Open access • Posted Content • DOI:10.1101/2020.06.26.173328

Urbanization promotes specific bacteria in freshwater microbiomes including potential pathogens — Source link

Daniela Numberger, Luca Zoccarato, Jason N. Woodhouse, Lars Ganzert ...+6 more authors

Institutions: Leibniz Association, University of Göttingen, Max Delbrück Center for Molecular Medicine, University of Potsdam ...+1 more institutions

Published on: 26 Jun 2020 - bioRxiv (Cold Spring Harbor Laboratory)

Topics: Freshwater ecosystem

Urbanization promotes specific bacteria in freshwater microbiomes including potential pathogens

Running title: Urban freshwater microbiomes

Daniela Numberger[1,2,#], Luca Zoccarato[2], Jason Woodhouse[2], Lars Ganzert[2,3,4], Sascha Sauer[5], Hans-Peter Grossart[2,6,7]*, Alex Greenwood[1,8]*

1 Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke-Strasse 17, 10315 Berlin, Germany

2 Leibniz Institute of Freshwater Ecology and Inland Fisheries, Alte Fischerhütte 2, 16775 Stechlin, Germany

3 GFZ German Research Centre for Geosciences, Helmholtz Centre Potsdam, Section 3.7 Geomicrobiology, Telegrafenberg C-422, 14473 Potsdam, Germany

4 University of Göttingen, Experimental Phycology and Culture Collection of Algae (EPSAG), Nikolausberger Weg 18, 37073 Göttingen, Germany

5 Max Delbrück Center for Molecular Medicine in the Helmholtz Association, 16775, 13125 Berlin, Germany

6 University of Potsdam, Institute of Biochemistry and Biology, Karl-Liebknecht-Strasse 24-25, 14476 Potsdam, Germany

7 Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), Altensteinstrasse 32, 14195 Berlin, Germany

8 Freie Universität Berlin, Department of Veterinary Medicine, Institute for Virology, Robert von Ostertag-Strasse 7-13, 14163 Berlin, Germany

current address: Institute for Microbiology, University of Veterinary Medicine Hannover, Foundation, Bischofsholer Damm 15, 30173 Hannover, Germany

***Correspondence:**

Professor **Hans-Peter Grossart**, Leibniz Institute of Freshwater Ecology and Inland Fisheries, Alte Fischerhütte 2, 16775 Stechlin, Germany, Phone: +49 33082 69991, Fax: +49 33082 69917, E-mail: hgrossart@igb-berlin.de

Professor **Alex D. Greenwood**, Leibniz Institute for Zoo and Wildlife Research, Alfred Kowalke Strasse 17, 10315 Berlin, Germany, Phone: +49 30 5168 255, Fax: +49 30 5168 104, E-mail: greenwood@izw-berlin.de

ABSTRACT

Freshwater ecosystems are not closed or sterile environments. They support complex and highly dynamic microbiological communities strongly structured by their local environment. Growing city populations and the process of urbanization is predicted to strongly alter freshwater environments. To determine the changes in freshwater microbial communities associated with urbanization, full-length 16S rRNA gene PacBio sequencing was performed on DNA from surface water and sediments from five lakes and a wastewater treatment plant in the Berlin-Brandenburg region of Germany. Water samples exhibited highly environment specific bacterial communities with multiple genera showing clear urban signatures. We identified potential harmful bacterial groups that were strongly associated with environmental parameters specific to urban environments such as *Clostridium*, *Neisseria*, *Streptococcus*, *Yersinia* and the toxic cyanobacterial genus *Microcystis*. We demonstrate that urbanization can alter natural microbial communities in lakes and promote specific bacterial genera which include potential pathogens. Urbanization, creates favourable conditions for pathogens that might be introduced by sporadic events or shift their proportions within the ecosystem. Our findings are of global relevance representing a long-term health risk in urbanized waterbodies at a time of global increase in urbanization.

INTRODUCTION

The process of urbanization leads to changes in land-cover and -use, hydrological systems, local climate and biodiversity [1]. Urbanization is predicted to continue to strongly increase over the coming decades [2, 3]. Expansion rates of urban land area are higher than or equal to population growth rates resulting in more expansive than compact urban growth [3]. While urban land area increased 58,000 km^2 worldwide from 1970 to 2000, an increase of an average of 1,527,000 km^2 urban land cover is predicted by 2030 [3]. Massive concentrations of people challenge freshwater hygiene and as a consequence human health [4–6]. Anthropogenic activities, such as introducing faecal bacteria into water systems, causing eutrophication and introducing other forms of pollution, have the potential to alter the natural microbial community composition of freshwater. This could create new communities that are favourable to proliferation of pathogens that enter water bodies sporadically, whereas natural communities may restrict pathogen growth [7]. For example, the increasing frequency and dominance of toxic cyanobacterial blooms and other pathogens are of particular concern since they directly affect human and animal health and are found to be associated with anthropogenic pollution resulting in eutrophication [5, 8, 9]. Yet, how and to what extent human activities impact the general microbial community structure of freshwater systems remains largely unknown [6, 10–13].

Wastewater treatment plants (WWTPs) serve the principle function of maintaining water hygiene by reducing nutrients and pathogenic microorganisms [14–17]. However, they represent one of the major sources of environmental freshwater pollution including pathogenic microorganisms and antibiotic resistant microbes or pharmaceuticals. Wastewater effluents strongly contribute to the humanization of natural microbial communities, creating water communities that can “resemble” enteric bacterial communities [6, 13, 16]. Urban lakes that otherwise are not affected by treated wastewater, remain susceptible to anthropogenic influence associated with intense recreational activity and urban storm water inflow [18, 19]. Rural lakes, when not influenced by agricultural activities and other land-use, should exhibit natural bacterial communities, where most spatio-temporal variability may be in response to environmental factors such as pH, calcium carbonate and nutrient content, organic matter availability and temperature differences [18–23].

Urbanization can cause multiple simultaneous disturbances of microbial communities, which in some cases, may favour the proliferation of pathogenic microbes. For example, human-introduced microplastics can serve as a preferential habitat for pathogens by enabling biofilm formation in freshwater [24, 25]. In addition, urban areas have higher temperatures than their rural surrounding landscapes, e.g. 4.6 °C difference in the mean air temperature in Beijing [26–28] and increasing water temperatures are known to stimulate growth of some pathogenic species [29–31]. Alterations in microbial communities that favour groups of bacteria containing pathogenic species increases the likelihood that such pathogens may emerge [31–33]. In urban areas, water can be easily contaminated with pathogens by humans and pets during recreational activity [34–36], wildlife [37, 38], storm water/runoffs [39–41], agriculture [5, 42] and wastewater effluent [17, 43, 44]. Although there are hints that lake trophy and anthropogenic activity drive microbial community composition and function [45], it remains unclear which bacterial phyla are indicative for increasing urbanization and hence are indicators for human health risks.

Best practice for identifying pathogenic organisms in aquatic environments remains the utilisation of selective culture media, or molecular detection by qPCR targeting specific markers of pathogenicity [46, 47]. Such approaches are typically laborious, requiring multiple assays targeting distinct pathogens. Furthermore, these techniques presume a specific target and are not convenient, if one wants an overview of what bacteria are present. In contrast, while amplicon sequencing has been proposed as a more cost-effective method for profiling microbial communities for the presence of potentially pathogenic organisms, short-read sequencing often falls short of classification to a family or genus level [48]. Several studies have proposed specific primer pairs, or increasing the number of targeted variable regions [49] for bacterial community structure determination, but these suffer from the same pitfalls. Full-length sequencing of the entire 16S ribosomal RNA gene can provide a comprehensive profile of the entire microbial community with taxonomic resolution to the species level in many cases [48, 50].

To compare urban and rural bacterial communities we investigated five lakes and a wastewater treatment plant at four time points over one year in the Berlin-Brandenburg region. The Berlin-

Brandenburg area serves as a model region with steep gradients of urbanization from a densely populated and rapidly growing city (ca. 3.7 Mio. inhabitants) to a hinterland with one of the lowest population density in Germany (85 people per km^2). Therefore, we expected a clear impact of increasing urbanization on microbial community structure of the studied aquatic systems greatly differing in anthropogenic influence. To improve the phylogenetic resolution and better characterize community composition, we sequenced the full-length 16S rRNA gene by high throughput long read sequencing on the PacBio Sequel I platform [51].

MATERIALS AND METHODS

Sampling

Berlin, the capital of Germany was selected to study urban lakes and a wastewater treatment plant. Berlin is a metropole with an area of 891.1 km^2 and 3.7 Mio inhabitants. In addition, a lake in the smaller city Feldberg in Mecklenburg-Vorpommern with 4,000 inhabitants was selected as another urban lake since it shows a pronounced anthropogenic impact due to previous wastewater input and thus is comparable to lakes in bigger cities. The rural lakes are located in a forested natural reserve area in Northern Brandenburg and have little anthropogenic impact, surrounded by only 1,200 inhabitants in total. All lakes originate from the last ice age, but vary in their present environmental status. Characteristics of all five lakes and the wastewater treatment plant are shown in **Suppl. Table S1**.

Surface water and sediment samples were taken every three months in 2016 from two (small lake 'Weißer See') to three different locations in each lake of in total five lakes in Northeast Germany (**Suppl. Fig. S1**). Water was collected in 2 L bottles and filtered through 0.22 µm Sterivex® filters (EMD Millipore, Darmstadt, Germany) connected to a peristaltic pump (Model XX8200115 6-600 with XX80EL004 head, EMD Millipore, Germany) to collect bacteria. In addition, the first centimetre of sediment was sampled using a plexiglas tube (length 50 cm, Ø 44 mm) and a ruler as a sediment

corer. After slicing the cores, samples were frozen immediately at -20°C until DNA extraction in the lab.

Measurement of nutrients and dissolved organic carbon

For measurement of orthophosphate, nitrate, nitrite, ammonium and dissolved organic carbon (DOC) 200 mL water was filtered through 0.45 µm cellulose acetate filters (Sartorius Stedim Biotech GmbH, Göttingen, Germany) after pre-flushing. The filtrate was frozen at -20°C prior to analyses. Dissolved nutrients were analysed spectrophotometrically using a flow injection analyzer (FOSS, Hilleroed, Denmark), while DOC was analysed with a Shimadzu TOC-5050 total organic carbon analyser (Duisburg, Germany). All analyses were conducted according to Wetzel and Likens [52].

DNA extraction

The QIAamp DNA mini kit (Qiagen, Hilden, Germany) was used for DNA extraction from Sterivex® filters (EMD Millipore, Darmstadt, Germany) following the protocol for tissue with some modifications. Prior to extraction the filters were cut into small pieces and placed into a 2 mL tube. After the addition of 200 µm low binding zirconium glass beads (OPS Diagnostics, NJ, USA) and 360 µL of buffer ATL, the samples were vortexed for 5 min at 3,000 rpm with an Eppendorf MixMate® (Eppendorf, Hamburg, Germany). For lysis, 40 µL of proteinase K was added and incubated at 57°C for 1 h. Then, the samples were centrifuged for 1 min at 11,000 rpm and the supernatant was transferred to a new 2 mL tube. The extraction was then continued following the manufacturer's protocol. DNA from sediment samples was extracted using the NucleoSpin® Soil kit (Macherey Nagel, Düren, Germany), according to the manufacturer's instructions.

Amplification of the full-length 16S rRNA genes

For each sample a unique symmetric set of 16 bp barcodes designed by Pacific Biosciences (CA, USA) was coupled with the primers (27F: 5'-AGRGTTYGATYMTGGCTCAG-3' and 1492R: 5'-RGYTACCTTGTTACGACTT-3'). PCR was performed in a total volume of 25 µL containing 12.5 µL MyFi^{TM} Mix (Bioline, London, UK), 9.3 µL water, 0.7 µL of 20 mg mL^{-1} bovine serum albumin (New England Biolabs, MA, USA), 0.75 µL of each primer (10 µM) and 1 µL of DNA. Denaturation

159 occurred at the following steps: 95°C for 3 min, 25 cycles of 95°C for 30 s, 57°C for 30 s and 72°C for
160 60 s with a final elongation step at 72°C for 3 min. The concentration and quality of 16S rRNA gene
161 amplicons were measured using a TapeStation 4200 system with D5000 tapes and reagents (Agilent
162 Technologies, CA, USA). Equimolar pools of samples were generated before sequencing.

163 **Library building, purification and sequencing**

164 Samples were purified with an Agencourt AMPure XP kit (Beckman Coulter, USA) and sequencing
165 libraries including DNA damage repair, end-repair and ligation of hairpin adapters were built using the
166 SMRTbell Template Prep Kit 1.0-SPv3 following the instructions in the amplicon template protocol
167 (Pacific Biosciences, USA). The Sequel Binding Kit 2.0 (Pacific Biosciences, USA) was used to bind
168 DNA template libraries to the Sequel polymerase 2.0. The data were collected in a single Sequel
169 SMRT Cell 1M v2 with 600 min movie time on the Sequel system I (Pacific Biosciences, USA). The
170 Diffusion Loading mode was used in combination with a 5 pM on-plate loading concentration on the
171 Sequel Sequencing Plate 2.0 (Pacific Biosciences, USA). The SMRT Analysis Software (Pacific
172 Biosciences, USA) generated Circular Consensus Sequences (CCS) for each multiplexed sample that
173 was used for further downstream analyses.

174 **Bioinformatics and statistics**

175 We obtained an average of 7 Gb total output per SMRT cell. The average CCS read length was 17 kb
176 with a mean amplicon lengths of 1,500 bp. Circular consensus sequences (CCS) for each multiplexed
177 sample were generated from the raw reads using the SMRT Analysis Software (Pacific Biosciences,
178 USA) setting subhead length range to 1400-1600 and stringent accuracy to 0.999. Quality scores of the
179 CCS were scaled with the function *reformat.sh* in bbmap (BBMap - Bushnell B. -
180 sourceforge.net/projects/bbmap/). De-replicated and sorted sequences were de-noised by mean of the
181 UNOISE3 algorithms built into USEARCH v11 [53]. A *de-novo* chimera detection step was
182 implemented in the de-noising algorithm and also in the following OTU clustering step at 99%
183 sequence similarity (best threshold that approximates species for full-length sequences; [54]). An OTU
184 table (**Suppl. Table S2**) was generated mapping the CCS to the OTU centroid sequences and the

taxonomic classification was performed with SINA v1.6 against the SILVA reference database (SSU NR 99 v138) [55–57]. Downstream analyses were performed in R [58].

Weighted correlation network analysis (WGCNA package [59]) was carried out to identify modules of bacterial community OTUs associated with the presence of potential pathogenic taxa. Briefly, noisy signal from rare OTUs was removed from the OTU table retaining only OTUs which occurred with 10 or more sequences in at least 3 samples. An adjacency matrix was computed using the function *adjacency* on the centred log-ratio transformed OTU sequence counts (*clr* function, package compositions; [60]) to ensure sub-compositional coherence. The function infers OTUs connectivity by calculating an OTU similarity matrix (based on Pearson correlation) and apply soft thresh-holding to empathize the strongest correlations. The soft threshold value 6 was picked with the function *pickSoftThreshold* as it was the smallest values achieving a $R^2 > 0.9$ for a scale-free topology fit. Topological overlap dissimilarity was calculated with the function *TOMdist* on the adjacency matrix and fed into a hierarchical clustering (*hclust* function, ward.D2 agglomeration method). OTU modules were automatically identified on the clustering by mean of the function *cutreeDynamic* to identify branch boundaries for modules (deepSplit = 4 and minClusterSize = 20). The OTU modules were summarized by their first principal component (function *moduleEigengenes*) which was correlated against vectors of relative abundance of the potential pathogenic groups. The latter were obtained summing the sequence counts of the OTUs classified as belonging to either one of the potential pathogenic taxa across all samples; these relative abundance vectors were then centred log-ratio transformed. Correlations and *p*-values were obtained from a univariate regression model between each module principal component and each vector of potential pathogens and results were visualized as heatmaps using the package ComplexHeatmap [61].

For the ternary plots, the probability of the presence of each OTU in the different habitats was calculated with the function *multipatt* (func = "*IndVal.g*", duleg = F, max.order = 3; package vegan; [62]) and only OTUs with a *p*-value < 0.05 from a permutation test (n=1000) were displayed. Ternary plots were plotted using the function *ggtern* (package ggtern; [63]).

Non-metric multidimensional scaling analyses were performed by using package vegan in R version 3.5 and Bray-Curtis as dissimilarity index. Constrained correspondence analysis (CCA) was also performed in R using the package vegan and the function *cca* followed by an one-way analysis of variance (ANOVA) with the function *anova* and 'n perm=999' [58, 62, 64].

RESULTS

Between and among lake bacterial community heterogeneity

Sediment samples had a significantly higher bacterial diversity than water samples with an average Shannon-Wiener index of 5.08 for water and 7.35 for sediment samples. The five most abundant phyla in the sediment samples were Gammaproteobacteria (34.1 ± 7.1%), Bacteroidota (14.4 ± 3.7%), Cyanobacteria (9.8 ± 6.2%), Alphaproteobacteria (7.6 ± 3.3%) and Verrucomicrobiota (6.8 ± 3.1%). Water samples were dominated by Gammaproteobacteria (29.7 ± 10.8%), Cyanobacteria (18.9 ± 17.1%), Bacteroidota (11.5 ± 5.3%), Actinobacteriota (10.1 ± 7.4%) and Alphaproteobacteria (9.8 ± 4.3%). We defined less, but more abundant OTUs in surface water than in sediment samples. Dominant OTUs (average relative abundance >1.0% in surface water and WWTP, and >0.1% in sediment samples) are listed in **Suppl. Table S3**.

Non-metric multidimensional scaling (NMDS) analyses of water and sediment samples showed two main clusters in each lake: water and sediment. Furthermore, water samples showed a higher variance than the sediment samples, which were more similar to each other. Within the water samples we observed a clustering of samples by season, whereas the sediment samples revealed either random or spatial patterns (**Fig. 1**).

A constrained correspondence analysis of the surface water samples in combination with an analysis of variance (ANOVA) showed that pH, temperature, orthophosphate, nitrate, nitrite, ammonium and dissolved organic carbon (DOC) concentration had a significant (all $p \leq 0.001$) correlation with the composition of the lake bacterial communities (**Fig. 2a**). Temperature ($\chi^2 = 0.4425$) and the concentration of orthophosphate ($\chi^2 = 0.4026$) had the strongest impact. Cyanobacteria were positively

correlated with DOC, Alphaproteobacteria, Bacteroidota and Verrucomicrobiota with temperature, Actinobacteriota, Firmicutes and Gammaproteobacteria with orthophosphate and Acidobacteriota, Chloroflexi and Planctomycetota with nitrogen-based nutrients (**Fig. 2b**). Only temperature ($p=0.04$, $\chi^2=2.61$), orthophosphate ($p=0.03$, $\chi^2=2.90$), nitrate ($p=0.002$, $\chi^2=5.99$) and nitrite ($p=0.04$, $\chi^2=2.58$) were statistically significant in correlation with the bacterial phyla.

Habitat-specific bacterial communities in rural and urban freshwater habitats

Fig. 3 shows the relative abundances of the bacterial phyla contributing more than 1.0% to the bacterial community and the differences between wastewater, urban and rural lakes.

Lakes were characterized by significant higher fractions of Actinobacteriota, Alphaproteobacteria, Planctomycetota and Verrucomicrobiota, while wastewater had significant higher levels of Firmicutes and Gammaproteobacteria (without the order Burkholderiales). Urban lakes differed significantly from rural lakes having higher relative abundance of Actinobacteria, Burholderiales and Firmicutes.

Among all defined OTUs from water, sediment and wastewater samples (total = 112,133 OTUs) only 1.1% were shared between all three environments, i.e. wastewater, urban and rural lakes (**Fig. 4**). 10.2% of OTUs were unique to wastewater, 38.4% were unique to urban lakes and 30.0% were unique to rural lakes. Wastewater shared 0.9% of the OTUs with urban lakes and 0.3% with rural lakes, respectively. Urban and rural lakes shared 19.0% of OTUs. The percentages of OTUs unique to the respective lakes ranged from 8.9-16.2%. Among all lakes 53.9% of OTUs were unique to sediment and 20.3% to surface water.

The ternary plots in **Fig. 5a** show the distribution of all OTUs of a certain bacterial taxon in the three different habitats: rural lake water, urban lake water and wastewater. We excluded the sediment from this analysis as we only had water samples from the wastewater treatment plant. Each dot indicates an OTU and the position in the ternary plot reflects its percentage presence in each of the three habitats. The percentages of the number of OTUs for each and shared habitats are shown in **Fig. 5b**.

Wastewater and rural lakes shared only very few OTUs. Most OTUs were shared between rural and urban lakes except for the phylum Firmicutes that showed the highest prevalence of OTUs unique to

wastewater followed by OTUs shared by wastewater and urban lakes such as OTUs belonging to the genera *Acinetobacter*, *Bacteroides*, *Bifidobacterium* and *Enterococcus*. The Actinobacteriota and Gammaproteobacteria showed high OTU numbers in urban waters and shared between urban and rural lake water. All bacterial phyla showed higher numbers of OTUs in urban lake water than in rural lake water alone. An indicator species analysis (ISA) identified in total ~2,600 OTUs as significant indicators for urban waters (urban lakes and wastewater) including *Acinetobacter* (Gammaproteobacteria), *Aeromonas* (Gammaproteobacteria), *Bacteroides* (Bacteroidota), *Bifidobacterium* (Actinobacteriota), *Blautia* (Firmicutes), *Clostridium* sensu-stricto (Firmicutes), *Comamonas* (Burkholderiales), *Enterococcus* (Firmicutes), *Lachnospira* (Firmicutes), *Paracoccus* (Alphaproteobacteria) and *Uruburuella* (Burkholderiales).

Bacterial genera including (known) potential pathogenic species

The prevalence of the most relevant genera, which include species that are known human pathogens are shown in **Fig. 6a**. A weighted correlation network analyses (WGCNA) identified 13 bacterial sub-communities that were consistent throughout the sampled environments and significantly correlated with bacterial genera containing potential pathogenic species (**Fig. 6b**). Some of those genera were correlated significantly and positively with sub-communities that were only defined for urban waters such as *Aeromonas, Alistipes, Clostridium* (sensu-stricto), *Enterococcus*, *Escherichia/Shigella* *Staphylococcus*, *Streptococcus* and *Yersinia*. We could not find any significant correlation between potential pathogenic groups and sub-communities that were only present in rural waters.

A CCA analysis (**Fig. 7**) shows those potential pathogenic genera and how they were correlated with the measured environmental parameters. ANOVA revealed that all parameters, except nitrite concentration, were significant. While *Enterococcus* was not correlated with any of the measured environmental factors, other groups showed clear correlations such as *Microcystis* with orthophosphate, *Legionella* with DOC and ammonium, *Rickettsia* with nitrite, *Neisseria* with nitrate and *Peptoclostridium* with temperature.

DISCUSSION

Urbanisation represents a multifaceted stressor that impacts the quality of freshwater systems, promoting eutrophication [8, 65] and contributing to the accumulation of emerging pollutants [9, 66]. Eutrophication has long been recognised as a major driver of microbial community composition with high loads of organic matter leading to increased bacterial activity and creating opportunities for the proliferation of copiotrophs, including many pathogens [67, 68]. **Eutrophication** however, is not strictly an urban problem. Rural freshwater, particularly in close proximity to agricultural lands, can also be affected. However, our results revealed significant differences in the microbial community composition of sediments and water from rural and urban lakes, and wastewater. Sampled urban lakes, though not directly connected to wastewater effluents, showed a higher similarity to wastewater samples than rural lakes. Sewage generally reflects the human faecal microbiome [13], suggesting urbanization might have led to a humanization of freshwater bacterial communities [6]. In addition, it is known that bathers release bacteria from the skin during recreational water activity [34, 35] and animal or human urine could also be a source of bacterial contamination [69, 70].

The presence of **habitat specific bacterial communities** was supported for wastewater, urban and rural lake water. The differences between rural and urban lake communities appear to be mainly driven by the prevalence of specific dissolved nutrients (**Fig. 2**). Increased availability of orthophosphate and ammonium coincided with an increase in the relative abundance (**Fig. 3**) and diversity (**Fig. 5a**) of most bacterial phyla in urban lakes, particularly the Actinobacteriota, Alphaproteobacteria, Bacteroidota, Firmicutes and Gammaproteobacteria.

Actinobacteriota and **Alphaproteobacteria** are typically oligotrophic members of freshwater systems, notably represented by the genera *Planktophila* (acI) and *Fonsibacter* (LD12), respectively. These two groups alone can account for up to 50% of the bacterial community composition in lakes outside of periods of high phytoplankton biomass [71, 72]. In addition, there is a high capacity for organic matter utilisation within Actinobacteriota and Alphaproteobacteria, in particular by the genera *Planktoluna* (acIV) and *Sphingomonas,* respectively. A greater abundance of these latter phyla in urban landscapes reflects the enrichment of these copiotrophic taxa at the expense of other

316 oligotrophic taxa. A clear indicator of anthropogenic impacts in the phylum Actinobacteriota is the
317 genus *Bifidobacterium* [73] that was only present in urban water.

318 **Bacteroidota** are well established components of freshwater systems [18]. They perform important
319 roles in the degradation of organic matter, in particular complex biopolymers [18, 74]. Typically, in
320 freshwater systems Bacteroidota dominance and diversity are driven by increasing concentrations of
321 either autochthonous, in the form of algal or zooplankton biomass, or allochthonous, in the form of
322 terrestrial detritus, particulate organic matter. A recent study demonstrated that Bacteroidota strains
323 are highly specific to individual polymeric substrates [75], suggesting that diversity of Bacteroidota
324 scales with diversity of the organic matter pool. A higher diversity of Bacteroidota in urban lakes
325 would be supportive of this. The high terrestrial-aquatic coupling and dynamic nature of urban
326 landscapes would imply a greater diversity of organic matter including faecal contamination [48, 76,
327 77] than in the rural lakes where cyanobacterial derived autochthonous organic matter is dominant.
328 The genus *Bacteroides*, a known faecal contamination indicator [78, 79] showing a clear urban
329 signature in our study, but also Prevotellacae, Rikenellaceae, Tannerellaceae and Weeksellaceae were
330 significantly enriched in urban waters. In rural lakes the families Chitinophagaceae, Flavobacteriaceae,
331 Saprospiraceae and Spirosomaceae, well-known freshwater taxa and decomposers of complex carbon
332 sources such as from phytoplankton [18, 80, 81], were enriched.

333 The **Firmicutes**, usually not abundant in lake water [18], but in faeces and wastewater [13, 17, 48, 82]
334 were highly abundant in the wastewater samples, particularly the inflow samples and showed an
335 enrichment in urban lakes, but not in rural lakes. The enrichment in urban lakes is explained by an
336 increase of typical human derived groups such as Enterococcaceae, Eubacteriaceae,
337 Peptostreptococcaceae, Ruminococcaceae, Streptococcaceae and Veillonellaceae. This “human
338 footprint” also includes potential pathogens such as from the genera *Bacillus*, *Clostridium*,
339 *Peptoclostridium* and *Staphylococcus*, *Streptococcus*. Furthermore, toxigenic *C. difficile*, a well-
340 known human pathogen, was isolated previously from one sample obtained in summer from the urban
341 lake “Weißer See” [46]. This supports the hypothesis that urbanization creates favourable bacterial

342 communities for the growth of potential pathogens and thus, constitutes a higher risk for waterborne or
343 –transmitted infections.

344 The **Gammaproteobacteria**, occur at low abundance in natural freshwater lakes [18, 19, 83]. The
345 increased relative abundance of Gammproteobacteria in urban and rural lakes was due to the
346 abundance of members of Burkholderiaceae, Comamonadaceae and Methylophilaceae, all belonging
347 to the order Burkholderiales. Although the relative abundance of Gammaproteobacteria as a whole did
348 not increase in urban lakes, clear urban lake signatures were observed, represented by
349 Aeromonadaceae, Enterobacteriaceae, Moraxellaceae, Pseudomonadaceae, Succinivibrionaceae,
350 Xanthomonadaceae and Yersiniaceae that were enriched in urban water. These bacterial families
351 include potential human pathogens such as *Aeromonas hydrophila*, *Pseudomonas aeruginosa* and
352 *Yersinia enterocolitica* and thus, their enrichment in our urban lakes constitutes a potential health risk.
353 A positive correlation of Gammaproteobacteria with lake eutrophication was recently observed [45]
354 and it has been shown that Gammaproteobacteria grow faster than the average lake bacterioplankton,
355 particularly when nitrogen and phosphorus levels are high [84, 85]. The CCA (**Fig. 2b**) showed a clear
356 positive correlation of Gammaproteobacteria with orthophosphate and to a lesser degree, nitrogen-
357 based nutrients. Even if we cannot distinguish between urbanization or eutrophication as driver for this
358 enrichment, urbanization also leads to an unavoidable eutrophication of environmental water [86, 87]
359 and hence, could lead indirectly to favourable conditions for those potential pathogens. In addition,
360 Aeromonadaceae and Pseudomonadaceae, in particular, have been identified as the most likely
361 reservoirs for antibiotics resistance genes in aquatic environments and hence, constitute a further
362 potential threat to human health by the ability to spread these genes to harmful microorganisms [88–
363 90].

364 Urban lakes contained a higher proportion of taxa, which include **potentially pathogenic organisms**.
365 Urbanization may favour taxa that include potential pathogens indicating that if pathogenic bacteria
366 contaminate urban waters, they will find a favourable environment in which to proliferate [5, 6, 34, 46,
367 91]. We found statistically significant correlations between the occurrence of some potential
368 pathogenic groups and sets of bacterial sub-communities of which four were only present in urban

water. These sub-communities were strongly correlated with the presence of potential pathogenic genera such as *Acinetobacter*, *Aeromonas*, *Alistipes*, *Clostridium* (sensu-stricto), *Klebsiella*, *Neisseria*, *Staphylococcus*, *Streptococcus* and *Yersinia* suggesting that urbanization favours the presence of these potential pathogenic groups. Nevertheless, while the occurrence of pathogenic species was rare in this study the enrichment of the taxonomic groups to which they belong was constant among all urban samples. This could favour stochastic and sudden outbreaks of pathogenic bacteria in urban settings that may be less likely to occur in rural settings, where environmental conditions are less favourable for such copiotrophic, pathogenic bacteria. The potential health risk of urban water bacterial communities may need to be accounted for in future urban lake management [92, 93]. Furthermore, potential pathogenic groups are also present in coastal marine waters in the proximity of wastewater output confirming urbanization as health risk for waterborne or –transmitted diseases [48].

Within lakes, bacterial communities were more stable over time in the sediment than in surface water. Sediment samples showed a higher bacterial diversity than in the water column. Sediment seems to be more stable in environmental variables and might have a protective effect on microbes against environmental changes, UV radiation, drifting and grazing. Furthermore, sediment grains can be used as a substrate for microbial biofilms, which may enhance microbial stability and persistence in the system [39, 94, 95]. Some bacterial groups that include potential pathogens were also present in sediment samples such as *Acinetobacter*, *Aeromonas*, *Legionella*, *Leptospira*, *Streptococcus* and *Treponema*. Toxigenic *C. difficile* was isolated from the sediment of urban lake 'Weißer See' [46] and other studies demonstrated an extended persistence of faecal indicator bacteria such as *Enterococcus* associated with sediment representing a reservoir function [94, 95].

In conclusion, increased urbanization will accelerate the humanization of aquatic bacterial communities. A better understanding of the ecological and functional consequences of urbanization and the roles of habitat specific bacterial groups is needed to mitigate potential health impacts of urban bacterial communities. We identified specific taxa that can exploit niches in urban water (i.e. human-derived bacterial groups such as *Alistipes*, *Bifidobacterium, Bacteroides, Enterococcus, Streptococcus and Yersinia*), and demonstrated that specific environmental conditions and the presence of specific

396 sub-communities of bacteria represent risk factors for the emergence and spread of pathogenic taxa.
397 Urbanization may create aquatic microbiomes that favour the growth of pathogens and antibiotic-
398 resistant bacteria that sporadically enter urban water systems that would otherwise face barriers to
399 grow in rural water bodies and may represent an underestimated risk of urban associated pathogen and
400 antibiotic resistance propagation and transmission. Beyond the increased proliferation of pathogenic
401 and antibiotic-resistant microorganisms in urban waters, urbanization is likely to have additional
402 impacts on aquatic biodiversity and biogeochemical cycling. Additional research is required to fully
403 explore the impacts of urbanization, and action will need to be taken to reduce the impact of
404 urbanization on aquatic ecosystems and offset harmful effects for both humans and the environment.

405

406 **ACKNOWLEDGEMENTS**

407 We sincerely thank Claudia Quedenau, Michael Sachtleben, Uta Mallock and Elke Mach for technical
408 and Dr. Daniele Y. Sunaga-Franze for computational assistance. This work was supported by a project
409 grant (IRG 3 - Water) from the Leibniz Research Alliance "INFECTIONS'21 - Transmission Control
410 of Infections in the 21st Century" as funded by the Leibniz Association, Germany (SAS-2015-FZB-
411 LFV) and the Leibniz SAW project "MycoLink" (SAW-2014-IGB). The German Federal Ministry of
412 Education and Research (BMBF) supported this project within the Collaborative Project "Bridging in
413 Biodiversity Science - BIBS" (funding number 01LC1501G).

414

415 **COMPETING INTERESTS**

416 The authors confirm that they have no conflicts of interest related to the content of this article.

417

418

419

420

REFERENCES

1. Grimm NB, Faeth SH, Golubiewski NE, Redman CL, Wu J, Bai X, et al. Global change and the ecology of cities. *Science* 2008; **319**: 756–760.
2. Gardner GT, Prugh T, Renner M, Worldwatch Institute (eds). Can a city be sustainable? 2016. Island Press, Washington.
3. Seto KC, Fragkias M, Güneralp B, Reilly MK. A meta-analysis of global urban land expansion. *PLoS ONE* 2011; **6**.
4. Vörösmarty CJ, Green P, Salisbury J, Lammers RB. Global water resources: Vulnerability from climate change and population growth. *Science* 2000; **289**: 284–288.
5. Walters SP, Thebo AL, Boehm AB. Impact of urbanization and agriculture on the occurrence of bacterial pathogens and *stx* genes in coastal waterbodies of central California. *Water Res* 2011; **45**: 1752–1762.
6. McLellan SL, Fisher JC, Newton RJ. The microbiome of urban waters. *Int Microbiol Off J Span Soc Microbiol* 2015; **18**: 141–149.
7. Wang H, Edwards MA, Falkinham JO, Pruden A. Probiotic approach to pathogen control in premise plumbing systems? A Review. *Environ Sci Technol* 2013; **47**: 10117–10128.
8. Lapointe BE, Herren LW, Debortoli DD, Vogel MA. Evidence of sewage-driven eutrophication and harmful algal blooms in Florida's Indian River Lagoon. *Harmful Algae* 2015; **43**: 82–102.
9. Zinia NJ, Kroeze C. Future trends in urbanization and coastal water pollution in the Bay of Bengal: The lived experience. *Environ Dev Sustain* 2015; **17**: 531–546.
10. Ibekwe AM, Ma J, Murinda SE. Bacterial community composition and structure in an urban river impacted by different pollutant sources. *Sci Total Environ* 2016; **566–567**: 1176–1185.
11. Hall RI, Leavitt PR, Quinlan R, Dixit AS, Smol JP. Effects of agriculture, urbanization, and climate on water quality in the northern Great Plains. *Limnol Oceanogr* 1999; **44**: 739–756.
12. Rizzo L, Manaia C, Merlin C, Schwartz T, Dagot C, Ploy MC, et al. Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: A review. *Sci Total Environ* 2013; **447**: 345–360.

448 13. Newton RJ, McLellan SL, Dila DK, Vineis JH, Morrison HG, Eren AM, et al. Sewage reflects the
449 microbiomes of human populations. *mBio* 2015; **6**: e02574-14.
450 14. Asano T, Levine AD. Wastewater reclamation, recycling and reuse: past, present, and future.
451 *Water Sci Technol* 1996; **33**: 1–14.
452 15. Al-Jassim N, Ansari MI, Harb M, Hong P-Y. Removal of bacterial contaminants and antibiotic
453 resistance genes by conventional wastewater treatment processes in Saudi Arabia: Is the treated
454 wastewater safe to reuse for agricultural irrigation? *Water Res* 2015; **73**: 277–290.
455 16. Wakelin SA, Colloff MJ, Kookana RS. Effect of wastewater treatment plant effluent on microbial
456 function and community structure in the sediment of a freshwater stream with variable seasonal
457 flow. *Appl Environ Microbiol* 2008; **74**: 2659–2668.
458 17. Numberger D, Ganzert L, Zoccarato L, Mühldorfer K, Sauer S, Grossart H-P, et al.
459 Characterization of bacterial communities in wastewater with enhanced taxonomic resolution by
460 full-length 16S rRNA sequencing. *Sci Rep* 2019; **9**: 9673.
461 18. Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S. A guide to the natural history of
462 freshwater lake bacteria. *Microbiol Mol Biol Rev* 2011; **75**: 14–49.
463 19. Zwart G, Crump BC, Agterveld MPK, Hagen F, Han S-K. Typical freshwater bacteria: An
464 analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat Microb*
465 *Ecol* 2002; **28**: 141–155.
466 20. Güde H. Participation of bacterioplankton in epilimnetic phosphorus cycles of Lake Constance.
467 *SIL Proc 1922-2010* 1991; **24**: 816–820.
468 21. Bloem J, Albert C, Bär-Gillissen M-JB, Berman T, Cappenberg TE. Nutrient cycling through
469 phytoplankton, bacteria and protozoa, in selectively filtered Lake Vechten water. *J Plankton Res*
470 1989; **11**: 119–131.
471 22. Allgaier M, Grossart H-P. Diversity and seasonal dynamics of actinobacteria populations in four
472 lakes in Northeastern Germany. *Appl Environ Microbiol* 2006; **72**: 3489–3497.
473 23. Allgaier M, Grossart H-P. Seasonal dynamics and phylogenetic diversity of free-living and
474 particle-associated bacterial communities in four lakes in Northeastern Germany. *Aquat Microb*
475 *Ecol* 2006; **45**: 115–128.

24. Viršek MK, Lovšin MN, Koren Š, Kržan A, Peterlin M. Microplastics as a vector for the transport of the bacterial fish pathogen species *Aeromonas salmonicida*. *Mar Pollut Bull* 2017; **125**: 301–309.
25. Kirstein IV, Kirmizi S, Wichels A, Garin-Fernandez A, Erler R, Löder M, et al. Dangerous hitchhikers? Evidence for potentially pathogenic *Vibrio* spp. on microplastic particles. *Mar Environ Res* 2016; **120**: 1–8.
26. Armson D, Stringer P, Ennos AR. The effect of tree shade and grass on surface and globe temperatures in an urban area. *Urban For Urban Green* 2012; **11**: 245–255.
27. Sundborg Å. Local climatological studies of the temperature conditions in an urban area. *Tellus* 1950; **2**: 222–232.
28. Tan J, Zheng Y, Tang X, Guo C, Li L, Song G, et al. The urban heat island and its impact on heat waves and human health in Shanghai. *Int J Biometeorol* 2010; **54**: 75–84.
29. Baker-Austin C, Trinanes JA, Taylor NGH, Hartnell R, Siitonen A, Martinez-Urtaza J. Emerging *Vibrio* risk at high latitudes in response to ocean warming. *Nat Clim Change* 2013; **3**: 73–77.
30. Charron DF, Thomas MK, Waltner-Toews D, Aramini JJ, Edge T, Kent RA, et al. Vulnerability of waterborne diseases to climate change in Canada: A review. *J Toxicol Environ Health A* 2004; **67**: 1667–1677.
31. Hunter PR. Climate change and waterborne and vector-borne disease. *J Appl Microbiol* 2003; **94**: 37–46.
32. Andersson Y, De Jong B, Studahl A. Waterborne *Campylobacter* in Sweden: The cost of an outbreak. *Water Sci Technol* 1997; **35**: 11.
33. Clark RM, Geldreich EE, Fox KR, Rice EW, Johnson CH, Goodrich JA, et al. A waterborne *Salmonella typhimurium* outbreak in Gideon, Missouri: Results from a field investigation. *Int J Environ Health Res* 1996; **6**: 187–193.
34. Plano LR, Garza AC, Shibata T, Elmir SM, Kish J, Sinigalliano CD, et al. Shedding of *Staphylococcus aureus* and methicillin-resistant *Staphylococcus aureus* from adult and pediatric bathers in marine waters. *BMC Microbiol* 2011; **11**: 5.

35. Elmir SM, Wright ME, Abdelzaher A, Solo-Gabriele HM, Fleming LE, Miller G, et al. Quantitative evaluation of bacteria released by bathers in a marine water. *Water Res* 2007; **41**: 3–10.
36. Gerba CP. Assessment of enteric pathogen shedding by bathers during recreational activity and its impact on water quality. *Quant Microbiol* 2000; **2**: 55–68.
37. Markwell DD, Shortridge KF. Possible waterborne transmission and maintenance of influenza viruses in domestic ducks. *Appl Environ Microbiol* 1982; **43**: 110–115.
38. Babudieri B. Animal reservoirs of leptospires. *Ann N Y Acad Sci* 1958; **70**: 393–413.
39. Schillinger JE, Gannon JJ. Bacterial adsorption and suspended particles in urban stormwater. *J Water Pollut Control Fed* 1985; **57**: 384–389.
40. Ward MP. Seasonality of canine leptospirosis in the United States and Canada and its association with rainfall. *Prev Vet Med* 2002; **56**: 203–213.
41. Kupek E, De MSSF, De JSP. The relationship between rainfall and human leptospirosis in Florianópolis, Brazil, 1991-1996. *Braz J Infect Dis Off Publ Braz Soc Infect Dis* 2000; **4**: 131–134.
42. Givens CE, Kolpin DW, Borchardt MA, Duris JW, Moorman TB, Spencer SK. Detection of hepatitis E virus and other livestock-related pathogens in Iowa streams. *Sci Total Environ* 2016; **566–567**: 1042–1051.
43. Cai L, Zhang T. Detecting human bacterial pathogens in wastewater treatment plants by a high-throughput shotgun sequencing technique. *Environ Sci Technol* 2013; **47**: 5433–5441.
44. Steyer A, Gutiérrez-Aguirre I, Rački N, Beigot Glaser S, Brajer Humar B, Stražar M, et al. The detection rate of enteric viruses and *Clostridium difficile* in a waste water treatment plant effluent. *Food Environ Virol* 2015; **7**: 164–172.
45. Kiersztyn B, Chróst R, Kaliński T, Siuda W, Bukowska A, Kowalczyk G, et al. Structural and functional microbial diversity along a eutrophication gradient of interconnected lakes undergoing anthropopressure. *Sci Rep* 2019; **9**: 1–14.
46. Numberger D, Riedel T, McEwen G, Nübel U, Frentrup M, Schober I, et al. Genomic analysis of three *Clostridioides difficile* isolates from urban water sources. *Anaerobe* 2019.

47. Aw TG, Rose JB. Detection of pathogens in water: From phylochips to qPCR to pyrosequencing. *Curr Opin Biotechnol* 2012; **23**: 422–430.
48. Buccheri MA, Salvo E, Coci M, Quero GM, Zoccarato L, Privitera V, et al. Investigating microbial indicators of anthropogenic marine pollution by 16S and 18S High-Throughput Sequencing (HTS) library analysis. *FEMS Microbiol Lett* 2019; **366**.
49. Wang J, Jia H. Metagenome-wide association studies: fine-mining the microbiome. *Nat Rev Microbiol* 2016; **14**: 508–522.
50. Conlan S, Kong HH, Segre JA. Species-level analysis of DNA sequence data from the NIH Human Microbiome Project. *PloS One* 2012; **7**: e47075.
51. Mosher JJ, Bowman B, Bernberg EL, Shevchenko O, Kan J, Korlach J, et al. Improved performance of the PacBio SMRT technology for 16S rDNA sequencing. *J Microbiol Methods* 2014; **104**: 59–60.
52. Wetzel RG, Likens GE. Inorganic nutrients: Nitrogen, phosphorus, and other nutrients. *Limnological analyses*. 1991. Springer, pp 81–105.
53. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinforma Oxf Engl* 2010; **26**: 2460–2461.
54. Edgar RC. Updating the 97% identity threshold for 16S ribosomal RNA OTUs. *Bioinformatics* 2018; **34**: 2371–2375.
55. Pruesse E, Peplies J, Glöckner FO. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 2012; **28**: 1823–1829.
56. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013; **41**: D590–D596.
57. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer K-H, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 2014; **12**: 635–645.
58. R Core Team. R: A language and environment for statistical computing. *R Found Stat Comput Vienna Austria* .

559 59. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC*
560 *Bioinformatics* 2008; **9**: 559.
561 60. Parent S-É, Parent LE. Biochemical fractionation of soil organic matter after incorporation of
562 organic residues. *Open J Soil Sci* 2015; **5**: 135–143.
563 61. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in
564 multidimensional genomic data. *Bioinforma Oxf Engl* 2016; **32**: 2847–2849.
565 62. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara RB, et al. Package 'vegan'.
566 *Community Ecol Package R Package Version* 2014; **2**.
567 63. Hamilton NE, Ferry M. ggtern: Ternary diagrams using ggplot2. *J Stat Softw* 2018; **87**: 1–17.
568 64. Dixon P. VEGAN, a package of R functions for community ecology. *J Veg Sci* 2003; **14**: 927–
569 930.
570 65. Taylor SL, Roberts SC, Walsh CJ, Hatt BE. Catchment urbanisation and increased benthic algal
571 biomass in streams: linking mechanisms to management. *Freshw Biol* 2004; **49**: 835–851.
572 66. Pal A, Gin KY-H, Lin AY-C, Reinhard M. Impacts of emerging organic contaminants on
573 freshwater resources: Review of recent occurrences, sources, fate and effects. *Sci Total Environ*
574 2010; **408**: 6062–6069.
575 67. Wu RSS. Eutrophication, water borne pathogens and xenobiotic compounds: Environmental risks
576 and challenges. *Mar Pollut Bull* 1999; **39**: 11–22.
577 68. Smith VH, Schindler DW. Eutrophication science: where do we go from here? *Trends Ecol Evol*
578 2009; **24**: 201–207.
579 69. Lewis DA, Brown R, Williams J, White P, Jacobson SK, Marchesi J, et al. The human urinary
580 microbiome: Bacterial DNA in voided urine of asymptomatic adults. *Front Cell Infect Microbiol*
581 2013; **3**.
582 70. Rojas P, Monahan AM, Schuller S, Miller IS, Markey BK, Nally JE. Detection and quantification
583 of leptospires in urine of dogs: A maintenance host for the zoonotic disease leptospirosis. *Eur J*
584 *Clin Microbiol Infect Dis* 2010; **29**: 1305–1309.

71. Salcher MM, Pernthaler J, Posch T. Seasonal bloom dynamics and ecophysiology of the freshwater sister clade of SAR11 bacteria 'that rule the waves' (LD12). *ISME J* 2011; **5**: 1242–1252.

72. Woodhouse JN, Kinsela AS, Collins RN, Bowling LC, Honeyman GL, Holliday JK, et al. Microbial communities reflect temporal changes in cyanobacterial composition in a shallow ephemeral freshwater lake. *ISME J* 2016; **10**: 1337–1351.

73. Bonjoch X, Ballesté E, Blanch AR. Multiplex PCR with 16S rRNA gene-targeted primers of *Bifidobacterium* spp. to identify sources of fecal pollution. *Appl Environ Microbiol* 2004; **70**: 3171–3175.

74. Kirchman DL. The ecology of Cytophaga–Flavobacteria in aquatic environments. *FEMS Microbiol Ecol* 2002; **39**: 91–100.

75. Krüger K, Chafee M, Ben Francis T, Glavina del Rio T, Becher D, Schweder T, et al. In marine Bacteroidetes the bulk of glycan degradation during algae blooms is mediated by few clades using a restricted set of genes. *ISME J* 2019; **13**: 2800–2816.

76. Krentz CA, Prystajecky N, Isaac-Renton J. Identification of fecal contamination sources in water using host-associated markers. *Can J Microbiol* 2013; **59**: 210–220.

77. Dick LK, Bernhard AE, Brodeur TJ, Domingo JWS, Simpson JM, Walters SP, et al. Host distributions of uncultivated fecal Bacteroidales bacteria reveal genetic markers for fecal source identification. *Appl Env Microbiol* 2005; **71**: 3184–3191.

78. Kabiri L, Alum A, Rock C, McLain JE, Abbaszadegan M. Isolation of *Bacteroides* from fish and human fecal samples for identification of unique molecular markers. *Can J Microbiol* 2013; **59**: 771–777.

79. Hong P-Y, Wu J-H, Liu W-T. Relative abundance of *Bacteroides* spp. in stools and wastewaters as determined by hierarchical oligonucleotide primer extension. *Appl Environ Microbiol* 2008; **74**: 2882–2893.

80. McIlroy SJ. The family Saprospiraceae. *Prokaryotes* 2014; 863–889.

81. Raj HD, Maloy SR. Family Spirosomaceae: Gram-negative ring-forming aerobic bacteria. *Crit Rev Microbiol* 1990; **17**: 329–364.

613 82. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The Human
614 Microbiome Project. *Nature* 2007; **449**: 804–810.
615 83. Lindström ES, Leskinen E. Do neighboring lakes share common taxa of bacterioplankton?
616 Comparison of 16S rDNA fingerprints and sequences from three geographic regions. *Microb Ecol*
617 2002; **44**: 1–9.
618 84. Šimek K, Horňák K, Jezbera J, Nedoma J, Vrba J, Straškrábová V, et al. Maximum growth rates
619 and possible life strategies of different bacterioplankton groups in relation to phosphorus
620 availability in a freshwater reservoir. *Environ Microbiol* 2006; **8**: 1613–1624.
621 85. Gasol JM, Comerma M, García JC, Armengol J, Casamayor EO, Kojecká P, et al. A transplant
622 experiment to identify the factors controlling bacterial abundance, activity, production, and
623 community composition in a eutrophic canyon-shaped reservoir. *Limnol Oceanogr* 2002; **47**: 62–
624 77.
625 86. Bowen JL, Valiela I. The ecological effects of urbanization of coastal watersheds: Historical
626 increases in nitrogen loads and eutrophication of Waquoit Bay estuaries. *Can J Fish Aquat Sci*
627 2001; **58**: 1489–1500.
628 87. Yu S, Yu GB, Liu Y, Li GL, Feng S, Wu SC, et al. Urbanization impairs surface water quality:
629 Eutrophication and metal stress in the Grand Canal of China. *River Res Appl* 2012; **28**: 1135–
630 1148.
631 88. Figueira V, Vaz-Moreira I, Silva M, Manaia CM. Diversity and antibiotic resistance of
632 *Aeromonas* spp. in drinking and waste water treatment plants. *Water Res* 2011; **45**: 5599–5611.
633 89. Stalder T, Press MO, Sullivan S, Liachko I, Top EM. Linking the resistome and plasmidome to
634 the microbiome. *ISME J* 2019; **13**: 2437–2446.
635 90. Bert F, Maubec E, Bruneau B, Berry P, Lambert-Zechovsky N. Multi-resistant *Pseudomonas*
636 *aeruginosa* outbreak associated with contaminated tap water in a neurosurgery intensive care unit.
637 *J Hosp Infect* 1998; **39**: 53–62.
638 91. Wiedenmann A, Krüger P, Dietz K, López-Pila JM, Szewzyk R, Botzenhart K. A randomized
639 controlled trial assessing infectious disease risks from bathing in fresh recreational waters in

relation to the concentration of *Escherichia coli*, intestinal enterococci, *Clostridium perfringens*, and somatic coliphages. *Environ Health Perspect* 2006; **114**: 228–236.

92. Naselli-Flores L. Urban lakes: Ecosystems at risk, worthy of the best care. *Proc. Taal2007 12th World Lake Conf.* 2008. p 1337.
93. Hipsey MR, Brookes JD. Pathogen management in surface waters: Practical considerations for reducing public health risk. *Curr Top Public Health* 2013.
94. Walters E, Graml M, Behle C, Müller E, Horn H. Influence of particle association and suspended solids on uv inactivation of fecal indicator bacteria in an urban river. *Water Air Soil Pollut* 2013; **225**: 1822.
95. Haller L, Amedegnato E, Poté J, Wildi W. Influence of freshwater sediment characteristics on persistence of fecal indicator bacteria. *Water Air Soil Pollut* 2009; **203**: 217–227.

664 **FIGURE LEGENDS**

665 **Figure 1: Similarity studies of lake samples.** Non-metric multidimensional scaling (NMDS)
666 analyses based on Bray-Curtis dissimilarity index of water (empty symbols) and sediment samples
667 (filled symbols) for each lake. Colours indicate the season and the different symbols represent the
668 sampling site of each lake.

669 **Figure 2: Influence of environmental parameters on the bacterial communities in surface water.**
670 **[a]** A constrained correspondence analysis (CCA) of all water samples and their corresponding
671 environmental measurements: ammonium, dissolved organic carbon (DOC), nitrite, nitrate,
672 orthophosphate (OP), pH, and temperature. Colours indicate the season (blue: winter, green: spring,
673 pink: summer and brown: autumn) and the symbols show the different lakes (square: Stechlinsee,
674 diamond: Dagowsee, circle: Feldberger Haussee, triangle: Müggelsee and inverse triangle: Weißer
675 See). **[b]** A constrained correspondence analysis (CCA) showing the most abundant bacterial phyla
676 Acidobacteriota, Actinobacteriota (Actino), Alphaproteobacteria (α), Bacteroidota (Bacter),
677 Chloroflexi (Chloro), Cyanobacteria (Cyano), Firmicutes, Gammaproteobacteria (γ), Planctomycetota
678 (Plancto) and Verrucomicrobiota (Verruco) and their correlations with the environmental
679 measurements in water samples.

680 **Figure 3: Differences in relative abundance of dominant bacterial phyla between wastewater,**
681 **urban and rural lakes** (summing all seasons and sites)**.** Boxplots showing the relative abundance of
682 the most abundant bacterial phyla for wastewater inflow (IN), wastewater outflow (OUT), urban lakes
683 (U: Weisser See, Müggelsee, Feldberger Haussee), and rural lakes (R: Dagowsee, Stechlinsee).
684 Significant differences are indicated by brackets based on pairwise Mann-Whitney *U* test.
685 Gammaproteobacteria do not include any members of the order Burkholderiales, which have been
686 analysed separately. Rel. abundance – relative abundance.

687 **Figure 4: Core, shared and unique OTUs of lake water and sediments. [a]** Venn diagram showing
688 core, shared and unique OTUs of the wastewater treatment plant (WWTP), urban (Weißer See - WS,
689 Müggelsee - MS, Feldberger Haussee - FHS) and rural lakes (Dagowsee - DS, Stechlinsee - SS). **[b]**

690 Bars showing OTUs that were unique either for each lake, WWTP plus all lakes together, or lake
691 water and lake sediment.

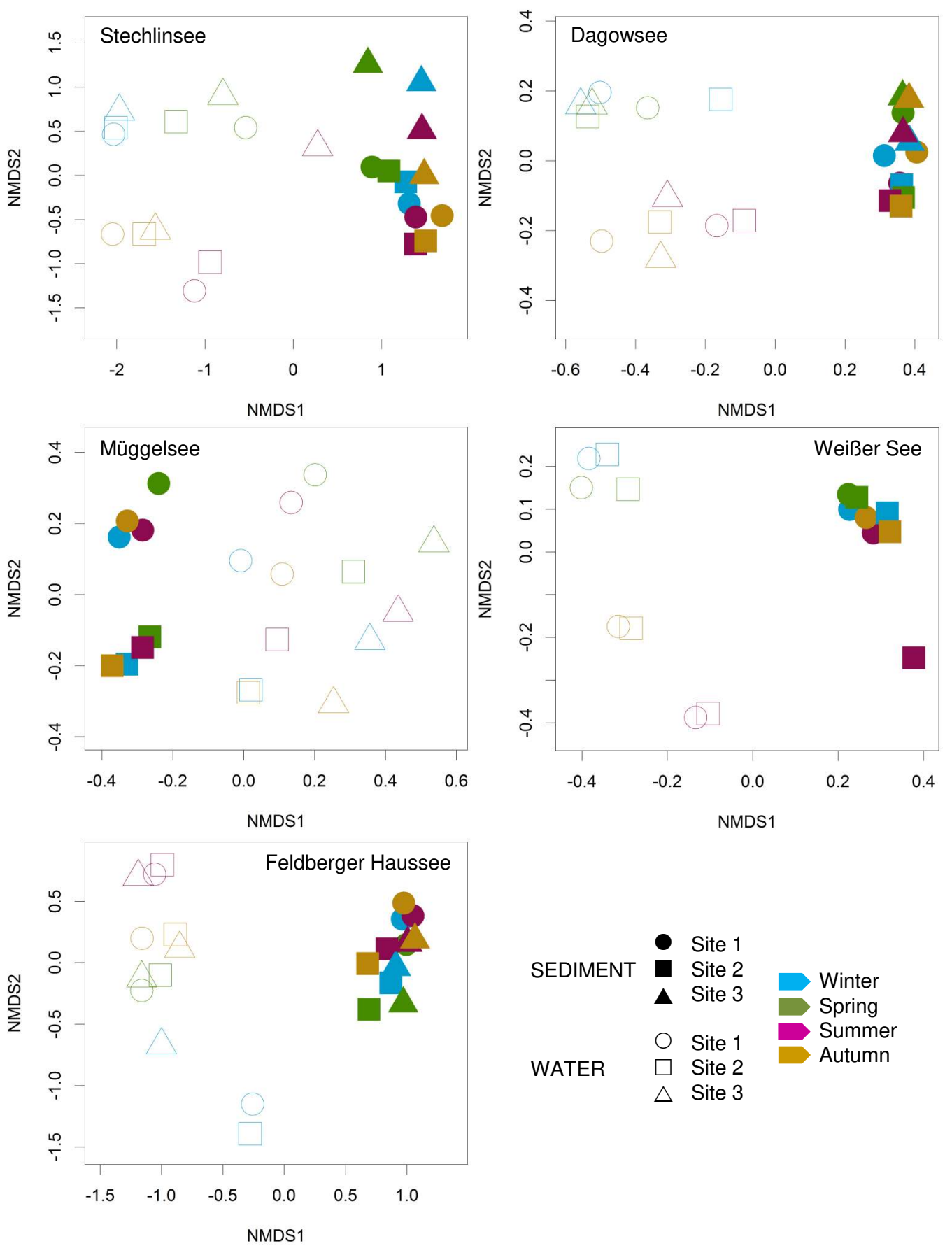
692 **Figure 5: Habitat specific bacterial communities. [a]** Ternary plots showing the number and relative
693 abundance of OTUs (dots) that had 10 or more sequences in at least 3 samples and their occurrence in
694 rural freshwater, urban freshwater and wastewater. Only the most abundant bacterial phyla/groups are
695 shown. Colours in the plots indicate the number of OTUs (log-transformed) and the size of the dots
696 indicate the maximum relative abundance for each OTU. Points close to the corners of the plots
697 represent either OTUs that occur more often or that are specific for that given habitat, while points
698 between two vertexes or in the middle of the plots have similar occurrence or are specific for the
699 combination of the related habitat. Max. RA – maximum relative abundance. [**b**] Relative proportion
700 of OTUs [%] present in one or more habitats. Coloured parts of the triangles correspond to the region
701 in the ternary plots. Alphaproteo – Alphaproteobacteria, Gammaproteo. – Gammaproteobacteria.

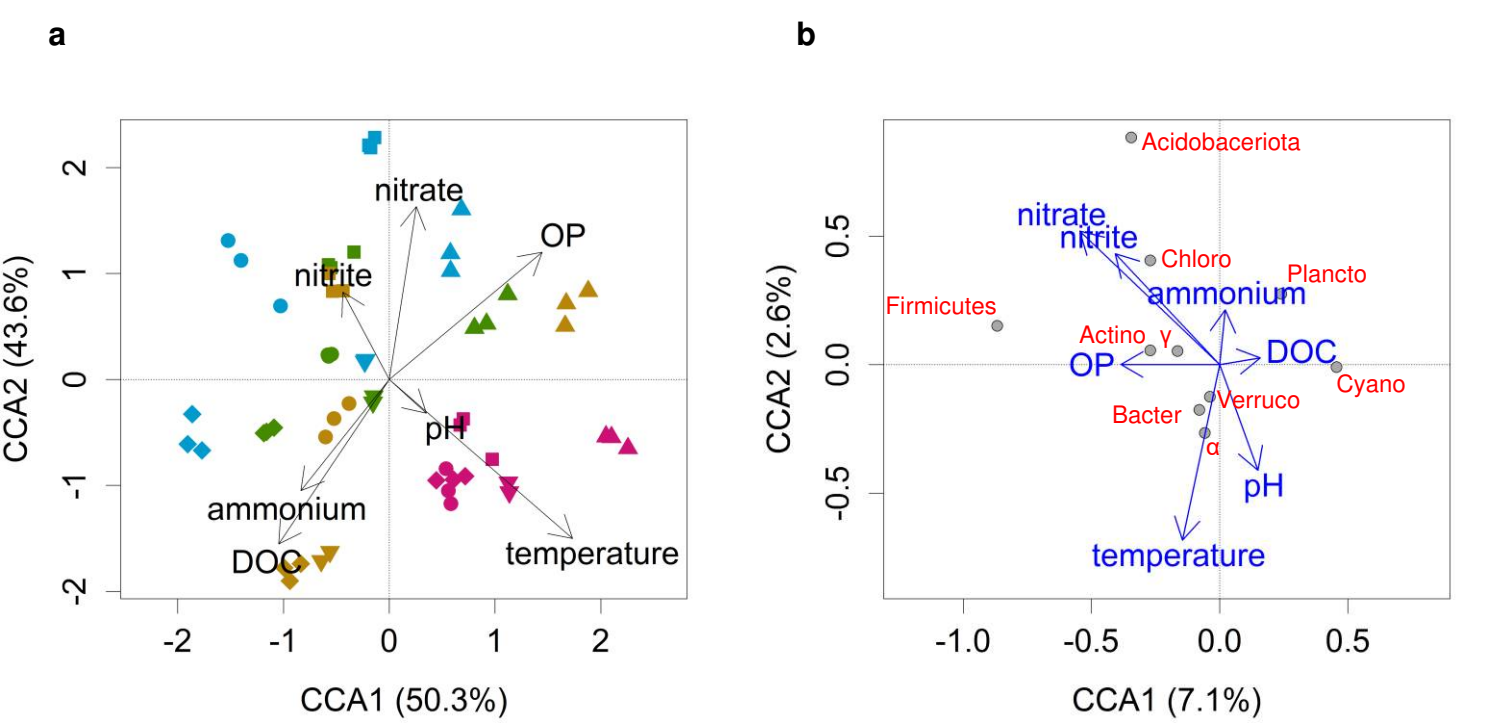
702 **Figure 6: The prevalence of genera that are known to contain potential human pathogens [a]**
703 **and their correlation with specific sub-communities** (Sub-Com.) **[b].** Heatmap **[a]** shows the
704 average relative abundance in the wastewater treatment plant (WWTP), lake water, and lake sediment.
705 Heatmap **[b]** shows the results of a weighted correlation network analyses (WGCNA). Only
706 significant correlations ($p<0.05$) of the potential pathogenic genera and specific sub-community
707 structures are shown. The composition and detailed occurrence of the sub-community OTUs can be
708 found in **Suppl. Figure S2**. Alphaproteo – Alphaproteobacteria, Gammaproteo. –
709 Gammaproteobacteria.

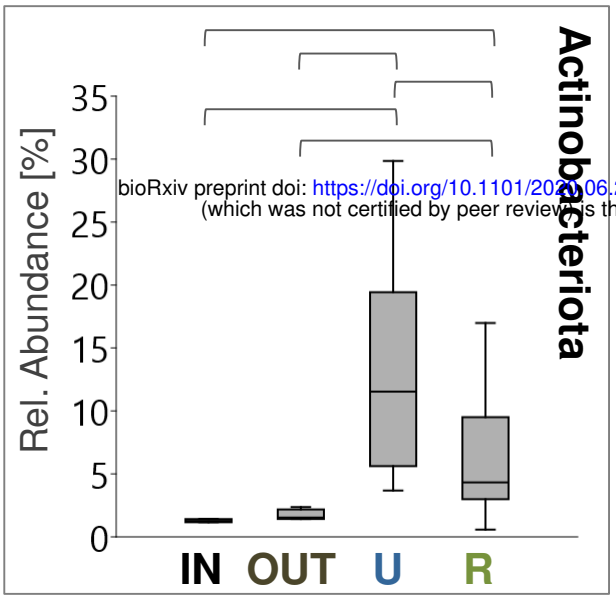
710 **Figure 7: Correlation of potential pathogenic genera and environmental measurements of the**
711 **water samples.** A constrained correspondence analysis (CCA) of genera that are known to contain
712 potential human pathogens from all water samples and the measured environmental measurements:
713 ammonium, dissolved organic carbon (DOC), nitrite, nitrate, orthophosphate (OP), pH, and
714 temperature. *Ac. Acinetobacter, Ae. Aeromonas, Al. Alistipes, Ba. Bacillus, Ca. Campylobacter, Cl.*
715 *Clostridium (sensu-stricto), En. Enterococcus, ES Escherichia/ Shigella, Kl. Klebsiella, Lg.*
716 *Legionella, Lp. Leptospira, Mi. Microcystis, Mb. Mycobacterium, Mp. Mycoplasma, Ne. Neisseria, Pe.*

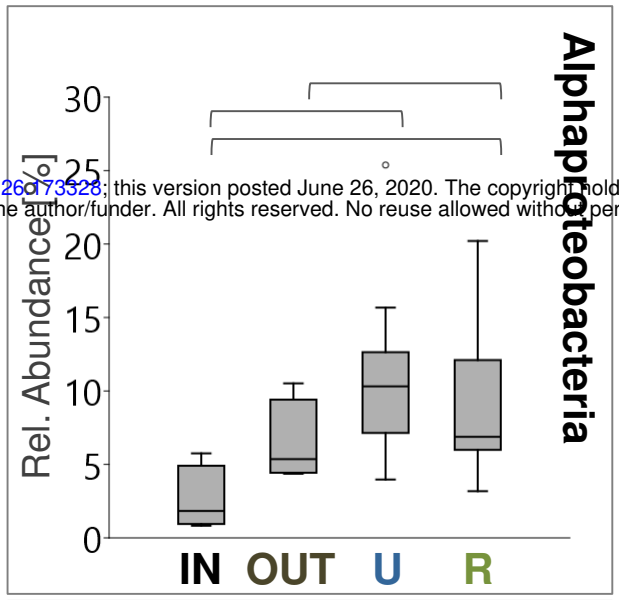
717 *Peptoclostridium, Ps Pseudomonas, Ri. Rickettsia, Sa. Staphylococcus, Se. Streptococcus, Tr.*

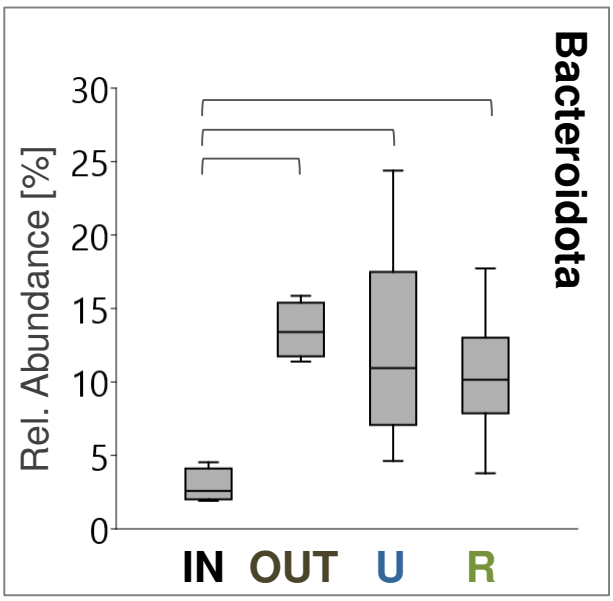
718 *Treponema, Vi. Vibrio* and *Ye. Yersinia.*

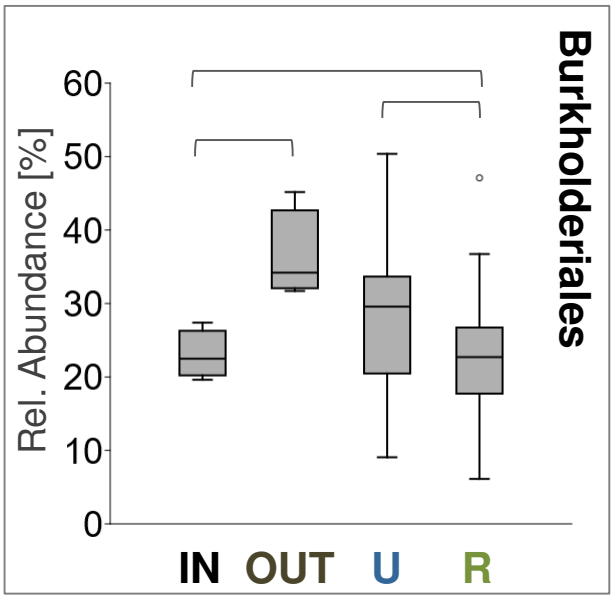

Stechlinsee
Dagowsee
Müggelsee
Weißer See
Feldberger Haussee
NMDS1
NMDS2
SEDIMENT
Site 1
Site 2
Site 3
WATER
Site 1
Site 2
Site 3
Winter
Spring
Summer
Autumn

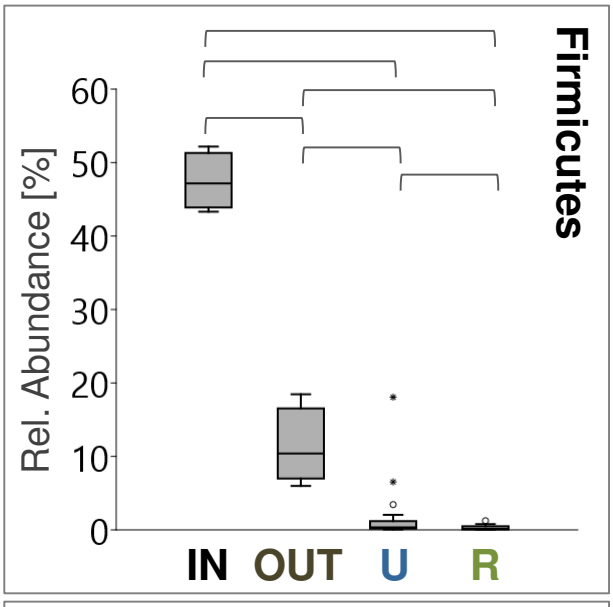

a
nitrate
OP
nitrite
pH
ammonium
DOC
temperature
CCA2 (43.6%)
CCA1 (50.3%)
b
Acidobaceriota
nitrate
nitrite
Chloro
Plancto
ammonium
Firmicutes
Actino
γ
OP
DOC
Cyano
Verruco
Bacter
α
pH
temperature
CCA2 (2.6%)
CCA1 (7.1%)

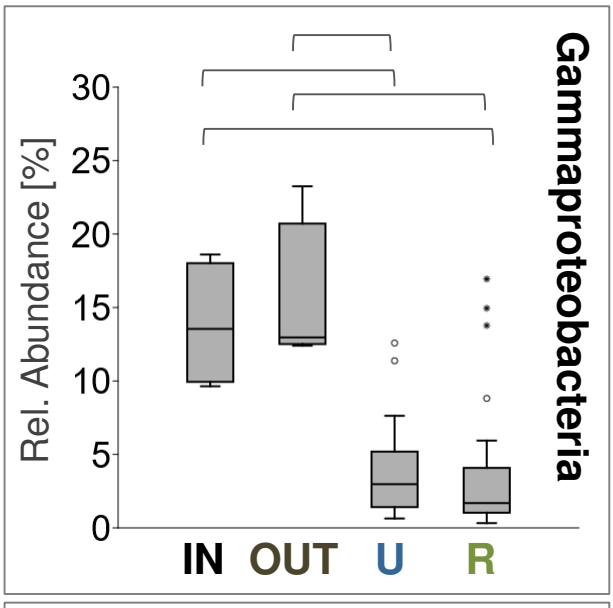


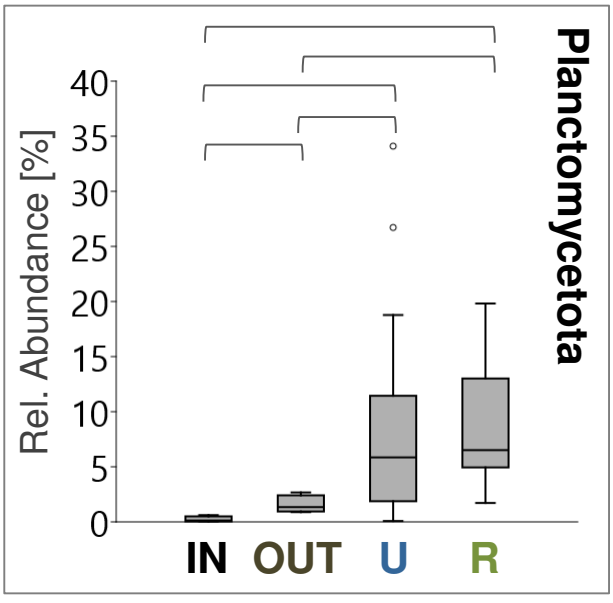


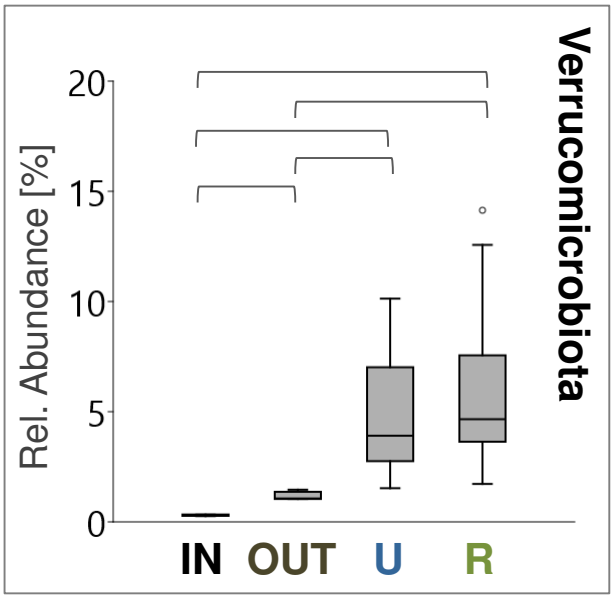




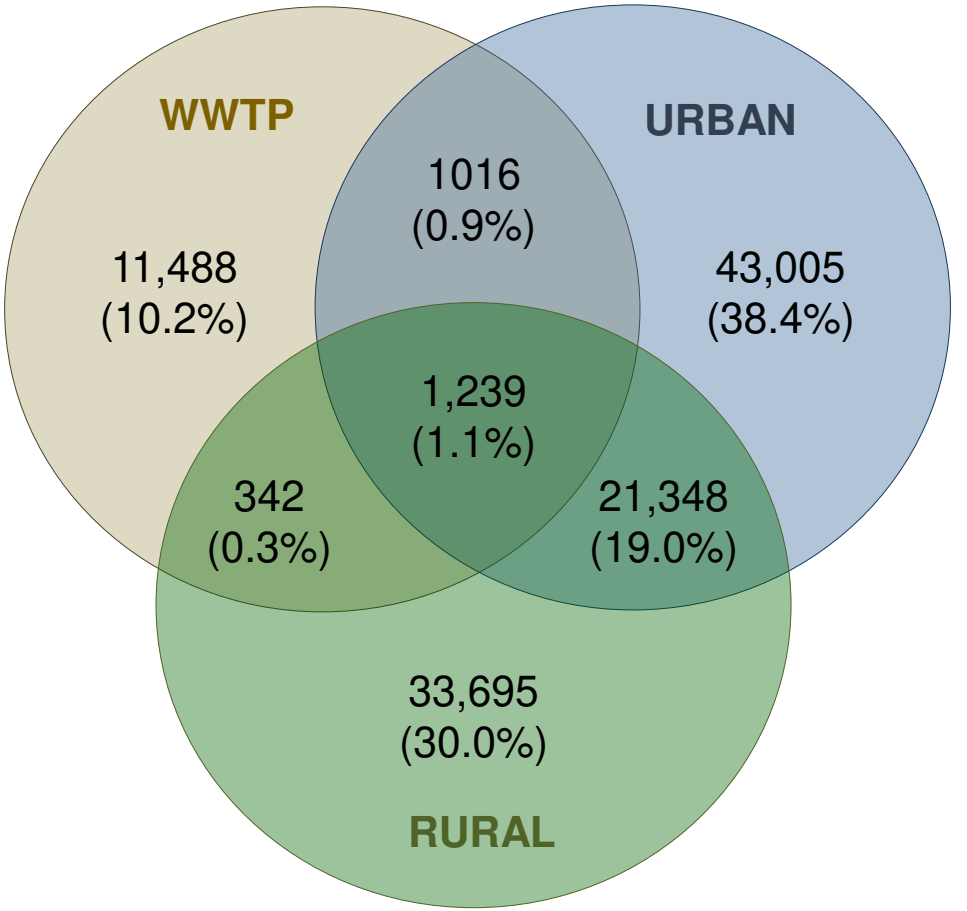




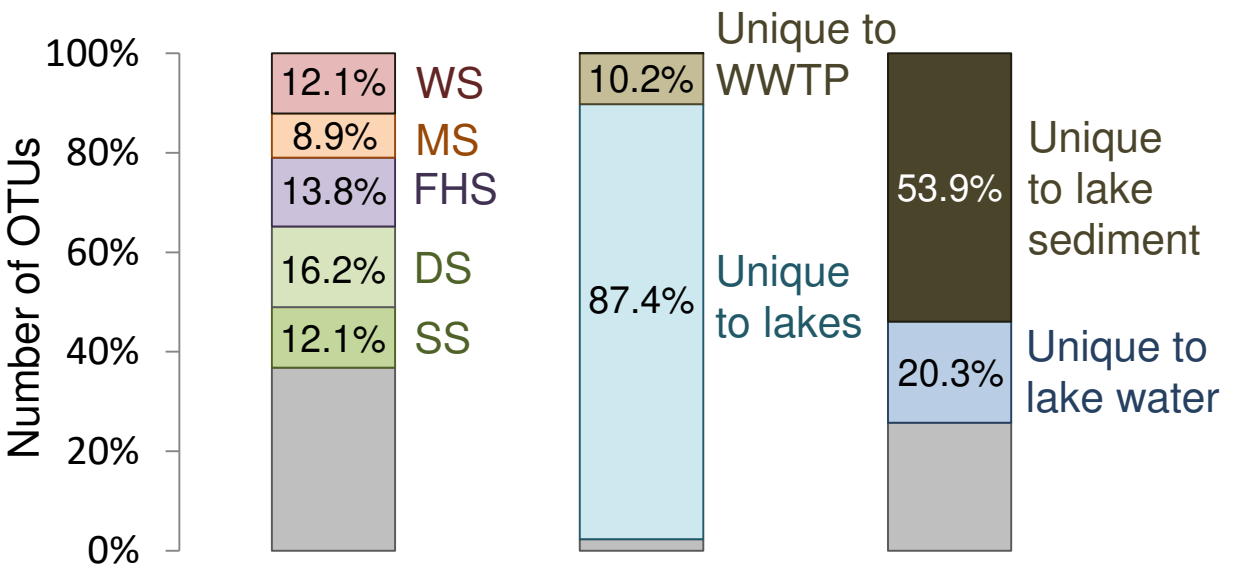




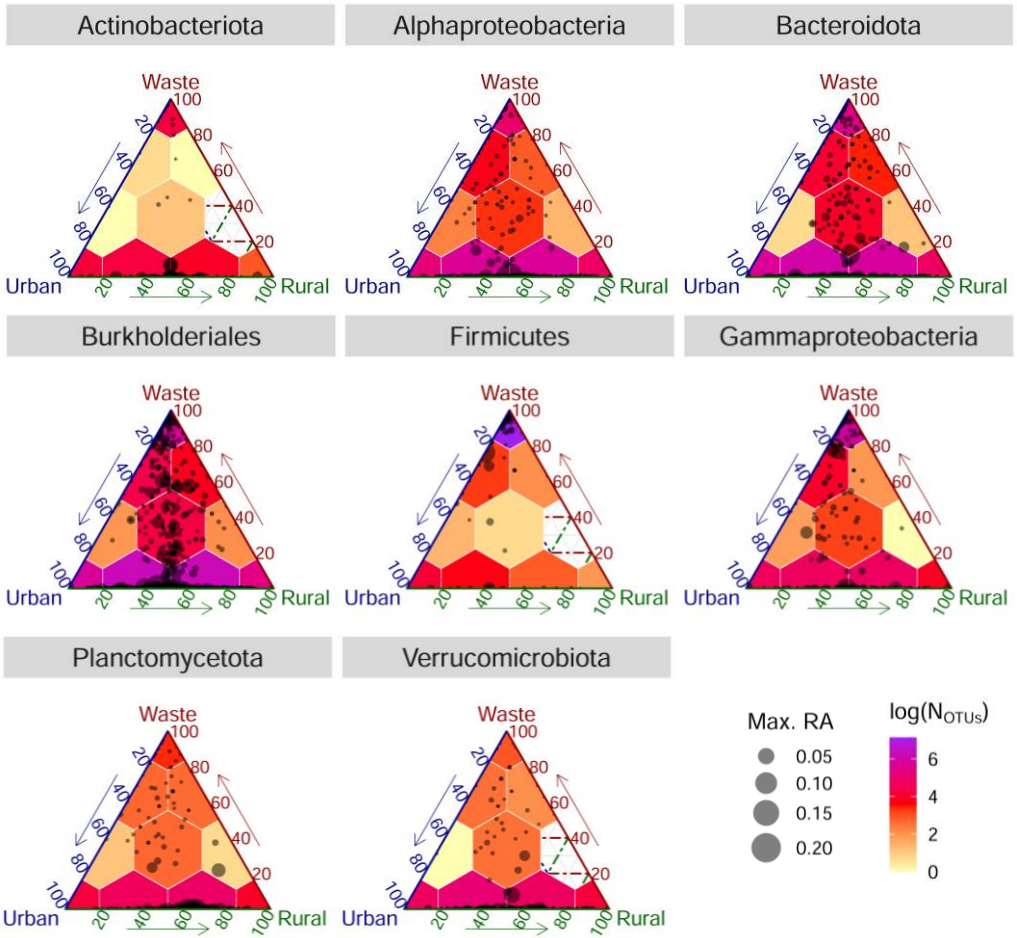
A


WWTP
URBAN
RURAL
11,488
(10.2%)
1016
(0.9%)
43,005
(38.4%)
1,239
(1.1%)
342
(0.3%)
21,348
(19.0%)
33,695
(30.0%)

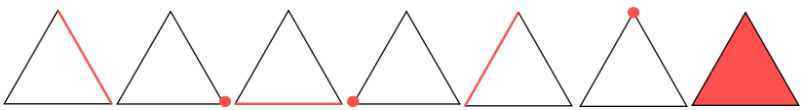
B


Number of OTUs
100%
80%
60%
40%
20%
0%
12.1%
WS
8.9%
MS
13.8%
FHS
16.2%
DS
12.1%
SS
Unique to WWTP
10.2%
Unique to lakes
87.4%
53.9%
Unique to lake sediment
20.3%
Unique to lake water

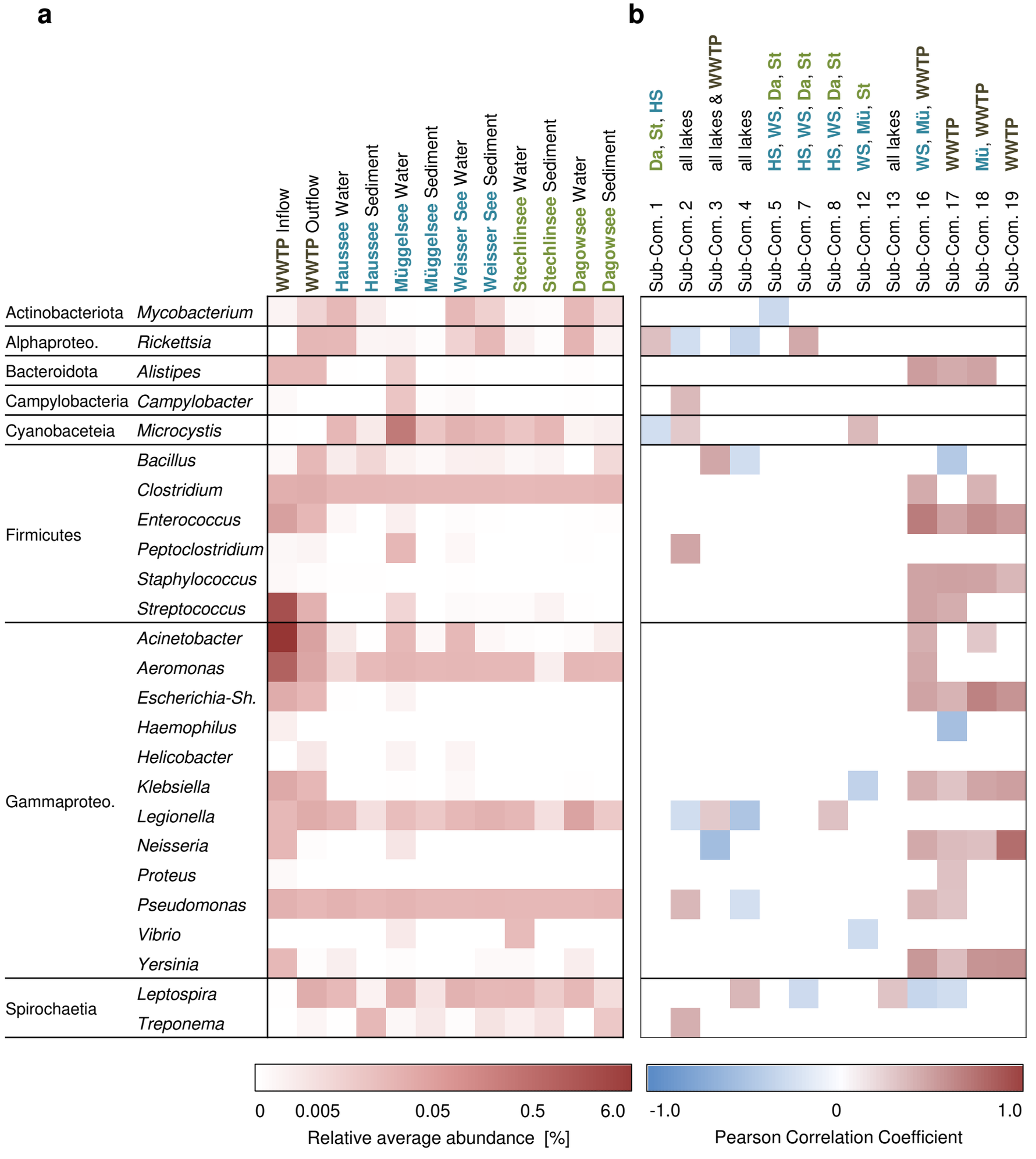
a

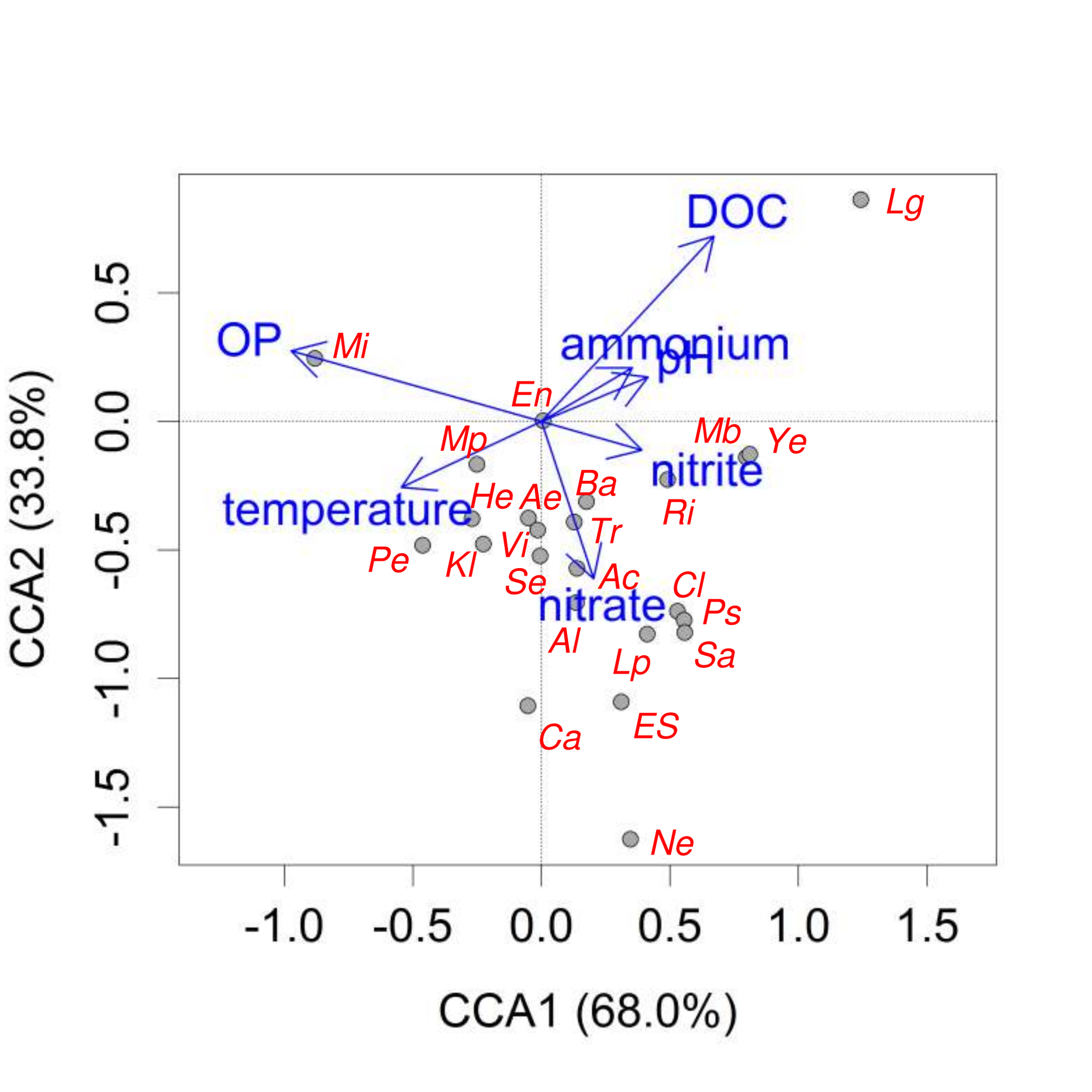


b



	waste & rural	rural	urban & rural	urban	waste & urban	waste	all
Actinobacteriota	0.0	5.7	43.6	22.1	6.1	18.9	3.6
Alphaproteo.	1.1	9.5	54.4	14.1	5.8	7.7	7.4
Bacteroidota	1.5	6.5	50.3	17.9	6.0	10.1	7.6
Burkholderiales	1.6	7.3	52.1	14.1	6.3	9.6	9.0
Firmicutes	0.6	0.4	4.2	4.3	12.4	76.8	1.3
Gammaproteo.	1.4	4.4	29.4	12.4	15.3	27.9	9.2
Planctomycetota	1.4	8.2	59.1	14.8	4.3	2.3	10.0
Verrucomicrobiota	1.1	9.2	57.8	22.1	2.3	1.9	5.5


a
b
WWTP Inflow
WWTP Outflow
Haussee Water
Haussee Sediment
Müggelsee Water
Müggelsee Sediment
Weisser See Water
Weisser See Sediment
Stechlinsee Water
Stechlinsee Sediment
Dagowsee Water
Dagowsee Sediment
Sub-Com. 1 Da, St, HS
Sub-Com. 2 all lakes
Sub-Com. 3 all lakes & WWTP
Sub-Com. 4 all lakes
Sub-Com. 5 HS, WS, Da, St
Sub-Com. 7 HS, WS, Da, St
Sub-Com. 8 HS, WS, Da, St
Sub-Com. 12 WS, Mü, St
Sub-Com. 13 all lakes
Sub-Com. 16 WS, Mü, WWTP
Sub-Com. 17 WWTP
Sub-Com. 18 Mü, WWTP
Sub-Com. 19 WWTP
Actinobacteriota
Alphaproteo.
Bacteroidota
Campylobacteria
Cyanobaceteia
Firmicutes
Gammaproteo.
Spirochaetia
Mycobacterium
Rickettsia
Alistipes
Campylobacter
Microcystis
Bacillus
Clostridium
Enterococcus
Peptoclostridium
Staphylococcus
Streptococcus
Acinetobacter
Aeromonas
Escherichia-Sh.
Haemophilus
Helicobacter
Klebsiella
Legionella
Neisseria
Proteus
Pseudomonas
Vibrio
Yersinia
Leptospira
Treponema
0
0.005
0.05
0.5
6.0
Relative average abundance [%]
-1.0
0
1.0
Pearson Correlation Coefficient


DOC
Lg
OP
Mi
ammonium
pH
En
Mp
Mb
Ye
nitrite
temperature
He
Ae
Ba
Ri
Tr
Pe
Kl
Vi
Se
Ac
Cl
nitrate
Ps
Al
Sa
Lp
ES
Ca
Ne
CCA2 (33.8%)
CCA1 (68.0%)
0.5
0.0
-0.5
-1.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5