Importance of environmental factors over habitat connectivity in shaping bacterial communities in microbial mats and bacterioplankton in an Antarctic freshwater system

Ramoneda J.[1,2], Hawes I.[3], Pascual-García A.[4], Mackey T.J.[5,6], Sumner D.Y.[5], Jungblut A.D.[1]*

[1] Life Sciences Department, Natural History Museum, Cromwell Road, London, SW7 5BD, UK.

[2] Present affiliation: Department of Environmental Microbiology, Swiss Federal Institute of Aquatic Science and Technology (Eawag), 8600 Dübendorf, Switzerland.

[3] Coastal Marine Field Station, University of Waikato, 58 Cross Road, Tauranga 3110, New Zealand.

[4] Theoretical Biology, Institute of Integrative Biology, ETH Zürich, Universitätstrasse 16, Zürich, Switzerland.

[5] Department of Earth and Planetary Sciences, University of California–Davis, 1 Shields Avenue, Davis, CA 95618, United States of America

[6] Present affiliation: Department of Earth and Planetary Sciences, University of New Mexico, 221 Yale Boulevard NE, Albuquerque, NM 87131, United States

* Corresponding author:

Anne D. Jungblut, Life Sciences Department, Natural History Museum, Cromwell Road, 18 SW7 5BD, United Kingdom, email: a.jungblut@nhm.ac.uk; phone: +44 (0) 20 7242 5285

Keywords: biofilm, bacterioplankton, freshwater, 16S rRNA gene, McMurdo Dry Valleys, Antarctica

Abstract

Freshwater ecosystems are considered hotspots of biodiversity in Antarctic polar deserts. Anticipated warming is expected to change the hydrology of these systems due to increased meltwater and reduction of ice cover, with implications for environmental conditions and physical connectivity between habitats. Using 16S rRNA gene sequencing, we evaluated microbial mat and planktonic communities within a connected freshwater system in the McMurdo Wright Valley, Antarctica, to determine the roles of connectivity and habitat conditions in controlling microbial assemblage composition. We examined communities from glacial Lake Brownworth, the perennially ice-covered Lake Vanda, and the Onyx River, which connects the two. In Lake Vanda, we found distinct microbial assemblages occupying sub-habitats at different lake depths, while the communities from Lake Brownworth and Onyx River were structurally similar. Despite the higher physical connectivity and dispersal opportunities between bacterial communities in the shallow parts of the system, environmental abiotic conditions dominated over dispersal in driving community structure. Functional metabolic pathway predictions suggested differences in the functional gene potential between the microbial mat communities located in shallower and deeper water depth. The findings suggest that increasing temperatures and meltwater due to future climate change will affect bacterial diversity and functioning in Antarctic freshwater ecosystems.

Introduction

Environmental change poses a major threat to the maintenance of terrestrial aquatic ecosystems, particularly due to alterations in water flow and distribution (Woodward *et al.*, 2010). This is particularly acute for the cryosphere, where liquid water is primarily derived from melting snow and ice, and where winter-summer ice dynamics play a major role in driving ecosystem function. In the Antarctic Peninsula, significant increases in temperature over the last 50 years have corresponded with increases in melting of ice and liquid water supply (Convey & Peck, 2019). In the McMurdo Dry Valleys, complex processes have resulted in increases in lowland glacier melting (Doran *et al.*, 2002; Fountain *et al.*, 2014), which in turn has led to water level rise in endorheic lakes occupying the valley floors (Bomblies *et al.*, 2001; Castendyk *et al.*, 2016).

Freshwater systems in the polar deserts of continental Antarctica are particularly important hosts of inland biodiversity, the bulk of which is microbial (Chown *et al.*, 2015). In these meltwater-derived systems, primary productivity and biomass generation relies on phototrophic microbial communities, since allochthonous loadings of organic carbon are negligible from surrounding soils. To understand the implications of changing hydrology for biodiversity and ecosystem function in Antarctic aquatic ecosystems, it is important to investigate how environmental factors drive the spatial distribution and diversity of these microbial communities across habitats and depths (Chown *et al.*, 2015; Convey *et al.*, 2014).

While a large number of microbial taxa are common across polar regions (Jungblut *et al.*, 2010; Kleinteich *et al.*, 2017), on a local scale variation in the taxonomic identity of microorganisms has been ascribed to environmental selection. A number of studies have investigated how Antarctic freshwater microbial diversity is affected by environmental

drivers, in particular its responses to pH, salinity, light and water availability across a range of aquatic habitats (Comeau et al. 2012; Glatz *et al*., 2016; Hawes *et al*., 2013; Dolhi *et al*., 2015; Jungblut *et al*., 2016; Valdespino-Castillo *et al*., 2018; Sutherland *et al*., 2020). While most of these studies found links between environmental factors and microbial diversity and functionality at local scales, a significant amount of variation remains unexplained (Wilkins *et al*., 2013). For example, a study comparing samples from a range of depths in three lakes in the McMurdo Dry Valleys found high levels of similarity in benthic microbial mat community composition, with only a very low influence of irradiance and conductivity in structuring communities (Zhang *et al*., 2015). The study ascribed this to wide environmental tolerance and slow turnover of most cyanobacterial taxa. It is therefore conceivable that factors related to the dispersal of bacterial propagules between habitats also play a role in defining local bacterial assemblages.

Several specific conditions that characterize Antarctic freshwater ecosystems influence the degree of connectivity between habitats (Hawes *et al*., 1999; McKnight *et al*., 1999). Water flow is strongly seasonal, and the endorheic (closed basin), meromictic lakes (perennially stratified lakes) have water columns stabilised by salinity gradients, within which there may be little, or unidirectional (downstream) connectivity (Vincent & Laybourn-Parry, 2008). The Lake Brownworth - Onyx River - Lake Vanda (BOV) system in the McMurdo Wright Valley represents a natural laboratory to evaluate the roles of connectivity and habitat conditions on microbial community assembly across three distinct types of freshwater ecosystems in the McMurdo Dry Valleys. At its eastern end it contains a closed freshwater system formed by Lake Brownworth (an exorheic, proglacial lake), which is fed by seasonal runoff from the Lower Wright Glacier, the melt rate of which has increased in recent years (Fountain *et al*., 2014). During the austral summer, Lake Brownworth feeds water into the Onyx River, which

flows inland and discharges into Lake Vanda, an endorheic and meromictic lake stratified by a salinity gradient (Castendyk *et al*., 2016; Fig. 1A, B, and Figure S1). The BOV system thus provides a range of aquatic habitats with likely different levels of connectivity. The system is also subject to ongoing change, with flows in the Onyx River currently exceeding ablation from the lake surface, resulting in water level rise and the formation of new aquatic habitat, and subsequent colonisation by new microbial mat communities at the lake margins.

In this study, we evaluated the bacterial diversity and community structure of the BOV system by sequencing the V4 region of the 16S rRNA gene of microbial mat and planktonic bacterial communities across geographical units and water depths. We also used 16S rRNA gene sequences to predict metabolic pathway potential across habitats and lake depths. The study includes two 16S rRNA gene datasets that are treated separately throughout the study such as 1) benthic and planktonic microbial communities along the BOV transect, and 2) benthic microbial mats along a transect at water edge to 2 m depth in the moat of Lake Vanda, with samples taken every 20 cm. We hypothesized that benthic habitats with microbial mats, and planktonic habitats with bacterial communities in the water column would be compositionally distinct, but that local abiotic factors would strongly and consistently affect bacterial diversity. We also expected that the degree of connectivity between geographical units would be reflected in similarities in bacterial community composition, and therefore the potential role of dispersal in structuring local assemblages was explicitly tested. By comparing benthic and planktonic communities along the BOV system, we address the role of dispersal and environmental filtering in structuring the bacterial communities. This is relevant for assessing the effects of climate-driven hydrological changes on bacterial diversity in Antarctic freshwater ecosystems.

Materials and Methods

Study site and sample collection

The Wright Valley is located in the McMurdo Dry Valleys, in Southern Victoria Land, Antarctica (Fig. 1A, Fig. S1). Characterized by low annual average temperature (-19.8°C) and precipitation (below 100 mm year^{-1} water equivalent) (Doran *et al.*, 2002), it is a closed freshwater basin. The low point of the valley is inland; glacial meltwater and groundwater flow towards this low point and evaporation and sublimation are the only water outputs from the system (Spigel *et al.*, 1998). Meltwater from the western side of the Lower Wright Glacier (77°25′S 163°0′E) feeds Lake Brownworth (77°26'4.96"S, 162°45'48.22"E, Fig. S1A and B), which is connected to Lake Vanda (77°31′S 161°34′E, Fig. S1C and F) through the seasonal Onyx River (Fig. S1D and E), a 32 km stream flowing through a polar desert landscape. Lake Vanda is perennially covered by a 3.5-4 m ice sheet, and in 2010 it was over 74 m deep (Hawes *et al.*, 2013; Castendyk *et al.*, 2016).

The structure of Lake Vanda reflects its current and historical climatic context, and is fully described in Castendyk *et al.* (2016). In brief, the lake is understood to have undergone a series of climate-related filling and drying events (Fig. S2). The most recent high stand, some 50 m above the current level, was around 3000 years ago during a period of high meltwater production. Between 3000 and 2000 years ago, meltwater inflow declined and the lake evaporated to a shallow, hypersaline pool. Approximately 1000 years ago, renewed flow into the lake resulted in a freshwater layer from 27 to 55 m depth, mixed by double diffusion convection, overlying a salinity gradient. Additionally, approximately 100 years ago a new, discreet freshwater layer formed at the lake surface, which has since increased in thickness and now occupies the zone from 4 to 24 m depth. The lake continues to rise at approximately 0.25 m y^{-1} (Doran *et al.*, 2002), and the shallowest depths sampled represent an inundation

time series that can be dated from lake level records. The current lake structure is illustrated in Fig. S2.

A total of 29 cyanobacterial mat and 11 water samples were collected across the three main geographical units of the Wright Valley, namely Lake Brownworth, the Onyx River and Lake Vanda (collectively the BOV system, Fig. 1A and 1B, Table 1). The Lake Brownworth samples were collected in January 2012, and the samples from the Onyx river, Lake Vanda (Fig. 1B) Vanda Moat (located at the interface between the Onyx River and Lake Vanda), and the Vanda Island in December 2014. These littoral sites are frozen for much of the year and hence have short growth periods, whereby the high biomass accumulations of microbial mats accumulate slowly (Howard-Williams *et al.*, 1997). The succession of microbial communities is therefore thought to be limited. The upper cohesive layer of the moat and river mats was sampled from the shore with a sterile spatula, and the deeper mats by a SCUBA diver using a 38 mm-diameter coring device from sites between 0.1-31 m depth. Mat samples were transferred into sterile plastic containers and immediately frozen at -20°C after collection, shipped frozen and stored at -80°C at the Natural History Museum (NHM, London, UK) until further processing. Water samples were collected from the shore (shallow samples from all sites) or using a Kemmerer water sampler through a hole in the ice of Lake Vanda (8-63 m). Biomass was collected from water samples by filtering 1-2 L through sterile Sterivex-GP® filter units (Merck Millipore, Darmstadt, Germany) with 0.22 µm pore size. Biomass was stored in lysis buffer at -20°C immediately after processing, shipped frozen and stored at -80°C at the NHM until further processing. The environmental variables temperature (°C), pH and conductivity (µS/cm) were measured *in situ* at the site of each benthic and water sample replicate per location using a portable instrument (pH/Con 10 Series; Oakton Instruments,

Vernon Hills, IL). The measurements were taken at the same and single time as the sampling of the benthic mat and water samples.

Additionally, in January 2017 microbial mats that were collected in 20 cm intervals along the 2 m transect in the moat of Lake Vanda, close to the entry of the Onyx River, were also evaluated by 16S rRNA gene sequencing, with the aim of obtaining more depth resolution in cyanobacterial mat community structure (Table 1). At each site, a depth transect was obtained, sampling at least 3 mat samples at depths of 0.1, 0.5, 0.8, 1.1, 1.5 and 1.9 m using a 25 mm diameter rod-mounted piston corer. These depths correspond to inundation dates of 2017, 2016, 2014, 2013, 2012, 2011, 2010 based on lake level data in Castendyk *et al.* (2016) and the McMurdo LTER data repository (http://mcm.lternet.edu/). The corer was cleaned with 70% ethanol between samplings. Cores were placed in 15 ml Falcon tubes and flash frozen in liquid nitrogen. Thereafter they were stored at -20°C until analysed four months after collection.

DNA extraction, Polymerase Chain Reaction (PCR) and Illumina sequencing
DNA was extracted from 0.3-0.5 g of mat material using a MoBio PowerBiofilm DNA Isolation kit (Carlsbad, CA) following the manufacture's protocol, and quantified using NanoDrop ND8000 (Labtech International, UK). The DNA was extracted from the sterivex filters using a salt (NaCl)-based method modified from Aljanabi and Martinez (1997) with lysozyme and proteinase K (Diez *et al.*, 2001) for planktonic communities from polar freshwater ecocystems as detailed by Charvet *et al.* (2012; 2014). The V4 variable region of the 16S rRNA gene was PCR-amplified using 8.84 μl of PCR grade water, 5 μl of 5x GoTaq Flexi buffer, 2 μl of 25 μM magnesium chloride ($MgCl_2$), 0.8 μl of 20 mg/ml Bovine Serum

Albumin (BSA), 0.16 μl of 200 μM dNTPs (deoxynucleoside triphosphate) and 0.2 μl of 5 μg/μl GoTaq polymerase. The reaction mix was completed with 1 μl of the 515F forward primer (10 μM) and 1 μl of the 806R barcoded reverse primer (Caporaso *et al*., 2010; Caporaso *et al*., 2011) (10 μM). For the samples collected in 2017, the same procedures were followed, except that the amplification was made with primers 341F and 805R as the work was carried out in a different laboratory and therefore the two dataset were processed in Qiime separately and not directly compared in any of the statistical analyses (Caporaso *et al*., 2010). A volume of 1 μl of template DNA completed the 20 μl reaction. PCR conditions involved an initial denaturation of 2 minutes at 94°C, followed by 45 seconds at 94°C, 1 minutes at 50°C of annealing temperature and 90 seconds at 72°C during 35 cycles. A final elongation of 72°C for 10 minutes followed and the products were kept at 10°C until removal.

The PCR products were visualised using electrophoresis in a 1% agarose gel. The amplicons were purified following an AxyPrep Mag PCR clean-up magnetic protocol (Axygen, New York, NY). The purified PCR amplicons were quantified using a Qubit 2.0 fluorometer (Life Technologies, Glasgow, UK). The multiplexed pooled amplicons were sequenced in an Illumina MiSeq platform at the Natural History Museum (NHM) sequencing facility.

16S rDNA gene sequence processing

The results from moat samples collected in January 2017 were processed separately but following the same pipelines from those taken in 2012-2013. The majority of the sequence processing was conducted in QIIME (Caporaso *et al*., 2010; Caporaso *et al*., 2011). For the 16S rRNA gene sequences, FastQC was used to verify the expected quality and read length of the initially demultiplexed 16S rRNA sequences. Primers and Illumina adaptors were trimmed using Cutadapt v1.14. The paired ends were stitched together using Flash v1.2.11, setting a

minimum and maximum read overlap of 15 bp and 300 bp respectively, and a maximum

mismatch density of 0.15. Sequences were quality and size filtered using Prinseq v0.20.4,

with a maximum phred quality threshold of 20 and a size range of 250-300bp. The functions

UPARSE and UTAX from USEARCH v9.2were used for OTU clustering, *de novo* chimera

removal and taxonomic assignment using Silva (16S_128) as a reference database, setting a

97% similarity threshold and 80% sequence coverage. Chloroplast and mitochondrial

sequences were excluded from the data using the *phyloseq* package v1.19.1 in R v3.5.2

Oksanen *et al.*, 2019), as the primers were not designed for good coverage of plastid 16S

rRNA genes as well as the 16S rRNA gene provides limited taxonomic resolution for

eukaryotes. Sample coverage ranged between 96% and 99%, so diversity analyses were

conducted on un-rarefied data. A total of 4,900,431 16S rRNA gene sequences were finally

processed for the 2012-2013 datasets, ranging from 60,676 to 197,711 sequences per sample,

with an average ($\pm$ SD) of 122,511 ($\pm$25,529) sequences. From the 2017 dataset, an average ($\pm$

SD) of 7,472 ($\pm$168) sequences per sample were processed, totalling 194,283 sequences.

Before analysis of the diversity and community structure of these datasets, singletons were

removed and only for beta diversity analyses sequence reads were relative abundance-

transformed.   The sequences were submitted to Genbank SRA (Bioproject ID:

PRJNA638378).

Analysis of the diversity of bacterial communities

The description of the composition, diversity and community structure of the BOV system

bacterial communities was performed in R (Team R Core, 2018), with the packages *phyloseq*

v1.19.1 (McMurdie *et al.*, 2013) and *vegan* v2.5-5 (Oksanen *et al.*, 2019). We analysed the

whole communities and also the diversity and composition of the cyanobacterial subset

separately, given the latter's importance for the productivity of the system. To estimate the

alpha diversity, we considered the number of OTUs in each community (Observed species index), and the Chao1 and Simpson's (D) diversity indices. Beta-diversity was estimated by computing the Bray-Curtis dissimilarity, and ordination was performed by computing Non-Metric Multidimensional Scaling (NMDS) and Constrained Correspondence Analysis (CCA). The ANOSIM test with 999 permutations was used to test the similarity between bacterial communities from different geographic units and habitats in the BOV system.

Identification of environmental drivers of bacterial community structure

Relationships between measured environmental parameters (i.e. temperature, conductivity, and pH) and bacterial community structure for the BOV system were statistically assessed using an ANOVA-like permutation test on the CCA model, with 999 permutations. Using the *permutest.cca* function in *vegan v2.5-5* (Oksanen *et al.*, 2019), the test computes the significance of the constraints on the ordination. Geographical distances were included in the CCA analysis as Principal Coordinates of Neighbour Matrices (PCNM)-transformed distances. Both habitats and geographical locations were included as categorical variables. To obtain the variance explained by each of the aforementioned variables, linear regressions on the CCA ordinations were tested with the function *envfit* of *vegan*.

In order to obtain mechanistic insights into the processes of assembly in the bacterial communities in the BOV system, we calculated the Mean Nearest Taxon Distance (MNTD) of each sample, and computed a z-score by comparing to a null distribution, the so-called β-nearest taxon index (βNTI) following Stegen et al., (2013), and using the package *picante (v1.8.1)* (Kembel *et al.,* 2019). We identified the presence of phylogenetic signal in the communities by means of a Mantel correlogram for correlation between the Unifrac distance matrix of the samples and the Euclidean distance matrices of the measured environmental

factors. The MNTD is the mean phylogenetic distance of each taxon to its closest relative of a given sample, and the comparison to a random draw of taxa from the same community gives a measure of phylogenetic dispersion. βNTI values higher than +2 indicate significantly more than expected phylogenetic turnover and is interpreted as predominantly variable selection processes in driving community composition, while βNTI values below -2 indicate significantly less than expected turnover and predominance of homogenizing selection (Stegen *et al.*, 2012, 2013). Values falling between -2 and +2 are indicative of ecological drift and/or dispersal. The input phylogenetic tree for the analysis was built from a multiple sequence alignment using MUSCLE (Edgar, 2004). In order to understand the potential role of dispersal between bacterial communities in the BOV system, a combination of the ANOSIM test and calculation of βNTI values in different geographical units were performed.

Prediction of functional potential using PiCRUST

We predicted functional gene potential from the 16S rRNA sequences using PiCRUST (Langille *et al.*, 2013) for the BOV system. The method uses reference genomes that are biased towards human-gut taxa, and hence predictions should be taken with caution in other environments. We addressed the quality of the prediction by computing the Nearest Sequenced Taxon Index (NSTI), which indicates the mean similarity of the relatives used in the prediction of a given community (e.g. a NSTI of 0.05 indicates a 95% similarity on average). In Langille *et al.* (2013) the authors showed that the quality of the prediction decreases with increasing NSTI values depending on the environment. In particular, for soil environments (including 6 Antarctic samples taken from Fierer *et al.* (2012), they showed that the accuracy of the prediction is high (Spearman correlation coefficients with respect to metagenome sequencing around 0.8) for NSTI values as high as 0.20. Indeed, previous predictions in mat samples were considered high quality for mean values of 0.11 (Koo *et al.,*

2017). Our mean NSTI values were 0.18, which lie within the high-quality boundaries for mat samples, and we also verified that the NSTI values were uncorrelated to depth (Fig. S3). Nevertheless, special care was taken in the interpretation of water samples, especially for those in the deepest part of Lake Vanda exposed to higher salinity, which is known to decrease the accuracy of the prediction.

Predictions were first analysed by reducing the dimensionality using Principal Component Analysis (PCA) with STAMP v2.1.3 (Parks *et al.*, 2014), to investigate which environmental variables were most closely associated with the clustering of the communities. We then merged the genes into genetic pathways within the highest resolution level in the KEGG (Kyoto Encyclopaedia of Genes and Genomes) classification (Kanehisa & Goto, 2000), and tested which pathways showed significant differences across the identified environmental variables. To address this question, we tested if the mean proportions of genes belonging to a given pathway were significantly different between communities belonging to two different conditions (e.g. water vs. mat), performing Games-Howell tests. We considered significant those tests with Bonferroni-corrected P-values below 0.05 and effect sizes larger than 0.4.

Results

16S rRNA gene composition of microbial mats and water habitats

For the samples across the BOV system, 6347 different bacterial OTUs were found across all microbial mat and water habitats investigated in the BOV system, of which 170 OTUs belonged the phylum Cyanobacteria (Table 2). In microbial mats, the highest bacterial richness was found in Lake Vanda (average Observed Species index ± SE = 1782 ± 136 OTUs), followed by Lake Brownworth (1591 ± 14 OTUs), and Onyx river (1525 ± 313 OTUs) (Table 2), while the highest cyanobacterial richness occurred in Lake Brownworth (65

± 3 OTUs). In the water habitat Onyx river hosted on average five times more bacterial OTUs than Lake Vanda's water column (Onyx river: 3098 ± 106 OTUs; Lake Vanda: 626 ± 302 OTUs), while the sample from Lake Brownworth had richness values closer to Onyx river (2112 OTUs) (Table 2). The same pattern was observed for Cyanobacteria.

At the phylum level, there were compositional differences between the microbial mat and water habitats, particularly in Lake Vanda, as well as a difference by depth. Dominance of Cyanobacteria and Proteobacteria in microbial mats contrasted with high relative abundance of Actinobacteria in water (Fig. 1C). Within the mat habitat, cyanobacterial abundance dropped from an average 29.1 % in shallow water samples less than 1 m depth (i.e. Brownworth-Onyx and the shores of Lake Vanda), to 11.8 % in the samples from Lake Vanda below ice cover (11-31m deep). This change in relative abundance did not reflect a clear compositional change at the cyanobacterial genus level, which was dominated by the filamentous oscillatorians *Phormidium, Tychonema* and *Leptolyngbya* (Fig. 1D). Within the water column, location and depth also played a role on bacterial taxonomic composition. At the phylum level, communities shifted from Bacteroidetes- and Planctomycetes-dominated assemblages in Lake Brownworth and Onyx river, towards Actinobacteria-dominated communities in samples under ice cover in Lake Vanda (Fig. 1C). A strong shift in cyanobacterial community composition was also observed, as water samples under ice cover were dominated by unicellular *Chamaesiphon*, while *Leptolyngbya* dominated shallow water samples (Fig. 1D).

Comparison of bacterial community structure and environmental drivers

Microbial mat and water habitats contained distinct bacterial assemblages (Fig. 2A). The community structure of microbial mat samples from shallow locations (i.e. Lake Brownworth,

Onyx River, Vanda Island and Vanda Moat) clustered together, and all were more similar to their corresponding water samples, than was the case in the deeper water column samples in Lake Vanda (Fig. 2A). The CCA showed that habitat type (i.e. mat vs. water), and geographical location and distance were important factors determining bacterial community structure (Fig. 2B; Table 3). The constrained ordination with pH, conductivity and temperature values for the sites (Table S1), explained 55.7% of variation in bacterial community structure in total (24.8% in the first 2 axes), and identified temperature and conductivity as the most important factors of bacterial community variation, with no influence of pH (Fig. 2B; Table 3). Within Lake Vanda, the importance of depth, a covariate of the environmental factors, was shown by correlating NMDS1 and depth for microbial mat (Fig. 2C) and water (Fig. 2D) bacterial communities. This analysis revealed both mat and water communities gradually shifted in structure with increasing depth.

Comparison of bacterial community structure in microbial mats along 2 m transect

The microbial mat assemblages were dominated by Cyanobacteria, Bacteroidetes, and Proteobacteria (Fig. S4), and a shift in bacterial community structure was observed at OTU level along the gradient by NMDS analysis within the first 2 m depth (Fig. S4) with the shallowest samples being most distinct (Fig. S5A) as also observed in the alpha diversity (Chao1 and Simpson index, Fig. S5B). The importance of depth, a covariate of the environmental factors, was shown by correlating NMDS1 and depth for the microbial mats along the 2 m transect (Fig. 3).

Potential role of dispersal between geographical units and habitats of the BOV system

Microbial mat and water samples were classified into units by their physical connectivity (i.e. Lake Brownworth-Onyx river, Vanda Moat-Vanda Island, Lake Vanda at 8-15 m, Lake

Vanda at 19-23 m, and Lake Vanda > 27 m). The ANOSIM test revealed bacterial community structure was distinct at each of the habitat-geographical unit combinations (Fig. 4A), with the exception of the comparison between water communities of Lake Brownworth-Onyx river unit (n = 3), and Lake Vanda below 27 m deep (i.e. "Lake Vanda deep", n = 3) (R = 1, P = 0.100), which may be attributed to the smaller sample size for these sites. The MNTD randomization analysis supported the results of the ANOSIM, as the βNTI values calculated in the different geographical units were equal to or lower than -2 (Fig. 4B), indicating a strong role for homogeneous selection (Stegen *et al.*, 2013), and little influence of dispersal on the structure of the bacterial communities studied. The most negative βNTI values were calculated for the mat communities, while water communities had higher βNTI values (Fig. 4B), indicating higher potential for dispersal in water. Note the significant phylogenetic signal detected in response to temperature, conductivity and pH (Fig. S6).

Profiling of metabolic potential

Predictions of metabolic pathways in bacterial communities can point at potential bacterial adaptations to the environmental conditions across the BOV system, and reveal potentially unrecorded functions in natural communities. The gene inference analysis suggested the presence of distinct metabolic functional profiles between microbial mat and water bacterial communities, (Fig. 5A), whereas the microbial mat and water environments were distinguishable by only a few metabolic functions (Fig. S7). Instead, predicted metabolic functional profiles for microbial mat communities gradually changed with increasing depth within Lake Vanda (Fig. 5B). The mats in shallower zones were enriched in potential functions associated with glutathione and tyrosine metabolism, as well as degradation of halogenated aromatic hydrocarbon compounds such as xylene (Fig. 5C-D; Fig. S8). The predicted metabolism of such hydrocarbon compounds was mostly associated to

Sphingomonadaceae and Comamonadaceae (alpha- and beta-Proteobacteria), which were predominant contributors in the shallow mats of Lake Vanda (Fig. 5D; Fig. S8). Xylene degradation was particularly influenced by the abundance of Erythrobacteraceae (alpha-Proteobacteria), whereas the predicted degradation of halogenated hydrocarbon compounds was mostly influenced by the cyanobacterial family Pseudanabaenaceae. Deeper mat communities had a higher proportion of predicted functions related to amino acid related enzymes and aminoacyl-rRNA, valine biosynthesis as well as leucine isoleucine biosynthesis (Fig. S9).

Discussion

In this study, we compared the 16S rRNA gene community structure and evaluated the predicted metabolic potential of microbial mat and water bacterial assemblages across depths in an Antarctic freshwater system. Our results revealed a highly compartmentalized system in which high diversity is achieved through the strong environmental selection of taxa in different habitats. Although some of these habitats are physically connected, our study suggests that bacterial dispersal between them had a limited role in structuring bacterial communities.

The composition and relative abundance of dominant bacterial and cyanobacterial taxa were distinct between the microbial mat and water habitats in the BOV system (McMurdo Wright Valley, Antarctica), which agrees with findings from Lakes Bonney Fryxell, Hoare and Miers (Kwon *et al*., 2017, Dillon *et al.,* 2019) suggesting differences between the bacterial benthos and plankton are widespread in Antarctic Dry Valley lakes. The cyanobacterial composition of the mats in the system was also comparable to previous studies on microbial mats in Southern Victoria land (Zhang *et al*., 2015; Jungblut *et al*., 2012). Interestingly, there was a

dominance of unicellular *Chamaesiphon* cyanobacteria in the deeper layers of Lake Vanda. Dominant filamentous oscillatorian morphotypes had been recorded in Lake Vanda's bottom of the euphotic zone (Vincent & Vincent; 1982), but not unicellular types. In Antarctica, *Chamaesiphon* had been previously reported in microbial mats and sediment from glacial cryoconite holes in the Southern Victoria land and the McMurdo Dry Valleys (Zhang *et al*., 2015; Martineau *et al*. 2013; Webster-Brown *et al*., 2015).

Microbial mats and bacterioplankton in the shallower parts of the system displayed higher structural similarity to each other than to sites under the perennial ice cover of Lake Vanda, suggesting common responses to similar environmental conditions, for example freezing during winter, and a possible degree of dispersal among them. However, the higher cyanobacterial and bacterial diversity detected in Onyx River suggests the river is important in propagule dispersal downstream, since taxa from the surrounding soils can be mobilized, and because Onyx river is the physical connector between Lakes Brownworth and Vanda. The dominance of filamentous cyanobacterial *Leptolyngbya* in the water samples of Onyx river, most resembling those found in mat communities, could have originated from the mats in the Onyx river, and might have been suspended through disturbance by freeze-thawing and water currents during the rewetting of the dried mat in the river bed (Cullis *et al*., 2014). Despite this role as the most likely dispersal pathway in the system, communities from Onyx river were markedly different from those under perennial ice cover in Lake Vanda, suggesting that despite bacterial propagule transport might be commonplace from the shallow parts of the BOV system in the Onyx river, propagules do not successfully establish in the locations they are transported to.

The cyanobacterial mat communities in Lake Vanda moat along the 2 m long transect further supports the idea that bacterial communities rapidly shift towards a common lake underwater assemblage. Considering the timespan each depth had been inundated, the convergence of sample composition at 1 m suggests that assemblages from the sampling locations on the shore of Lake Vanda achieved a common moat configuration within 3-4 years (Convey & Peck, 2019). Divergence of the most recently inundated communities from those below 1 m may reflect the early influence of terrestrial microbes pre-existing in soils, prior to colonists arriving from inflowing water and from deeper moat mats via mixing of moat water.

The possibility of dispersal as an effective driver of bacterial diversity in the BOV system was discarded by the finding that selective environmental filtering was homogenising bacterial community structure in each habitat, as it has been reported in previous studies in vertically stratified lakes (Comeau *et al*., 2012; Logares *et al*., 2013). The observed gradual divergence of microbial mat and water communities with depth further supports the conjecture that deeper environments in Lake Vanda select for increasingly distinct bacterial assemblages (Kwon *et al.,* 2017). Previous studies evaluated the role of wind as a mechanisms of dispersal of biomass in the McMurdo Dry Valleys, and it was shown that strong local winds with prevailing down valley orientation contribute to dispersal of organic matter and differences in cyanobacteria and diatom communities between the McMurdo Dry Valleys (Šabacká *et al*., 2010; Michaud *et al*., 2012; Sakaeva *et al*. 2016). At local scale, Michaud *et al*. (2012) suggested that prevailing winds play a role in cyanobacteria richness whereas diatom assemblages were found to be more driven by environmental conditions in the different habitats (Sakaeva *et al*. 2016).

Environmental filtering of bacterial assemblages could also be reflected in the metabolic functional potential in the specific habitats under study. Prediction of metabolic functions with PiCRUST (Langille *et al.*, 2013) showed different potential metabolic pathways across habitats and depths. The biggest changes in the predicted metabolic functionality for microbial mats happened with increasing depth in Lake Vanda. Predicted pathways included tyrosine and glutathione metabolism as well as anaerobic degradation of xenobiotic halogenated aromatic hydrocarbon compounds, which characterized shallow mats. While anaerobic oxidation of acetate for sulphur reduction is characteristic of the light and oxygen-depleted lower microbial mat layers (Dillon *et al.*, 2017; Mountfort *et al.*, 2003), these particular pathways had never been predicted in living Antarctic microbial mats. A recent metagenomic survey on microbial paleomats buried in soils around Lake Vanda identified proportions below 1% of xenobiotics metabolism genes (Zaikova *et al.*, 2019). Since these functions are more commonly found in soil-borne bacteria (Oliveira *et al.*, 2017), a potential explanation is that shallow microbial mats may be composed of a number of soil-borne taxa recruited from the sediment and surrounding soil environment, which gradually decrease with lake depth. The bacterial families contributing most to xenobiotic degradation in the system were Sphingomonadaceae and Pseudanabaenaceae, which contain members commonly found in soils (Alwathnani & Johansen, 2011; Glaeser & Kämpfer, 2014), and some are of interest to bioremediation (Kertesz & Kawasaki, 2010). An alternative is that these functions are remnants of microbial activity happening after contamination by oil burning and chronic spills around the lake in the 1970s and 1990s (Webster *et al.*, 2003). Hydrocarbon patches remain in soils close to previous human habitations around the shores of the lake, at up to 9 g kg^{-1} soil (total petroleum hydrocarbon), and it is likely that similarly contaminated soils have been inundated during lake level rise (Taylor, 2015). Future work using metagenomics and transcriptomics are however needed to confirm the presence and active transcription of

functional gene diversity in the microbial mats and bacterioplankton across the BOV systems in the McMurdo Dry Valleys.

Conclusions

Environmental change is expected to lead to a redistribution of water and its availability in the McMurdo Dry Valleys (Fountain *et al*., 2014), which will impact the connectivity between microbial habitats. This study shows that the Wright Valley contains distinct benthic and planktonic bacterial communities, which are primarily structured through environmental filtering. While superficial water flow seems to play a role in dispersing bacterial taxa, strong environmental filtering is a dominant process and leads to a vertical structuring of bacterial communities. Likewise, we show that within 3-4 years of inundation, bacterial communities achieve a common configuration in the shallow parts of Lake Vanda. These observations suggest that future impacts of hydrological changes on the bacterial communities will likely be driven firstly by changes in the physicochemical water properties, followed by an increasing role of connectivity as system stratification disappears. This implies that a homogenization of the water conditions could compromise the persistence of bacterial communities not adapted to those conditions. In Antarctica, understanding the effects of future environmental change on freshwater microbial diversity requires a detailed knowledge of how hydrological changes disrupt the diverse suite of environmental conditions that sustain bacterial diversity in the present.

Acknowledgements

by a British Phycological Society Project grant. APG was funded by the Simons Collaboration: Principles of Microbial Ecosystems (PriME), award number 542381. The field work and logistics were supported by Antarctica New Zealand. We thank Hannah Christenson and Devin Castendyk for their help with the field work. We also thank the two Reviewers for their insightful comments and suggestions

Compliance with ethical standards

Conflict of interest: The authors declare that they have no conflict of interest.

References

Aljanabi SM, and Martinez I (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res* 25:4692-4693.

Alwathnani H, Johansen JR (2011). Cyanobacteria in soils from a Mojave Desert ecosystem. *Monographs of the Western North American Naturalist 5*: 71-89.

Bomblies A, McKnight DM, Andrews ED (2001) Retrospective simulation of lake-level rise in Lake Bonney based on recent 21-year record: indication of recent climate change in the McMurdo Dry Valleys, Antarctica. *J Paleolimnol* 25: 477-492.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 2010 7: 335.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ *et al*. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* 108: 4516-4522.

Castendyk DN, Obryk MK, Leidman SZ, Gooseff M, Hawes I (2016) Lake Vanda: A sentinel for climate change in the McMurdo Sound Region of Antarctica. *Glob Planet Change 144*: 213-227.

Charvet S, Vincent WF, Lovejoy C (2012) Chrysophytes and other protists in high Arctic lakes: molecular gene surveys, pigment signatures and microscopy. *Polar Biol* 35: 733-748.

Charvet S, Vincent WF, Lovejoy (2014) Effects of light and prey availability on Arctic freshwater protist communities examined by high-throughput DNA and RNA sequencing, *FEMS Microbiol Ecol* 88: 550–564

Chown S L, Clarke A, Fraser CI, Cary SC, Moon KL, McGeoch MA (2015) The changing form of Antarctic biodiversity. *Nature 522*: 431-438.

Comeau AM, Harding T, Galand PE, Vincent WF, Lovejoy C (2012) Vertical distribution of microbial communities in a perennially stratified Arctic lake with saline, anoxic bottom waters. *Sci Rep* 2: 604.

Convey P, Peck LS (2019) Antarctic environmental change and biological responses. *Sci Adv* 5: eaaz0888.

Convey P, Chown SL, Clark A, Barnes DK, Bokhorst S, Cummings V et al. (2014) The spatial structure of Antarctic biodiversity. *Ecoll Monographs 84*:203-244.

Cullis JD, Stanish LF, McKnight DM (2014) Diel flow pulses drive particulate organic matter transport from microbial mats in a glacial meltwater stream in the McMurdo Dry Valleys. *Water Resour Res* 50: 86-97.

Diez B., Pedròs-Aliò C, Massana R (2001) Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Appl Environ Microbiol* 67: 2932-2941.

Dillon ML, Hawes I, Jungblut, AD, Mackey TJ, Eisen JA, Doran PT, Sumner DY (2019) Energetic and environmental constraints on the community structure of benthic microbial mats in Lake Fryxell, Antarctica. *FEMS Microbiol Ecol* 96: fiz207.

Dolhi JM, Teufel AG, Kong W, Morgan-Kiss RM (2015) Diversity and spatial distribution of autotrophic communities within and between ice-covered Antarctic lakes (McMurdo Dry Valleys). *Limn Oceano* 60: 977-991.

Doran PT, McKay CP, Clow GD, Dana GL, Fountain AG, Nylen T, Lyons WB (2002) Valley floor climate observations from the McMurdo Dry Valleys, Antarctica, 1986–2000. *J Geophys Res Atmos* 107: ACL-13.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792-1797.

Fierer N, Leff JW, Adams BJ, Nielsen UN, Bates ST, Lauber CL et al. (2012) Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc Natl Acad Sci* 109: 21390-21395.

Fountain AG, Levy JS, Gooseff MN, Van Horn D (2014) The McMurdo Dry Valleys: a landscape on the threshold of change. *Geomorph* 225: 25-35.

Glatz RE, Lepp PW, Ward BB, Francis CA (2016) Planktonic microbial community composition across steep physical/chemical gradients in permanently ice-covered Lake Bonney, Antarctica. *Geobiol* 4: 53-67.

Glaeser SP, Kämpfer P (2014) The Family *Sphingomonadaceae*. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F. (eds) The Prokaryotes. Springer, Berlin, Heidelberg.

Hawes I, Sumner D, Andersen D, Jungblut AD, Mackey JT (2013) Timescales of growth response of microbial mats to environmental change in an ice-covered Antarctic lake. *Biol* 2: 151-176.

Hawes, I., Smith, R., Howard-Williams, C., & Schwarz, A. M (1999) Environmental conditions during freezing, and response of microbial mats in ponds of the McMurdo Ice Shelf, Antarctica. *Antarc Sci* 11: 198-208.

Jungblut AD, Lovejoy C, Vincent WF (2010) Global distribution of cyanobacterial ecotypes in the cold biosphere. *ISME J* 4: 191-202.

Jungblut AD, Wood SA, Hawes I, Webster-Brown J, Harris C (2012) The Pyramid Trough Wetland: environmental and biological diversity in a newly created Antarctic protected area. *FEMS Microbiol Ecol* 82: 356-366.

Jungblut AD, Hawes I, Mackey TJ, Krusor M, Doran PT, Sumner DY et al. (2016) Microbial mat communities along an oxygen gradient in a perennially ice-covered Antarctic lake. *Appl Environ Microbiol* 82: 620-630.

Kembel S, Ackerly DD, Blomberg SP, Cronwell WK, Cowan PD, Helmus MR, Morlon H, Webb CO (2019) Package 'picante- Integrating Phylogenies and Ecology.

Kertesz MA, Kawasaki A (2010) Hydrocarbon-Degrading *Sphingomonads: Sphingomonas, Sphingobium, Novosphingobium*, and *Sphingopyxis*. In: Timmis KN (eds) Handbook of Hydrocarbon and Lipid Microbiology. Springer, Berlin, Heidelberg.

Kleinteich J, Hildebrand F, Bahram M, Voigt AY, Wood SA, Jungblut AD et al. (2017) Pole-to-pole connections: similarities between Arctic and Antarctic microbiomes and their vulnerability to environmental change. *Front Ecol Evol* 5: 137.

Koo H, Mojib N, Hakim JA, Hawes I, Tanabe Y, Andersen DT et al. (2017) Microbial communities and their predicted metabolic functions in growth laminae of a unique large conical mat from Lake Untersee, East Antarctica. *Front Microbiol* 8: 1347.

Kwon M, Kim M, Takacs-Vesbach C, Lee J, Hong SG, Kim SJ et al. (2017) Niche specialization of bacteria in permanently ice-covered lakes of the McMurdo Dry Valleys, Antarctica. *Environ Microbiol* 19: 2258-2271.

Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA et al. (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotech* 31: 814.

Logares R, Lindström ES, Langenheder S, Logue JB, Paterson H, Laybourn-Parry J et al. (2013) Biogeography of bacterial communities exposed to progressive long-term environmental change. *ISME J* 7: 937-948.

Martineau E, Wood SA, Miller MR, Jungblut AD, Hawes I, Webster-Brown J et al. (2013) Characterisation of Antarctic cyanobacteria and comparison with New Zealand strains. *Hydrobiol* 711: 139-154.

McKnight DM, Niyogi DK, Alger AS, Bomblies A, Conovitz PA, Tate CM (1999) Dry valley streams in Antarctica: ecosystems waiting for water. *Biosci* 49: 985-995.

McMurdie PJ, Holmes S (2013) Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* 8.

Michaud AB, Šabacká M, Priscu JC (2012) Cyanobacterial diversity across landscape units in a polar desert: Taylor Valley, Antarctica. *FEMS Microbiol Ecol* 82: 268-78.

Mountfort D, Kaspar H, Asher RA, Sutherland D (2003) Influence of pond geochemistry, temperature and freeze–thaw on terminal anaerobic processes occurring in sediments of six ponds of the McMurdo Ice Shelf, near Bratina Island, Antarctica.; *Appl Environ Microbiol* 69: 583-592

Obryk MK, Doran PT, Priscu JC (2019) Prediction of ice-free conditions for a perennially ice-covered Antarctic lake. *JGR Earth Surface* 124: 686-694.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB et al. (2019) Package 'vegan'-Community Ecology Package.

Oliveira JS, Araújo WJ, Figueiredo RM, Silva-Portela RC, de Brito Guerra A, da Silva Araújo SC et al. (2017) Biogeographical distribution analysis of hydrocarbon degrading and biosurfactant producing genes suggests that near-equatorial biomes have higher abundance of genes with potential for bioremediation. *BMC Microbiol* 17: 168.

Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 30: 3123-3124.

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28: 27-30.

Šabacká M, Priscu JC, Basagic HJ, Fountain AG, Wall DH, Virginia RA, Greenwood MC (2010 Aeolian flux of biotic and abiotic material in Taylor Valley, Antarctica. *Geomorph* 155-156:102-111.

Sakaeva A, Sokol ER, Kohler TJ, Stanish LF, Spaulding SA, Howkins A, et al. (2016) Evidence for dispersal and habitat controls on pond diatom communities from the McMurdo Sound Region of Antarctica. *Polar Biol* 39: 2441-2456

Spigel RH, Priscu JC (1998) Physical Limnology of the McMurdo Dry Valleys Lakes, In: Priscu JC (ed.), Ecosystem Dynamics in a Polar Desert: The McMurdo Dry Valleys, Antarctica, American Geophysical Union, Washington DC, 153–187.

Stegen JC, Lin X, Konopka AE, Fredrickson JK (2012) Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J* 6: 1653–1664.

Stegen JC, Lin X, Fredrickson JK, Chen X, Kennedy DW, Murray CJ, Rockhold ML, Konopka AE (2013) Quantifying community assembly processes and identifying features that impose them. *ISME J* 7(11): 2069–2079.

Sutherland D, Howard-Williams C, Hawes I (2020) Environmental drivers that influence microalgal species in meltwater pools on the McMurdo Ice Shelf, Antarctica. *Polar Biol* 32: 1-16.

Taylor, P (2015) Residual contamination and environmental effects at the former Vanda Station, Wright Valley, Antarctica. MSc Thesis, University of Canterbury, New Zealand. http://dx.doi.org/10.26021/9184.

Team R Core. R: A Language and Environment for Statistical Computing. Vienna, Austria. https://www.r-project.org/. 2018.

Valdespino-Castillo PM, Cerqueda-García D, Espinosa AC, Batista S, Merino-Ibarra M, Taş N et al. (2018) Microbial distribution and turnover in Antarctic microbial mats highlight the relevance of heterotrophic bacteria in low-nutrient environments. FMES Microbiol Ecol 94: fiy129.

Vincent WF, Laybourn-Parry J (eds.) (2008) Polar lakes and rivers: limnology of Arctic and Antarctic aquatic ecosystems. Oxford University Press, p. 1-320.

Vincent WF, Vincent CL (1982) Factors controlling phytoplankton production in Lake Vanda (77 S). *Ca J Fish Aquat Sci* 39: 1602-1609.

Webster-Brown JG, Hawes I, Jungblut AD, Wood SA, Christenson HK (2015) The effects of entombment on water chemistry and bacterial assemblages in closed cryoconite holes on Antarctic glaciers. *FEMS Microbi Ecol* 91: fiv144.

Webster J, Webster K, Nelson P, Waterhouse E (2003) The behaviour of residual contaminants at a former station site, Antarctica. *Environ Pollut* 123: 163-179.

Wilkins D, Yau S, Williams TJ, Allen MA, Brown MV, DeMaere MZ, et al. (2013) Key microbial drivers in Antarctic aquatic environments. *FEMS Microbiol Rev* 37: 303-335.

Woodward G, Perkins DM, Brown LE (2010*)* Climate change and freshwater ecosystems: impacts across multiple levels of organization. *Phil Trans R Soc B: Biol Sci 365*: 2093-2106.

Zaikova E, Goerlitz DS, Tighe SW., Wagner NY, Bai Y, Hall BL, et al. (20019) Antarctic Relic Microbial Mat Community Revealed by Metagenomics and Metatranscriptomics. *Frontiers Ecol Evol* 7. fevo.2019.00001.

Zhang L, Jungblut AD, Hawes I, Andersen DT, Sumner DY, Mackey TJ (2015) Cyanobacterial diversity in benthic mats of the McMurdo Dry Valley lakes, Antarctica. *Polar Biol* 38: 1097-1110.
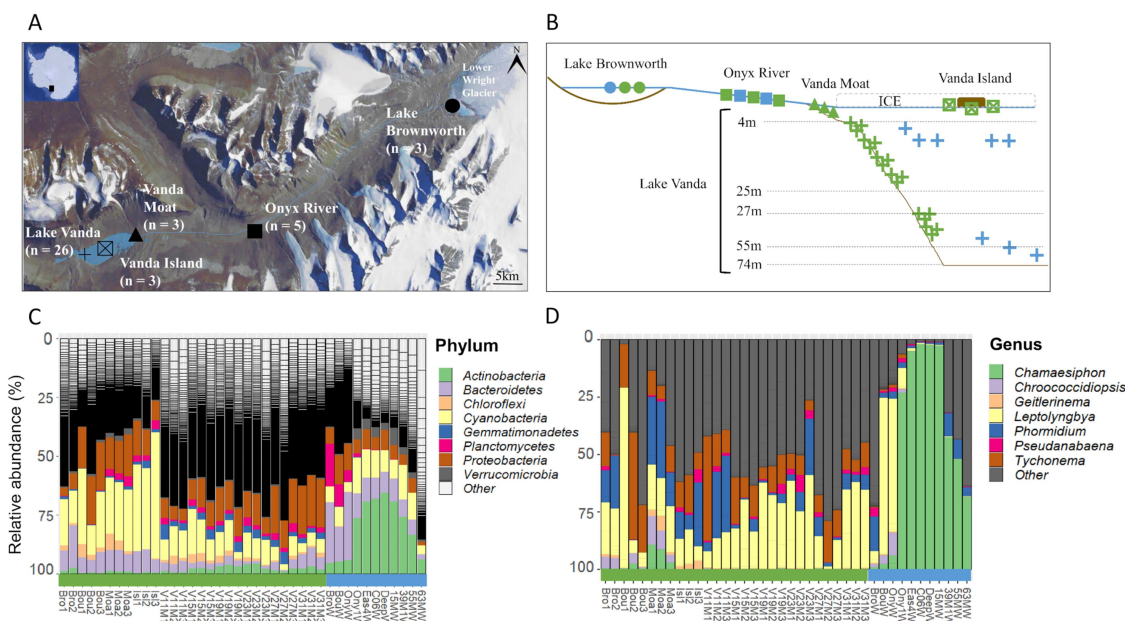
Figure 1: Location of the geographical units and bacterial community composition in the Wright Valley (Antarctica). (A) Aerial picture of the Wright Valley, from right to left: Lake Brownworth (circle), Onyx River (square), Vanda moat (triangle), Vanda island (open cross-square), and Lake Vanda (cross). The direction of flow of Onyx River is from Lake Brownworth to Lake Vanda. (B) Cross-section of the sampling sites including the depth profile of Lake Vanda, divided by meromictic water layers. Green symbols = mats, blue symbols = water samples. (C) Relative abundance of dominant bacterial taxa (Phylum level), and (D) cyanobacterial taxa (Genus level), assessed using 16S rDNA sequencing. Samples are sorted left to right according to habitat (mat, in green; water, in blue) and by increasing depth. Black areas correspond to overlapping edges of taxa with very low abundance ("Other").
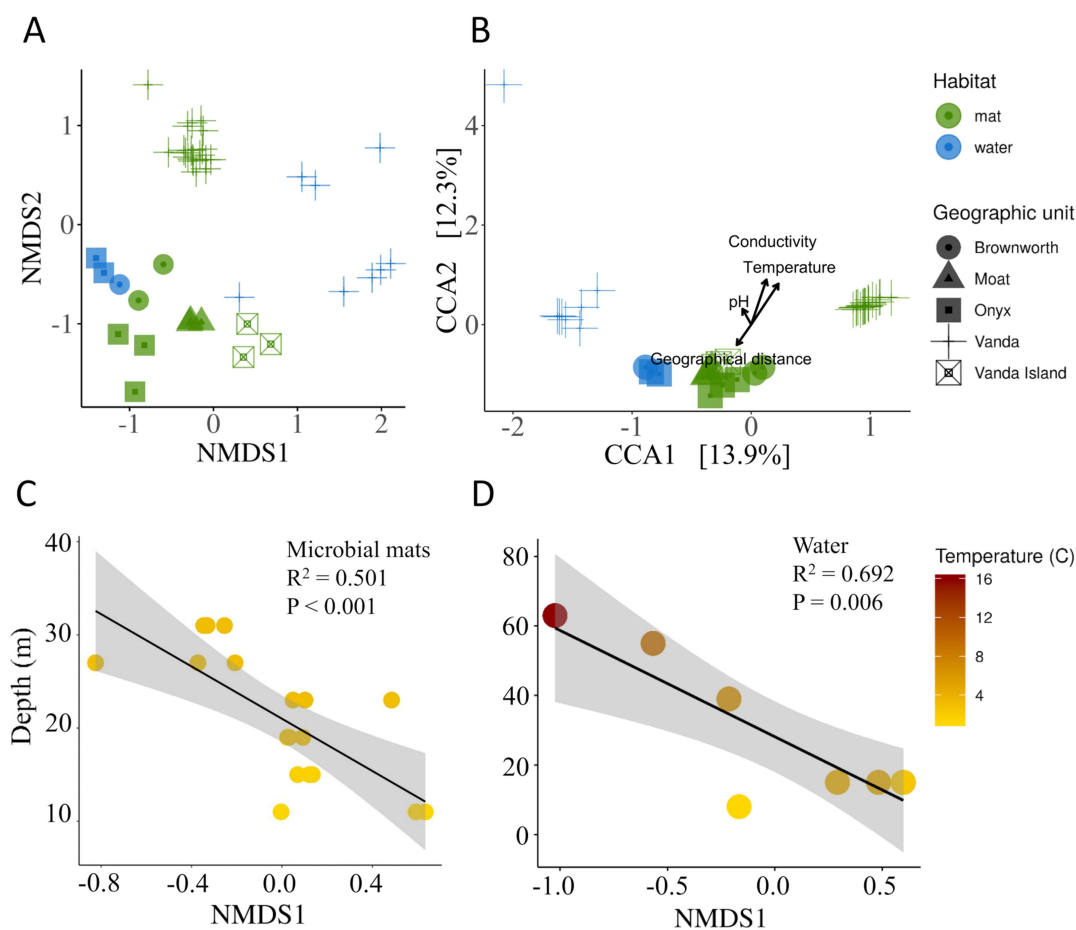
Figure 2: Evaluation of relationship between geographic units, environmental conditions and the bacterial and cyanobacterial communities in microbial mat and water habitats. (A) Non-metric multidimensional scaling (NMDS) ordination based on Bray-Curtis dissimilarities between bacterial communities across habitats and geographical units. (B) Constrained Correspondence Analysis (CCA) of bacterial community structure by temperature, conductivity, pH and geographic distance. Model outputs and statistical significance are reported in the table (new table number). Correlation between NMDS axis 1 (obtained from (A)) and depth within Lake Vanda, depicting the associated temperature gradient for microbial mat (C), and water (D) communities. Note that the x-axes of diagrams C and D do not match the range of values in A because independent ordinations were created.
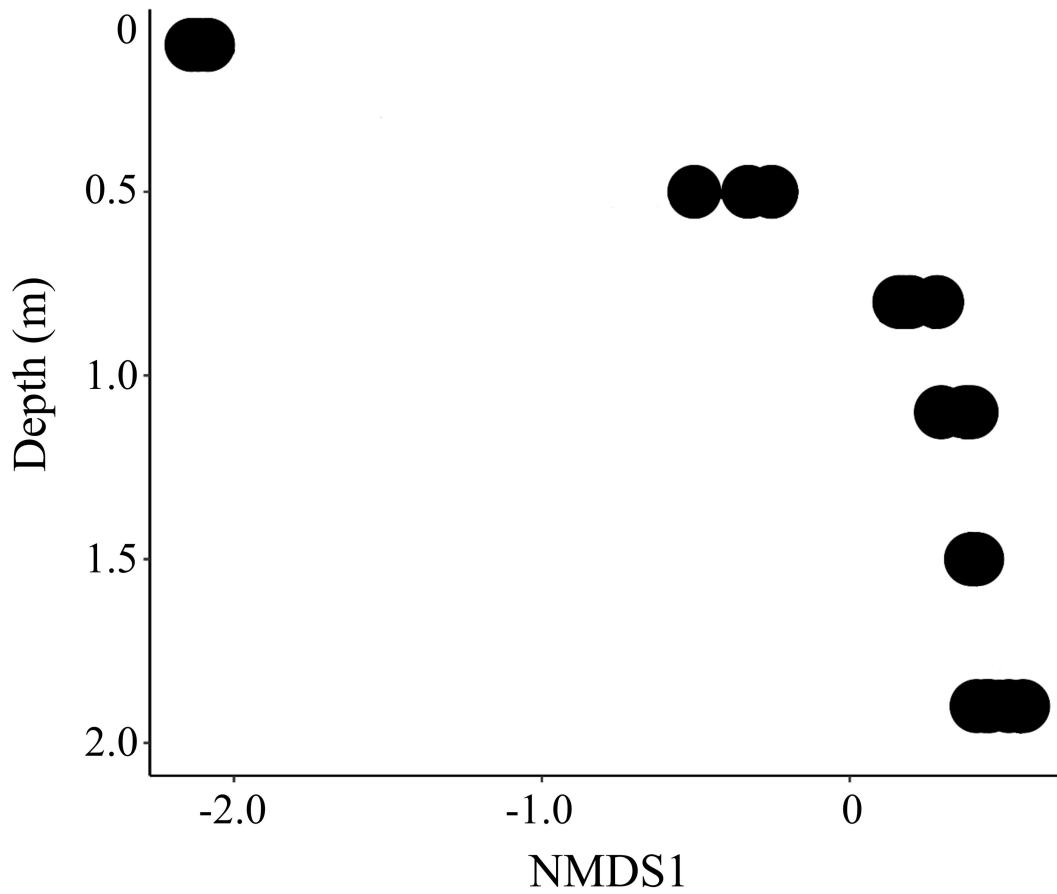
Figure 3. Correlation between NMDS axis 1 and depth of 16S rRNA gene communities mat communities along the 2 m transect in Lake Vanda moat with samples collected every 20 cm. The collection is based on the NMDS is reported in Figure S4.
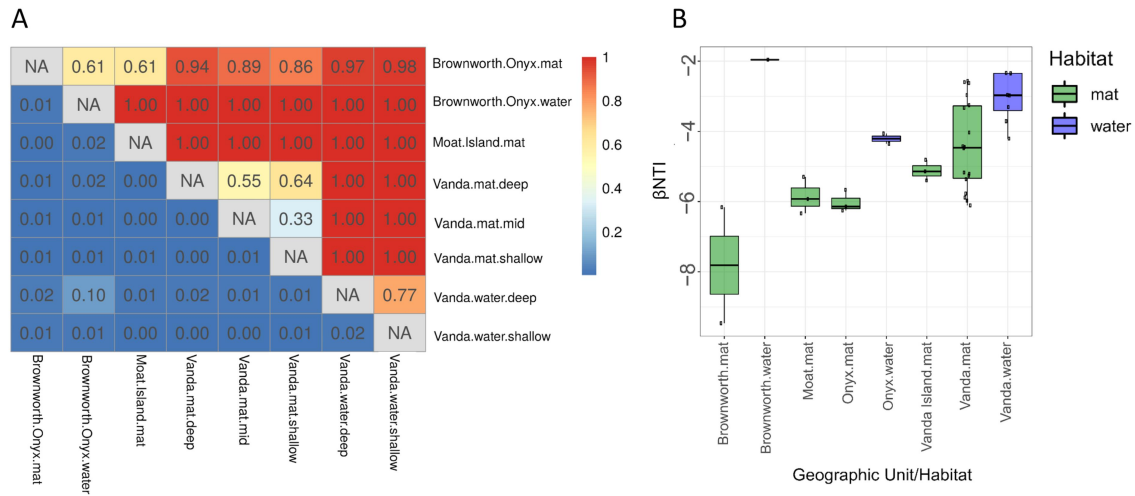
Figure 4: Assessment of bacterial community similarity and inference of environmental selection across habitats and geographical units in the BOV (Brownworth-Onyx-Vanda) system. (A) Analysis of similarities test (ANOSIM) between bacterial communities by geographical unit and habitat (mat vs. water) combinations. The upper half of the matrix depicts the ANOSIM R statistic, indicating higher community dissimilarity for values closer to 1; and the lower half depicts the P-values, with statistical significance set for $P \leq 0.05$. (B) Mean Nearest Taxon Difference z-scores for bacterial communities across geographical units and habitats. Values below -2 are indicative of phylogenetic underdispersion within the community, indicating a role of homogenizing selection in driving community structure.
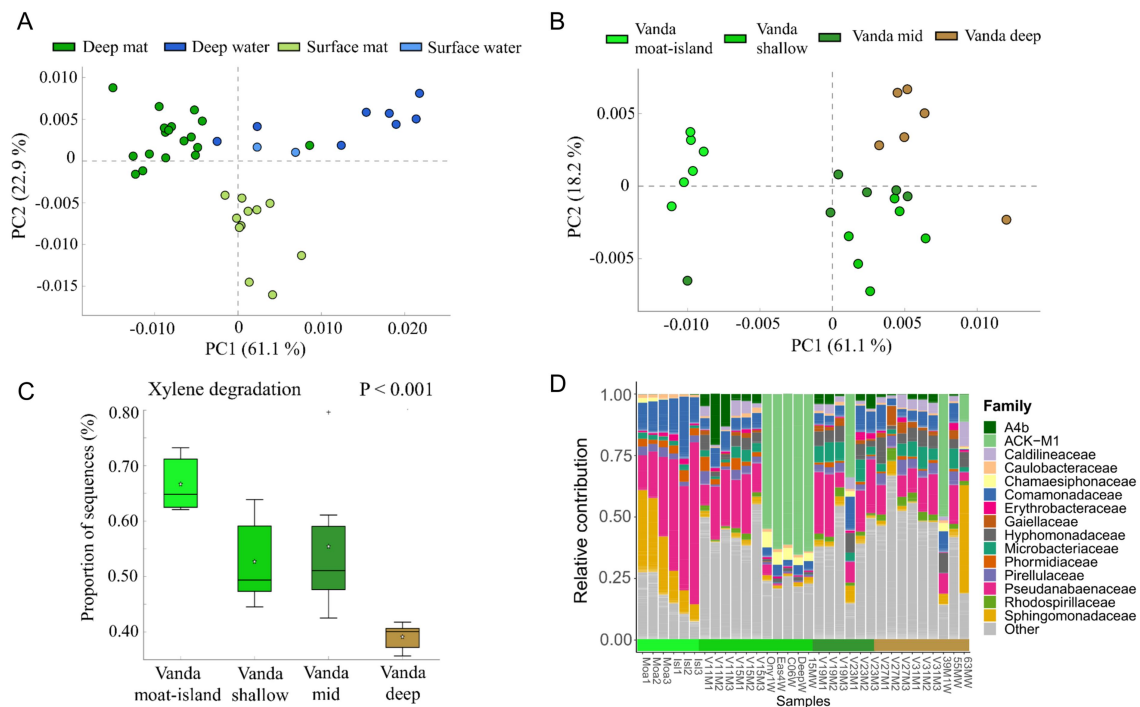
Figure 5: Principal component analysis (PCA) on the predicted bacterial metabolic gene composition between microbial mat and water communities at different depths in the BOV (Brownworth-Onyx-Vanda) system. (A) PCA on differences between the metabolic profiles of mat and water bacterial communities from the surface (<0.5 m) and deep (>4 m) environments, which correspond with different hydrological and temperature regimes. (B) PCA on differences between the metabolic profiles of bacterial communities in Lake Vanda, including its moat and island shores, and a division into three depth layers (shallow = 8-15 m, mid = 19-23 m, and deep = 27-63 m; for more information, see Fig. 3C). (C) Boxplot comparing the proportion of sequences predicting genes for xylene degradation. (D) Predicted relative contribution of OTUs at the family level to the metabolism for xylene degradation across depths in Lake Vanda (sample colour coding according to C).

Table 1: Sample types and water depths sampled at Lake Brownworth, Onxy River and Lake Vanda (McMurdo Wright Valley, Antarctica).

Geographic location	Sample	Sample type (Habitat)	Water depth (m)	Replicates
Lake Brownworth	BRO	Mat	0.1	2
	BROW	Water	0.1	1
Onyx River	BOU	Mat	0.1	3
	BOUW	Water	0.1	1
	ONYW	Water	0.1	1
Lake Vanda moat	MOA	Mat	0.1	3
Lake Vanda Island	ISL	Mat	0.1	3
	V11M	Mat	11	3
	V15M	Mat	15	3
	V19M	Mat	19	3
	V23M	Mat	23	3
	V27M	Mat	27	3
	V31M	Mat	31	3
Lake Vanda	ONY1W	Water	8	1
	EAS4W	Water	15	1
	C06W	Water	15	1
	DEEPW	Water	15	1
	15MW	Water	15	1
	39MW	Water	39	1
	55MW	Water	55	1

	63MW	Water	63	1
Lake Vanda moat transect (water edge to 2 m depth, samples taken every 20 cm)	VMOAT	Mat	0.1-1.9	3-5 per depth (26 samples)

ORIGINAL UNEDITED MANUSCRIPT

Table 2: Average (±SE) Chao1 diversity estimate and Simpson's diversity index (D) per sample for bacterial and cyanobacterial microbial mat and water communities in the BOV (Brownworth-Onyx-Vanda) system (microbial mats along water edge – 2 m depth transect from Lake Vanda not shown).

Geographical unit	Microbial mat				Water			
	Prokaryotes		Cyanobacteria		Prokaryotes		Cyanobacteria	
	Chao1	D	Chao1	D	Chao1	D	Chao1	D
Lake Brownworth (n=3)	1839 (37)	0.987 (0.004)	73 (2)	0.843 (0.060)	2464	0.961	78	0.576
Onyx river (n=5)	1802 (324)	0.950 (0.025)	69 (11)	0.551 (0.123)	3444 (98)	0.969 (0.009)	95 (4)	0.596 (0.085)
Vanda moat (n=3)	1725 (80)	0.978 (0.003)	67 (8)	0.848 (0.037)	-	-	-	-
Vanda island (n=3)	1112 (141)	0.945 (0.003)	52 (5)	0.722 (0.076)	-	-	-	-
Lake Vanda (n=18)	2038 (139)	0.981 (0.008)	51 (7)	0.807 (0.086)	849 (378)	0.935 (0.016)	39 (24)	0.284 (0.232)

Variable	F-value	P-value	*envfit* r^2	*envfit* P-value
Area	5.40	0.001	0.291	0.011
Geographical distance	2.09	0.004	0.411	0.002
Habitat	10.21	0.001	0.373	0.001
Temperature	4.13	0.001	0.884	0.001
Conductivity	3.24	0.001	0.769	0.001
pH	0.847	0.596	0.104	0.137

Table 3. Model outputs and statistical significance of the explanatory factors of the CCA, as

analysed using ANOVA and linear regressions using *envfit* (Figure 2 B).