Pollution shapes the microbial communities in river water and sediments from the Olifants River catchment, South Africa

Angel Valverde[1,2,3*], Errol D. Cason[1,2], Alba Gómez-Arias[1], Derya Bozkale[4], Danny Govender[5,6], Eddie Riddell[5,7], Don Cowan[4]

[1]Department of Microbial, Biochemical and Food Biotechnology, University of the Free State, Bloemfontein, South Africa

[2]Department of Animal-, Wildlife-, and Grassland Sciences, University of the Free State, Bloemfontein, South Africa

[3]Instituto de Recursos Naturales y Agrobiología de Salamanca, Consejo Superior de Investigaciones Científicas, Salamanca, Spain

[4]Centre for Microbial Ecology and Genomics, Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa

[5]South African National Parks, Conservation Management, Skukuza, South Africa

[6]Department of Paraclinical Sciences, University of Pretoria, Onderstepoort, South Africa

[7]Centre for Water Resources Research, University of KwaZulu-Natal, Pietermaritzburg, South Africa

*Corresponding author: IRNASA-CSIC, C/ Cordel de Merinas, 37008 Salamanca, Spain

E-mail address: angel.valverde@csic.es

Abstract

Human activities such as agriculture and mining are leading causes of water pollution worldwide. Individual contaminants are known to negatively affect microbial communities. However, the effect of multifaceted pollution on these communities is less well understood. We investigated, using next-generation sequencing of the 16S rRNA genes, the effects of multisource (i.e., fertilizer industry and mining) chronic pollution on bacterial and archaeal communities in water and sediments from the Olifants River catchment, South Africa. Water samples showed less microbial species diversity than sediments and both habitats displayed different microbial communities. Within each of these habitats, pollution had no effect on alpha diversity but shaped the microbial composition and taxonomy-based predicted functions. Certain prokaryotic taxa and functional groups were indicative of different degrees of pollution. Heterotrophic taxa (e.g., *Flavobacterium* sp.) and sulphur-oxidizing bacteria (i.e., *Thiobacillus* sp.) were indicators of pollution in water and sediments, respectively. Ultimately, this information could be used to develop microbial indicators of water quality degradation.

Keywords: Chronic pollution, Microbial communities, river sediments, 16S rRNA gene sequencing

Introduction

Pollution has been identified as the main pressure affecting freshwater systems and resources around the world (Vörösmarty et al., 2010). Water pollution can result from different anthropogenic activities, including mining and fertilizer industry. Discharged mine water negatively impacts aquatic environments by increasing the levels of suspended solids, leading to mobilization of elements such as iron, copper, manganese and zinc and also altering the pH of the receiving water (Baker and Banfield, 2003). On the other hand, chemical fertilizers containing phosphorus and nitrogen contributes to eutrophication, one of the leading causes of degraded water quality worldwide (Dodds, 2006).

Microorganisms play important roles in freshwater ecosystems, such as the fixation of carbon through the process of photosynthesis and participation in the release of nutrients (Cotner and Biddanda, 2002), which support aquatic food webs (Eiler et al., 2014). The activities of aquatic microbial communities are affected, both positively and negatively, by a wide variety of pollutants. Trace amounts of transition metals such as iron, copper, zinc and manganese contribute to enzyme activation (Samanovic et al., 2012 and references therein). Phosphorus levels are important for the population growth rates of phytoplankton (Reynolds, 2009). Nevertheless, in large amounts most pollutants, including transition metals, are toxic.

In addition to affecting the activity of microorganisms, many studies have also demonstrated that water pollution affects the diversity and composition of microbial communities. For example, it was found that the relative abundance of *Betaproteobacteria* decreased across an alkaline contamination gradient in Central Appalachian streams (Bier et al., 2015), whereas *Alphaproteobacteria* increased in relative abundance in acidic waters across Southeast China (Kuang et al., 2013). Thus, it is postulated that some microbial taxa might be indicative of pollution; that is, they can be used as environmental bioindicators. Nowadays, the diversity and composition of microbial communities can be easily studied using next generation

sequencing techniques and they have been shown to respond quickly to pollution (Feris et al., 2009). Therefore, the use of microorganisms as pollution indicators might be a more sensitive option compared to the use of other organisms such as invertebrates and fish.

Most of the studies have evaluated the effects of a single source or type of pollution. Furthermore, most of these investigations have focused on the water column and not on sediments (but see García-Moyano et al., 2012), although sediments tend to accumulate higher amounts of contaminants than water masses (Salomons and Stigliani, 1995). High sediment contaminant loads can have a significant impact on the entire ecosystem and on human health, as metals in sediments can be remobilized and transported downstream, causing secondary contamination. Conversely, sediment microbial assemblages seem to be more diverse than those of any other environment (Lozupone and Knight, 2007). This may result from the high heterogeneity of the sediment, both in terms of environmental gradients and biogeochemical processes. Sediment heterogeneity may provide a larger number of niches, allowing the coexistence of more diversified assemblages of organisms.

Here, using Illumina sequencing of the 16S rRNA genes and chemical analyses, we investigated the concomitant effects of different pollutants (i.e., phosphorus and heavy metals) on bacterial and archaeal communities in river water and sediment samples. To this end, we used a section of the Olifants River network (South Africa) affected by the activities of the Phalaborwa industrial complex (PIC), which includes two mines and a fertilizer industry (Gomez-Arias et al., 2016; Heath et al., 2010; Marr et al., 2017). The mines extract metals and phosphate from alkaline rocks; whereas the fertilizer industry produces phosphoric acid attacking the phosphate with sulphuric acid. Thus, the mines produce alkaline mine drainage; while the fertilizer industry generates acidic wastewater. We posed the following questions. To what extent do the water and sediment bacterial communities differ in terms of taxonomy and predicted function? does contamination (both in water and

sediments) significantly alter microbial community composition and function? and, if so, which taxa and functions can be used as bioindicators of such pollution?

Materials and methods

Sample collection

Samples were collected in April 2016 at 14 sampling points (Supplementary meterial Fig. S1) and grouped *a priori* in three categories ('high contamination', 'mid contamination' and 'low contamination') based on previous data (Gomez-Arias et al., 2016) and on the assumption that samples near the PIC should be more contaminated than samples distant from the complex. The six 'high contamination' sites were located in the Selati River, a tributary of the Olifants River. The four 'mid contamination' sites were located upstream (1 site) and downstream (3 sites) the PIC. The five 'low contamination' sites were situated in the Olifants River, after the confluence with the Selati River, further down the complex. At each sampling site, two different habitat types (water and sediments) were sampled, resulting in a total of 28 samples. Surface water (top 10 cm, 500 ml per sample) was filtered through 0.22 µm sterile nitrocellulose membranes (Nalgene, Rochester, NY, USA). Sediment samples (250 g per sample) were collected with a soil probe from the top 10 cm (probably containing both oxic and anoxic sediments) of the streambed and stored in sterile Whirl-Pak sampling bags. Water and sediment samples were kept at 4°C during transport. Once in the laboratory sediments were sieved using sterile 2 mm sieves to obtain a homogeneous sediment particle size and then stored, together with nitrocellulose membranes, at -80 °C prior to DNA extraction.

Sediment and water chemistry

We measured water pH *in situ* using a water probe (Campbell Scientific, South Africa) and analysed sediment and water samples for several major and trace elements (Table 1) using

standard procedures. Trace elements (e.g., Cu, Pb) were measured with a VG PlasmaQuad-3 (Thermo Fisher Scientific Inc.) inductively coupled plasma mass spectrometry (ICP-MS) and major elements (e.g., Ca, Na) with an ARL SpectraSpan 7 (Thermo Fisher Scientific Inc.) direct current plasma optical emission spectrometry (DCP-OES). Total P was measured using the P Bray method.

DNA extraction and amplicon sequencing

The DNA extraction for both sample types was carried out using the Power Soil DNA Isolation Kit as per the standard protocol (MoBio Laboratories, Carlsbad, CA, USA). For water samples, half of the nitrocellulose filters were cut into small pieces with a sterile blade. In the case of sediment samples, genomic DNA was directly extracted from 250 mg of sediment. DNA was amplified via a single step PCR using the HotStarTaq Plus Master Mix Kit (Qiagen, Valencia, CA). The primer pair 515F (5'-GTGYCCAGCMGCCGCGGTA-3') and 909R (5'-CCCCGYCAATTCMTTTRAGT-3') (Tuan *et al*., 2014) was used for the amplification of the 16S rRNA genes. PCR was performed in triplicate for each sample under the following conditions: 94°C for 3 minutes, followed by 28 cycles of 94°C for 30 seconds, 53°C for 40 seconds and 72°C for 1 minute and a final elongation step at 72°C for 5 minutes. Amplicon products containing sample-specific barcodes were pooled together in equal concentrations (5 ng/µl) and purified using Agencourt Ampure XP beads (Agencourt Bioscience Corporation, MA, USA). The PCR product was then used to prepare a DNA library following the Illumina TruSeq DNA library preparation protocol. Sequencing was performed at MRDNA (www.mrdnalab.com, Shallowater, TX, USA) on an Ilumina MiSeq2000, using a paired-end approach, following the manufacturer's guidelines.

Sequence data processing

The raw Illumina sequence data was analysed using QIIME v1.9.0 (Caporaso et al., 2010b). Briefly, demultiplexing and quality filtering was performed using split_libraries_fastq.py with a phred_quality_threshold of 25. Chimeric sequences were identified using usearch 6.1.544 (Edgar, 2010) against the RDP v16 database (Cole et al., 2009) and filtered out by running the identify_chimeric_seqs.py and filter_fasta.py commands, respectively. Open reference OTU picking was performed and taxonomy assigned to representative OTUs using the pick_open_reference_otus.py script, at 97% sequence identity against the SILVA v128 database (Pruesse et al., 2007). For phylogenetic analysis, OTUs were aligned using PyNAST (Caporaso et al., 2010a) and a phylogenetic tree constructed with fastTree (Price et al., 2010) implemented in Qiime. Any OTU classified as chloroplast or mitochondria was excluded from further analysis.

Statistical analysis

All statistical analyses were performed with R version 3.6.3 (Team, 2011). OTU richness and phylogenetic diversity (PD) were obtained with the package picante (Kembel et al., 2010). Rarefaction curves and Chao1 were calculated using phyloseq (McMurdie and Holmes, 2013) in R (Team, 2011). Sequencing data were also used to predict potential functional capacity of the OTUs using FAPROTAX (Louca et al., 2016). Briefly, an OTU is associated with a particular metabolic function if all cultured representatives within that OTU have been reported to exhibit that function.

Abiotic data were standardized and pair-wise distances computed based on Euclidean distances. Normalized weighted UniFrac distances were obtained with the taxonomic data matrix. Bray-Curtis distances were used with the functional data matrix after Hellinger-transformation. The environmental variables were visualized using principal component analysis (PCA) and the taxonomic and functional structures of the microbial community

principal coordinate analysis (PCoA). The effect of abiotic data in explaining variations in bacterial community composition was assessed by distance-based redundancy analysis (db-RDA). A permutational analysis of variance (PERMANOVA) was used to test for differences in composition and function between and within habitats, whereas permutation dispersion (PERMDISP) was used to test for differences in habitat dissimilarity; both analyses were performed with the “adonis" and “betadisper” functions in vegan (Oksanen et al., 2013) for R. Kruskal-Wallis test were used to determine significant differences in alpha diversity, chemistry and phyla relative abundances between habitats (sediment *vs* water). Within habitat differences were assessed using Wilcox tests following significant Kruskal-Wallis tests. The Benjamini-Hochberg FDR correction was applied to adjust the P value for multiple comparisons in R.

To identify which prokaryotic phyla, families, genera and functions were indicative of the six sub-habitats, we used species indicator analyses (Dufrene and Legendre, 1997). Indicator species are species that are found mostly in a single habitat and are present in most sites or samples from that habitat (Legendre and Legendre, 1998). Only taxa and functions with significant ($P < 0.01$) indicator values that were > 0.3 were considered, as this latter value can be regarded as a good threshold for habitat specialization (Dufrene and Legendre, 1997). For all analyses the OTU table was rarefy to 22,697 sequences per sample to equalize sequencing depth.

The raw sequences are available in the NCBI database under the BioProject accession number PRJNA485640.

Results and discussion

Environmental characterization

The chemistry of the sediments was more variable (PERMDISP $F_{1,26}$= 7.8, P=0.001) but clearly distinct from that of the water samples (PERMANOVA $F_{1,26}$= 30.3, R^2=53.8%, P=0.001) (Figure 1). Furthermore, sediment samples showed higher levels of contaminants than water samples (Table 1); i.e., Al, Cu, Fe, Pb, and P concentrations were higher in sediments than in water (Kruskal-Wallis test < 0.05). In contrast, pH values were higher in water (Kruskal-Wallis test < 0.05). Variables such as organic matter, mineralogy and pH differ between water and sediments, and these factors are well known to affect the solubility of metals (Calmano et al., 1993). For example, organic matter sedimentation plays a major role in the accumulation of metals in sediments (Hsu et al., 2016).

Within sediment samples, the chemistry of low contamination samples was significantly different from those of mid and high contamination samples (PERMANOVA $P < 0.05$, in both cases). Aluminium, copper and phosphorus showed higher values (Wilcox test, $P < 0.05$) in high contamination sediments when compared to low contamination sediments (Table 1). The concentration of these contaminants differ in sediments around the world, but for example, the levels of copper are below the guideline sediment value of 65 mg/kg used by regulatory bodies internationally (Sutcliffe et al., 2019 and references therein). Similar PERMANOVA results were found for water. Phosphorus showed higher values (Wilcox test, $P < 0.05$) in high contamination water when compared to low and mid contamination water, while pH was lower in high contamination water compared to low and mid contamination water (Table 1). In general, the physicochemical parameters did not exceed the prescribed South African water quality standards for livestock watering; that is, the water was less contaminated than expected. In general for both sediment and water samples, individual contaminant concentrations decreased substantially after the confluence with the Olifants River, but did not show a consistent decline with distance from the PIC. Particularly, copper and lead sediment levels, were higher in some samples from sites downstream and distant

from the PIC (Supplementary meterial Table S1). These patterns suggest that the levels of contaminants in aquatic environments are controlled not only by anthropogenic sources but possibly also by other factors, such as natural geological weathering of rocks and soils exposed to surface water, and by the preferential accumulation of the contaminants by suspended particles and sediments (Jackson et al., 2015).

Alpha-diversity patterns

A total of 3,750 bacterial and archaeal OTUs (97% similarity cut-off), ranging from 673 to 1,932 per sample, were found in water and sediments using identical sequencing depth. Rarefaction curves, Chao1 and Good's coverage estimates suggest that this sequencing depth was adequate to capture most of the prokaryotic diversity in each sample (Supplementary meterial Fig. S2). Of the total number of OTUs, 555 (representing 1.3 % of the total number of sequences) were unique to water, 1,332 (7.6%) were unique to sediments and 1,863 (91.1%) were shared between the two habitats.

Alpha diversity (richness and PD) was higher in sediments than in water samples (Kruskal-Wallis test, $P < 0.05$) (Figure 2). This was expected as sediments are known to harbour highly diverse microbial communities (Lozupone and Knight, 2007) and some of the sediment samples likely contained both oxic and anoxic layers. River water bodies are thought to be more homogeneous habitats that maintain relatively low and constant microbial populations, while sediments are considerably more complex environments, which can lead to spatially/resource-driven niche partitioning (Crump et al., 2012). Niche partitioning commonly increases microbial alpha diversity (Gibbons and Gilbert, 2015). In contrast, alpha diversity was not affected by the level of contamination (Kruskal-Wallis test, $P > 0.05$), either in the sediment or in the water samples (Figure 2). It is possible that the concentration of the contaminants was too low to reduce the microbial diversity. For instance, lab studies have

demonstrated that diversity decrease in sediments containing 46 mg/kg of copper (Sutcliffe et al., 2019), while here high contamination sediments contained ~30 mg/kg.

The number of the functions identified using FAPROTAX correlated with both richness and phylogenetic diversity ($r^2 > 0.86$, $P < 0.001$ in both water and sediments). Similar results have been found in other microbial communities (Bryant et al., 2012), indicating that the overall functional diversity found in a sample is, to a certain degree, predictable from the taxonomic and phylogenetic diversity of the microbial communities in that sample.

Beta diversity patterns: Microbial community composition

Distinct prokaryotic communities (OTU level) were detected in water and sediment samples (PERMANOVA $F_{1,26}= 23.7$, $R^2=47.7\%$, $P=0.001$) using normalized weighted UniFrac dissimilarities. Microbial communities were also distinct between high contamination, mid contamination and low contamination sediments (PERMANOVA $P < 0.05$, in all three comparisons). In contrast the water communities from low and mid contamination did not differ significantly (PERMANOVA $P > 0.05$). There were no differences in compositional heterogeneity within high contamination and low contamination communities for both water and sediment samples after adjusting the P value for multiple comparisons (PERMDISPER $P > 0.05$, in both cases). Using distance redundancy analysis, Pb, Cu and pH levels were found to be the most important factors contributing to the overall differences in bacterial community composition between sediment and water samples, explaining 47% of the total variation (Figure 3). When water and sediment bacterial communities were analysed separately, Cu, Fe and pH were the factors that best explained the variation (43%) of sediment communities, while Al and P were the most important contributing factors (32%) for water communities. This suggests that sediment communities were primarily affected by the level of metal contamination, whereas water communities were affected by both metal

contamination (i.e., Al) and the supply of nutrients (i.e., phosphorus). Heavy metals such a copper, nutrients such as phosphorus and pH, have been shown to influence microbial community composition in aquatic environments (Kuan et al., 2013; Langenheder et al., 2012).

Altogether, these results suggest that both between and within habitat differences play a major role in shaping these microbial communities, which is consistent with the concept of environmental filtering (Lindstrom and Langenheder, 2012), where abiotic factors select against or favour certain species. Conversely, the fact that sediment microbial communities were more variable in composition than water microbial communities could also be explained by differences in dispersal rates, which are inherently higher in water than in sediment samples. Nevertheless, 23-40% of the total variation in microbial community composition could be explained by the chemistry of the samples, suggesting that unmeasured environmental variables, biotic interactions (Lima-Mendez et al., 2015) and demographic drift (Ofiteru et al., 2010) might also affect microbial community composition.

Contamination-indicator taxa

Overall, Proteobacteria (60% in sediments, 38% in water) (mean relative abundance), Bacteroidetes (19% in sediments, 33% in water), Actinobacteria (0.8% in sediments, 22% in water) and Cyanobacteria (1% in sediments, 4% in water) were the most abundant phyla in the samples. Six other phyla, with average relative abundances higher than 0.5 %, were also detected (Figure 4a). Most of these phyla (except Verrucomicrobia) showed relative abundances that were significantly different between water and sediments samples (Kruskal-Wallis test, $P < 0.05$). Acidobacteria, Chloroflexi, Nitrospirae, Planctomycetes, Proteobacteria and Euryarchaeota were more abundant in sediments, while Actinobacteria,

Bacteroidetes and Cyanobacteria were more abundant in water. Indeed, the phylum Bacteroidetes was found to be indicator of high contamination in waters (Figure 4b).

We found that 71 families were indicators of one of the six sub-habitats (Supplementary meterial Table S2). For instance, the families *Thiobacillaceae* (Betaproteobacteria) and *Saprospiraceae* (Bacteroidetes) were indicators of high contamination in sediments (Figures 4c, d). These are families typically found in freshwater and comprise members known to use elemental sulphur, sulphide, thiosulfate, or polythionates as energy sources (*Thiobacillaceae*) and to degrade complex organic compounds (*Saprospiraceae*). On the other hand, the families *Flavobacteriaceae* (Bacteroidetes), and *Methylophilaceae* (Betaproteobacteria) were indicators of high contamination in water (Figure 4d). The utilization of macromolecules such as polysaccharides and proteins is a common feature of many members of the family *Flavobacteriaceae.* Members of the family *Methylophilaceae* are methylotrophs that are specialized in using reduced one-carbon (C_1) compounds like methanol, methylamine, and formaldehyde as sole energy and carbon sources (Salcher et al., 2019).

Of a total of 196 microbial genera detected in the study, fifty-two were classified as possible indicators of one of the six sub-habitats (Supplementary meterial Table S3). For example, *Thiobacillus* (Betaproteobacteria) and *Cloacibacterium* (Bacteroidetes) were assigned as indicators of high contamination in sediments (Figures 4e, f), while *Flavobacterium* (Bacteroidetes) and *Polynucleobacter* (Betaproteobacteria) were likely indicators of high contamination in water (Figures 4e, f). Genera such as *Thiobacillus* and *Flavobacterium*, typical representatives of the above-mentioned families *Thiobacillaceae* and *Flavobacteriacea*, respectively, have previously been found to dominate in contaminated environments (Yergeau et al., 2012; Pei et al., 2018). Bacterial strains affiliated to *Polynucleobacter* were reported to occur both as obligate endosymbionts of ciliates and as free-living forms that perform the assimilatory reduction of nitrate and assimilate sulphur and

sulphate (Boscaro et al., 2013). Overall, the presence of microbial bioindicators are probably linked to differences in contaminant toxicity tolerance and nutritional preferences between the different taxa, although biological interactions cannot be ruled out. Experiments with different combinations and concentrations of pollutants, and different numbers of microbial species, are needed to disentangle the contribution of these factors.

Beta diversity patterns: Functional groups

Functional annotation of OTUs revealed a total of 77 metabolic functional groups in the samples (Supplementary meterial Table S4). There were several functional groups indicative of contamination (Figure 5, Supplementary meterial Table S5). For instance, 'high contamination' sediments were enriched in dark sulphide- and sulphur-oxidizers, which usually couple the oxidation of sulphur and sulphide to the reduction of oxygen or nitrate under microaerophilic conditions. This can probably be explained by the high concentration of sulphur compounds found in water bodies near the Phalaborwa industrial complex (Gomez-Arias et al., 2016) and the sampling of the sediments at the oxic-anoxic interfase. Chemoheterotrophs were overrepresented in 'high contamination' water. Presumably, some bacteria that decompose organic matter are better competitors for organic and inorganic nutrients than others, leading to shifts in community structure in response to nutrient (i.e., phosphorus) enrichment. The observed functional patterns suggest that microbial primary and secondary production in the Olifants River catchment might be affected by chronic contamination. Nevertheless, as the functional profiles were inferred from taxonomy, the next logical step is to investigate whether or not the compositional shifts observed translate to functional responses using more sophisticated tools such metatranscriptomics, and to quantify those productivity changes.

Conclusion

We have shown that water and sediment samples collected from the Olifants river catchment harbour significantly different microbial communities, and that the sediment communities are more diverse than water communities. Also, that chronic pollution did not affect alpha diversity (richness and phylogenetic diversity). In contrast, we demonstrate that chronic pollution shapes microbial community composition and, accordingly, taxonomy-based predicted function (beta diversity). Community shifts were found to occur at various taxonomic levels from phylum to species. Overall, these results indicate that several microbial taxa can be used as bioindicators of the underlying differences in pollution.

Acknowledgements

Funding for this research was provided by the National Research Foundation and SANPARKS, South Africa. We are grateful to the soil chemistry laboratory at the University of Pretoria for performing the chemical analysis, and SANPARKS for facilitating the collection of the samples.

References

Baker BJ, Banfield JF (2003) Microbial communities in acid mine drainage. FEMS Microbiol Ecol 44:139-152. https://doi.org/10.1016/S0168-6496(03)00028-X

Bier RL, Voss KA, Bernhardt ES (2015) Bacterial community responses to a gradient of alkaline mountaintop mine drainage in Central Appalachian streams. ISME J 9:1378-1390. https://doi.org/10.1038/ismej.2014.222.

Boscaro, V, Felletti, M, Vannini, C, Ackerman, MS, et al. (2013) *Polynucleobacter necessarius*, a model for genome reduction in both free-living and symbiotic bacteria. *Proc Natl Acad Sci USA* **110:** 18590–18595.

Bryant JA, Stewart FJ, Eppley JM, et al. (2012) Microbial community phylogenetic and trait diversity declines with depth in a marine oxygen minimum zone. Ecology 93:1659-1673. https://doi.org/10.1890/11-1204.1

Calmano W, Hong J, Forstner U (1993) Binding and mobilization of heavy metals in contaminated sediments affected by pH and redox potential. Wat Sci Tech 28:223-235. https://doi.org/10.15480/882.450

Caporaso, JG, Bittinger K, Bushman, FD, et al. (2010a) PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics 26:266-267. https://doi.org/10.1093/bioinformatics/btp636

Caporaso, JG, Kuczynski, J, Stombaugh, J, et al. (2010b) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335-336. https://doi.org/10.1038/nmeth.f.303.

Cole, JR, Wang, Q, Cardenas, E, et al. (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res 37:D141-D145. https://doi.org/10.1093/nar/gkn879.

Cotner, JB, Biddanda, BA (2002) Small players, large role: Microbial influence on biogeochemical processes in pelagic aquatic ecosystems. Ecosystems 5:105-121. https://doi.org/10.1007/s10021-001-0059-3.

Crump, BC, Amaral-Zettler, LA, Kling, GW (2012) Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. ISME J 6:1629-1639. https://doi.org/10.1038/ismej.2012.9

Dodds, WK (2006) Eutrophication and trophic state in rivers and streams. Limnol Oceanogr 51: 671-680.

Dufrene, M, Legendre, P (1997) Species assemblages and indicator species: The need for a flexible asymmetrical approach. Ecol Monograph 67:345-366. https://doi.org/10.2307/2963459.

Edgar, RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26,:2460-2461. https://doi.org/10.1093/bioinformatics/btq461.

Eiler, A, Zaremba-Niedzwiedzka, K, Martínez-García, M, et al. (2014) Productivity and salinity structuring of the microplankton revealed by comparative freshwater metagenomics. Environ Microbiol 16:2682-2698. https://doi.org/10.1111/1462-2920.12301.

Feris, KP, Ramsey, PW, Gibbons, SM, et al. (2009). Hyporheic microbial community development is a sensitive indicator of metal contamination. Environ Sci Technol 43: 6158-6163.

García-Moyano, A, González-Toril, E, Aguilera, A, et al. (2012) Comparative microbial ecology study of the sediments and the water column of the Río Tinto, an extreme acidic environment. FEMS Microbiol Ecol 81:303-314. https://doi.org/10.1111/j.1574-6941.2012.01346.x.

Gibbons, SM, Gilbert, JA (2015) Microbial diversity-exploration of natural ecosystems and microbiomes. Curr Opin Gen Dev 35:66-72. http://dx.doi.org/10.1016/j.gde.2015.10.003

Gomez-Arias, A, Castillo, J, van Heerden, E, et al. (2016) Use of alkaline mine waste as treatment for acid drainage, Proceedings IMWA. Freiberg/Germany.

Heath, R, Coleman, T, Engelbrecht, J (2010) Water quality overview and literature review of the ecology of the Olifants River, Pretoria: Water Research Commission.

Hsu, LC, Huang, CY, Chuang, YH, et al. (2016) Accumulation of heavy metals and trace elements in fluvial sediments received effluents from traditional and semiconductor industries. Sci Rep 6:34250. https://doi.org/10.1038/srep34250.

Jackson, TA, Vlaar, S, Nguyen, N, et al. (2015) Effects of Bioavailable Heavy Metal Species, Arsenic, and Acid Drainage from Mine Tailings on a Microbial Community Sampled Along a Pollution Gradient in a Freshwater. Ecosystem Geomicrobiol J 32:724-750. https://doi.org/10.1080/01490451.2014.969412.

Kembel, SW, Cowan, PD, Helmus, MR, Cornwell, WK, et al. (2010) Picante: R tools for integrating phylogenies and ecology. Bioinformatics 11:1463-1464. https://doi.org/10.1093/bioinformatics/btq166.

Kuang, JL, Huang, LN, Chen, LX, et al. (2013) Contemporary environmental variation determines microbial diversity patterns in acid mine drainage. ISME J 7:1038-50. https://doi.org/10.1038/ismej.2012.139.

Langenheder, S, Berga, M, Östman, O, et al. (2012) Temporal variation of β-diversity and assembly mechanisms in a bacterial metacommunity. ISME J 6:1107-1114.

Legendre P, Legendre L (1998) Numerical Ecology, 2nd edn. Elsevier Science BV: Amsterdam.

Lima-Mendez, G, Faust, K, Henry, N, et al. (2015) Determinants of community structure in the global plankton interactome. Science 348:1262073. https://doi.org/10.1126/science.1262073

Lindstrom, ES, Langenheder, S (2012) Local and regional factors influencing bacterial community assembly. Environ Microbiol Rep 4:1-9. https://doi.org/10.1111/j.1758-2229.2011.00257.x.

Louca, S, Parfrey, LW, Doebeli, M (2016) Decoupling function and taxonomy in the global ocean microbiome. Science 353:1272-1277. https://doi.org/10.1126/science.aaf4507.

Lozupone, CA, Knight, R (2007) Global patterns in bacterial diversity. Proc Natl Acad Sci USA 104:11436-11440. https://doi.org/10.1073/pnas.0611525104.

Marr, SM, Mohlala, TD, Swemmer, A (2017) The ecological integrity of the lower Olifants River, Limpopo province, South Africa: 2009–2015–Part B: Tributaries of the Olifants River. Afr J Aquat Sci 42:171-179.

McMurdie, PJ, Holmes, S (2013) Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE 8:e61217. https://doi.org/10.1371/journal.pone.0061217.

Ofiteru, ID, Lunn, M, Curtis, TP, et al. (2010) Combined niche and neutral effects in a microbial wastewater treatment community. Proc Natl Acad Sci USA 107:15345-15350. https://doi.org/10.1073/pnas.1000604107.

Oksanen, J , Blanchet, FG, Kindt, R, et al. (2013) Vegan: Community Ecology Package.

Pei, Y, Yu, Z, Ji, J, et al. (2018) Microbial community structure and function indicate the severity of chromium contamination of the Yellow River. Front Microbiol 9:38. https://doi.org/10.3389/fmicb.2018.00038.

Price, MN, Dehal, PS, Arkin, AP (2010) FastTree 2-Approximately Maximum-Likelihood Trees for Large Alignments. Plos One 5:e9490. https://doi.org/10.1371/journal.pone.0009490.

Pruesse, E, Quast, C, Knittel, K, et al. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188-7196. https://doi.org/10.1093/nar/gkm864.

Reynolds, CS (2009) Phytoplankton population dynamics in natural environments. In: Likens GE (ed.), Encyclo- pedia of Inland Waters. Elsevier: Amsterdam, pp 197-203.

Salcher, MM, Schaefle, D, Kaspar, M et al. (2019) Evolution in action: habitat transition from sediment to the pelagial leads to genome streamlining in *Methylophilaceae*. ISME J 13:2764-2777.

Salomons, W, Stigliani, W (1995) Biogeodynamics of pollutants in soils and sediments. Springer-Verlag.

Samanovic, MI, Ding, C, Thiele, DJ, et al. (2012) Copper in microbial pathogenesis: Meddling with the metal. Cell Host Microbe 11:106-115. https://doi.org/10.1016/j.chom.2012.01.009.

Team, RDC (2011) R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing; http://www.R-project.org/.

Tuan, NN, Chang, YC, Yu, CP, et al. (2014) Multiple approaches to characterize the microbial community in a thermophilic anaerobic digester running on swine manure: a case study. Microbiol Res 169:717–724. https://doi: 10.1016/j.micres.2014.02.003.

Thompson, RM, Brose, U, Dunne, JA, et al. (2012) Food webs: Reconciling the structure and function of biodiversity. Trends Ecol Evol 27:689-697. 697. https://doi.org/10.1016/j.tree.2012.08.005.

Vörösmarty, CJ, McIntyre, PB, Gessner, MO, et al. (2010) Global threats to human water security and river biodiversity. Nature 467:555-561. https://doi.org/10.1038/nature09440.

Yergeau, E, Lawrence, JR, Sanschagrin, S, et al. (2012) Next-generation sequencing of microbial communities in the Athabasca River and its tributaries in relation to oil sands mining activities. Appl Environ Microbiol 78:7626-37. https://doi.org/10.1128/AEM.02036-12.

Tables

Table 1. Average chemistry values for sediment and water samples.

	Sediment (n=14)			Water (n=14)		
	LC (n=5)	MC (n=4)	HC (n=5)	LC (n=5)	MC (n=4)	HC (n=5)
Parameters	Mean ± SE	Mean ± SE	Mean ± SE	Mean ± SE	Mean ± SE	Mean ± SE
Al	9.2 ± 1.2**a***	19.7 ± 10.2**ab**	42.4 ± 17.8**b**	0.4 ±0.0	0.5 ± 0.1	0.4 ± 0.0
Ca	565.0 ± 242.0	3048.0 ± 1102.0	532.0 ± 153.0	37.9 ± 1.5**a**	34.4 ± 5.6**a**	73.0 ± 10.3**b**
Cu	0.45 ± 0.1**a**	29.1 ± 9.0**ab**	28.1 ± 5.1**b**	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Fe	75.1 ± 12.2	159.0 ± 40.7	333.0 ± 98.2	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0
K	21.2 ± 4.2	143.0 ± 74.5	57.3 ± 19.0	3.4 ± 0.2**a**	3.4 ± 0.6**a**	16.7 ± 5.0**b**
Mg	134.0 ± 35.8	730.0 ± 333.0	143.0 ± 33.1	29.4 ± 1.3**a**	29.4 ± 2.3**a**	88.3 ± 25.0**b**
Mn	29.6 ± 8.8	83.5 ± 25.9	39.4 ± 15.6	0.2 ± 0.0**a**	0.2 ± 0.0**ab**	0.3 ± 0.0**b**
Na	18.6 ± 2.2**a**	82.4 ± 30.1**ab**	61.8 ± 11.7**b**	2.7 ± 0.1**a**	2.7 ± 0.6**a**	12.4 ± 1.9**b**
P	1.68 ± 0.4**a**	2.4 ± 0.7**ab**	27.2 ± 10.3**b**	0.1 ± 0.0**a**	0.1± 0.0**a**	2.4 ± 0.3**b**
Pb	0.2 ± 0.0	0.5 ± 0.2	0.5 ± 0.1	0.0 ± 0.0	0.0± 0.0	0.0 ± 0.0
pH	8.1 ± 0.2	7.8 ± 0.2	7.8 ± 0.1	8.8 ± 0.0**a**	8.8 ± 0.1**a**	8.4 ± 0.0**b**

pH values represent pH units. Other values are concentrations in mg/l for water and mg/kg dry weight for sediment samples.
*****Different letters indicate significant differences (Wilcox test, $P < 0.05$) in chemistry between low, mid and high contamination within sediments and water samples (Supplementary data Table S1).
SE, standard error.
LC, low contamination; MC, mid contamination; HC, high contamination

Figures

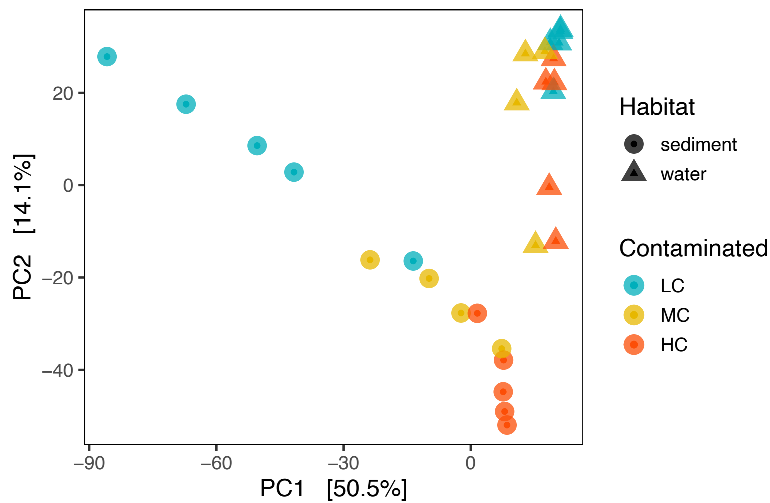

Figure 1. Principal component analysis (PCA) depicting the degree of similarity (standardized Euclidean distance) in chemical composition between the different samples. HC, high contamination; MC, mid contamination; LC, low contamination.

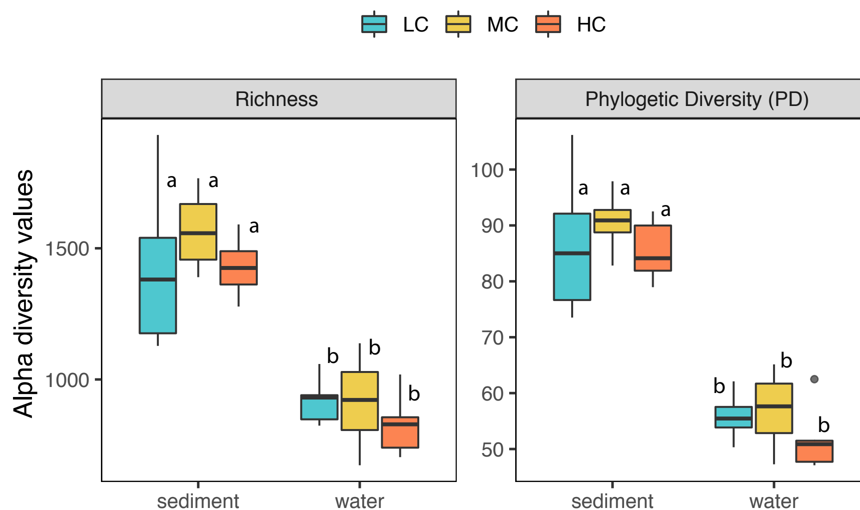

Figure 2. Alpha diversity metrics of microbial communities. Different letters next to the boxplots indicate significant differences in means (Wilcox test, $P < 0.05$). HC, high contamination; MC, mid contamination; LC, low contamination. PD, phylogenetic diversity.

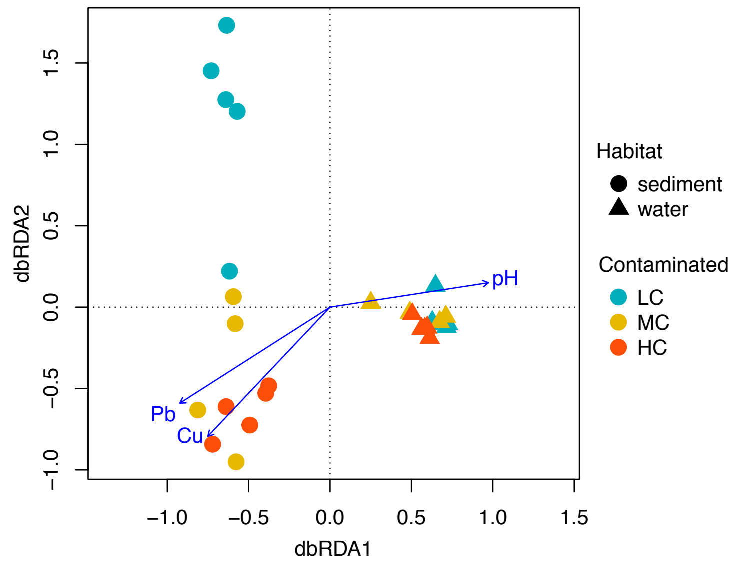

Figure 3. Distance-based redundancy analysis (dbRDA) biplot of microbial communities and microenvironmental parameters. Only the environmental variables that significantly explained variability in microbial community structure are depicted (arrows). The direction of the arrow indicates the direction of maximum change of that variable, whereas the length of the arrow is proportional to the magnitude of change. Symbols are as in figure 4.

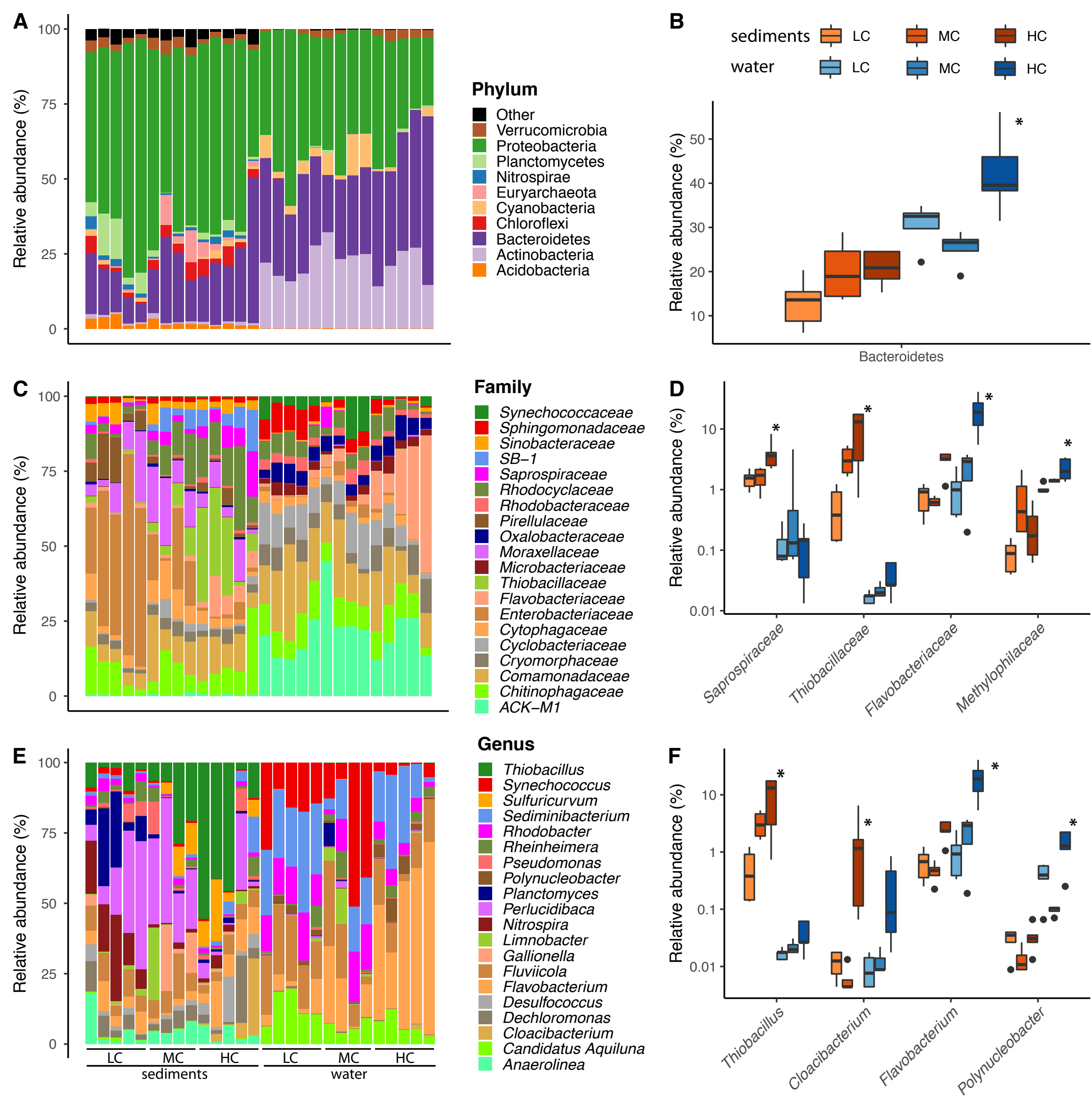

Figure 4. Comparative analysis of the prokaryotic communities. **a, c, e** Relative abundance of the most abundant microbes across sites at the phylum, family, and genus levels, respectively. **b, d, f** Relative abundance of the most abundant taxa classified as indicators for the different sub-habitats (marked with asterisk). HC, high contamination; MC, mid contamination; LC, low contamination.

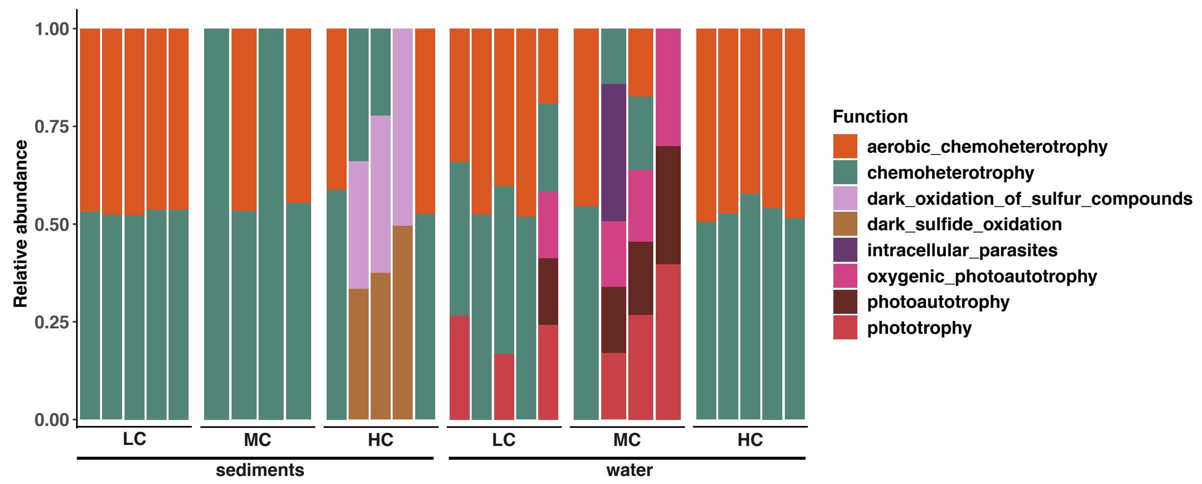

Figure 5. Functional structure (based on FAPROTAX analysis) of the prokaryotic communities. HC, high contamination; MC, mid contamination; LC, low contamination.