synthase [EC:1.14.13.83], in 44% and 26%);
266 K00520 and K08365 (mercuric reductase [EC:1.16.1.1] and MerR family transcriptional
267 regulator, mercuric resistance operon regulatory protein, in 44% and 26%); K18833 (MFS
268 transporter, DHA3 family, multidrug efflux protein, in 35%); K09819 (manganese/iron transport
269 system permease protein, in 26%); K00596 (2,2-dialkylglycine decarboxylase (pyruvate), in
270 30%); genes related to fructose/fructan (K10554, *frcA* fructose transport system ATP-binding
271 protein and K00895 diphosphate-dependent phosphofructokinase [EC:2.7.1.90], in 30% and
272 30%), K16876 (2-furoate---CoA ligase [EC:6.2.1.31], in 26%); genes related to glycerol
273 contamination (K17325 glycerol transport system ATP-binding protein in 30%, K00712 *tagE*
274 poly(glycerol-phosphate) alpha-glucosyltransferase [EC:2.4.1.52] in 26%, K17322 *glpP* glycerol
275 transport system permease protein in 26%), signatures of viral infection (K01186, sialidase-1
276 [EC:3.2.1.18] in 26%, K19046, *casB*, CRISPR system Cascade subunit CasB, in 39%), some
277 uncharacterized functions, and several functions related to amino acid transport and
278 metabolism.

Despite having the fewest auxiliary functional genes, Baltimore soils harbored a distinctly unique microbiome enriched in genes associated with heavy metal resistance and toxin processing, in contrast to other cities where auxiliary functions were dominated by common cellular processes. Excluding reference soils, Lahti soils had the most auxiliary microbial KO at 2895, followed by Budapest (2190), Helsinki (2104), Potchefstroom (2053), and Baltimore (1122). While the auxiliary microbiome of Baltimore soils contained the fewest potential functions of any city, it was also mostly unique in comparison to other cities. The most abundant of these auxiliary potential functions is thought to convey heavy metal resistance (K15725, outer membrane protein, heavy metal efflux system). Other highly abundant auxiliary functions in Baltimore soils were also associated with metals or other toxins, such as: K01578 MLYCD, malonyl-CoA decarboxylase [EC:4.1.1.9], K15855, exo-1,4-beta-D-glucosaminidase [EC:3.2.1.165], K01577 *oxc*, oxalyl-CoA decarboxylase [EC:4.1.1.8], K04108 4-hydroxybenzoyl-CoA reductase subunit gamma [EC:1.1.7.1],and K01301 N-acetylated-alpha-linked acidic dipeptidase [EC:3.4.17.21]. In contrast, the auxiliary microbiome of other cities was mostly comprised of common cellular processes (K01223, K02761, K07345, K02073, K03475).

Discussion

Urban centers are increasing in extent and population density, in turn disturbing soil microbiomes and the biogeochemical processes that they mediate (Christel *et al.* 2023). Soil microbial functions are an essential component of global biogeochemical cycles, as microorganisms regulate soil C storage and respiration (Gougoulias, Clark and Shaw 2014). Here, we used 89 soil metagenomes from the GLUSEEN network to evaluate patterns in the taxonomic and functional potential of urban soil microbiomes. We extend the work of Epp Schmidt by revealing a suite of core urban microorganisms and metabolisms related to (1) trace and heavy metals, (2) fructose/fructan, (3) viral infection, and (4) glycerol; in addition to N

303 cycling processes that have been reported previously. These novel findings present a basis for
304 understanding and modeling the ways in which urban biogeochemical cycles are distinct from
305 natural ecosystems and for sustainably managing growing population centers.

306 *Taxonomy and Functional Potential of Urban Soil Microbiomes*

307 We found biogeographical patterns in the urban soil microbiome that align with
308 ecological processes such as microbial niche diversification (urban mosaic hypothesis) and
309 environmental filtering (urban convergence hypothesis). There was a great array of functional
310 potential relative to the amount of taxonomic diversity present in the urban soil microbiome
311 (Figure 2). High functional diversity lends support for the urban mosaic hypothesis, which
312 proposes that microbial diversity increases with urbanization as a result of interacting selective
313 pressures like land use, nutrients, pollution, and other factors on microbiomes (Christel *et al.*
314 2023; Scholier *et al.* 2023). Under the urban mosaic hypotheses, potential microbial functions,
315 and in some cases taxonomic composition, expand to occupy new niche space opened by
316 anthropogenic activity. For example, Mo, et al. and Scholier, et al. (2024; 2023) reported
317 increased soil microbiome alpha diversity in anthropogenic areas compared to natural areas.
318 Taxonomic similarities, on the other hand, have been more widely reported across cities
319 (McKinney 2006; Luck and Smallbone 2010; Danko *et al.* 2021; Lokatis and Jeschke 2022).
320 This provides evidence towards the urban convergence hypothesis, where physical
321 characteristics of the urban environment drive microbial similarities, and supports the existence
322 of a core urban microbiome. In our study, we observed comparatively low taxonomic diversity in
323 urban soils (Figure 2) and a clustering of microbiomes from highly urbanized land uses across
324 cities (Figure 2, NMDS) that lends support for common microbial taxa and functional potential
325 across five vastly different cities. This is consistent with findings in urban greenspaces where
326 despite greater functional diversity, microbial taxonomies are relatively homogenized (Delgado-
327 Baquerizo, Eldridge and Yu-Rong 2021; Mo *et al.* 2024).

328 Interestingly, Baltimore appeared to have a distinct soil microbiome relative to all other
329 urban centers. It had the highest taxonomic diversity but also the least functional diversity as
330 compared to other cities (which had similar diversity, Figure 2). Baltimore was also the only city
331 that showed a sizable increase in taxonomic diversity in more urban (turf and ruderal) vs. more
332 pristine (reference and remnant) soils (Figure 2). We posit that the influence of environmental
333 factors on soil microbiomes may be particularly important in Baltimore, which has a long history
334 of industrialization, leaky sewage infrastructure, and high pollution levels (Yesilonis, Pouyat and
335 Neerchal 2008; Dickerson *et al.* 2024; *Baltimore's Wastewater Treatment Plants*). For example,
336 Baltimore has been shown to display a dramatic urban heat island effect with increased
337 temperatures along an urban gradient (Corpuz *et al.* 2024) and has reportedly high levels of
338 lead and other pollutants (Yesilonis, Pouyat and Neerchal 2008; Schwarz, Pouyat and Yesilonis
339 2016). Each of these factors may both exert unique selective pressures on microbial functions
340 and vary substantially at the local scale. They may also lead to the emergence of genetic
341 potential that is not well-described in microbial communities (i.e., microbial dark matter), which
342 would result in the suppression of detectable functional potential observed in this study. While
343 the reasons driving different patterns in microbial taxonomic and functional diversity in Baltimore
344 are unknown, its distinctive patterns suggest further investigation into the unique taxonomic and
345 functional potential its soil microbiomes contain.

346 *Spatial variation in functional potential and taxonomic composition across the urban soil*
347 *microbiome*

348 While the mixture of diverse land uses within cities may obscure patterns based on
349 space alone, the low magnitude of compositional changes within and across cities further
350 underscores the existence of a core urban microbiome. Differences between urban and
351 reference soils suggest that these core organisms and functions are distinct from more natural
352 environments (Figure 3). We found that urban soil microbiome composition and functional

353 diversity tended to be consistently structured through space, consistent with previous work on
354 this dataset (Epp Schmidt *et al.* 2019). For each individual city, neither taxonomic nor functional
355 diversity varied significantly across spatial distances. Across cities, both taxonomic and
356 functional dissimilarity showed some correlation with spatial distances, 33.7% and 34.6%
357 respectively (Figure 3). This variation in microbial community composition across cities was
358 likely due to overarching drivers such as soil mineralogy, pH, weather, and other abiotic factors
359 that are known to influence soil microbiomes (Brockett, Prescott and Grayston 2012; Cao *et al.*
360 2016; Wang *et al.* 2019; Widdig *et al.* 2020).

361 Interestingly, microbial taxonomic composition and functional potential were partially
362 decoupled from each other and showed different biogeographical relationships (Figure 3).
363 Taxonomic composition was more variable than functional potential through space, both within
364 and across cities (as indicated by flatter slopes on Figure 3), and was virtually uncorrelated to
365 functional potential in Lahti and Baltimore. Microbial functional potential is expected to be
366 somewhat correlated to taxonomic composition due to phylogenetic conservation of microbial
367 traits, which can lead to similar responses to environmental factors like substrate availability and
368 climatic variables (Martiny, Treseder and Pusch 2013). However, if ecological selection
369 operates on microbial genomic elements that are exchanged between organisms or if selection
370 operates on unrelated genes and organisms, shifts in microbial taxonomy may not be
371 coordinated with changes in genomic potential (Knelman and Nemergut 2014; Bier *et al.* 2015;
372 Graham *et al.* 2016; Graham and Knelman 2023). We posit that the diversity of factors such as
373 available substrates, carbon sources, and/or pollutants in the urban environment may lead to
374 variation in gene presence and abundance despite taxonomic convergence. Given the usage of
375 microbial taxa as biomarkers of ecosystem health (Dubey *et al.* 2019) and as engineering
376 solutions (Bala *et al.* 2022), the disconnect between taxonomy and potential function observed
377 here for urban soils is an important consideration in sustainable management practices.

Ubiquitous (core) vs. auxiliary taxa and metabolisms in global cities

Examining the core urban microbiome provides insight into the microorganisms that are likely to mediate the biogeochemical cycling of key nutrients, greenhouse gasses, and pollutants in urban soils. Bacterial orders that were ubiquitous across cities and urban land covers relative to reference soils were typified by microorganisms usually associated with plant health, rhizosphere functions, and antibiotic and/or contaminant production and consumption (Figure 5, Figure S1). For example, Enterobacterales and Rhizobiales reside in the root systems of plants and some members have been found to maintain plant health under abiotic stress (Garrido-Oter *et al.* 2018; Andrés-Barrao *et al.* 2021). *Bacillus*, a common microorganism in the Bacillales order, is also ubiquitous in soils and promotes plant health through nutrient cycling and pathogen protection (Garbeva, van Veen and van Elsas 2003; Liu *et al.* 2019). With regards to antibiotic and contaminant metabolism, some members in the Enterobacterales order (e.g., *Enterobacteria*) harbor increased antibiotic resistance genes in urban centers, while others are human pathogens (e.g., *Salmonella*, *Escherichia*, *Shigella*). Burkholderiales have also been shown to participate in pentachlorophenol degradation, a chemical used as a pesticide that persists in soil and has been coupled to nitrogen, sulfate, and iron reduction (Tong *et al.* 2015; Kang *et al.* 2018; Fresia *et al.* 2019); while Clostridiales and Caulobacterales are commonly associated with a wide range of contaminants (Abbasian *et al.* 2016; Thomas and Cébron 2016). Like bacteria, ubiquitous urban archaea have been previously associated with nitrogen and methane cycling and/or salt tolerance (Fenn *et al.* 2003; Wang *et al.* 2013; Asabere *et al.* 2018; Strokal *et al.* 2021). While we cannot directly assess the coupling of these urban organisms with trace metals, pollutants, and/or harsh conditions using the data available in this study, their high plasticity and potential to acquire new genes (Brunder and Karch 2000; Justice *et al.* 2008; Liu *et al.* 2019) may make them a key focus for future research.

402 We also report core soil microorganisms with greater than 50% of their abundance
403 contained in ruderal soils and corresponding low abundance in reference soils as organisms for
404 deeper investigation (Figure 5, Figure S1). We focus on ruderal soils because they represent
405 the most urbanized land cover type investigated here. First, Saccharibacteria has been reported
406 to be a key member isolated from activated sludge (Ferrari *et al.* 2014; Kindaichi *et al.* 2016)
407 and is often found in soil and soil-like environments. Members of Saccharibacteria are able to
408 metabolize diverse compounds, putatively conveying an ability to thrive in a polluted urban
409 environment (Kindaichi *et al.* 2016). Saccharibacteria are a common constituent of the human
410 oral microbiome which may also explain its presence in urban environments that have high
411 human populations (Bor *et al.* 2019). Members of Clostridiales and other Firmicutes are also
412 commonly associated with contaminant degradation, including polycyclic aromatic hydrocarbons
413 (PAHs), polychlorinated biphenyls (PCBs), and phthalates (PAEs) that are of key interest to
414 urban ecosystems because of their detrimental impacts on human health and long residence
415 times (Wang *et al.* 2013; Quero *et al.* 2015; Kumar *et al.* 2020; Liu *et al.* 2021; Lin *et al.* 2023).
416 Finally, many core urban soil microorganisms that were enriched in ruderal soils are thought to
417 be thermophilic, including Caldisericales, Desulfurellales, and Petrotogales, in addition to having
418 associations with various environmental contaminants. These organisms may therefore be more
419 tolerant to higher temperatures associated with urban heat islands than their mesophilic
420 counterparts (Heaviside, Macintyre and Vardoulakis 2017). The prevalence of these taxa and
421 other taxa with related reported functions in urban soils are suggestive of an urban soil
422 microbiome that may be at least partially adapted to pollutants and extreme temperatures
423 common to urban areas.

424 When extending our analysis to include urban turf (50% of total abundance contained in
425 turf soils with low corresponding abundance in reference soils), we note an overrepresentation
426 of archaea, along with some of the previously discussed bacterial clades associated with

427 contaminant cycling (e.g., Caulobacterales, Burkholderiales). Archaea may be particularly well-
428 suited to turf soils due to their abilities to thrive in extreme environments. Urban turf is subjected
429 to a unique set of stressors including fertilization, irrigation, mowing, chemical applications for
430 weed control and other uses, and high temperatures from lack of tree canopy cover. Indeed,
431 many of the core archaeal clades enriched in turf are associated with nitrogen cycling
432 (Candidatus Nitrosocaldales, Nitrosopumilales, unclassified Thaumarchaeota), anaerobic
433 metabolisms (Methanopyrales), and/or heat tolerance (unclassified Thermoplasmata,
434 Thermoplasmatales, unclassified Aigarchaeota; Pester, Schleper and Wagner 2011; Hedlund *et*
435 *al.* 2015; Qin *et al.* 2017; Kallistova *et al.* 2023; Sriaporn *et al.* 2023; Zheng *et al.* 2024).

436 When considering core functions encoded by the urban soil microbiome, we found a
437 plethora of metal-dependent genes and/or genes related to metal cycling (Table S1). We
438 highlight a suite of genes that were found in all cities, no reference soils, and at least 25% of
439 ruderal soils (Table S2). Many iron-related functions fell in this category including: K05297,
440 rubredoxin---NAD+ reductase, containing a central iron atom, and K07231, a putative iron-
441 regulated protein, K09819, a manganese/iron transport system permease protein. In a study by
442 Goff, et al. (2022), multi-metal contamination led to disruption in iron homeostasis, essential for
443 most microbes, which could be a consequence of interest for urban environments (Andrews,
444 Robinson and Rodríguez-Quiñones 2003). Furthermore, although N cycling genes have been
445 reported previously, we note the dependence of N-cycling genes in the core urban soil
446 microbiome on molybdenum, in particular, *nifK*, *nifD*, and *nifE.* Molybdenum is a known urban
447 contaminant with sources including industrialization and fertilization (Lane *et al.* 2013; Wong *et*
448 *al.* 2021; Tepanosyan, Yenokyan and Sahakyan 2023). These genes display a functional
449 potential for adaptive response to an anthropogenic environment where greater concentrations
450 of metals and contaminants are present.

When we extended our analysis to genes found in all cities and present in more than 25% of ruderal soils and less than 5% of reference soils, we found further evidence for the role of industrial processes in structuring the urban soil microbiome. Both *phsA* and *psrA*, met this threshold and are involved in thiosulfate reductase, which is important in cycling of metals in the environment (Burns and DiChristina 2009). Specifically, *phsA* is required for hydrogen sulfide production, which is toxic to both humans and the environment (Kourtidis, Kelesis and Petrakakis 2008; Quist and Johnston 2023). Additionally, MerA (K00520 mercuric reductase [EC:1.16.1.1]), encoded by the *mer* operon, was found in 43% of ruderal samples. MerA is a protein encoding a subunit in the mercuric reductase enzyme, allowing organisms to survive in the presence of high mercury (Freedman, Zhu and Barkay 2012). Within the same operon, MerR (K08365), a mercuric resistance operon regulatory protein, was found in ruderal soils at 26% (Boyd and Barkay 2012). Fossil fuel combustion has led to increased mercury pollution in the environment, which is expected to continue to rise, and is also highly toxic to humans and the environment (United Nations 2019; Gworek, Dmuchowski and Baczewska-Dąbrowska 2020; Zhang *et al.* 2021).

We also found a suite of genes associated with fructose and glycerol metabolism, and with cellular infection and replication, in more than 25% of ruderal soils and less than 5% of reference soils. While these genes are central to basic metabolic functions in many microorganisms, their disproportionate presence in urban soils raises questions surrounding additional functions these genes may serve in more polluted environments. Fructose, for example, is commonly implicated in plant-microbiome interactions in agricultural and natural soils (Klepek *et al.* 2010; Sasse, Martinoia and Northen 2018; Chen *et al.* 2023; Tharanath, Upendra and Rajendra 2024). However, the urban environment includes additional sources of fructose, such as food waste enriched in high fructose corn syrup (Duffey and Popkin 2008; Pang *et al.* 2021; Martianto *et al.* 2024; Parsa *et al.* 2024) and industrial byproducts including

476 furfural (Chemical Substances Bureau 2008), that may place a greater importance on fructose
477 cycling in urban soils than more natural soils. We specifically highlight *fruA* (K03332), *frcA*
478 (K10554), *pfp* (K00895), and *hmfD* (K16876), as potential areas for future research. Likewise,
479 glycerol is a major component of manufactured products like fertilizers, soaps, cosmetics,
480 antifreeze solutions, and pharmaceuticals, in addition to its natural occurrences (Gerpen 2005;
481 Hájek and Skopal 2010; Pinto and De Araujo Mota 2014). In particular, *glpT* (K17325), *glpP*
482 (K17322), and *tagE* (K00712), may be disproportionately involved in the cycling of these urban
483 waste products in urban soils versus central metabolic processes. Finally, we elucidate three
484 cellular maintenance genes that could be signatures of heightened viral activity in urban soils –
485 K19123 *casA*, CRISPR system Cascade subunit CasA, K03041, *rpoA1*, DNA-directed RNA
486 polymerase subunit A' [EC:2.7.7.6], K10726, replicative DNA helicase Mcm [EC:5.6.2.3]. While
487 the exact reason for the enrichment of these genes in urban soils is unknown, soil viral ecology
488 is rapidly progressing to the forefront of microbial ecology due to the potential for soil viral
489 impacts on human health and soil biogeochemistry (Neri *et al.* 2022; Roux and Emerson 2022;
490 Graham *et al.* 2023; Jansson and Wu 2023; Zimmerman *et al.* 2024). The enrichment of
491 putative viral signatures in urban soils further underscores the importance of new investigations
492 into soil viral ecology within urban centers.

493 Despite many commonalities in urban soil microbiomes, Baltimore soils strikingly
494 contained some unique potential functions that may signify a greater impact of anthropogenic
495 processes on microbial communities in this city. Baltimore's unique functional potential
496 supported diverse genes associated with metal cycling and aromatic degradation (Table S2).
497 Genes unique to Baltimore include K15725, the *czcCBA* efflux system, which is found to play a
498 role in resistance to cobalt, zinc, and cadmium (Nies 2003), and K04108 4-hydroxybenzoyl-CoA
499 reductase, which is found in anaerobic phenol metabolism for benzene degradation (Breese and
500 Fuchs 1998). We therefore posit that cities that have a relatively recent legacy of industrial

501 pollution and fossil fuel combustion, like Baltimore, may harbor organic contaminants (An *et al.*
502 2013; Gieg, Fowler and Berdugo-Clavijo 2014; Tan *et al.* 2015) that drive distinct soil functional
503 potential relative to cities with more distant industrial eras or stronger environmental regulations.

504 ***Future Directions.***

505 Our research brings to light several exciting avenues for future research into urban
506 ecology and biogeochemistry. Efforts to understand core urban microorganisms and functions
507 will benefit from additional sampling campaigns to capture spatial and temporal fluctuations
508 within microbial communities and from methodologies that describe soil microbiome
509 composition and function in more detail. In particular, transitioning from read-based taxonomy to
510 amplicon-based methodologies, or increasing sequencing depth, will likely improve the breadth
511 and resolution of taxonomic identification, providing deeper insights into the microorganisms
512 driving urban soil biogeochemistry. Investigating microbial gene expression through
513 metatranscriptomics or metaproteomics, particularly focusing on the functions identified in this
514 study, will also deepen our understanding of microbial roles within urban ecosystems. Finally,
515 expanding sample collection to encompass a broader range of cities and intra-city locations, as
516 well as temporal fluctuations, will ensure a more comprehensive and representative dataset.
517 These advances can not only refine current ecological models but also provide a robust
518 foundation for hypothesis-driven research and ecosystem management strategies.

519 ***Conclusion***

520 Studies into the unique organisms and functions contained within the urban soil
521 microbiome are essential as the human population continues to expand. Here, we conducted a
522 re-analysis of 66 urban and reference soil microbiomes using metagenomic sequences from the
523 Global Urban Soil Environment Ecology Network (GLUSEEN). Building upon previous work by
524 Epp Schmidt, et al. (2017, 2019), we revealed a set of core microorganisms associated with the

urban soil environment and uncovered a set of potential biogeochemical functions that were overrepresented in urban soils relative to more pristine environments. These included genes associated with industrial byproducts, trace and heavy metal cycling, and glycerol and fructose/fructan metabolism; as well as N-cycling genes previously suggested by Epp Schmidt, et al. (2017, 2019). While we hypothesized that changes in urban soil microbiome taxonomic and functional diversity would be correlated to each other, our results suggested that soil microbial taxonomy was more likely to vary within and across cities than functional potential. Baltimore soils, in particular, had a distinct soil microbiome with functions related to contaminants and heavy metal cycling, that may reflect a more recent history of industrial activity and could be a key insight in maintaining healthy urban ecosystems (Pataki *et al.* 2011). Overall, our synthesis reveals unique and ubiquitous features of the urban soil microbiome that provide a basis for further studies on the effects urbanization has on soil and human health.

Funding.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research program Early Career award to EBG.

Acknowledgements.

The work was performed by Pacific Northwest National Laboratory, operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract DE-AC05-76RL01830. We thank the Global Urban Soil Environment Ecology Network for providing publicly available data upon which this publication is based. We also thank Caroline G. Simonsen, William L. Peterson, Saige G. Smith, and James Kourum for guidance in drafting this publication.

Conflict of Interest.

The authors report no conflicts of interest.

548 ***Figure Captions.***

549 **Figure 1.** Geographic distribution and land use/land cover of urban soil microbiomes. All soils
550 are from locations with a built area land use/land cover classification according to ESRI (Kontgis
551 et al., 2017; 2021). Cities are colored by land use/land cover categories, and sampling locations
552 are denoted by black dots. Land use/land cover distributions for each city are derived from
553 Environmental Systems Research Institute, ESRI (2017 data; Kontgis et al., 2017; 2021). Land
554 use analyses were performed in QGIS 3.32.3-Lima equipped with a semi-automatic
555 classification plug-in.

556 **Figure 2.** Alpha and beta diversity of microbial taxonomy and KO. Shannon Diversity Index (A)
557 and Non-metric MultiDimensional Scaling (NMDS) for each sampling location for taxonomic
558 order (B) and KO (C), colored by city and with distinct shapes for treatment. The top 10 most
559 abundant (D) metabolic pathways are shown with stacked bar plots. In (A), letters represent
560 statistically distinct groupings (P=<0.05). In (B) geographic location accounts for 43% of the
561 variation seen for taxonomic order and 33% of the variation seen for KO (P=0.001,
562 respectively).

563 **Figure 3.** Correlations between urban soil microbiome taxonomy, functional potential, and
564 geographic distance. (A-C) show correlations within individual cities, denoted by color. (A) and
565 (B) show dissimilarity in the composition of KO's and microbial order, respectively, vs.
566 geographic distance. (C) shows KO vs. taxonomic dissimilarity. (D) and (E) show dissimilarity in
567 the composition of KO's and microbial order, respectively, vs. geographic distance across all
568 cities. For all plots, geographic dissimilarity is calculated using Haversine distances, and KO
569 and taxonomic dissimilarity are calculated with Bray-Curtis dissimilarity. Statistical significance
570 was determined using Mantel tests. Taxonomy is grouped at the order-level for all analyses.

571 **Figure 4.** Microbial organisms and functional potential across cities. (A) and (B) show UpSet
572 plots of microbial taxonomy (order-level) and functional potential (KO) respectively. Sets of
573 intersections between cities are shown by dots on the x-axis. The number of shared (A)
574 microbial orders or (B) KO for each intersection is denoted by bar height. Insets show the total
575 number of microbial orders or KO per city.

576 **Figure 5.** Ubiquitous urban soil bacteria. (A) Phylogenetic tree of soil bacteria found across all
577 five cities. Color bars denote phyla. Grayscale heatmap shows the normalized abundance of
578 each bacterial order in urban soils (i.e., all soils except reference soils). (B) Cleveland dot plot
579 showing the percent of soils containing each microbial order in each soil type. The mean across
580 all samples is set at zero (black line) and over/underrepresentation is denoted by dots further
581 away from the center line. Land use type is depicted by colors, warmer colors indicate more
582 urban land cover. (C) Histograms showing the distribution of data in (B) for each land cover
583 type. The vertical gray line represents zero deviation from the mean of all samples (equivalent
584 to the black line in C). The mean of each land cover type is shown with a dashed vertical line of
585 that color. Histograms with means to the right of 0 suggest an overrepresentation of selected
586 bacterial orders in the associated land cover type.

587 **Figure 6.** Selected metabolic pathways of urban soil microbiomes that are found in less than
588 5% of reference soils. Selected pathways were manually reconstructed from KEGG pathways
589 and are color-coded as indicated in the legend.

590

591 ***References.***

592 [dataset]*
593
594 *2-FURALDEHYDE (Furfural) SUMMARY RISK ASSESSMENT REPORT*. The Netherlands:
595 Chemical Substances Bureau, 2008.

596 Abbasian F, Lockington R, Megharaj M *et al.* The Biodiversity Changes in the Microbial
597 Population of Soils Contaminated with Crude Oil. *Curr Microbiol* 2016;**72**:663–70.

598 An D, Caffrey SM, Soh J *et al.* Metagenomics of Hydrocarbon Resource Environments Indicates
599 Aerobic Taxa and Genes to be Unexpectedly Common. *Environ Sci Technol*
600 2013;**47**:10708–17.

601 Andrés-Barrao C, Alzubaidy H, Jalal R *et al.* Coordinated bacterial and plant sulfur metabolism
602 in Enterobacter sp. SA187–induced plant salt stress tolerance. *Proc Natl Acad Sci U S A*
603 2021;**118**:e2107417118.

604 Andrews SC, Robinson AK, Rodríguez-Quiñones F. Bacterial iron homeostasis. *FEMS Microbiol*
605 *Rev* 2003;**27**:215–37.

606 Asabere SB, Zeppenfeld T, Nketia KA *et al.* Urbanization Leads to Increases in pH, Carbonate,
607 and Soil Organic Matter Stocks of Arable Soils of Kumasi, Ghana (West Africa). *Front*
608 *Environ Sci* 2018;**6**, DOI: 10.3389/fenvs.2018.00119.

609 Bala S, Garg D, Thirumalesh BV *et al.* Recent Strategies for Bioremediation of Emerging
610 Pollutants: A Review for a Green and Sustainable Environment. *Toxics* 2022;**10**:484.

611 *Baltimore's Wastewater Treatment Plants*. https://www.cbf.org/about-
612 cbf/locations/maryland/issues/baltimores-wastewater-treatment-plants.html (February
613 21, 2025, date last accessed)

614 Benson AK, Kelly SA, Legge R *et al.* Individuality in gut microbiota composition is a complex
615 polygenic trait shaped by multiple environmental and host genetic factors. *Proc Natl*
616 *Acad Sci U S A* 2010;**107**:18933–8.

617 Bier RL, Bernhardt ES, Boot CM *et al.* Linking microbial community structure and microbial
618 processes: an empirical and conceptual overview. *FEMS Microbiol Ecol* 2015;**91**:fiv113.

619 Bittinger K. usedist: Distance Matrix Utilities. 2020.

620 Bor B, Bedree JK, Shi W *et al.* Saccharibacteria (TM7) in the Human Oral Microbiome. *J Dent*
621 *Res* 2019;**98**:500–9.

622 Boyd E, Barkay T. The Mercury Resistance Operon: From an Origin in a Geothermal
623 Environment to an Efficient Detoxification Machine. *Front Microbiol* 2012;**3**, DOI:
624 10.3389/fmicb.2012.00349.

625 Breese K, Fuchs G. 4-Hydroxybenzoyl-CoA reductase (dehydroxylating) from the denitrifying
626 bacterium Thauera aromatica. *Eur J Biochem* 1998;**251**:916–23.

627 Brockett B, Prescott C, Grayston S. Soil Moisture is the Major Factor Influencing Microbial
628 Community Structure and Enzyme Activities Across Seven Biogeoclimatic Zones in
629 Western Canada. *Soil Biol Biochem* 2012;**44**:9–20.

630 Brunder W, Karch H. Genome plasticity in Enterobacteriaceae. *Int J Med Microbiol IJMM*
631 2000;**290**:153–65.

632 Burns JL, DiChristina TJ. Anaerobic Respiration of Elemental Sulfur and Thiosulfate by
633 Shewanella oneidensis MR-1 Requires psrA, a Homolog of the phsA Gene of
634 Salmonella enterica Serovar Typhimurium LT2. *Appl Environ Microbiol* 2009;**75**:5209–
635 17.

636 Bushnell B. BBMap. 2023.

637 Campitelli E. ggnewscale: Multiple Fill and Colour Scales in “ggplot2”. 2020.

638 Cao H, Chen R, Wang L *et al.* Soil pH, total phosphorus, climate and distance are the major
639 factors influencing microbial activity at a regional spatial scale. *Sci Rep* 2016;**6**:25815.

640 Chaumeil P-A, Mussig AJ, Hugenholtz P *et al.* GTDB-Tk: a toolkit to classify genomes with the
641 Genome Taxonomy Database. *Bioinformatics* 2020;**36**:1925–7.

642 Chen J, Sun M, Xiao G *et al.* Starving the enemy: how plant and microbe compete for sugar on
643 the border. *Front Plant Sci* 2023;**14**, DOI: 10.3389/fpls.2023.1230254.

644 Christel A, Dequiedt S, Chemidlin-Prevost-Bouré N *et al.* Urban land uses shape soil microbial
645 abundance and diversity. *Sci Total Environ* 2023;**883**:163455.

646 Clum A, Huntemann M, Bushnell B *et al.* DOE JGI Metagenome Workflow. *mSystems*
647 2021;**6**:e00804-20.

648 Corpuz B, Zaitchik B, Waugh D *et al.* Shifting Islands: How weather conditions and urban form
649 shape the spatiotemporal character of Baltimore’s urban heat island. *Urban Clim*
650 2024;**56**:102058.

651 Danko D, Bezdan D, Afshin EE *et al.* A global metagenomic map of urban microbiomes and
652 antimicrobial resistance. *Cell* 2021;**184**:3376-3393.e17.

653 Delegan Y, Sushkova S, Minkina T *et al.* Diversity and Metabolic Potential of a PAH-Degrading
654 Bacterial Consortium in Technogenically Contaminated Haplic Chernozem, Southern
655 Russia. *Processes* 2022;**10**:2555.

656 Delgado-Baquerizo M, Eldridge DJ, Yu-Rong L. *Global Homogenization of the Structure and*
657 *Function in the Soil Microbiome of Urban Greenspaces | Science Advances*.
658 https://www.science.org/doi/full/10.1126/sciadv.abg5809 (December 18, 2023, date last
659 accessed)

660 Dickerson RR, Stratton P, Ren X *et al.* Mobile laboratory measurements of air pollutants in
661 Baltimore, MD elucidate issues of environmental justice. *J Air Waste Manag Assoc*
662 2024.

663 Dubey A, Malla MA, Khan F *et al.* Soil microbiome: a key player for conservation of soil health
664 under changing climate. *Biodivers Conserv* 2019;**28**:2405–29.

665 Duffey KJ, Popkin BM. High-Fructose Corn Syrup: Is this what's for dinner? *Am J Clin Nutr*
666 2008;**88**:1722S-1732S.

667 Epp Schmidt DJ, Kotze DJ, Hornung E *et al.* Metagenomics Reveals Bacterial and Archaeal
668 Adaptation to Urban Land-Use: N Catabolism, Methanogenesis, and Nutrient
669 Acquisition. *Front Microbiol* 2019;**10**.

670 Epp Schmidt DJ, Pouyat R, Szlavecz K *et al.* Urbanization erodes ectomycorrhizal fungal
671 diversity and may cause microbial communities to converge. *Nat Ecol Evol* 2017;**1**:0123.

672 Fenn ME, Baron JS, Allen EB *et al.* Ecological Effects of Nitrogen Deposition in the Western
673 United States. *BioScience* 2003;**53**:404–20.

674 Fernández-Luqueño F, Valenzuela-Encinas C, Marsch R *et al.* Microbial communities to
675 mitigate contamination of PAHs in soil—possibilities and challenges: a review. *Environ*
676 *Sci Pollut Res* 2011;**18**:12–30.

677 Ferrari B, Winsley T, Ji M *et al.* Insights into the distribution and abundance of the ubiquitous
678 Candidatus Saccharibacteria phylum following tag pyrosequencing. *Sci Rep*
679 2014;**4**:3957.

680 Freedman Z, Zhu C, Barkay T. Mercury Resistance and Mercuric Reductase Activities and
681 Expression among Chemotrophic Thermophilic Aquificae. *Appl Environ Microbiol*
682 2012;**78**:6568–75.

683 Fresia P, Antelo V, Salazar C *et al.* Urban metagenomics uncover antibiotic resistance
684 reservoirs in coastal beach and sewage waters. *Microbiome* 2019;**7**:35.

685 Garbeva P, van Veen JA, van Elsas JD. Predominant Bacillus spp. in Agricultural Soil under
686 Different Management Regimes Detected via PCR-DGGE. *Microb Ecol* 2003;**45**:302–16.

687 Garrido-Oter R, Nakano RT, Dombrowski N *et al.* Modular Traits of the Rhizobiales Root
688 Microbiota and Their Evolutionary Relationship with Symbiotic Rhizobia. *Cell Host*
689 *Microbe* 2018;**24**:155-167.e5.

690 Gerpen JV. Biodiesel processing and production. *Fuel Process Technol* 2005;**86**:1097–107.

691 Gieg LM, Fowler SJ, Berdugo-Clavijo C. Syntrophic biodegradation of hydrocarbon
692 contaminants. *Curr Opin Biotechnol* 2014;**27**:21–9.

693 Gill AS, Purnell K, Palmer MI *et al.* Microbial Composition and Functional Diversity Differ Across
694 Urban Green Infrastructure Types. *Front Microbiol* 2020;**11**:912.

695 Goff JL, Chen Y, Thorgersen MP *et al.* Mixed heavy metal stress induces global iron starvation
696 response. *ISME J* 2022;**17**, DOI: 10.1038/s41396-022-01351-3.

697 Gougoulias C, Clark JM, Shaw LJ. The role of soil microbes in the global carbon cycle: tracking
698 the below-ground microbial processing of plant-derived carbon for manipulating carbon

699 dynamics in agricultural systems. *J Sci Food Agric* 2014;**94**:2362–71.

700 Graham EB, Camargo AP, Wu R *et al.* Global Biogeography of the Soil Virosphere.
701 2023:2023.11.02.565391.

702 Graham EB, Knelman JE. Implications of Soil Microbial Community Assembly for Ecosystem
703 Restoration: Patterns, Process, and Potential. *Microb Ecol* 2023;**85**:809–19.

704 Graham EB, Knelman JE, Schindlbacher A *et al.* Microbes as Engines of Ecosystem Function:
705 When Does Community Structure Enhance Predictions of Ecosystem Processes? *Front*
706 *Microbiol* 2016;**7**, DOI: 10.3389/fmicb.2016.00214.

707 Gworek B, Dmuchowski W, Baczewska-Dąbrowska AH. Mercury in the terrestrial environment:
708 a review. *Environ Sci Eur* 2020;**32**:128.

709 Hájek M, Skopal F. Treatment of glycerol phase formed by biodiesel production. *Bioresour*
710 *Technol* 2010;**101**:3242–5.

711 Hamady M, Knight R. Microbial community profiling for human microbiome projects: Tools,
712 techniques, and challenges. *Genome Res* 2009;**19**:1141–52.

713 Heaviside C, Macintyre H, Vardoulakis S. The Urban Heat Island: Implications for Health in a
714 Changing Environment. *Curr Environ Health Rep* 2017;**4**:296–305.

715 Hedlund BP, Murugapiran SK, Alba TW *et al.* Uncultivated thermophiles: current status and
716 spotlight on 'Aigarchaeota.' *Curr Opin Microbiol* 2015;**25**:136–45.

717 Herrmann DL, Schifman LA, Shuster WD. Urbanization drives convergence in soil profile texture
718 and carbon content. *Environ Res Lett* 2020;**15**:114001.

719 Jansson JK, Wu R. Soil viral diversity, ecology and climate change. *Nat Rev Microbiol*
720 2023;**21**:296–311.

721 Justice SS, Hunstad DA, Cegelski L *et al.* Morphological plasticity as a bacterial survival
722 strategy. *Nat Rev Microbiol* 2008;**6**:162–8.

723 Kallistova AYu, Kosyakova AI, Rusanov II *et al.* Methane Production in a Temperate Freshwater
724 Lake during an Intense Cyanobacterial Bloom. *Microbiology* 2023;**92**:638–49.

725 Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Oxf Acad* 2000;**Nucleic**
726 **Acids Research**:27–30.

727 Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Wiley Online*
728 *Libr* 2019;**Protein Science**:1947–51.

729 Kanehisa M, Furumichi M, Y. Sato *et al.* KEGG for taxonomy-based analysis of pathways and
730 genomes. *Oxf Acad* 2023;**Nucleic Acids Research**:D587–92.

731 Kang E, Crouse A, Chevallier L *et al.* Enterobacteria and host resistance to infection. *Mamm*
732 *Genome Off J Int Mamm Genome Soc* 2018;**29**:558–76.

733 Karhu K, Auffret MD, Dungait JAJ *et al.* Temperature sensitivity of soil respiration rates

734 enhanced by microbial community response. *Nature* 2014;**513**:81–4.

735 Kettering J, Ruidisch M, Gaviria C *et al.* Fate of fertilizer 15N in intensive ridge cultivation with
736 plastic mulching under a monsoon climate. *Nutr Cycl Agroecosystems* 2013;**95**:57–72.

737 Kindaichi T, Yamaoka S, Uehara R *et al.* Phylogenetic diversity and ecophysiology of Candidate
738 phylum Saccharibacteria in activated sludge. *FEMS Microbiol Ecol* 2016;**92**:fiw078.

739 Klepek Y-S, Volke M, Konrad KR *et al.* Arabidopsis thaliana POLYOL/MONOSACCHARIDE
740 TRANSPORTERS 1 and 2: fructose and xylitol/H+ symporters in pollen and young
741 xylem cells. *J Exp Bot* 2010;**61**:537–50.

742 Knelman JE, Nemergut DR. Changes in community assembly may shift the relationship
743 between biodiversity and ecosystem function. *Front Microbiol* 2014;**5**, DOI:
744 10.3389/fmicb.2014.00424.

745 Kolde R. pheatmap: Pretty Heatmaps. 2019.

746 Kourtidis K, Kelesis A, Petrakakis M. Hydrogen sulfide (H2S) in urban ambient air. *Atmos*
747 *Environ* 2008;**42**:7476–82.

748 Krassowski M. ComplexUpset. 2020.

749 Kumar B, Verma VK, Joshi D *et al.* Polycyclic aromatic hydrocarbons in urban and rural
750 residential soils, levels, composition profiles, source identification and health risk &
751 hazard. *SN Appl Sci* 2020;**2**:2007.

752 Lane S, Proemse BC, Tennant A *et al.* Concentration measurements and isotopic composition
753 of airborne molybdenum collected in an urban environment. *Anal Bioanal Chem*
754 2013;**405**:2957–63.

755 Lex A, Gehlenborg N, Strobelt H *et al. UpSet: Visualization of Intersecting Sets*.
756 IEEE Transactions on Visualization and Computer Graphics (InfoVis '14), 2020.

757 Lin W, Fan F, Xu G *et al.* Microbial community assembly responses to polycyclic aromatic
758 hydrocarbon contamination across water and sediment habitats in the Pearl River
759 Estuary. *J Hazard Mater* 2023;**457**:131762.

760 Liu J, Cui X, Liu Z *et al.* The Diversity and Geographic Distribution of Cultivable Bacillus-Like
761 Bacteria Across Black Soils of Northeast China. *Front Microbiol* 2019;**10**.

762 Liu Y, Huang Y-H, Lü H *et al.* Persistent contamination of polycyclic aromatic hydrocarbons
763 (PAHs) and phthalates linked to the shift of microbial function in urban river sediments. *J*
764 *Hazard Mater* 2021;**414**:125416.

765 Lokatis S, Jeschke JM. Urban biotic homogenization: Approaches and knowledge gaps. *Ecol*
766 *Appl* 2022;**32**:e2703.

767 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq
768 data with DESeq2. *Genome Biol* 2014;**15**:550.

769 Luck G, Smallbone L. The impact of urbanization on taxonomic and functional similarity among

770 bird communities. *J Biogeogr* 2010;**38**:894–906.

771 Martianto D, Diana R, Baliwati YF *et al.* The quantity and composition of household food waste:
772 Implications for policy. *PLOS ONE* 2024;**19**:e0305087.

773 Martiny AC, Treseder K, Pusch G. Phylogenetic conservatism of functional traits in
774 microorganisms. *ISME J* 2013;**7**:830–8.

775 McClure R, Farris Y, Danczak R *et al.* Interaction Networks Are Driven by Community-
776 Responsive Phenotypes in a Chitin-Degrading Consortium of Soil Microbes. *mSystems*
777 2022;**7**:e0037222.

778 McKinney ML. Urbanization as a major cause of biotic homogenization. *Biol Conserv*
779 2006;**127**:247–60.

780 Mo L, Zanella A, Squartini A *et al.* Anthropogenic vs. natural habitats: Higher microbial
781 biodiversity pays the trade-off of lower connectivity. *Microbiol Res* 2024;**282**:127651.

782 Nagendra H, Sudhira HS, Katti M *et al.* Urbanization and its Impacts on Land Use, Biodiversity
783 and Ecosystems in India. *INTER Discip* 2014;**2**, DOI:
784 10.22201/ceiich.24485705e.2014.2.46532.

785 Naylor D, Sadler N, Bhattacharjee A *et al.* Soil Microbiomes Under Climate Change and
786 Implications for Carbon Cycling. *Annu Rev Environ Resour* 2020;**45**:29–59.

787 Neri U, Wolf Y, Roux S *et al.* Expansion of the global RNA virome reveals diverse clades of
788 bacteriophages. *Cell* 2022;**185**:4023–37.

789 Neu AT, Allen EE, Roy K. *Defining and Quantifying the Core Microbiome: Challenges and*
790 *Prospects | PNAS*. https://www.pnas.org/doi/abs/10.1073/pnas.2104429118 (January
791 29, 2025, date last accessed)

792 Nies DH. Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiol Rev*
793 2003;**27**:313–39.

794 Nurk S, Meleshko D, Korobeynikov A *et al.* metaSPAdes: a new versatile metagenomic
795 assembler. *Genome Res* 2017;**27**:824–34.

796 Oksanen J, Simpson GL, Blanchet FG *et al.* vegan: Community Ecology Package. 2022.

797 Pang S, Song P, Sun X *et al.* Dietary fructose and risk of metabolic syndrome in Chinese
798 residents aged 45 and above: results from the China National Nutrition and Health
799 Survey. *Nutr J* 2021;**20**, DOI: 10.1186/s12937-021-00739-9.

800 Parks DH, Chuvochina M, Waite DW *et al. A Standardized Bacterial Taxonomy Based on*
801 *Genome Phylogeny Substantially Revises the Tree of Life | Nature Biotechnology*.
802 https://www.nature.com/articles/nbt.4229#Sec27 (December 18, 2023, date last
803 accessed)

804 Parsa A, Van De Wiel M, Schmutz U *et al.* Balancing people, planet, and profit in urban food
805 waste management. *Sustain Prod Consum* 2024;**45**:203–15.

Pataki DE, Carreiro MM, Cherrier J *et al. Coupling Biogeochemical Cycles in Urban Environments: Ecosystem Services, Green Solutions, and Misconceptions - Pataki - 2011 - Frontiers in Ecology and the Environment - Wiley Online Library*. https://esajournals.onlinelibrary.wiley.com/doi/full/10.1890/090220 (March 21, 2024, date last accessed)

Pauchard A, Aguayo M, Peña E *et al.* Multiple effects of urbanization on the biodiversity of developing countries: The case of a fast-growing metropolitan area (Concepción, Chile). *Biol Conserv* 2006;**127**:272–81.

Pereira MB, Wallroth M, Jonsson V *et al.* Comparison of normalization methods for the analysis of metagenomic gene abundance data. *BMC Genomics* 2018;**19**:274.

Pester M, Schleper C, Wagner M. The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. *Curr Opin Microbiol* 2011;**14**:300–6.

phyloT : a phylogenetic tree generator.

Pickett STA, Cadenasso ML, Grove JM *et al.* Urban ecological systems: Scientific foundations and a decade of progress. *J Environ Manage* 2011;**92**:331–62.

Pinto BP, De Araujo Mota CJ. Developments in glycerol byproduct-based biorefineries. In: Waldron K (ed.), *Advances in Biorefineries*. Woodhead Publishing, 2014, 364–85.

Pouyat RV, Setälä H, Szlavecz K *et al.* Introducing GLUSEEN: a new open access and experimental network in urban soil ecology. *J Urban Ecol* 2017;**3**:jux002.

Pouyat RV, Yesilonis ID, Dombos M *et al.* A Global Comparison of Surface Soil Characteristics Across Five Cities: A Test of the Urban Ecosystem Convergence Hypothesis. *Soil Sci* 2015;**180**:136.

Qin W, Jewell TNM, Russell VV *et al.* Candidatus Nitrosocaldales. *Bergeys Man Syst Archaea Bact* 2017.

Quero GM, Cassin D, Botter M *et al.* Patterns of benthic bacterial diversity in coastal areas contaminated by heavy metals, polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs). *Front Microbiol* 2015;**6**, DOI: 10.3389/fmicb.2015.01053.

Quist AJL, Johnston JE. Malodors as environmental injustice: health symptoms in the aftermath of a hydrogen sulfide emergency in Carson, California, USA. *J Expo Sci Environ Epidemiol* 2023:1–6.

Reinmann AB, Smith IA, Thompson JR *et al. Urbanization and Fragmentation Mediate Temperate Forest Carbon Cycle Response to Climate - IOPscience*. https://iopscience.iop.org/article/10.1088/1748-9326/abbf16/meta (December 18, 2023, date last accessed)

Ritchie H, Samborska V, Roser M. Urbanization. *Our World Data* 2024.

Rousk J, Bååth E, Brookes PC *et al.* Soil bacterial and fungal communities across a pH gradient

843 in an arable soil. *ISME J* 2010;**4**:1340–51.

844 Roux S, Emerson JB. Diversity in the soil virosphere: to infinity and beyond? *Trends Microbiol*
845 2022;**30**:1025–35.

846 Sasse J, Martinoia E, Northen T. Feed Your Friends: Do Plant Exudates Shape the Root
847 Microbiome? *Trends Plant Sci* 2018;**23**:25–41.

848 Scholier T, Lavrinienko A, Brila I *et al.* Urban forest soils harbour distinct and more diverse
849 communities of bacteria and fungi compared to less disturbed forest soils. *Mol Ecol*
850 2023;**32**:504–17.

851 Schwarz K, Pouyat RV, Yesilonis I. Legacies of Lead in Charm City's Soil: Lessons from the
852 Baltimore Ecosystem Study. *Int J Environ Res Public Health* 2016;**13**:209.

853 Shade A, Handelsman J. Beyond the Venn diagram: the hunt for a core microbiome. *Environ*
854 *Microbiol* 2012;**14**:4–12.

855 Sharma M, Sudheer S, Usmani Z *et al.* Deciphering the Omics of Plant-Microbe Interaction:
856 Perspectives and New Insights. *Curr Genomics* 2020;**21**:343–62.

857 Sommer F, Bäckhed F. The gut microbiota--masters of host development and physiology. *Nat*
858 *Rev Microbiol* 2013;**11**:227–38.

859 Sriaporn C, Campbell KA, Van Kranendonk MJ *et al.* Bacterial and archaeal community
860 distributions and cosmopolitanism across physicochemically diverse hot springs. *ISME*
861 *Commun* 2023;**3**:80.

862 Strokal M, Bai Z, Franssen W *et al.* Urbanization: an increasing source of multiple pollutants to
863 rivers in the 21st century. *Npj Urban Sustain* 2021;**1**:1–13.

864 Tan B, Jane Fowler S, Laban NA *et al.* Comparative analysis of metagenomes from three
865 methanogenic hydrocarbon-degrading enrichment cultures with 41 environmental
866 samples. *ISME J* 2015;**9**:2028–45.

867 Tenenbaum D, Maintainer B. KEGGREST: Client-side REST access to the Kyoto Encyclopedia
868 of Genes and Genomes (KEGG). 2023.

869 Tepanosyan G, Yenokyan T, Sahakyan L. Geospatial patterns and geochemical compositional
870 characteristics of molybdenum in different mediums of an urban environment. *Environ*
871 *Res* 2023;**239**:117340.

872 Tharanath AC, Upendra RS, Rajendra K. Soil Symphony: A Comprehensive Overview of Plant–
873 Microbe Interactions in Agricultural Systems. *Appl Microbiol* 2024;**4**:1549–67.

874 Thomas F, Cébron A. Short-Term Rhizosphere Effect on Available Carbon Sources,
875 Phenanthrene Degradation, and Active Microbiome in an Aged-Contaminated Industrial
876 Soil. *Front Microbiol* 2016;**7**, DOI: 10.3389/fmicb.2016.00092.

877 Tong H, Hu M, Li F *et al.* Burkholderiales participating in pentachlorophenol biodegradation in
878 iron-reducing paddy soil as identified by stable isotope probing. *Environ Sci Process*

879 *Impacts* 2015;**17**:1282–9.

880 Toor GS, Occhipinti ML, Yang Y-Y *et al.* Managing urban runoff in residential neighborhoods:
881 Nitrogen and phosphorus in lawn irrigation driven runoff. *PLoS ONE* 2017;**12**:e0179151.

882 Turnbaugh PJ, Gordon JI. The core gut microbiome, energy balance and obesity. *J Physiol*
883 2009;**587**:4153–8.

884 Turnbaugh PJ, Ley RE, Hamady M *et al.* The human microbiome project. *Nature* 2007;**449**:804–
885 10.

886 United Nations. *World Urbanization Prospects - Population Division - United Nations*.
887 https://population.un.org/wup/Download/ (March 21, 2024, date last accessed)

888 United Nations Environment. Global Mercury Assessment 2018: Key Findings. 2019.

889 Vael C, Desager K. The importance of the development of the intestinal microbiota in infancy.
890 *Curr Opin Pediatr* 2009;**21**:794–800.

891 Wang H, He P, Shen C *et al.* Effect of irrigation amount and fertilization on agriculture non-point
892 source pollution in the paddy field. *Environ Sci Pollut Res* 2019;**26**:10363–73.

893 Wang H, Marshall CW, Cheng M *et al.* Changes in land use driven by urbanization impact
894 nitrogen cycling and the microbial community composition in soils. *Sci Rep*
895 2017;**7**:44049.

896 Wang L-G, Lam TT-Y, Xu S *et al.* Treeio: An R Package for Phylogenetic Tree Input and Output
897 with Richly Annotated and Associated Data. *Mol Biol Evol* 2020;**37**:599–603.

898 Wang X-T, Miao Y, Zhang Y *et al.* Polycyclic aromatic hydrocarbons (PAHs) in urban soils of the
899 megacity Shanghai: Occurrence, source apportionment and potential human health risk.
900 *Sci Total Environ* 2013;**447**:80–9.

901 Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

902 Widdig M, Heintz-Buschart A, Schleuss P-M *et al.* Effects of nitrogen and phosphorus addition
903 on microbial community composition and element cycling in a grassland soil. *Soil Biol*
904 *Biochem* 2020;**151**:108041.

905 Wilpiszeski RL, Aufrecht JA, Retterer ST *et al.* Soil Aggregate Microbial Communities: Towards
906 Understanding Microbiome Interactions at Biologically Relevant Scales. *Appl Environ*
907 *Microbiol* 2019;**85**:e00324-19.

908 Wong MY, Rathod SD, Marino R *et al.* Anthropogenic Perturbations to the Atmospheric
909 Molybdenum Cycle. *Glob Biogeochem Cycles* 2021;**35**:e2020GB006787.

910 Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol*
911 2019;**20**:257.

912 Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact
913 alignments. *Genome Biol* 2014;**15**:R46.

914 *Yarwood SA, Epp Schmidt DJ. MG-RAST.

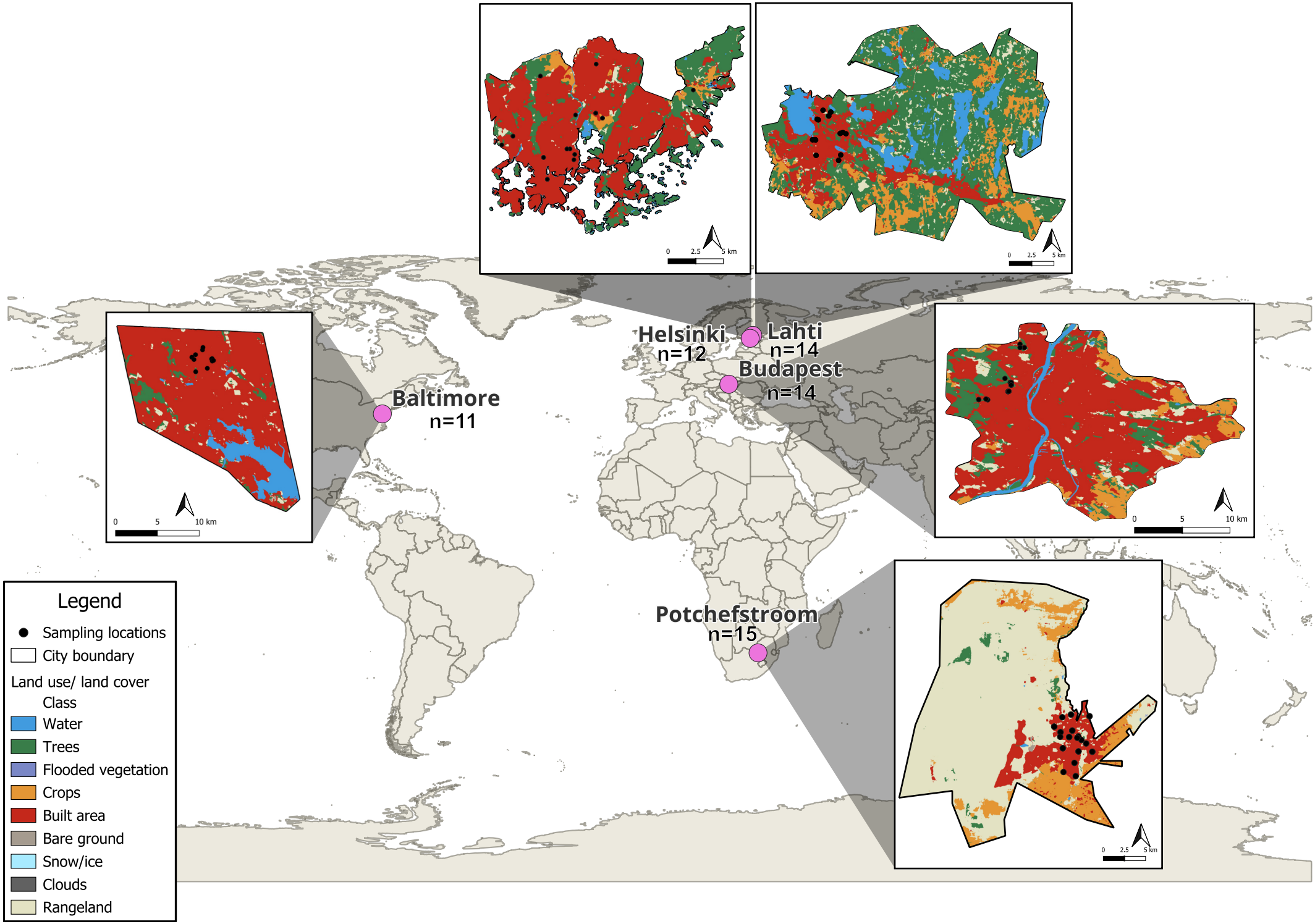
915 Yesilonis ID, Pouyat RV, Neerchal NK. Spatial distribution of metals in soils in Baltimore,
916 Maryland: Role of native parent material, proximity to major roads, housing age and
917 screening guidelines. *Environ Pollut* 2008;**156**:723–31.

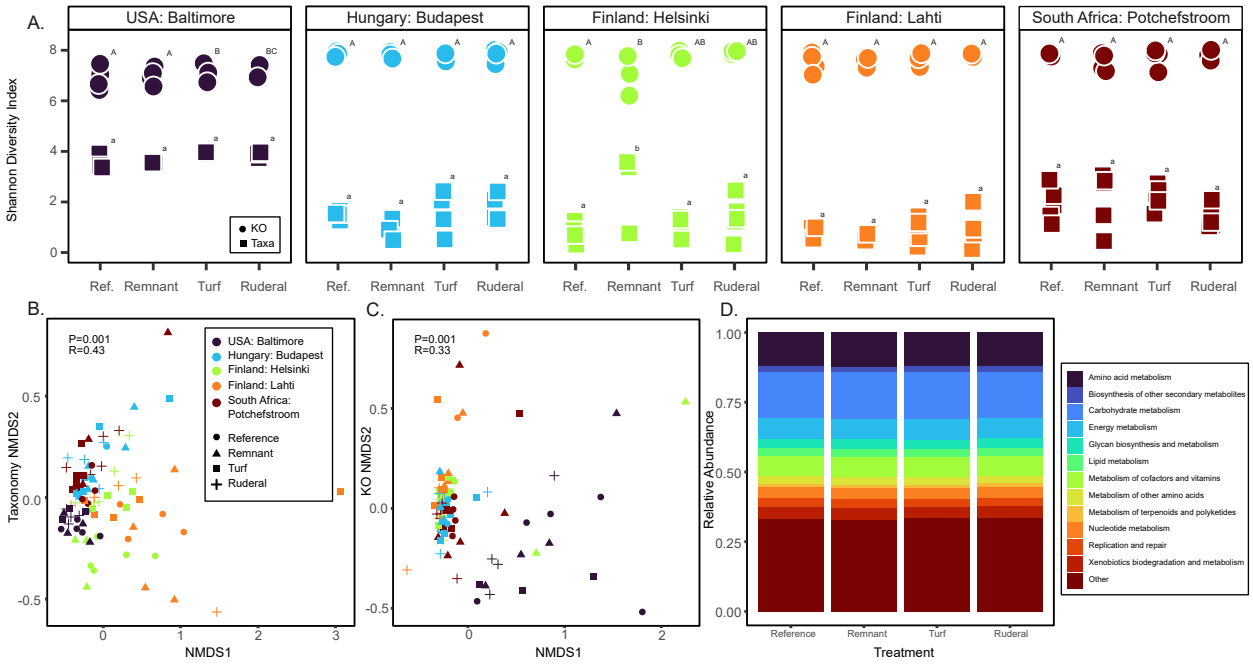
918 Yu G, Smith DK, Zhu H *et al.* ggtree: an r package for visualization and annotation of
919 phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol*
920 2017;**8**:28–36.

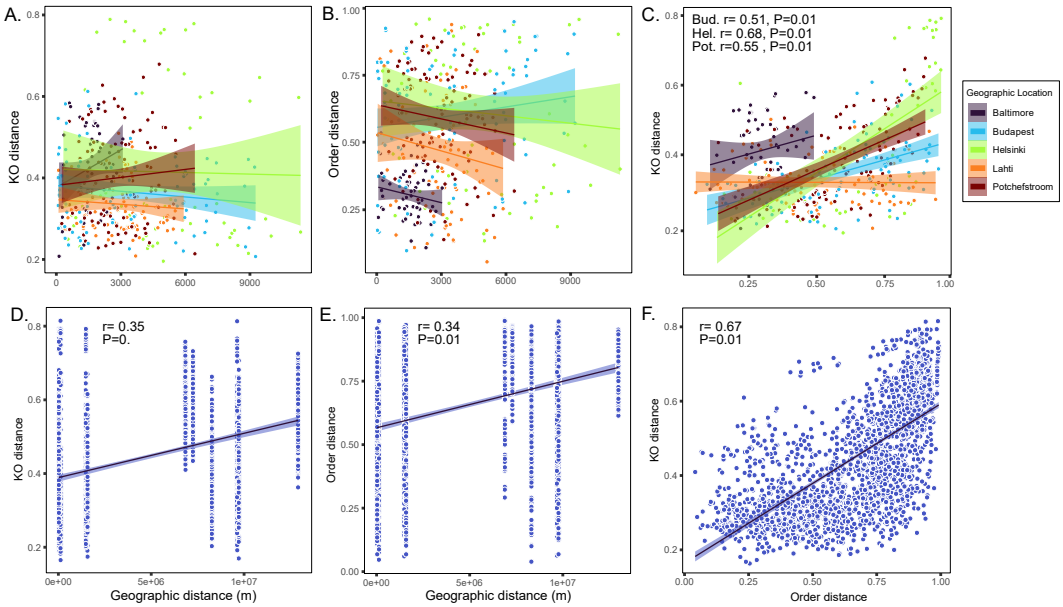
921 Zhang Y, Song Z, Huang S *et al.* Global health effects of future atmospheric mercury emissions.
922 *Nat Commun* 2021;**12**:3035.

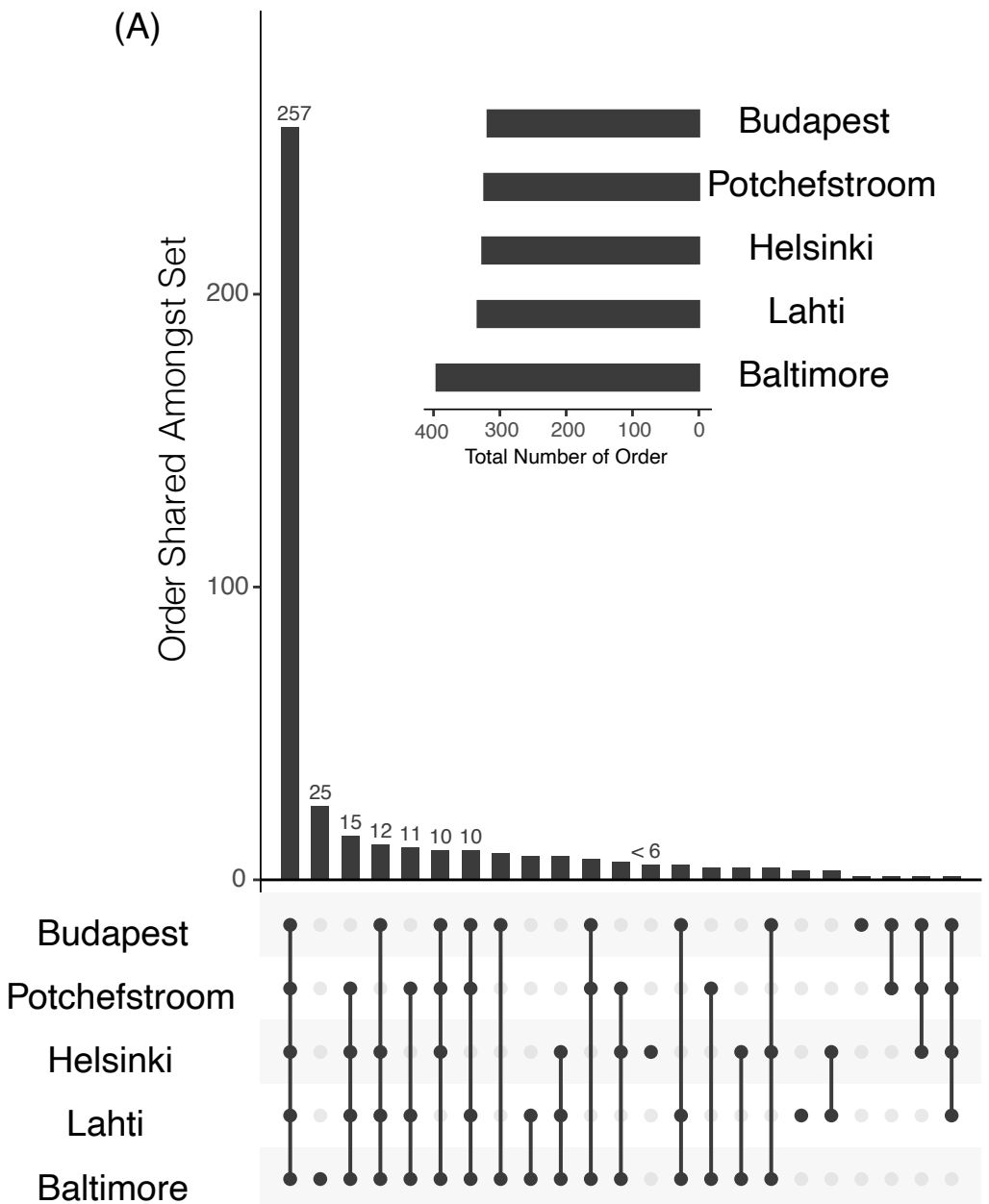
923 Zheng Y, Wang B, Gao P *et al.* Novel order-level lineage of ammonia-oxidizing archaea
924 widespread in marine and terrestrial environments. *ISME J* 2024;**18**:wrad002.

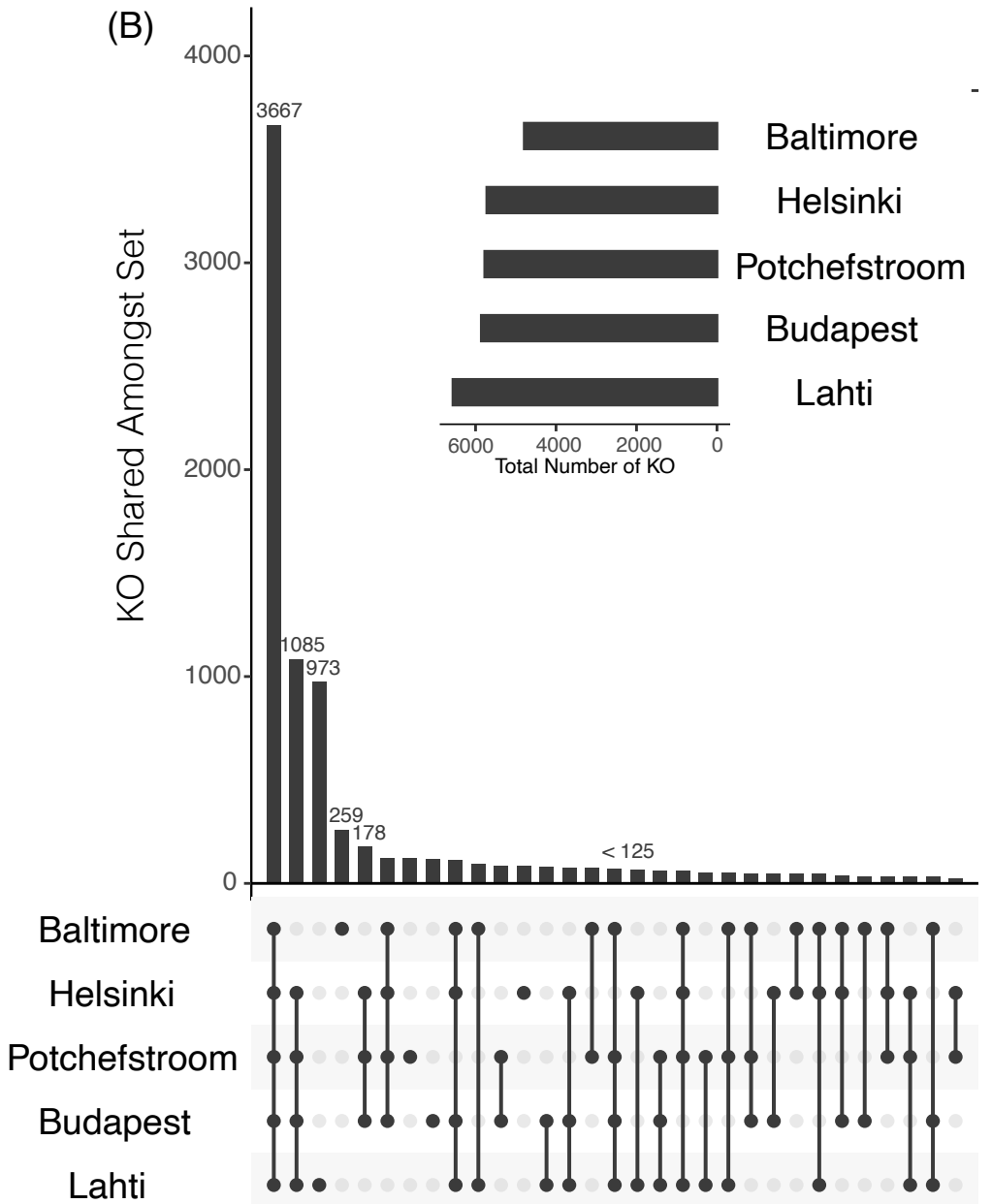
925 Zimmerman AE, Graham EB, McDermott J *et al.* Estimating the Importance of Viral
926 Contributions to Soil Carbon Dynamics. *Glob Change Biol* 2024;**30**:e17524.

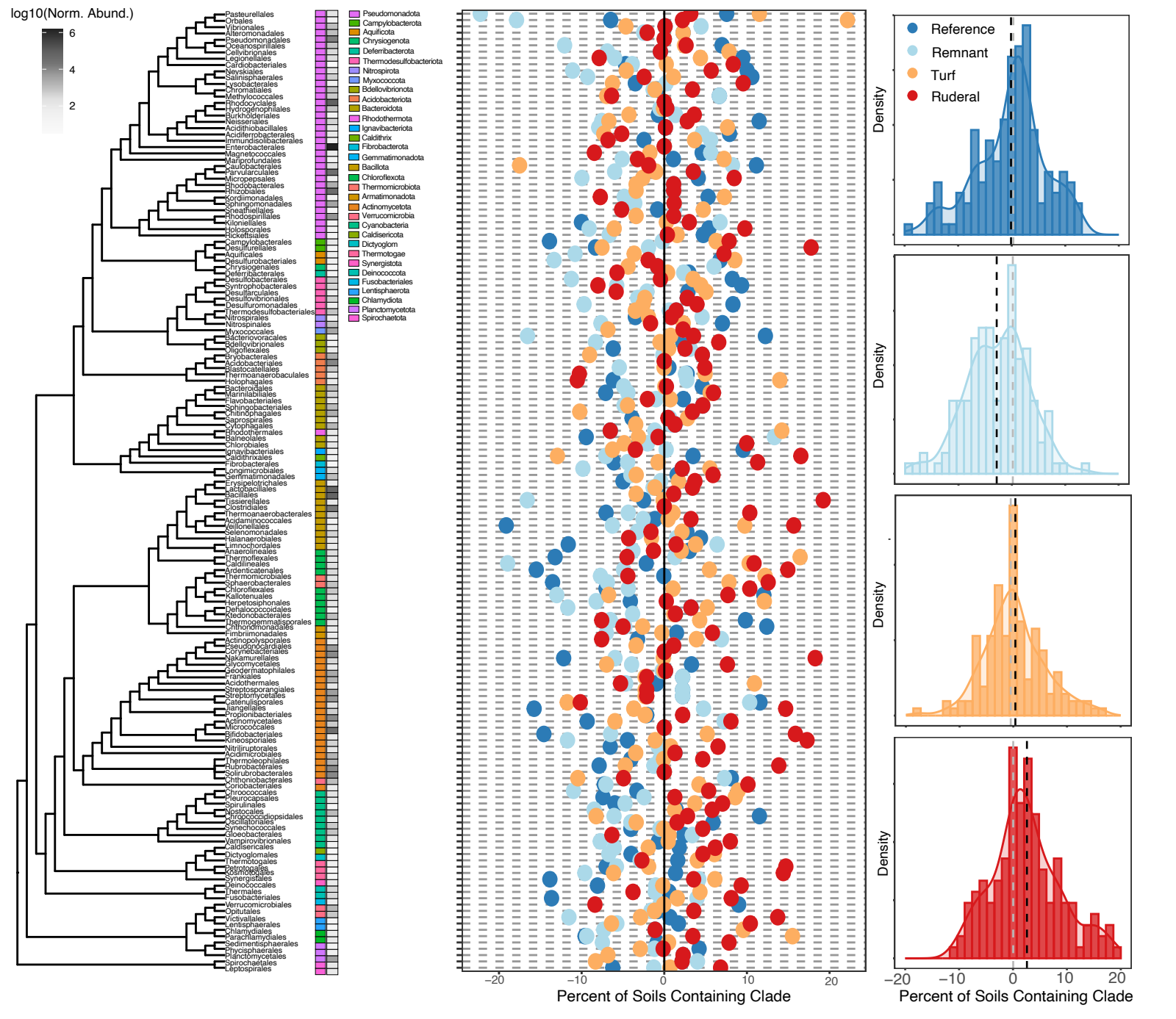

Helsinki
n=12
Lahti
n=14
Budapest
n=14
Baltimore
n=11
Potchefstroom
n=15
0 2.5 5 km
0 2.5 5 km
0 5 10 km
0 5 10 km
0 2.5 5 km
Legend
Sampling locations
City boundary
Land use/ land cover
Class
Water
Trees
Flooded vegetation
Crops
Built area
Bare ground
Snow/ice
Clouds
Rangeland

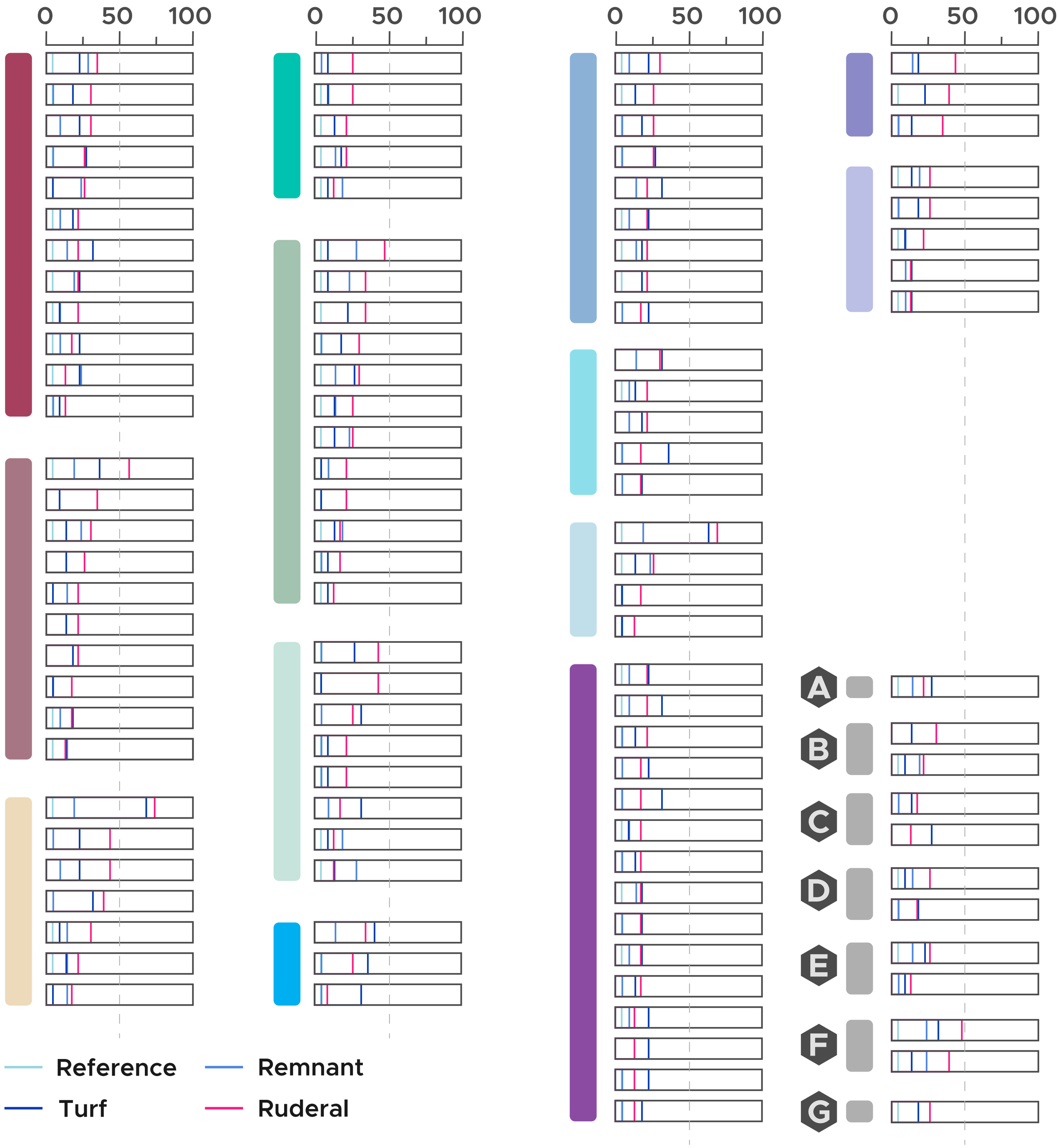

A.
USA: Baltimore
Hungary: Budapest
Finland: Helsinki
Finland: Lahti
South Africa: Potchefstroom
Shannon Diversity Index
KO
Taxa
Ref.
Remnant
Turf
Ruderal
B.
P=0.001
R=0.43
Taxonomy NMDS2
NMDS1
USA: Baltimore
Hungary: Budapest
Finland: Helsinki
Finland: Lahti
South Africa: Potchefstroom
Reference
Remnant
Turf
Ruderal
C.
P=0.001
R=0.33
KO NMDS2
NMDS1
D.
Relative Abundance
Treatment
Reference
Remnant
Turf
Ruderal
Amino acid metabolism
Biosynthesis of other secondary metabolites
Carbohydrate metabolism
Energy metabolism
Glycan biosynthesis and metabolism
Lipid metabolism
Metabolism of cofactors and vitamins
Metabolism of other amino acids
Metabolism of terpenoids and polyketides
Nucleotide metabolism
Replication and repair
Xenobiotics biodegradation and metabolism
Other


A.
KO distance
B.
Order distance
C.
Bud. r= 0.51, P=0.01
Hel. r= 0.68, P=0.01
Pot. r=0.55 , P=0.01
KO distance
Geographic Location
Baltimore
Budapest
Helsinki
Lahti
Potchefstroom
D.
r= 0.35
P=0.
KO distance
Geographic distance (m)
E.
r= 0.34
P=0.01
Order distance
Geographic distance (m)
F.
r= 0.67
P=0.01
KO distance
Order distance


(A)
Order Shared Amongst Set
0
100
200
257
25
15
12
11
10
10
< 6
Budapest
Potchefstroom
Helsinki
Lahti
Baltimore
400
300
200
100
0
Total Number of Order


(B)
KO Shared Amongst Set
0
1000
2000
3000
4000
3667
1085
973
259
178
< 125
Baltimore
Helsinki
Potchefstroom
Budapest
Lahti
6000
4000
2000
0
Total Number of KO


log10(Norm. Abund.)
6
4
2
Pseudomonadota
Campylobacterota
Aquificota
Chrysiogenota
Deferribacterota
Thermodesulfobacteriota
Nitrospirota
Myxococcota
Bdellovibrionota
Acidobacteriota
Bacteroidota
Rhodothermota
Ignavibacteriota
Caldithrix
Fibrobacterota
Gemmatimonadota
Bacillota
Chloroflexota
Thermomicrobiota
Armatimonadota
Actinomycetota
Verrucomicrobia
Cyanobacteria
Caldisericota
Dictyoglom
Thermotogae
Synergistota
Deinococcota
Fusobacteriales
Lentisphaerota
Chlamydiota
Planctomycetota
Spirochaetota
Reference
Remnant
Turf
Ruderal
Density
−20
−10
0
10
20
Percent of Soils Containing Clade



Reference
Remnant
Turf
Ruderal

Amino acid metabolism
Carbohydrate metabolism
Energy metabolism
Lipid metabolism
Membrane transport
Metabolism: Unclassified
Nucleotide metabolism
Poorly characterized
Replication and repair
Transcription
Translation
Transport and catabolism
Xenobiotics biodegradation and metabolism

A Cell growth and death
B Cellular processes and signaling
C Enzyme families
D Glycan biosynthesis and metabolism
E Metabolism of cofactors and vitamins
F Metabolism of terpenoids and polyketides
G Signal transduction