

Tracing boundaries in Eastern Antarctica: Multi-scale drivers of soil microbial communities across the hyperarid Vestfold Hills

Eden Zhang^{a,b}, Paul Czechowski^c, Aleks Terauds^d, Sin Yin Wong^{a,b}, Devan S. Chelliah^{a,b}, Mark Raven^e, Mark M. Tanaka^a, Belinda C. Ferrari^{a,b,*}

^aSchool of Biotechnology and Biomolecular Sciences, UNSW Sydney, 2052, Australia.

^bEvolution and Ecology Research Centre, UNSW Sydney, 2052, Australia.

^cDepartment of Anatomy, University of Otago, Dunedin, 9054, New Zealand.

^dAustralian Antarctic Division, Department of Agriculture, Water and the Environment, 203 Channel Highway, Kingston, 7050, Australia.

^eCommonwealth Scientific and Industrial Research Organisation, Mineralogical Services, Urrbrae, SA, 5064, Australia.

Author contributions: BCF, AT, MMT, DSC and EZ designed the study. AT collated the environmental predictor data and coordinated sample collection. MR conducted XRF analyses. PC developed the analysis code then both EZ and SYW revised it. EZ processed the sequencing data and performed analyses with assistance from SYW. EZ wrote the manuscript with input from BCF, DSC, PC and AT. All authors read, collaborated and approved the final manuscript.

*To whom correspondence should be addressed. E-mail: b.ferrari@unsw.edu.au

Running Title: Key edaphic drivers of the Vestfold Hills soil microbiome

Keywords

Antarctica | Soil | Microbiome | Gradient Forest | Spatial Ecology | Conservation

Abstract

Microorganisms are key to sustaining core ecosystem processes across terrestrial Antarctica but they are rarely considered in conservation frameworks. Whilst greater advocacy has been made towards the inclusion of microbial data in this context, there is still a need for better tools to quantify multispecies responses to environmental change. Here, we extend the scope of Gradient Forest modelling beyond macroorganisms and small datasets to the comprehensive polar soil microbiome encompassing >17, 000 sequence variants for bacteria, micro-eukarya and archaea throughout the hyperarid Vestfold Hills of Eastern Antarctica. Quantification of microbial diversity against 79 physiochemical variables revealed that whilst rank-order importance differed, predictors were broadly consistent between domains, with greatest sharing occurring between bacteria and micro-eukarya. Moisture was identified as the most robust predictor for shaping the regional soil microbiome, with highest compositional turnover or “splits” occurring within the 10 – 12 % moisture content range. Often the most responsive taxa were rarer lineages of bacteria and micro-eukarya with phototrophic and nutrient-cycling capacities such as *Cyanobacteria* (up to 61.81 % predictive capacity), *Chlorophyta* (62.17 %) and *Ochrophyta* (57.81 %). These taxa groups are thus at greater risk of biodiversity loss or gain to projected climate trajectories, which will inevitably disturb current ecosystem dynamics. Better understanding of these threshold tipping points will positively aid conservation efforts across Eastern Antarctica. Furthermore, the successful implementation of an improved Gradient Forest model also presents an exciting opportunity to broaden its use on microbial systems globally.

Main Text

Changing climate in the Antarctic drives the expansion and greater connectivity of its ice-free habitats, which predominantly consist of soil microbiota (Lee *et al.*, 2017). In response to warming, amelioration of environmental conditions such as increased soil moisture and nutrient content is expected to alter current ecosystem dynamics (Chown *et al.*, 2015; McGaughan, Laver and Fraser, 2021). However, it is difficult to predict the long-term impact of disturbance on community composition and functionality, especially given the immense diversity of microorganisms and their highly variable responses (Allison & Martiny, 2008; Cavicchioli *et al.*, 2019; Robinson *et al.*, 2018 and 2020). This is further exacerbated by a general lack of available studies, especially in Eastern Antarctica (Zhang *et al.*, 2020). The hyperarid Vestfold Hills comprise of low-lying hilly landscapes, built upon a mixture of unconsolidated rock and unsorted debris left behind by the retreating East Antarctic Ice Sheet (Hirvas, Nenonen and Quilty, 1993). Strewn throughout its glacially scoured basins are numerous bodies of water, which have been the primary focus for biodiversity, evolutionary and ecological studies (Laybourn-Parry and Bell, 2014; Wilkins *et al.*, 2013). In contrast, its soil microbiome remains poorly investigated (Zhang *et al.*, 2020). Moving beyond descriptive reports, better understanding on the relationship between soil microbial diversity and the edaphic environment will be crucial towards achieving sound conservation planning and management across terrestrial Antarctica (Bergstrom *et al.*, 2020; Groves *et al.*, 2002; Rath *et al.*, 2018).

Soil micro-niches are highly variable and complex (Fierer, 2017). Shifts in microbial community composition are expected to occur at extremely fine scales (Chong *et al.*, 2015; Cowan *et al.*, 2014; Ferrari *et al.*, 2016; Siciliano *et al.*, 2014). Great interest therefore lies in quantifying key drivers of diversity and threshold tipping points along edaphic gradients, which enables the formation of a solid basis for predicting shifts in response to complex

environmental change (Kennicutt II *et al.*, 2014; Rath *et al.*, 2018). This knowledge may then be used to critically inform the designation of protected areas in order to effectively manage Antarctica’s unique biodiversity (Lee *et al.*, 2017). Inclusion of microorganisms, however, has been rare in the development of Antarctica conservation frameworks as well as other ecosystems around the globe (Cavicchioli *et al.*, 2019; Hughes *et al.*, 2015 and 2018). As such, there is an increasingly urgent need to find or develop suitable methods to bridge this conservation gap – one option being Gradient Forest (Ellis, Smith and Pitcher, 2012).

Based on an aggregation of Random Forest (Breiman, 2001), which only considers the response of single species, Gradient Forest offers a flexible approach to explore multispecies responses between biodiversity and the environment (Ellis, Smith and Pitcher, 2012). Accounting for complex interactions by fitting multiple regression models, Gradient Forest quantifies the relative importance of different predictors and their non-linear correspondence with community compositional turnover or “splits” along predictor gradients, which can be upscaled across extensive geographic areas to define spatial groups that capture variation in species composition and turnover (Stephenson *et al.*, 2018). This output may be used to critically inform conservation planning and management, especially where biodiversity survey data is sparse (Pitcher *et al.*, 2012). However, the application of such modelling approaches has thus far been limited to studies on macroorganisms in forest or marine ecosystems (Maguire *et al.*, 2015; Pitcher *et al.*, 2012; Stephenson *et al.*, 2018).

For the first time, we employ a modified Gradient Forest framework on the polar soil microbiome (i.e., bacteria, micro-eukarya and archaea), using reanalysed 16S and 18S rRNA gene amplicon sequence variants (ASVs) obtained from four sites (Adam’s Flat “AF”, Heidemann Valley “HV”, Old Wallow “OW” and Rookery Lake “RL”) throughout the hyperarid Vestfold Hills of Eastern Antarctica (Zhang *et al.*, 2020). Doing so, we (1) quantify the relative importance of 79 measured physiochemical predictors on the local and regional

distribution of polar soil microbial communities across the region; and (2) identify where any major compositional shifts occur along important predictor gradients in order to define spatial boundaries and abiotic thresholds. These outcomes will be used to prioritize sensitive areas for further sampling and targeted microbial studies across the Vestfold Hills, with significant global outreach to other arid soil environments.

Results

Soil characteristics and phyla are distributed unevenly between and within sites along spatially explicit transects. Overlap of samples on the environmental nMDS indicate that measured soil physiochemical variables are generally conserved across the Vestfold Hills (Figure 1). However, significant (ANOSIM $p < 0.05$) local heterogeneity is still apparent, for example elevation ($R = 0.81$), aspect ($R = 0.45$) and moisture ($R = 0.27$) amongst other nutrient-related factors (Table 1). In terms of biodiversity, bubbleplots depicting the relative distribution and abundance of bacteria, micro-eukarya and archaea revealed that a handful of phyla dominate the regional soil microbiome (Figures 2 and S1). These include members belonging to the ubiquitous and metabolically diverse *Actinobacteria*, *Bacteroidetes*, *Proteobacteria* and *Chloroflexi* as well as a relatively high abundance (>60 %) of unclassified micro-eukarya (Figure S2a and b). Rarer phyla (<10 %) accounted for the main phylogenetic difference amongst sites (Figures 2 and S1), which include an uneven distribution of *Acidobacteria* (start of AF transect), *Cyanobacteria* (start of RL transect), *Ca. Eremiobacterota* (Start of HV transect) and micro-algae such as *Ochrophyta* (mid of OW transect) and *Chlorophyta* (start of RL transect). In contrast, archaea were generally consistent between and within sites (Figures 2c and S1c). Overall, greatest phylogenetic diversity was observed at OW (100/102 m) and RL (0/2 m) (Figure 2).

Moisture availability shapes the hyperarid Vestfold Hills soil microbiome. Overall importance plots generated from Gradient Forest analyses demonstrate that rank-order importance differed but predictors were generally consistent between domains (Figure 3). Moisture was identified as the most robust predictor in shaping the regional soil microbiome, followed by nutrient-related factors (e.g., Br and K), pH and aspect amongst other predictors of intermediate importance (Figure 3). On the local scale (Table 2), differences between sites reflect local physiochemical variations (Table 1). Some examples include calcium (Ca) at AF (Figure S2), phosphate-related factors (e.g., PO₄, P and TP) at HV (Figure S3), conductivity at OW (Figure S4) and boron (B) content at RL (Figure S5). In general, aspect and trace elements (e.g., B, Zinc “Zn”, iron “Fe” and manganese “Mn”) became more important as differentiating factors for local communities. Bacteria and micro-eukarya generally shared a similar set of predictors whereas archaea remained largely distinct, except at OW and RL (Figure 3 and Table 2). Overall, the suite of environmental variables predicted a relatively modest fraction of variation in the distribution of polar soil bacteria ($R^2 = 0.009 - 0.732$), micro-eukarya ($R^2 = 0.005 - 0.622$) and archaea ($R^2 = 0.031 - 0.300$) across the Vestfold Hills (Table S1). Of which these variables had some predictive capacity (i.e., positive R^2) for phyla on both regional (bacteria = 27, micro-eukarya = 21, archaea = 2) and local (bacteria = 4 – 21, micro-eukarya = 2 – 12, archaea = 1 – 3) scales (Tables S1 and S2).

Gradient responses and thresholds vary between domains and over different spatial scales. Cumulative importance plots reveal that changes in community composition or “splits” along edaphic gradients were non-uniform and responses largely varied between phyla over both regional and local scales (Figures 4 and S2 – 5). For example, along the moisture gradient, many important splits occurred in the range *c.* 10 – 12 % (Figure 4), indicating that high rates of compositional turnover (e.g., *Cyanobacteria*, *Ochrophyta* and *Chlorophyta*) corresponded with small changes in moisture content across the region (Figure 4a and b). All sites sampled

throughout the Vestfold Hills are extremely arid, with AF being the driest ($2.00\% \pm 2.00$) followed by HV ($7.00\% \pm 2.00$) whilst soils at OW ($8.00\% \pm 4.00$) and RL ($8.00\% \pm 4.00$) retain higher moisture content (Table 1). Higher abundances of *Cyanobacteria*, *Ochrophyta* and *Chlorophyta* were detected at OW and RL (Figures 2a/b and S1a/b). Similarly, along the Br gradient, many important splits occurred in the range *c.* 5 – 15 ppm affecting regional abundances of rarer phyla like *Kiritimatiellaeota* and *Ichthyosporea* (Figure 4a and b), which were primarily distributed towards the middle (17.71 Br ppm) of the OW transect (Figure 2a and b). Regarding archaea, particularly *Thermoproteota* (Figures 2c and S1c), its regional distribution and abundance did not change much along the K gradient despite exhibiting a steep response *c.* 600 – 700 ppm (Figure 4c).

In comparison, notable differences in compositional responses were observed at finer scales (Figures S2 – 5). For instance, along the Ca gradient at AF, important splits occurred in the range *c.* 100–1200 mg/kg (Figure S2), affecting *FBP* (*c.* 0 – 200 Ca mg/kg) at steeper rates than *Labyrinthulomycetes*, *Ochrophyta* and unclassified archaea (*c.* 1000 – 1200 Ca mg/kg), where higher abundances of *FBP* are distributed towards the start (329.34 Ca mg/kg) and end (316.90 Ca mg/kg) of the AF transect (Figure 2a). Whilst at HV, along the PO₄ gradient, important splits occurred in the range *c.* 2 – 8 ppm (Figure S3), mainly affecting *Chloroflexi* (*c.* 2 – 6 PO₄ ppm) and *Euryarchaeota* (*c.* 6 – 8 PO₄ ppm), with *Chloroflexi* present in higher abundances towards the middle (5.56 PO₄ ppm) of the HV transect (Figure 2a and c). Similarly, along the conductivity gradient at OW, important splits occurred in the range *c.* 0 – 15, 000 uS/cm (Figure S4), mainly affecting *Deinococcus-Thermus* (*c.* 7500 – 10, 000 uS/cm), *Ochrophyta* (*c.* 10, 000 – 15, 000 uS/cm) and *Euryarchaeota* (*c.* 2500 – 5000 uS/cm), where higher abundances of all three phyla were distributed towards the middle (14, 600 uS/cm) of the OW transect (Figure 2). At RL, important splits along the B gradient occurred in the range *c.* 0 – 12 mg/kg (Figure S5), affecting *Cyanobacteria* (*c.* 2 – 4 B mg/kg) and *Chlorophyta* (*c.*

0 – 2 B mg/kg) at steeper rates than *Euryarchaeota* (c. 6 – 10 B mg/kg), where higher abundances of *Cyanobacteria* and *Chlorophyta* were detected towards the start (0.97 B mg/kg) of the RL transect.

Classification of spatial groups for regional conservation planning. Mapping of expected biodiversity patterns based on information about gradient responses and thresholds previously stored in cumulative importance curves identified an average of two-to-four spatial groups on both regional and local scales (Figures 5 and S6 – 8). For the most part, the Vestfold Hills soil microbiome appeared to be a fairly homogenous environment though greater heterogeneity in compositional response and overall phylogenetic diversity was detected at OW and RL (Figures 2 and 5). In particular, their soil bacterial and micro-eukaryotic components situated towards the start (RL; Figure 5b) and middle (OW; Figure 5a and b) of the transects where retainment of soil moisture (11 – 12 %) is highest in the region (Table 1). On finer scales (Figures S6 – 8), biological responses were notably more heterogenous with visible boundaries between the start, middle and end of transects – with aspect playing a key role. However, when taking into consideration sample size and variation, predictive power remains overall more useful on the broad regional scale for identifying priority areas for further studies.

Discussion

Robust discrimination of key drivers underlying biogeographic patterns of polar soil microbial communities offered by Gradient Forest make it well suited as primary input towards systematic conservation planning across the hyperarid Vestfold Hills. Distinctive gradient responses and thresholds for bacterial and micro-eukaryotic communities at Old Wallow (OW) and Rookery Lake (RL) has led to the prioritisation of these two East Antarctic sites as conservation targets for further sampling (Figures 2, 4 and 5). More in-depth studies, however, are needed to critically inform potential management actions. Nonetheless, our findings provide a baseline for anticipating shifts in microbial response to ameliorated conditions as

consequence of regional warming (Robinson *et al.*, 2020; McGaughan, Laver and Fraser, 2021).

Environmental shaping of the hyperarid Vestfold Hills soil microbiome is primarily driven by variances in moisture throughout the region (Figure 3), which is largely regulated by extrinsic factors such as temperature (Delgado-Baquerizo *et al.* 2018). Along the regional moisture gradient (Figure 4), strongest rates of compositional turnover were observed for rarer lineages of bacteria (e.g., *Cyanobacteria*) and micro-eukarya (e.g., *Chlorophyta* and *Ochrophyta*) with phototrophic and/or nutrient-cycling capacities such as nitrogen and hydrogen metabolism (Bothe *et al.*, 2010; Fernandes *et al.*, 2016; Fuerst & Sagulenko 2012; Lea-Smith *et al.*, 2016). These taxa groups were found in relatively high abundances at OW and RL (Figures 2 and S1), likely due to the presence of meltwater lakes and stochastic nutrient inputs from wildlife (Huang & Zhu, 2009). Projected warming trends may favour the rapid growth of photosynthetic potential across terrestrial Antarctica (Bay, Greening and Ferrari, 2018). An example being in the form of microalgal blooms near bird or seal colonies, like those at OW and RL, which may also act as important terrestrial carbon sinks (Gray *et al.*, 2020). Cascading effects from this hypothetical scenario would inevitably disturb current ecosystem dynamics, especially when taking into consideration that <0.1 % of the frigid landmass is vegetated (Fretwell *et al.*, 2011). Therefore, placing extreme selective pressure against chemoautotrophic lineages (e.g., *Actinobacteria*, *Ca. Eremiobacterota* and *Ca. Dormibacterota*) and thereby leading to potential competition for dominance between primary production strategies across Eastern Antarctica (Bay *et al.*, 2021; Bay *et al.*, 2021; Ji *et al.*, 2017; Ray *et al.*, 2020).

Ecosystem resilience, however, remains a poorly understood concept due to the lack of empirical data (Allison & Martiny, 2008). As such, it is unknown whether compositional changes will also affect functionality nor the extent to which it drives biodiversity gain or loss over larger spatial scales, especially within the context of a warming Antarctica (Lee *et al.*,

2017). Outputs from Gradient Forest analyses provided some predictive capacity, up to 73.2 % in some cases for bacteria and micro-eukarya depending on spatial scale (Table S1), but offered limited resolution on archaea across the hyperarid Vestfold Hills. This suggests that the environment may not be the primary driver of distributions and compositions for some taxa. As such, spatial planning and management activities also need to consider these other processes that drive biodiversity patterns (Pitcher *et al.*, 2012). However, further sampling and individual site-level studies are necessary to draw more explicit conclusions – by starting at OW and RL. Building upon previous work highlighting the unique microbial biodiversity found in Eastern Antarctica (Zhang *et al.*, 2020), we advance methods to better apprehend their key edaphic drivers and threshold tipping points on multiple scales. Increased capacity of Gradient Forest to model strain-level data enables the integration of diverse microbial communities in Antarctic conservation and management actions, including the development of new protected areas (Hughes, Cowan and Wilmotte, 2015), thus taking a crucial step towards achieving sound conservation and management across terrestrial Antarctica (Cavicchioli *et al.*, 2019; Hughes, Cowan and Wilmotte, 2015). Moreover, our modified Gradient Forest framework will also be valuable for assessing disturbance, rehabilitation and restoration within other terrestrial environments around the globe.

233 **Methods**

234 Study area and sampling design. Sampling in the Vestfold Hills region (centred 68°55'S, 78°25'E) was performed in 2012 by expeditioners via the Australian Antarctic Program (AAP) across four locations within vicinity of the Australian Davis research station (68°35'S, 77°58'E) (Figure 1). These include: Adams Flat (AF: 68°33'S, 78°1'E); Heidemann Valley (HV: 68°35'S, 78°0'E); Old Wallow (OW: 68°36'S, 77°58'E) and Rookery Lake (RL: 68°36'S, 77°57'E). At each site, samples (n = 18) were taken from the top 10 cm of soil along three parallel transects at the following distance points of 0, 2, 100, 102, 200 and 202 m

(Siciliano *et al.*, 2014). For the purpose of this study, samples were not pooled but categorised into start (0/2 m), mid (100/102 m) and end (200/202 m) according to their location along the transect.

Soil physiochemical analysis. All soils (n = 72) included in this study had been submitted by the Australian Antarctic Division (AAD) for extensive physiochemical analysis using standard procedures as performed by Bioplatforms Australia (BPA). As described in Siciliano *et al.* (2014; Appendix A. Supplementary Data) and Table S3, this includes the following: pH, conductivity, moisture, aspect, elevation, slope, calcium “Ca”, magnesium “Mg”, potassium “K”, sulfur “S”, total carbon “TC”, total nitrogen “TN”, total phosphorous “TP”, aluminium “Al”, boron “B”, copper “Cu”, iron “Fe”, manganese “Mn”, zinc “Zn”, cation exchange capacity, chloride “Cl”, bromide “Br”, nitrate “NO₃”, nitrite “NO₂”, phosphate “PO₄”, phosphorous “PO₃”, sulphate “SO₄”, sorting order, gravel, sand, mud, min grain size, max grain size, aluminium oxide “Al₂O₃”, calcium oxide “CaO”, iron (III) oxide “Fe₂O₃”, magnesium oxide “MgO”, manganese oxide “MnO”, phosphorous pentoxide “P₂O₅”, potassium oxide “K₂O”, sodium oxide “Na₂O”, silicon oxide “SiO₂”, sulfur trioxide “SO₃” and titanium oxide “TiO₂”.

DNA extraction and Illumina amplicon sequencing. Soil samples were extracted in triplicate using the FASTDNA™ SPIN Kit for Soil (MP Biomedicals, Santa Ana, CA, US) and Qubit™ 4 Fluorometer (ThermoFisher Scientific, NSW, Australia) as per manufacturers’ instructions. Diluted DNA (1:10 using nuclease-free water) was submitted to the Ramaciotti Centre for Genomics (UNSW Sydney, Australia) for amplicon paired-end sequencing on the Illumina MiSeq platform (Illumina, California, US) with controls in accordance to protocols from the AusMicrobiome project (Bissett *et al.*, 2016). For bacteria and archaea, we respectively targeted the V1–3 hypervariable region of 16S rRNA gene sequences using the 27F/519R (Lane, 1991) and A2F/519R primer sets (Reysenbach *et al.*, 1995). For micro-eukarya, we

targeted the V9 hypervariable region of 18S rRNA gene sequences using the 1391f /EukBr primer set (Amaral-Zettler *et al.*, 2009).

Reprocessing of amplicon sequencing data. Forward and reverse primer sequences were removed from the raw fastq files using cutadapt v2.10 (Martin, 2011). The trimmed paired end reads were merged using FLASH 2 (Magoč & Salzberg, 2011) then converted into fasta format using SEQTK (<https://github.com/lh3/seqtk>). Individual files were concatenated and imported into MOTHUR (Schloss *et al.*, 2009) to screen the sequences, remove ambiguous bases and those with homopolymer runs >8 bp. UPARSE/UNOISE2 (Edgar, 2013 and 2016) was then used to dereplicate and denoise the screened sequences, yielding a set of amplicon sequence variants (ASVs – Callahan *et al.*, 2017) and abundance count tables. For both 16S and 18S rRNA amplicon gene datasets, taxonomy was classified against the SILVA database release 132 (Quast *et al.*, 2013).

Summary of input data for Gradient Forest analysis. In total, we recovered 51, 978, 909 bacterial 16S rRNA, 17, 994, 697 micro-eukaryotic 18S rRNA and 5, 150, 065 archaeal 16S rRNA gene sequences from the high-throughput amplicon sequencing of the Vestfold Hills soil microbiome (n = 72). After read-quality filtering, a total of 51, 026, 375 bacterial sequences remained, which clustered into 12, 658 ASVs (Table S2). In terms of micro-eukarya, a total of 9, 630, 123 sequences remained after read-quality filtering, which clustered into 1, 507 ASVs (Table S2). Meanwhile, a total of 2, 159, 440 high-quality archaeal sequences were left and these clustered into 367 ASVs (Table S2). Good overall coverage was obtained for bacteria (average reads per sample = 75, 811) micro-eukarya (average reads per sample = 52, 785) and archaea (average reads per sample = 69, 703). All environmental parameters (79) were checked for normality and no transformations were applied, though high-correlated variables ($R^2 > 0.7$) or those with incomplete data were manually removed, resulting in 43 complete predictors (Table S3).

The R environment. All multivariate and statistical analyses were carried out in R Studio version 4.0.3 (R Core Team, 2020). Unless specified otherwise, all plots were visualised using a combination of ggplot2 v3.1.0 (Wickham, 2016) and ggpubr v0.2 (Kassambara, 2018).

Gradient Forest analysis. Datasets were stored as individual phyloseq objects (McMurdie & Holmes, 2013). These were prepared by normalising the abundance data using DESeq2 (Love, Huber and Anders, 2014), then agglomerating taxa at the phylum level and extracting defined metadata variables as individual list elements for binary conversion and/or the removal of highly co-correlated variables ($R^2 > 0.7$) using the cor function in R. The R package gradientForest (Ellis *et al.*, 2010) was called upon to fit the model (rp1 = 1000, corr.threshold = 0.5) to our optimised list object and return a gradientForest object, which was used to generate a series of plots: (1) predictor overall importance plots (shows mean accuracy importance and the mean importance weighted by species R^2); (2) species cumulative plots (shows cumulative change in abundance of individual species, where changes occur on the gradient and the species changing most on each gradient) and; (3) bi-plots (shows transformed predictors that have been classified into spatial groups, which capture variation in species composition and turnover mapped onto biological and geographical space).

Deposition of data in an open-source database. All input files including associated environmental metadata and relevant code have been uploaded onto GitHub and is publicly available via <https://github.com/Eden3416110/gradientForest.git>.

Competing interests. The authors declare that they have no competing interests.

ACKNOWLEDGEMENTS

The authors would like to thank the Australian Antarctic Program and all expeditioners for the collection of soil samples (supported by Australian Antarctic Projects 3338 and 5097). We acknowledge funding support through Belinda Ferrari's

315 Australian Research Council Future Fellowship (FT170100341), the Australian Antarctic
316 Science Grants scheme (AAS 4406 – Ferrari, AAS 4296 – Terauds), and Ausmicrobiome for
317 the provision of the Vestfold Hills biodiversity data.

REFERENCES

1. Allison, S.D., Martiny, J.B.H. (2008). Resistance, resilience, and redundancy in microbial communities. *PNAS*, **105**, pages 11512-19.
2. Bay, S., Ferrari, B.C., Greening, C. (2018) “Life without water: How do bacteria generate biomass in desert ecosystems?”, *Microbiology Australia*, doi: 10.1071/MA18008.
3. Bay, S.K., Waite, D.W., Dong, X., Gillor, O., Chown, S.L., Hugenholtz, P., Greening, C. (2021) “Chemosynthetic and photosynthetic bacteria contribute differentially to primary production across a steep desert aridity gradient”, *The ISME Journal*, doi: 10.1038/s41396-021-01001-0.
4. Bergstrom, D.M., Wienecke, B.C., van den Hoff, J., Hughes, L., Lindenmayer, D.B., Ainsworth, T.D., Baker, C.M., Bland, L., Bowman, D.M.J.S., Brooks, S.T., Canadell, J.G., Constable, A.J., Dafforn, K.A., Depledge, M.H., Dickson, C.R., Duke, N.C., Helmstedt, K.J., Holz, A., Johnson, C.R., McGeoch, M.A., Melbourne-Thomas, J., Morgain, R., Nicholson, E., Prober, S.M., Raymond, B., Ritchie, E.G., Robinson, S.A., Ruthrof, K.X., Setterfield, S.A., Sgrò, C.M., Stark, J.S., Travers, T., Trebilco, R., Ward, D.F.L., Wardle, G.M., Williams, K.J., Zylstra, P.J., Shaw, J.D. (2021) “Combating ecosystem collapse from the tropics to the Antarctic”, *Global Change Biology*, doi: 10.1111/gcb.15539.
5. Bissett, A., Richardson, A.E., Baker, G., Wakelin, S., Thrall, P.H. (2010) “Life history determines biogeographical patterns of soil bacterial communities over multiple spatial scales”, *Molecular Ecology*, **19**, pp. 4315-4327.
6. Bothe, H., Schmitz, O., Yates, G.M., Newton, W.E. (2010) “Nitrogen fixation and hydrogen metabolism in Cyanobacteria”, *Microbiology and Molecular Biology Reviews*, doi: 10.1128/MMBR.00033-10.
7. Breiman, L. (2001) “Random Forests”, *Machine Learning*, **45**, pp. 5-32.
8. Callahan, B.J., McMurdie, P.J., Holmes, S.P. (2017) “Exact sequence variants should

- replace operational taxonomic units in marker-gene data analysis”, *ISME*, **11**, 113597.
9. Cavicchioli, R., Ripple, W.J., Timmis, K.N., Azam, F., Bakken, L.R., Baylis, M., Behrenfeld, M.J., Boetius, A., Boyd, P.W., Classen, A.T., Crowther, T.W., Danovaro, R., Foreman, C.M., Huisman, J., Hutchins, D.A., Jansson, J.K., Karl, D.M., Koskella, B., Welch, D.B.M., Martiny, J.B.H., Moran, M.A., Orphan, V.J., Reay, D.S., Remais, J.V., Rich, V.I., Singh, B.K., Stein, L.Y., Stewart, F.J., Sullivan, M.B., van Oppen, M.J.H., Weaver, S.C., Webb, E.A., Webster, N.S. (2019). Scientists’ warning to humanity: Microorganisms and climate change. *Nature Reviews Microbiology*, **17**, pages 569-586.
10. Chong, C., Pearce, D.A., Convey, P. (2015) “Emerging spatial patterns in Antarctic prokaryotes”, *Frontiers in Microbiology*, 6(1058), doi: 10.3389/fmicb.2015.01058.
11. Chown, S.L., Clarke, A., Fraser, C.I., Cary, S.C., Moon, K.L., and McGeoch, M.A. (2015) “The changing form of Antarctic biodiversity”, *Nature*, **552**, pp. 431-438.
12. Cowan, D.A., Makhallanyane, T.P., Dennis, P.G., Hopkins, D.W. (2014) “Microbial ecology and biogeochemistry of continental Antarctic soils”, *Frontiers in Microbiology*, **5(154)**, doi: 10.3389/fmicb.2014.00154.
13. Edgar, R.C. (2013) “UPARSE: Highly accurate OTU sequences from microbial amplicon reads”, *Nature Methods*, **10(10)**, pp. 996-998.
14. Edgar, R.C. (2016) “UNOISE2: Improved error-correction for Illumina 16S and ITS amplicon reads, *bioRxiv*, DOI: <https://doi.org/10.1101/081257>.
15. Ellis N., Smith S.J., and Pitcher C.R. (2012) “Gradient forests: calculating importance gradients on physical predictors”, *Ecology*, **93(1)**, pp. 156-168.
16. Fernandes, F., Barbosa, M., Oliveira, A.P., Azevedo, .C., Sousa-Pinto, I., Valentão, P., Andrade, P.B. (2016) “The pigments of kelps (Ochrophyta) as part of the flexible response to highly variable marine environments”, *Journal of Applied Phycol*, **28**, pp. 3689-3696.
17. Ferrari, B.C., Bissett, A., Snape, I., van Dorst, J., Palmer, A.S., Ji, M., Siciliano, S.D., Stark,

- J.S., Winsley, T., Brown, M.V. (2016) “Geological connectivity drives microbial community structure and connectivity in polar, terrestrial ecosystems”, *Environmental Microbiology*, **18(6)**, pp. 1834-1849.
18. Fierer, N. (2019) “Embracing the unknown: disentangling the complexities of the soil microbiome”. *Nature Reviews Microbiology*, **15**, pp. 579-590.
19. Fretwell, P.T., Convey, P., Fleming, A.H., Peat, H.J., Hughes, K.A. (2011) “Detecting and mapping vegetation distribution on the Antarctic Peninsula from remote sensing data”, *Polar Biology*, **34**, pp. 273-281.
20. Fuerst, J.A., Sagulenko, E. (2012) “Keys to eukaryality: Planctomycetes and ancestral evolution of cellular complexity”, *Frontiers in Microbiology*, **3(167)**, doi: 10.3389/fmicb.2012.00167.
21. Gray, A., Krolikowski, M., Fretwell, P., Convey, P., Peck, L.S., Mendelova, M., Smith, A.G., Davey, M.P. (2020) “Remote sensing reveals Antarctic green snow algae as important terrestrial carbon sink”, *Nature Communications*, doi: 10.1038/s41467-020-16018-w.
22. Groves, C.R., Jensen, D.B., Valutis, L.L., Redford, K.H., Shaffer, M.L., Scott, M., Baumgartner, J.V., Higgins, J.V., Beck, M.W., and Anderson, M.G. (2002) “Planning for biodiversity conservation: Putting conservation science into practice”, *BioScience*, **52(6)**, pp. 449-512.
23. Hirvas, H., Nenonen, K., and Quilty, P. (1993) “Till stratigraphy and glacial history of the Vestfold Hills area, east Antarctica”, *Quaternary International*, **18**, pp. 81-95.
24. Huang, T., Zhu, R. (2009) Penguin occupation in the Vestfold Hills, *Antarctic Science*, **21(2)**, pages 131-134.
25. Hughes, K.A., Cowan, D.A., Wilmotte, A. (2015) “Protection of Antarctic microbial communities – ‘Out of sight, out of mind’”, *Frontiers in Microbiology*, doi:

10.3389/fmicb.2015.00151.

26. Hughes K.A., López-Martínez J., Constable A., Njåstad B. (2018). Antarctic environmental protection: Strengthening the links between science and governance. *Environmental Science and Policy*, **83**, pp. 86-95.
27. Ji, M., Greening, C., Vanwonderghem, I., Carere, C.R., Bay, S.K., Steen, J.A., Montgomery, K., Lines, T., Beardall, J., van Dorst, J., Snape, I., Stott, M.B., Hugenholtz, P., and Ferrari, B.C. (2017) "Atmospheric trace gases support primary production in Antarctic desert surface soil", *Nature*, **552**, pp. 400-403.
28. Kassambara, A. (2018) "ggpubr: "ggplot2" based publication ready plots", R package version 0.1.8. <https://cran.r-project.org/web/package=ggpubr>.
29. Laybourn-Parry, J., and Bell, E.M. (2014) "Ace lake: three decades of research on a meromictic, Antarctic lake", *Polar Biology*, **37**, pp. 1685-1699.
30. Lea-Smith, D.J., Bombelli, P., Vasudevan, R., Howe, C.J. (2016) "Photosynthetic, respiratory and extracellular electron transport pathways in cyanobacteria", *BBA - Bioenergetics*, **1857(3)**, pp. 247-255.
31. Lee, J.R., Raymond, B., Bracegirdle, T.J., Chadès, I., Fuller, R.A., Shaw, J.S., and Terauds, A. (2017) "Climate change drives expansion of Antarctic ice-free habitat", *Nature*, **547(49)**, doi:10.1039/nature22996.
32. Love, M.I., Huber, W., Anders, S. (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." *Genome Biology*, **15**, 550. doi: 10.1186/s13059-014-0550-8.
33. Magoč T., and Salzberg S.L. (2011) "FLASH: fast length adjustment of short reads to improve genome assemblies", *Bioinformatics*, **27(21)**: pp. 2957-2963.
34. Maguire, K.C., Nieto-Lugilde, D., Fitzpatrick, M.C., Williams, J.W., Blois, J.L. (2015) "Modeling species and community responses to past, present, and future episodes of

- climatic and ecological change”, *Annual Review of Ecology, Evolution and Systematics*,
46, pp. 343-368.
35. Martin, M. (2011) “Cutadapt removes adapter sequences from high-throughput sequencing
reads”, *EMBnet.journal*, **17(1)**, p. 10. doi: 10.14806/ej.17.1.200.
36. McGaughan, A., Laver, R., Fraser, C. (2021) “Evolutionary responses to warming”,
Trends in Ecology and Evolution, doi: 10.1016/j.tree.2021.02.014.
37. McMurdie, P.J., Holmes, S. (2013) “phyloseq: An R package for reproducible interactive
analysis and graphics of microbiome census data”, *PLOS ONE*, **8(4)**, e61217.
38. O’Brien, P.E., Smith, J., Stark, J.S., Johnstone, G., Riddle, M., Franklin, D. (2015)
“Submarine geomorphology and sea floor processes along the coast of Vestfold Hills, East
Antarctica, from multibeam bathymetry and video data”, *Antarctic Science*, **27(6)**, pp. 566-
586.
39. Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al.
Vegan: Community ecology package. R package version. 2018; 25-2.
40. Pitcher, C.R., Lawton P., Ellis, N., Smith, S.J., Incze, L.S., Wei, C., Greenlaw, M.E., Wolff,
N.H., Sameoto, J.A., and Snelgrove, P.V.R. (2012) “Exploring the role of environmental
variables in shaping patterns of seabed biodiversity composition in regional-scale
ecosystems”, *Journal of Applied Ecology*, **49**, pp. 670-679.
41. Quast, C. *et al.* (2013) “The SILVA ribosomal RNA gene database project: improved data
processing and web-based tools”, *Nucleic Acids Research*, **41(D1)**, pp. D590–D596. doi:
10.1093/nar/gks1219.
42. Rath, K.M., Fierer, N., Murphy, D.V., Rousk, J. (2018). Linking bacterial community
composition to soil salinity along environmental gradients. *The ISME Journal*, doi:
10.1038/s41396-018-0313-8.
43. Ray, A.E., Zhang, E., Terauds, A., Ji, M., Kong, W., Ferrari, B.C. (2020) “Soil

microbiomes with the genetic capacity for atmospheric chemosynthesis are widespread across the poles and are associated with moisture, carbon and nitrogen limitation”, *Frontiers in Microbiology*, **11**, pp. 1936.

44. Reysenbach, A.L., Pace, N.R., Robb, F.T., Place, A.R., editors. (1995) “Archaea: A laboratory manual - Thermophiles. CSHLP. Protocol, **16**, pp. 101-107.

45. Robinson, S.A., King, D.H., Bramley-Alves, J., Waterman, M.J., Ashcroft, M.B., Wasley, J., Turnbull, J.D., Miller, R.E., Ryan-Colton, E., Benny, T., Mullany, K., Clarke, L.J., Barry, L.A., Hua, Q. (2018) “Rapid change in east Antarctic terrestrial vegetation in response to regional drying”, *Nature Climate Change*, **8**, pp. 879-884.

46. Robinson, S.A., Klekociuk, A.R., King, D.H., Rojas, M.P., Zúñiga, G.S., Bergstrom, D.M. (2020) “The 2019/2020 summer of Antarctic heatwaves”, *Global Change Biology*, doi: 10.1111/gcb.15083.

47. Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, R.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Horn, D.J.V., and Weber, C.F. (2009) “Introducing mothur: Open-Source, platform-independent, community-supported software for describing and comparing microbial communities”, *Applied and Environmental Microbiology*, DOI: 10.1128/AEM.01541-09.

48. Siciliano, S.D., Palmer, A.S., Winsley, T., Lamb, E., Bissett, A., Brown, M.V., van Dorst, J., Ji, M., Ferrari, B.C., Grogan, P., Chu, H., and Snape, I. (2014) “Soil fertility is associated with fungal and bacterial richness, whereas pH is associated with community composition in polar soil microbial communities”, *Soil Biology & Biochemistry*, **78**, pp. 10-20.

49. Stephenson, F. et al., (2018) “Using Gradient Forests to summarise patterns in species turnover across large spatial scales and inform conservation planning”, *Biodiversity Research*, **24**, doi: 10.1111/ddi.12787.

- 468 50. Wickham, H. (2016) “ggplot2: Elegant graphics for data analysis”, Springer-Verlag, New
469 York.
- 470 51. Wilkins, D., Yau, S., Williams, T.J., Allen, M.A., Brown, M.V., DeMaere, M.Z., Lauro,
471 F.M., and Cavicchioli, R. (2013) “Key microbial drivers in Antarctic aquatic
472 environments”, *FEMS Microbiology Reviews*, **37**, pp. 303-335.
- 473 52. Zhang E., Thibaut L.M., Terauds A., Raven M., Tanaka M.M., van Dorst J., Wong SY.,
474 Crane S., Ferrari B.C. (2020) Lifting the veil on arid-to-hyperarid Antarctic soil
475 microbiomes: A tale of two oases. *Microbiome*, **8(37)**, doi:10.1186/s40168-020-00809-w.

476 MAIN FIGURES AND TABLES

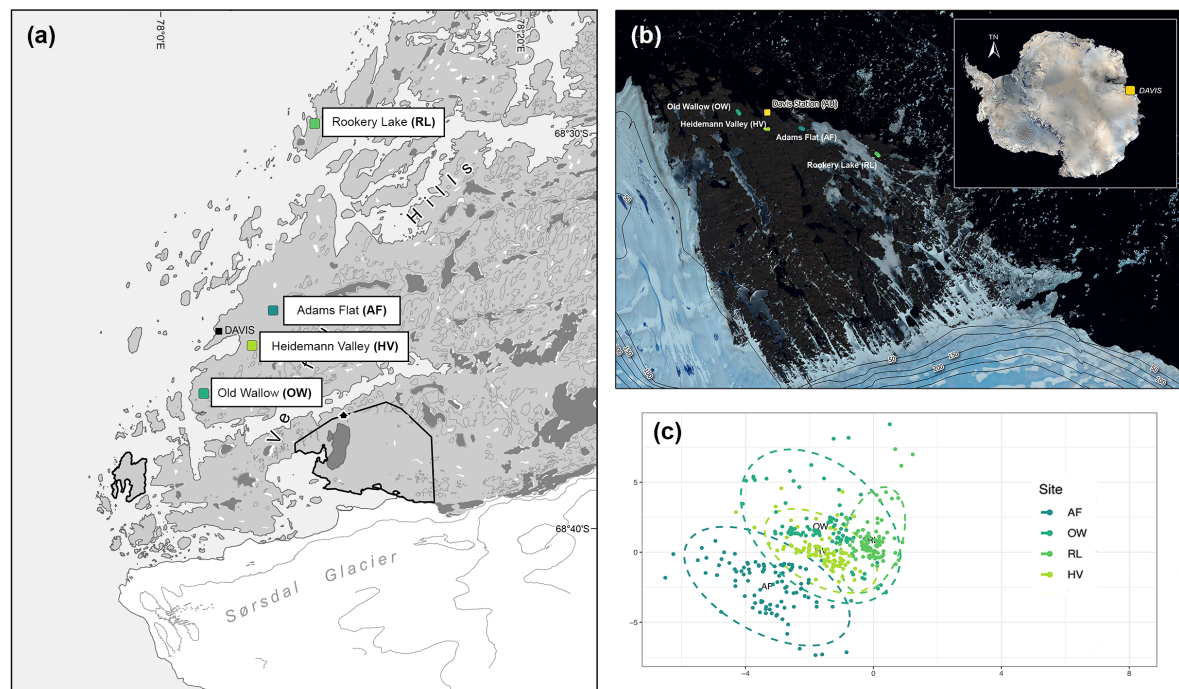


Figure 1. Overview of the four sampling sites (AF, HV, OW and RL) throughout the hyperarid Vestfold Hills in Eastern Antarctica as presented by (a) map (AAD map catalogue No. 14, 399), (b) satellite image (Quantarctica v3 GIS package in QGIS 3.4.7) and (c) environmental nMDS. Measured soil physiochemical variables (79) are generally conserved across the region with visible overlap amongst all sites, especially between OW, HV and RL whereas greater variability is observed for samples at AF.

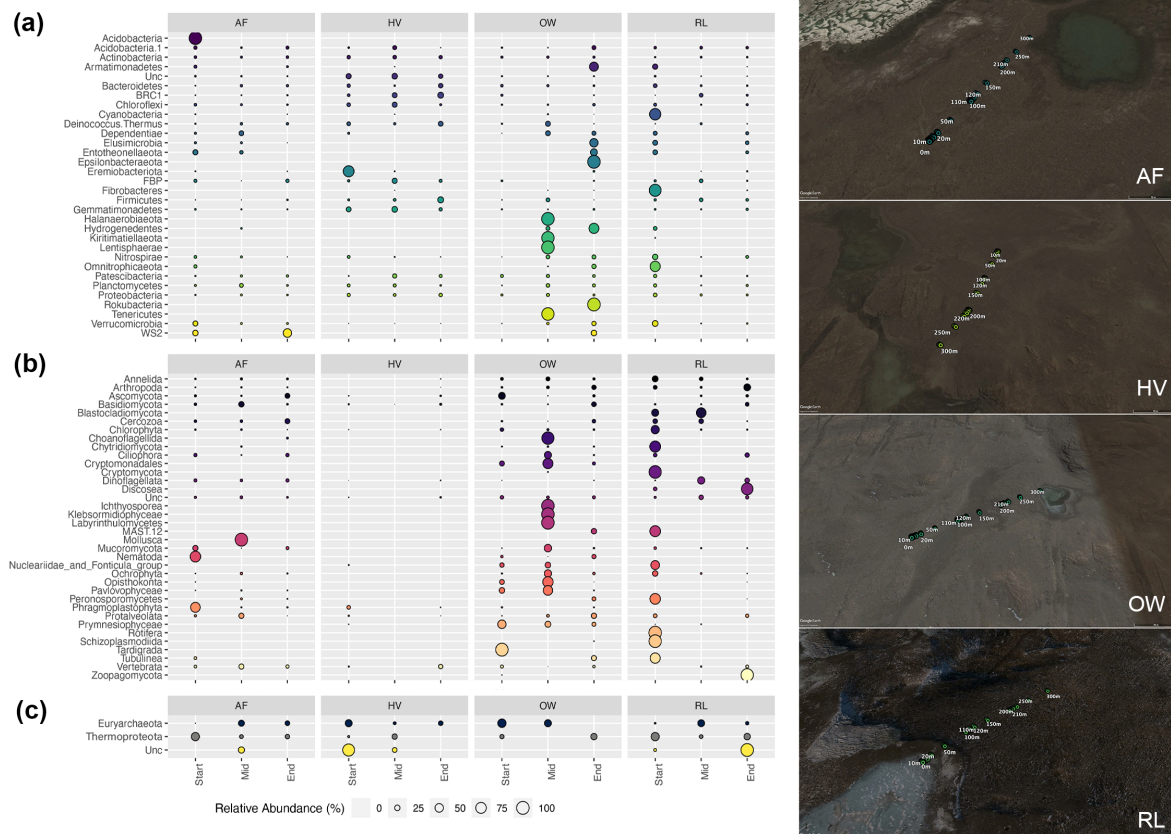


Figure 2. Relative distribution of each phylum across the Vestfold Hills region for polar soil (a) bacteria, (b) micro-eukarya and (c) archaea, where start (0/2 m), mid (100/102 m) and end (200/202 m) indicate the position of samples along transects at each site. Phyla are unevenly distributed between and within sites along their respective transects, with greatest phylogenetic diversity observed at OW (mid) and RL (start).

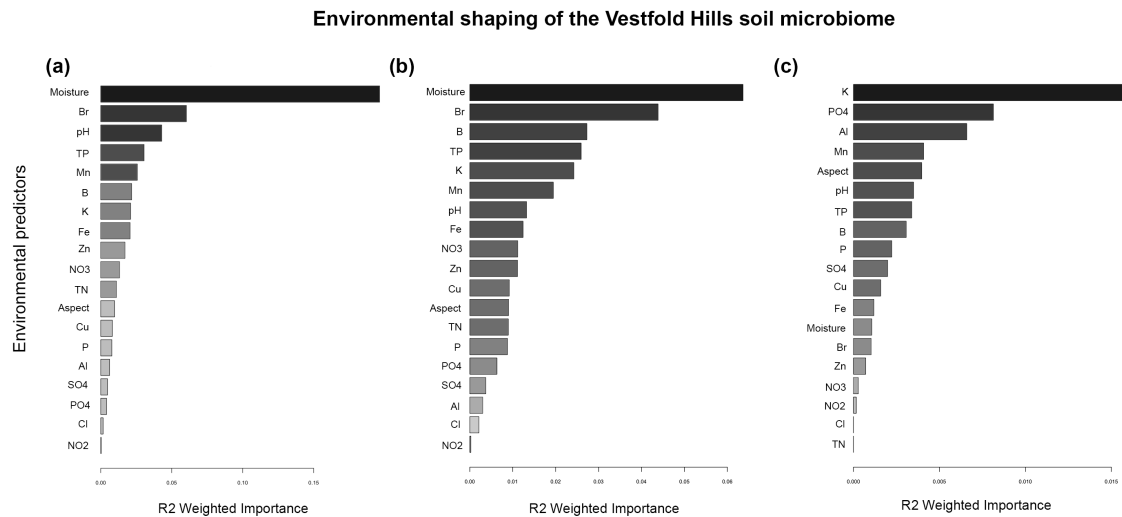


Figure 3. Overall importance (R^2 weighted) of predictors driving regional distributions of polar soil (a) bacteria, (b) micro-eukarya and (c) archaea across the Vestfold Hills. The suite of environmental variables predicted a moderate fraction of variation in regional species distribution. Although rank-order importance differed, most predictors were relatively consistent across the three domains, especially between bacteria and micro-eukarya. Soil moisture and nutrient-related factors (e.g., Br and K) along with pH were identified as the most robust predictors of regional diversity.

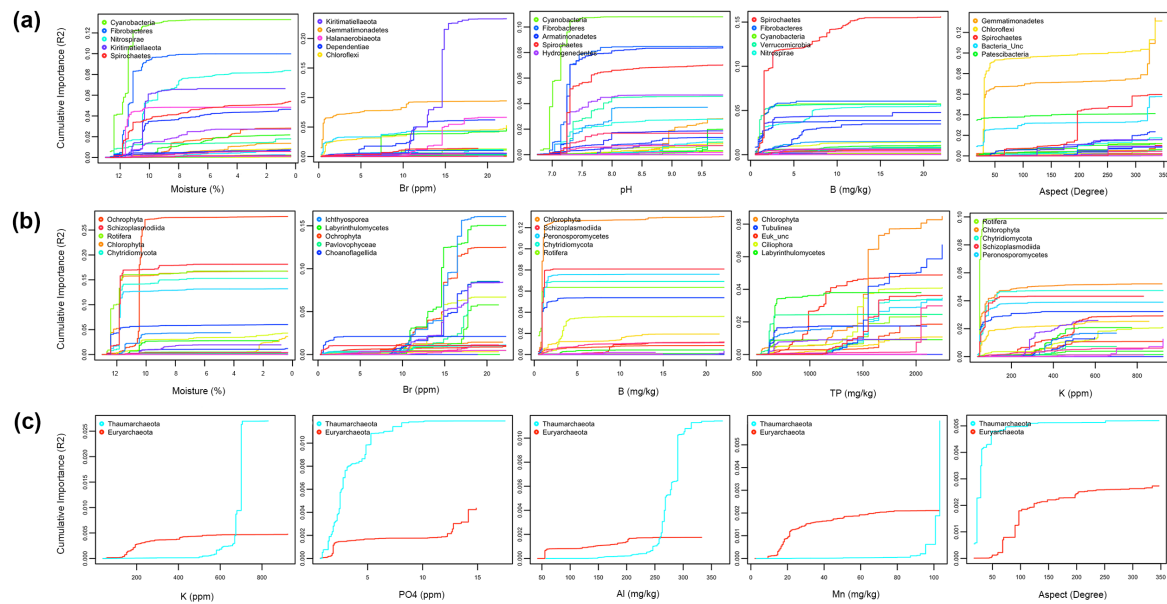


Figure 4. Cumulative plots visualising species turnover or “splits” along gradients of the top five most robust predictors for the regional distribution of polar soil (a) bacteria, (b) micro-eukarya and (c) archaea across the Vestfold Hills, where each line denotes a separate phylum and steeper slopes indicate higher rates of composition turnover. Splits along environments gradients were generally non-uniform, indicating variables rates of change in species composition. More pronounced transitions were observed for bacteria (e.g., *Kiritimatiellaeota*, *Cyanobacteria*, *Spirochetes* and *Gemmatimonadetes*) and micro-eukarya (e.g., *Ochrophyta*, *Ichthyosporae*, *Chlorophyta* and *Rotifera*). For bacteria, notable splits occurred between the following ranges: c. 5–15 bromide “Br” (ppm), c. 10–12 % moisture, c. 7.0–7.5 pH, c. 0–5 boron “B” (mg/kg) and at 50/200° aspect (degree). For micro-eukarya, notable splits occurred between the following ranges: c. 10–12 % moisture, c. 10–15 Br ppm, c. 2 B mg/kg, c. 1500–1600 total phosphorous “TP” (mg/kg) and c. 50 potassium “K” (ppm). For archaea, notable splits occurred between the following ranges: c. 600–700 K ppm, c. 0–5 phosphate “PO₄” (ppm), c. 250 aluminium “Al” (mg/kg), c. 100 manganese “Mn” (mg/kg) and c. 0/100° aspect.

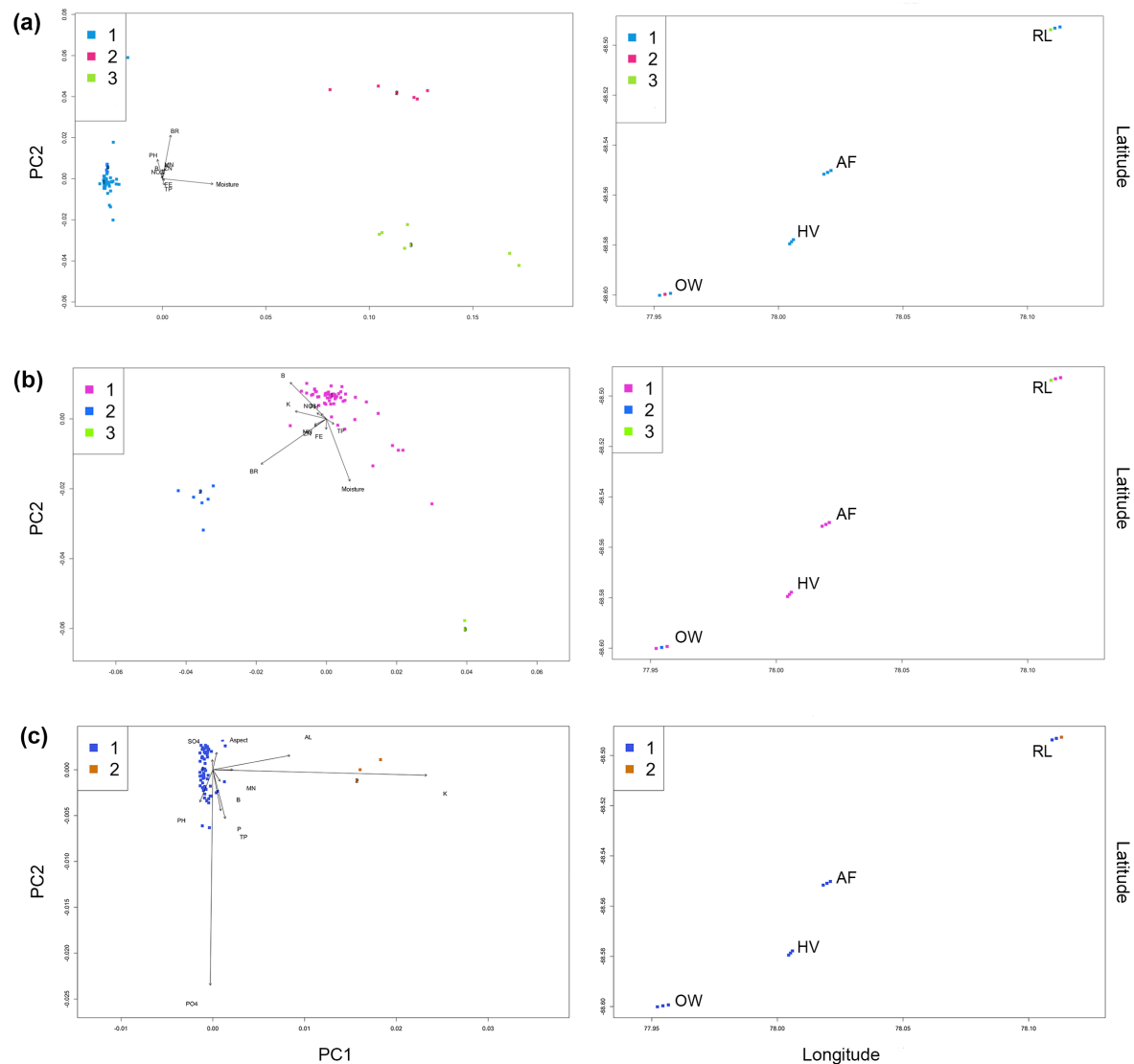


Figure 5. Biological and geographical bi-plots classifying spatial groups for polar soil (a) bacteria (3), (b) micro-eukarya (3) and (c) archaea (2) across the Vestfold Hills (n = 72), where different colours capture variations in species composition and turnover. Spatial patterns of turnover in soil biodiversity across the region described by cluster classification are strongly related to moisture content and nutrient-related factors, except for archaea. Greater heterogeneity in response groups were generally observed at OW (mid of transect for bacteria and micro-eukarya) and RL (start of transect for micro-eukarya and end of transect for archaea).

Table 1. Summary of complete environmental predictors averaged (n = 18) between sites. Where start (0/2 m), mid (100/102 m) and end (200/202 m) indicate position of samples (n = 6) along the transects and (*) denotes a significant ($p < 0.05$) difference for that variable between sites based on ANOSIM using 9999 permutations.

	AF			HV			OW			RL			ANOSIM	
	Start	Middle	End	Start	Middle	End	Start	Middle	End	Start	Middle	End	R Value	Significance
Aspect (Degree)	61.15	41.68	310.10	337.85	18.43	75.14	131.01	158.63	277.16	225.36	164.28	105.69	0.21	*
Conductivity (uS/cm)	2015.40	4820.50	1290.33	2768.50	583.33	1864.00	3958.78	14600.00	524.33	469.00	2917.50	5798.33	0.04	*
Elevation (m)	4.38	2.86	1.78	8.59	11.45	15.01	19.77	18.31	18.61	8.28	14.62	13.50	0.81	*
Moisture (%)	2.00	4.00	1.00	9.00	7.00	6.00	6.00	11.00	6.00	12.00	4.00	6.00	0.27	*
Slope (Degree)	0.83	0.97	0.37	1.79	2.94	1.11	3.54	0.29	2.01	2.48	3.96	3.30	0.45	*
pH	7.96	8.22	7.72	8.23	9.48	8.83	8.80	8.45	8.05	7.00	8.77	8.10	0.09	*
TC (mg/kg)	1564.30	1795.50	986.93	1545.83	1550.50	2044.83	4157.50	4447.50	3223.00	1195.12	1459.33	1589.33	0.21	*
TN (mg/kg)	171.60	120.17	116.83	136.33	145.50	138.33	285.50	266.67	358.33	124.83	221.67	158.33	0.15	*
TP (mg/kg)	1244.00	938.33	833.33	743.33	748.33	696.67	788.33	648.33	686.67	1900.00	1600.00	1466.67	0.45	*
Al (mg/kg)	96.03	135.43	82.80	197.25	194.66	159.78	147.95	188.51	216.94	173.02	151.55	273.42	0.20	*
B (mg/kg)	4.11	5.51	3.51	3.67	7.10	5.48	9.79	3.25	1.70	0.97	10.02	4.06	0.02	*
Ca (mg/kg)	329.34	2299.63	316.90	601.60	693.92	2219.56	5782.99	3670.79	396.78	173.90	1795.49	557.96	0.06	*
Cu (mg/kg)	5.33	4.87	11.44	6.59	8.25	11.44	8.99	13.35	11.87	7.81	7.40	6.32	0.16	*
Fe (mg/kg)	275.37	333.10	215.77	213.41	174.55	199.12	243.42	226.68	251.97	300.62	158.59	199.00	0.05	*
K (mg/kg)	118.58	217.38	80.40	164.00	261.35	191.21	195.96	440.01	88.79	103.36	460.62	692.49	0.15	*
Mg (mg/kg)	154.85	546.72	223.59	471.73	307.32	568.99	697.51	2963.35	383.81	193.26	484.50	750.64	0.11	*
Mn (mg/kg)	9.51	7.35	6.10	15.76	22.73	22.60	28.97	56.70	16.45	32.87	21.01	76.13	0.30	*
P (mg/kg)	9.85	5.20	12.30	20.76	21.16	13.72	9.53	10.58	6.78	25.85	63.11	19.30	0.30	*
S (mg/kg)	269.81	1024.98	233.21	196.87	99.92	1173.89	3881.96	264.35	35.21	38.12	1048.45	328.13	0.00	*
Zn (mg/kg)	0.82	1.83	0.89	2.26	1.21	1.29	1.52	2.97	1.29	1.66	1.38	1.73	0.12	*
Cation Exchange Capacity	34.32	58.21	31.30	42.22	33.33	43.13	70.45	117.35	30.57	28.18	48.71	61.59	0.06	*
Cl (ppm)	1531.91	5622.63	817.52	3196.87	317.26	945.63	2902.92	3449.04	79.32	332.59	868.48	3138.66	0.01	*
Br (ppm)	1.52	8.93	1.47	0.37	0.15	0.20	0.91	17.71	0.40	1.20	2.18	3.89	0.18	*
NO ₃ (ppm)	8.99	17.68	12.64	5.25	0.82	1.53	13.96	2.22	2.91	1.55	26.75	34.00	0.09	*
NO ₂ (ppm)	0.38	0.36	0.26	0.43	0.15	0.33	0.18	0.15	0.15	0.15	0.25	0.37	0.01	*
PO ₄ (ppm)	1.95	0.15	3.93	2.30	5.56	1.49	0.98	0.15	1.57	4.89	7.14	2.28	0.11	*
SO ₄ (ppm)	726.99	1836.16	578.58	606.20	242.95	947.53	3575.38	576.60	69.33	77.83	362.88	497.56	0.02	*
Sorting Order	0.42	0.45	0.48	0.30	0.25	0.29	0.25	0.29	0.29	0.34	0.26	0.25	0.37	*
Gravel (%)	7.22	7.82	11.12	11.17	9.22	16.47	8.26	5.44	9.44	6.27	14.90	15.78	0.04	*
Sand (%)	81.23	86.21	84.00	58.49	61.96	65.12	75.28	55.68	71.12	73.48	64.67	54.79	0.27	*
Mud (%)	11.55	5.98	4.88	30.34	28.82	18.41	16.45	38.88	19.45	20.25	20.42	29.43	0.26	*
Min Grain Size (uM)	8.67	12.31	19.17	2.23	1.85	2.57	4.00	1.31	3.80	5.28	2.13	1.47	0.22	*
Max Grain Size (uM)	934.31	1029.35	1301.10	838.88	831.73	1044.43	1082.24	520.00	902.19	936.33	1001.42	1030.22	0.10	*
Al ₂ O ₃ (%)	13.92	13.43	14.12	13.42	13.50	13.42	13.03	12.57	13.70	14.99	14.31	15.09	0.46	*
CaO (%)	5.28	5.54	5.07	5.16	5.07	5.51	5.61	5.38	4.73	5.65	5.50	4.81	0.05	*
Fe ₂ O ₃ (%)	8.13	8.19	7.28	8.96	9.59	9.72	9.73	8.18	10.44	9.73	9.53	9.93	0.35	*
MgO (%)	4.16	4.19	3.87	4.52	4.82	4.76	4.79	4.72	4.84	4.21	4.40	4.80	0.35	*
MnO (%)	0.12	0.12	0.11	0.13	0.14	0.15	0.14	0.13	0.14	0.12	0.12	0.13	0.19	*
P ₂ O ₅ (%)	0.24	0.19	0.19	0.20	0.19	0.18	0.20	0.18	0.18	0.46	0.36	0.36	0.41	*
K ₂ O (%)	1.53	1.41	1.50	1.57	1.63	1.47	1.45	1.63	1.57	2.05	2.25	2.53	0.42	*
Na ₂ O (%)	3.52	3.68	3.54	3.18	2.91	2.84	2.88	4.03	2.75	3.82	3.66	4.17	0.39	*
SiO ₂ (%)	62.17	62.02	63.71	61.97	61.13	60.95	59.75	60.84	60.64	58.21	58.47	56.83	0.44	*
TiO ₂ (%)	0.85	0.82	0.74	0.84	0.89	0.85	0.87	0.76	0.90	1.16	1.05	1.00	0.39	*
SO ₃ (%)	0.09	0.36	0.08	0.09	0.06	0.30	1.05	0.10	0.05	0.06	0.47	0.12	0.00	*

Table 2. Top five most robust predictors for local distributions of polar soil bacteria, micro-eukarya and archaea. Where shaded boxes indicate sharing of predictor(s) across two or more domains.

	Bacteria	Micro-Eukarya	Archaea
AF	Aspect	Ca	Zn
	Moisture	Aspect	pH
	Ca	Moisture	Ca
	Zn	S	NO ₂
	S	Fe	SO ₄
HV	PO ₄	TP	Aspect
	Mg	Cl	pH
	P	SO ₄	PO ₄
	Cation Exchange Capacity	Aspect	Mn
	TC	P	Fe
OW	Conductivity	Conductivity	Conductivity
	Aspect	Aspect	B
	pH	P	NO ₃
	P	NO ₃	P
	B	B	Aspect
RL	Fe	B	B
	B	S	NO ₃
	pH	Aspect	S
	Aspect	Fe	P
	S	pH	TP