environmental microbiology

Environmental Microbiology (2022) **00**(00), 00–00 doi:10.1111/1462-2920.15912

The ecological assembly of bacterial communities in Antarctic wetlands varies across levels of phylogenetic resolution

María V. Quiroga,[1*†] **Angel Valverde**,[2*†] **Gabriela Mataloni**,[3] **Valeria Casa**,[3] **James C. Stegen**[4] **and Don Cowan**[5]

[1]*Instituto Tecnológico de Chascomús (INTECH, UNSAM – CONICET), Chascomús, Argentina.*

[2]*Instituto de Recursos Naturales y Agrobiología de Salamanca (IRNASA-CSIC), Consejo Superior de Investigaciones Científicas, Salamanca, Spain.*

[3]*Instituto de Investigación e Ingeniería Ambiental (IIIA, UNSAM-CONICET), San Martín, Buenos Aires, Argentina.*

[4]*Pacific Northwest National Laboratory, Ecosystem Science Team, Richland, WA.*

[5]*Centre for Microbial Ecology and Genomics (CMEG), Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa.*

Summary

As functional traits are conserved at different phylogenetic depths, the ability to detect community assembly processes can be conditional on the phylogenetic resolution; yet most previous work quantifying their influence has focused on a single level of phylogenetic resolution. Here, we have studied the ecological assembly of bacterial communities from an Antarctic wetland complex, applying null models across different levels of phylogenetic resolution (i.e. clustering ASVs into OTUs with decreasing sequence identity thresholds). We found that the relative influence of the community assembly processes varies with phylogenetic resolution. More specifically, selection processes seem to impose stronger influence at finer (100% sequence similarity ASV) than at coarser (99%–97% sequence similarity OTUs) resolution. We identified environmental features related with the ecological processes and propose a conceptual model for the bacterial community assembly in this Antarctic ecosystem. Briefly, eco-evolutionary processes appear to be leading to different but very closely related ASVs in lotic, lentic and terrestrial environments. In all, this study shows that assessing community assembly processes at different phylogenetic resolutions is key to improve our understanding of microbial ecology. More importantly, a failure to detect selection processes at coarser phylogenetic resolution does not imply the absence of such processes at finer resolutions.

Received 22 September, 2021; revised 10 January, 2022; accepted 13 January, 2022. For correspondence. *E-mail mvquiroga@iib.unsam.edu.ar. **E-mail angel.valverde@csic.es; Tel:+34 923269606. †These authors contributed equally to this work.

Introduction

Understanding the processes governing community assembly is a key topic in microbial ecology (Vellend, 2016). It is now recognized that community assembly is dictated by the interaction of four major ecological and evolutionary processes (Vellend, 2010): selection, dispersal, drift and speciation, which collectively contribute to the assembly of microbial communities (Lindström and Langenheder, 2012; Vellend *et al*., 2014; Dini-Andreote *et al*., 2015; Stegen *et al*., 2015). Selection refers to deterministic changes in community structure due to fitness differences among organisms (Vellend, 2010; Stegen *et al*., 2015), and both abiotic features and biotic interactions relate to fitness. The type of selection will depend on the spatial pattern of environmental conditions. Homogeneous conditions will impose consistent selective pressure leading to low phylogenetic turnover, referred to as 'homogeneous selection' (Stegen *et al*., 2013, 2015; Dini-Andreote *et al*., 2015). In contrast, heterogeneous environmental conditions will promote variable selective pressures causing high phylogenetic turnover, referred to as 'variable selection'. High dispersal rates can potentially promote biotic homogenization (homogenizing dispersal) leading to low taxonomic turnover; whereas low dispersal rates (dispersal limitation) can in turn result in high taxonomic turnover due to ecological drift (Vellend, 2010; Stegen *et al*., 2013, 2015). Finally, speciation is the evolution of new species. Therefore, under this framework 'species are added to

communities via speciation and dispersal, and the relative abundances of these species are then shaped by drift and selection, as well as ongoing dispersal, to drive community dynamics' (Vellend, 2010).

One of the approaches most commonly used to investigate the relative influence of the ecological components of community assembly process is that developed by Stegen *et al*. (2013, 2015), which uses null models. However, it has been proposed that the ability to detect assembly processes can be conditional on the phylogenetic resolution (Hanson *et al*., 2012), because functional traits are conserved at different phylogenetic depths (Martiny *et al*., 2015). For example, pH and salinity optimum are usually shared among taxa within deep clades (Martiny *et al*., 2015), therefore selection pressures related to these features could be identified at coarser phylogenetic resolutions. In contrast, long-term drought response and temperature preference are shallowly conserved (Martiny *et al*., 2015), potentially allowing the detection of selection pressures imposed by these features at finer phylogenetic resolutions. Yet, only a single phylogenetic resolution [mainly ASVs or 97% similarity operational taxonomic units (OTUs)] is typically used for input data to investigate bacterial assembly processes (Langenheder *et al*., 2017; Logares *et al*., 2018, 2020; Allen *et al*., 2020; Danczak *et al*., 2020; Huber *et al*., 2020; Ji *et al*., 2020; Kraemer *et al*., 2020). For example, the use of null models with 97% similarity OTUs revealed a predominance of selection processes in grassland soils (Ji *et al*., 2020) and Antarctic lakes (Logares *et al*., 2018), while assembly processes in the Lake Kitkajärvi were apparently not dominated by any particular process (Langenheder *et al*., 2017). Conversely, the finer resolution of ASVs revealed strong selective pressures in the South Pacific Gyre marine microbiome (Allen *et al*., 2020), but weak selection and dispersal processes in global surface waters (Logares *et al*., 2020). Furthermore, using ASVs, selection was found to be the main assembly force in the floodplain of the Paraná River (Huber *et al*., 2020); while dispersal processes dominated in Eastern Canada lakes (Kraemer *et al*., 2020), and fractured shale ecosystems displayed scenarios not dominated by selection nor dispersal (Danczak *et al*., 2020). New frameworks have allowed to quantify community assembly processes within different phylogenetic groups (bins; Ning *et al*., 2020) or clades (Fodelianakis *et al*., 2021). However, it is still not clear whether the relative influence of these processes in bacterial community assembly changes with taxon phylogenetic resolution. Although previous studies of arbuscular mycorrhizal fungi (Roy *et al*., 2019) and vascular plants (Swenson *et al*., 2006) found that the relative influences of environmental selection can change with phylogenetic scale, studies are needed to evaluate if bacteria show similar ecological patterns.

Bacterial phylogenetic diversity studies have typically involved the clustering of sequences into OTUs with a fixed threshold of 97% sequence identity, considered to correspond approximately to species (Schloss and Handelsman, 2004), although several authors have proposed higher and sometimes dynamic cut-offs (Yarza *et al*., 2014; Mysara *et al*., 2017; Edgar, 2018). The definition of bacterial species and its relevance as the most significant unit in microbial ecology are still debated (Rosselló-Móra and Amann, 2015). More recently, to achieve a finer phylogenetic resolution, new methods have been developed for modelling and correcting Illumina-sequenced amplicon errors (Callahan *et al*., 2016, 2017) that allowed the discrimination of amplicon sequence variants (ASVs), which may diverge from one another in only one nucleotide. ASVs can then be clustered into OTUs using fixed sequence identity thresholds in order to study intra-species microdiversity (García-García *et al*., 2019). This approach provides an opportunity to evaluate the influences of ecological processes on bacterial community assembly at different phylogenetic resolutions.

The Antarctic continent is subjected to extreme climatic conditions (Hughes *et al*., 2015), potentially imposing strong selection pressures to its microbial communities. Cierva Point Wetland Complex (CPWC) is a macrobiodiversity hotspot on the north-west coast of the Antarctic Peninsula (Agraz *et al*., 1994; Antarctic Treaty Secretariat, 2013; Wilhelm *et al*., 2016). Its complex, fragmented topography defines a mosaic of distinct environmental units characterized by different combinations of land cover, slope and orientation, most hosting a large number of different environments (Agraz *et al*., 1994), which in this study have been grouped into lentic, lotic and terrestrial environments for simplicity. The complex is completely covered by snow from April to December but is mostly snow-free during the austral summer (Wilhelm *et al*., 2016). It has been shown that slope angle determines the extent and direction of hydrological connectivity during snow melt or rain events in this system (Mataloni *et al*., 2005, 2010). In addition, this protected area hosts an increasingly large colony of gentoo penguin (*Pygoscelis papua*) (González-Zevallos *et al*., 2013) that contribute to nutrient input and may contribute to the dispersal of microorganisms across the wetland complex.

Here, we aimed to study the relative influence of the ecological processes shaping the CPWC bacterial metacommunity and how the assessment of these processes may vary across different levels of phylogenetic resolution. We used high-throughput sequencing of 16S rRNA genes and applied the null models proposed by Stegen *et al*. (2013, 2015) to the phylogenetic data, implementing the approach of García-García *et al*. (2019) by clustering ASVs into OTUs with decreasing sequence identity thresholds

(i.e. 99%, 98%, 97% and 94% similarity OTUs). We hypothesize that the mosaic of different environments is subject to spatially varying environmental conditions, which will impose high variable selection in the different CPWC microbiomes; that is, each environment will select for different taxa. In addition, we expect dispersal to be enhanced across the different local communities during the austral summer period by snow melt, rain events and penguin movements, resulting in the homogenization of the metacommunity (i.e. homogenizing dispersal). We therefore expect to observe simultaneous influences of variable selection and homogenizing dispersal. These opposing forces may, however, lead to a situation in which individual processes cannot be discerned because assembly is not dominated by a single process. Furthermore, as Hanson *et al*. (2012) suggested that the ability to detect the ecological processes shaping microbial biogeographic patterns can be conditional on the phylogenetic resolution of the study, we hypothesize that finer phylogenetic resolutions might unveil selection patterns not detected at the coarser resolutions. Cohan (2001) defined an ecotype as a population of bacterial cells, corresponding to a DNA sequence cluster, adapted to a given ecological niche. As ASVs (i.e. 100% similarity OTUs) can detect ecotypes within the same species (García-García *et al*., 2019), and ecotypes may partition niche space within the environment (Martiny *et al*., 2006), we predict that ASV diversity patterns would be more influenced by selection processes.

Results

Bacterial community composition and structure

Sampling at CPWC (ca. area: 1 km^2) was carried out over the early 2018 Antarctic summer. A total of 64 samples were collected from three types of environments: lentic environments (ponds/lakes: 22 samples), lotic environments (streams, seepages and wet rocks: 18 samples) and terrestrial environments (soils, mosses and snow: 24 samples) (Fig. 1, Supporting Information Table S1). Since we clustered ASVs into OTUs with decreasing sequence identity thresholds, we will refer to 100% similarity OTUs as 'ASV', 99% similarity OTUs as 'OTU99', 98% similarity OTUs as 'OTU98', 97% similarity OTUs as 'OTU97' and 94% similarity OTUs as 'OTU94'.

Rarefaction curves for normalized ASVs counts for the majority of the samples (62 out of 64) reached a plateau, suggesting that the sequencing depth captured most of the diversity of the local communities (Supporting Information Fig. S1). The mean Shannon diversity values were similar in all three environments at each phylogenetic resolution (global Kruskal–Wallis tests: all $P > 0.05$, Supporting Information Fig. S2); whereas lentic environments showed higher mean number of taxa (richness) than lotic environments at OTU99, OTU98 and OTU94 resolutions (global Kruskal–Wallis tests: all $P < 0.05$; pairwise Wilcoxon Tests: all $P < 0.04$). The metacommunity of the CPWC was dominated by Proteobacteria (30%), Bacteroidota (27%), Actinobacteriota (21%) and Firmicutes (12%) (Supporting Information Fig. S3).

As required by the null modelling approach, we tested the phylogenetic signal at each taxonomic resolution before applying the null models. The OTU94 data did not show a phylogenetic signal (Supporting Information Fig. S4), and this resolution was therefore excluded from subsequent downstream analyses. The permutational analysis of variance (PERMANOVA) analyses based on phylogenetic dissimilarity (βMNTD) showed that the bacterial assemblages from lentic and lotic environments were

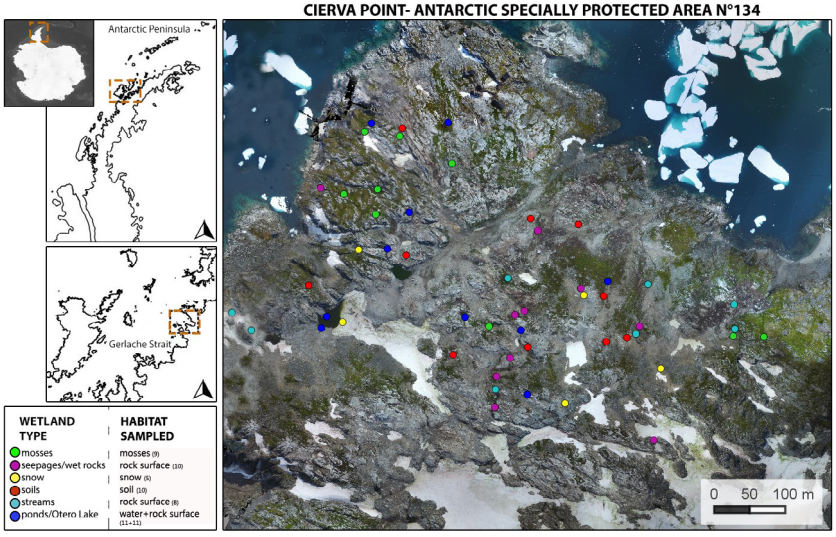

Fig. 1. Location of sampled sites within the CPWC. Modified from Ramos Marín (2018).

significantly different from those of terrestrial environments (all $P < 0.01$, detailed R^2 and P values are included in Supplementary Table S2) at each phylogenetic resolution, although this was not clearly visualized in the principal coordinate plots (Fig. 2). Overall, βMNTD values were quite low (0.08, 0.06, 0.05 and 0.04 on average for ASV, OTU99, OTU98 and OTU97 levels respectively, Supporting Information Fig. S5), reflecting the general low phylogenetic variability between the local communities. There were no significant differences in community heterogeneity between environments at each phylogenetic resolution (beta dispersion, all $P > 0.1$). In addition, no relationships were observed between phylogenetic dissimilarity (βMNTD) and geographic distance among samples (Mantel tests, $P > 0.1$ for all phylogenetic resolutions).

Assembly processes of the bacterial metacommunity

ASV, OTU99, OTU98 and OTU97 showed significant positive correlations over short phylogenetic distances (Mantel correlograms, Supporting Information Fig. S4), indicating that closely related taxa at these phylogenetic resolutions shared similar environmental optima.

The null models make use of βNTI (β-nearest taxon) and RC_{bray} (Raup–Crick metric using Bray–Curtis

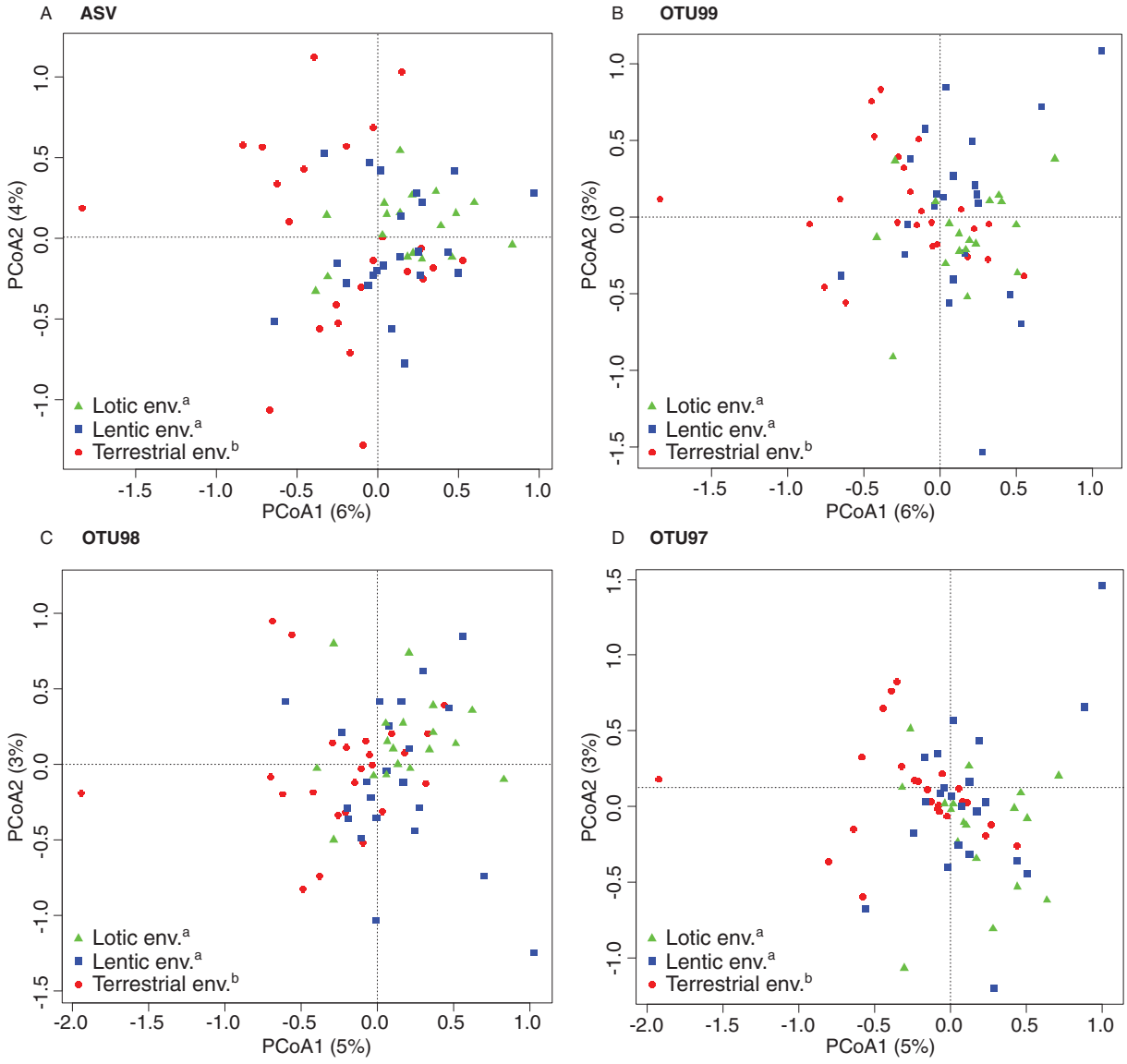

Fig. 2. PCoA based on phylogenetic dissimilarities (βMNTD) between bacterial communities at ASV (A), OTU99 (B), OTU98 (C) and OTU97 (D) phylogenetic resolutions in the CPWC. Different superscript letters in the legends of the ordinations indicate significant differences between environments (PERMANOVA, $P < 0.01$, detailed R^2 and P values are included in Supplementary Table S2). Env., environments.

dissimilarities) indices to quantify the degree to which observed phylogenetic and taxonomic turnover respectively, deviate from the null model expectation. Remarkably, a significantly lower mean βNTI value was observed for ASV data, an intermediate value for OTU99 data and higher mean βNTI values for OTU98 and OTU97 data (global Kruskal–Wallis test: $P < 0.001$; significant pairwise Wilcoxon Tests: $P < 0.001$, Fig. 3A). The opposite trend was observed for RC_{bray} values (global Kruskal–Wallis test: $P < 0.001$; significant pairwise Wilcoxon Tests: $P < 0.004$, Fig. 3B).

Based on these two indices we quantified the relative influence of homogeneous selection, variable selection, homogenizing dispersal and dispersal limitation for each phylogenetic resolution (Table 1). The term 'undominated' was used when neither selection nor

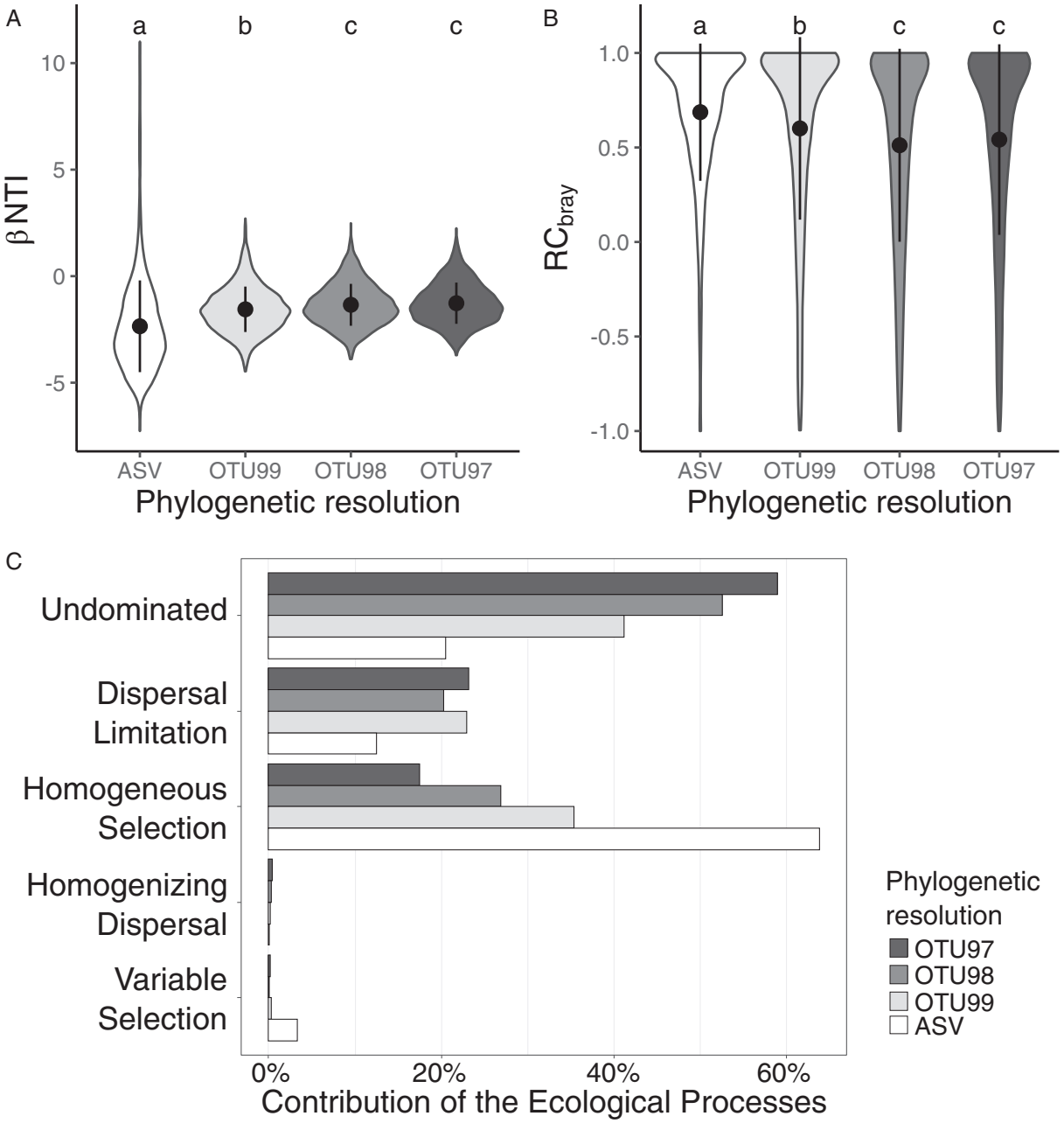

Fig. 3. Violin plots of (A) βNTI and (B) RC_{bray} indices across phylogenetic resolutions. (C) Contribution of the ecological processes shaping bacterial community structure across phylogenetic resolutions. (A, B): black points indicate the mean value; bars represent ±1 standard deviation. Differences between taxonomic resolutions were evaluated with global Kruskal–Wallis test and Mann–Whitney *post hoc* pairwise comparisons applying Bonferroni correction. Different letters indicate significant differences in mean between phylogenetic resolutions ($P < 0.05$).

Table 1. Microbial community assembly processes according to Stegen *et al*. (2013, 2015).

βNTI	RC_{bray}	Interpretation	Assembly process
<−2	−	Less than expected phylogenetic turnover	Homogeneous selection
>+2	−	Greater than expected phylogenetic turnover	Variable selection
<|2|	<−0.95	Less than expected taxonomic turnover	Homogenizing dispersal
<|2|	>+0.95	Greater than expected taxonomic turnover	Dispersal limitation
<|2|	<|0.95|	Neither selection nor dispersal is the dominant process	Undominated

dispersal was the dominant assembly process (Stegen *et al*., 2015). Strikingly, we observed different relative contributions of the ecological processes influencing bacterial community assembly, depending on the phylogenetic resolution (Fig. 3C). The relative influence of homogeneous selection increased at finer resolutions (i.e. increasing sequence identity thresholds), while the 'undominated' component showed the opposite pattern. Accordingly, the metacommunity structure based on the coarser resolution (i.e. OTU97) revealed a system not dominated by any particular process; while homogeneous selection (>60% contribution) strongly influenced the bacterial community structure at the ASV level.

Environmental factors contributing to assembly processes

The factors imposing selection and dispersal limitations were identified, independently for each phylogenetic resolution, using a distance-based redundancy analysis (dbRDA) model selection procedure with either βNTI (Fig. 4) or RC_{bray} (Fig. 5) matrices as the response variables (for environmental data see Supporting Information Table S3). Using βNTI distances, the contribution of the measured environmental factors in shaping the βNTI values increased toward finer phylogenetic resolutions, and pH was consistently identified as a system feature related to selection processes across all resolutions (Fig. 4). More specifically, PERMANOVA analysis showed that βNTI-OTU97 variation was significantly (though weakly) explained by pH ($R^2 = 0.08$, $P = 0.002$); the βNTI-OTU98 variation was significantly explained by pH ($R^2 = 0.08$, $P = 0.001$) and conductivity ($R^2 = 0.04$, $P = 0.022$); the βNTI-OTU99 variation was significantly explained by pH ($R^2 = 0.11$, $P = 0.001$), conductivity ($R^2 = 0.04$, $P = 0.021$) and slope ($R^2 = 0.03$, $P = 0.046$); and the βNTI-ASV variation was significantly explained by pH ($R^2 = 0.20$, $P = 0.001$), conductivity ($R^2 = 0.09$, $P = 0.001$) and penguin impact ($R^2 = 0.05$, $P = 0.006$).

Using RC_{bray} distances, we found that there was dispersal limitation between the different environments across all phylogenetic resolutions (Fig. 5). PERMANOVA analyses confirmed that RC_{bray} variation was significantly (though weakly) explained by the type of environment ($R^2 = 0.10$, $R^2 = 0.10$, $R^2 = 0.10$, $R^2 = 0.12$ (all $P = 0.001$) for OTU97, OTU98, OTU99 and ASV data respectively). Penguin impact ($R^2 = 0.06$, $R^2 = 0.04$, $R^2 = 0.05$ (all $P < 0.002$) for OTU97, OTU98 and OTU99 levels respectively) and conductivity ($R^2 = 0.04$, $P = 0.003$ for OTU97 data) significantly (though weakly) explained RC_{bray} variation.

Discussion

Community assembly processes vary across levels of phylogenetic resolution

Our work has demonstrated that the relative influence of different assembly processes changes across the ASV-OTU97 levels, in agreement with Hanson *et al*. (2012). These results are also in line with the work by Roy *et al*. (2019), which suggests that the relative influence of the ecological drivers of phylogenetic beta diversity patterns of arbuscular mycorrhizal fungi varies with taxon phylogenetic resolution. The shift in null model outcomes was most striking for the phylogenetically informed analyses (i.e. βNTI), with significant deviations from the stochastic expectation becoming much more common toward the finer levels of phylogenetic resolution. This corroborated our hypothesis, as finer phylogenetic resolutions unveiled selection processes within the CPWC not detected at the coarser resolutions.

The strong influence of selection processes at the ASV level appears to reflect environmental filters/constraints acting at this finer sub-species level, which is supported by the high phylogenetic signal of ASVs (Supporting Information Fig. S4). As ASVs can represent ecotypes within species (García-García *et al*., 2019) and display niche partitioning (Martiny *et al*., 2006), we propose that the CPWC ASV patterns may reflect a significant influence of microevolutionary processes in the assembly of this bacterial metacommunity. As the 16S rRNA genes are too conserved to detect recent evolutionary changes (see Chase *et al*., 2021 and references therein) further studies are needed to corroborate this.

Conversely, the differences in phylogenetic signal patterns across phylogenetic resolutions contradict the

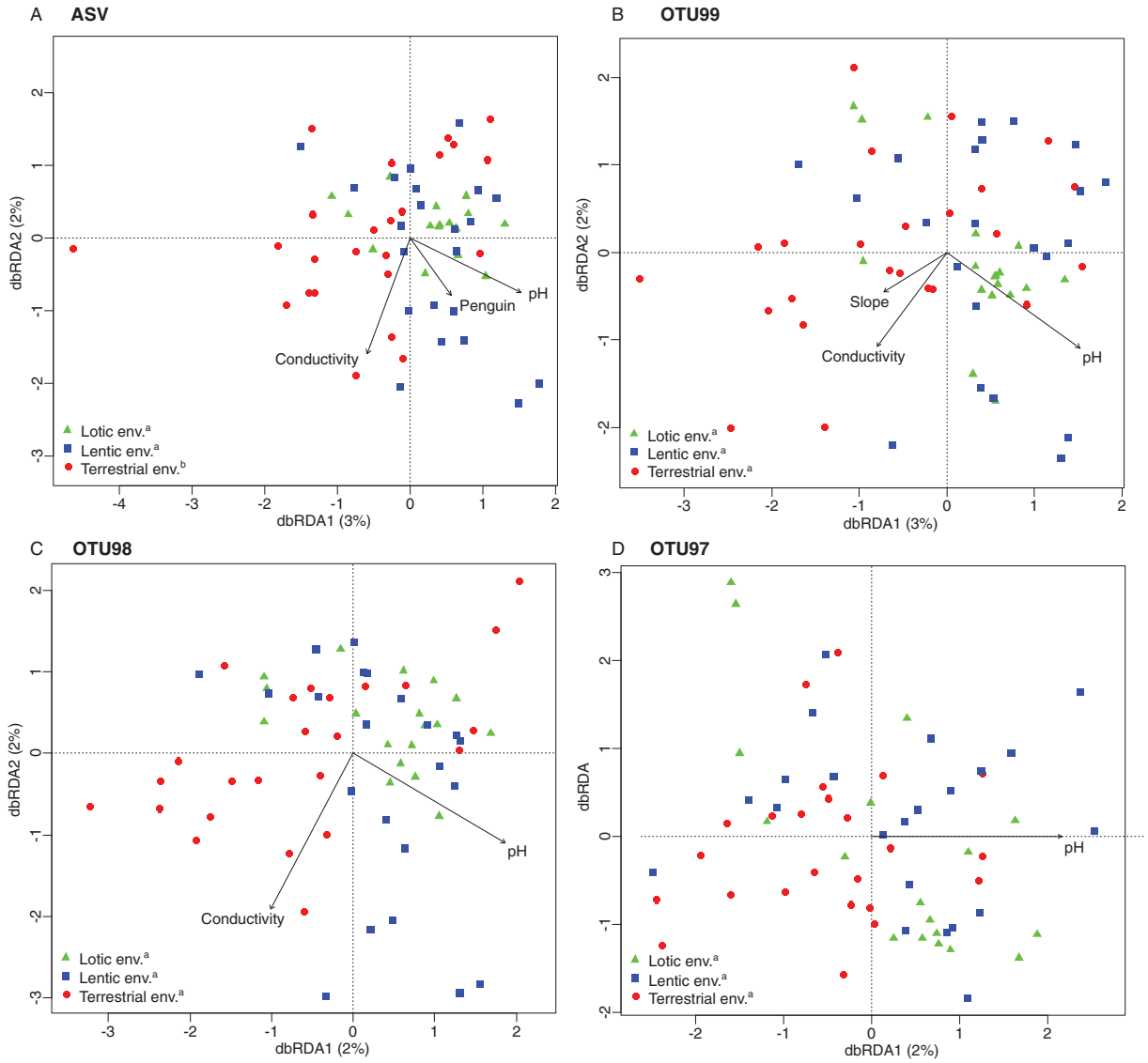

Fig. 4. Distance-based redundancy analyses (dbRDA) based on ßNTI matrices for ASV (A), OTU99 (B), OTU98 (C) and OTU97 (D) phylogenetic resolutions. Different superscript letters in the legends of the ordinations indicate significant differences between environments (PERMANOVA, $P < 0.02$, detailed R^2 and P values are included in Supplementary Table S2). Env., environments.

hypothesis of a consistent level of niche conservatism from finer (i.e. species) to broader (i.e. phylum) resolution (Lu *et al*., 2016). This indicates a lack of ecological coherence across deep prokaryotic evolutionary relationships, consistent with previous theoretical arguments (Stegen *et al*., 2012). The weaker phylogenetic signal at the OTU98 and OTU97 levels could have prevented the phylogenetic null models from detecting selection at these resolutions. However, the RC_{bray} null model is not sensitive to the phylogenetic signal, and we would expect that if selection was strong there would be a consistent deviation from the associated stochastic expectation (i.e. consistently significant values of RC_{bray} at the OTU98 and OTU97 levels). Therefore, the lack of consistent deviation from either βNTI or RC_{bray} null models at these levels indicates that no single process dominated community assembly when communities were analyzed at such coarser resolutions.

Environmental variables related to selection processes

We hypothesized that the mosaic of different environments from the CPWC would lead to spatially heterogeneous environmental conditions, and impose high variable selection, with each environment selecting for different taxa. However, we did not detect variable selection, and therefore rejected our hypothesis. Instead, we

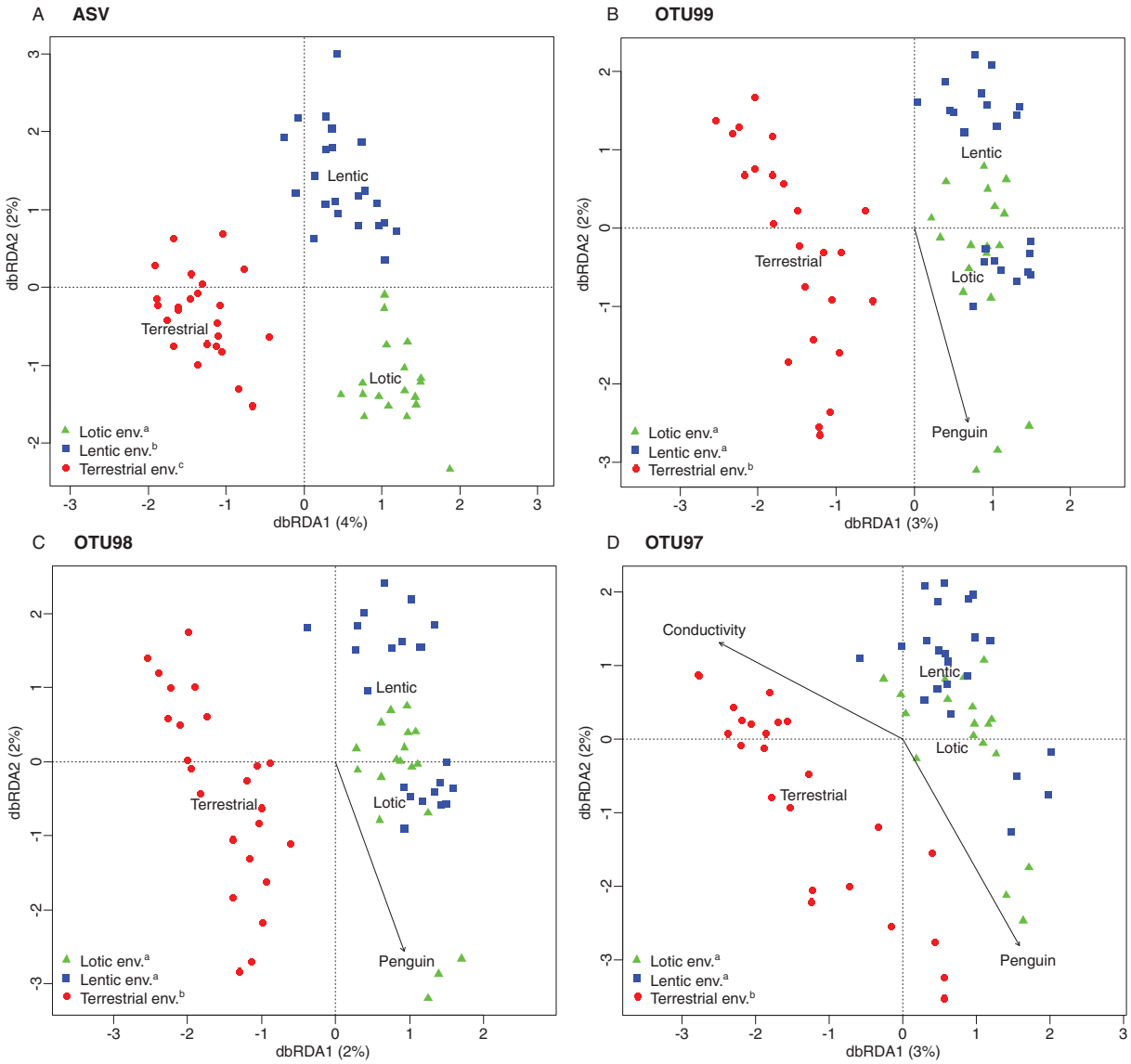

Fig. 5. Distance-based redundancy analyses (dbRDA) based on RC_{bray} matrices for ASV (A), OTU99 (B), OTU98 (C) and OTU97 (D) phylogenetic resolutions. Different superscript letters in the legends of the ordinations indicate significant differences between environments (PERMANOVA, $P < 0.01$, detailed R^2 and P values are included in Supplementary Table S2). Env., environments.

observed homogeneous selection at all phylogenetic resolutions in concordance with overlapping environmental conditions between the environments (Supporting Information Fig. S6).

Across all phylogenetic resolutions, pH was identified as a variable imposing selection, which agrees with previous observations showing that pH preference is a trait that is relatively deeply conserved (Martiny *et al*., 2015). At the ASV level, conductivity appeared to have a selective role influencing bacterial phylogenetic turnover. The presence of a colony of gentoo penguins in the vicinity also appeared to affect bacterial turnover. Penguin guano modifies the environment by increasing conductivity and nutrient content, especially ammonia-N inputs, and thus rising pH (Mataloni *et al*., 2005, 2010; Allende and Mataloni, 2013). These results highlight the importance of the interactions between the microbial communities and the macrofauna within this Antarctic wetland complex.

However, the low variance of βNTI explained by pH, conductivity and penguin impact in the dbRDA (Fig. 4A) and PERMANOVA tests suggest that, although these features impose selection on the metacommunity, one or more other environmental variables – not measured in this work – would be responsible for the strong homogeneous selection detected at the ASV level. Also, βNTI variance may be poorly explained due to the temporal

resolution of the environmental data. That is, single-date measurements (i.e. spatial sampling design) may provide an incomplete representation of selection pressures. Nonetheless, we observed a very strong signal of homogeneous selection. As Antarctica is the coldest continent (Kirby *et al*., 2014), the extremely low temperature could be imposing strong homogeneous selection in this system, preventing community phylogenetic divergence despite the heterogeneous landscape of the CPWC. Furthermore, temperature preference is a trait shallowly conserved (Martiny *et al*., 2015), and could be influencing community assembly processes at the finer phylogenetic resolution of ASVs.

A strong signal of homogenous selection was also observed in Antarctic bacterial communities from lakes on the Vestfold Hills region, Eastern Antarctica (Logares *et al*., 2018). In this case, despite the lakes having very heterogeneous physicochemical conditions, salinity was shown to be the homogenizing force. Salinity preference is a trait relatively deeply conserved (Martiny *et al*., 2015), such that microbial communities inhabiting marine and freshwater habitats are phylogenetically distinct at several phylogenetic resolutions (Paver *et al*., 2018). As more studies of bacterial metacommunities accumulate across Antarctica it will become possible to evaluate any dominant patterns in key environmental variables, types of influential assembly processes and how patterns change across levels of phylogenetic resolution.

Dispersal within the CPWC

We expected dispersal to be enhanced across the different local communities, resulting in the homogenization of the metacommunity. During the short austral summer, most of the area of the CPWC is snow-free due to glacial/snow melt and rain events (Wilhelm *et al*., 2016), contributing to the dissemination of microorganisms. Previous studies showed that slope angle and direction determine the shape of the hydrological network and the extent of the connectivity of this wetland complex (Mataloni *et al*., 2005, 2010). In turn, the presence of penguins over the breeding season (González-Zevallos *et al*., 2013) potentially adds to the dispersal of microorganisms. Yet, in contrast to our expectations, we did not detect homogenizing dispersal but a relatively constant contribution of dispersal limitation. Since CPWC is completely covered by snow from April to December, dispersal could be heavily restricted during most of the year between the three environments sampled (i.e. lotic, lentic and terrestrial). Thus, the high taxonomic turnover (high RC_{bray} values) and low phylogenetic turnover (low βNTI values) detected at the ASV level could be reflecting the action of dispersal limitation coupled with diversification processes (Zhou and Ning, 2017). In winter the consistent snow-cover insulates the ground surface from the colder air temperatures, which can reach down to approximately −20°C (Ramos Marín, 2018). This thermal insulation has been shown to allow bacteria growth below the snowpack (Brooks *et al*., 1998), and could enable microevolution in the isolated communities from CPWC. Previous experimental studies with bacteria from Arctic permafrost have demonstrated the physiological potential for genome replication at temperatures down to −20°C (Amato *et al*., 2010; Tuorto *et al*., 2014).

Conceptual model of ecological processes influencing community assembly in the CPWC

The observed differences in community assembly processes at different phylogenetic resolutions have led us to propose and discuss a conceptual model describing these differences (Fig. 6). Specifically, we hypothesize that the strong influence of homogeneous selection (i.e. low bacterial phylogenetic turnover) detected at the ASV level can potentially be interpreted as microevolutionary processes affecting community assembly through diversification (Nemergut *et al*., 2013; Zhou and Ning, 2017). Indeed, microevolution appears to occur within local communities with extremely low or zero dispersal rates (Leibold *et al*., 2004; Georgiades and Raoult, 2011; Stegen *et al*., 2013). Thus, dispersal limitation likely imposed by the snow-covered landscape could be acting in concert with drift and diversification (Nemergut *et al*., 2013; Stegen *et al*., 2013; Zhou and Ning, 2017) to generate different but very closely related ASVs across environments. This is supported by the lack

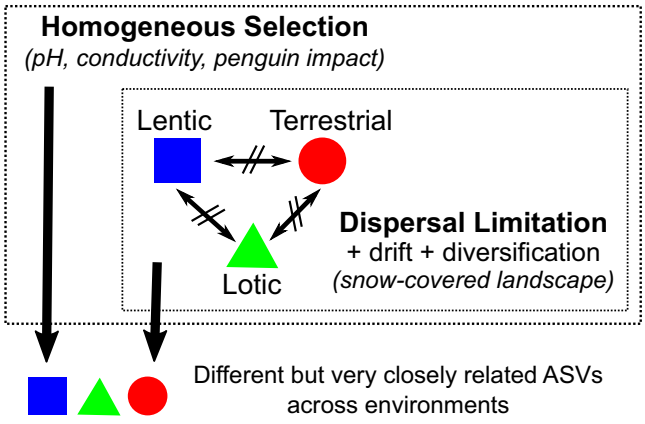

Fig. 6. A conceptual model of community assembly: dispersal limitation likely imposed by the snow-covered landscape could be acting in concert with drift and diversification, which together with the strong homogeneous selection generate different but very closely related ASVs across environments.

of separation between the three bacterial communities based on βNTI-ASV (Fig. 4A), the clear separation of these communities based on RC_{bray}-ASV (Fig. 5A), and the strong signal of homogeneous selection for ASVs. Moreover, in line with our conceptual model for CPWC, Cavicchioli (2015) suggested that the geographic isolation and strong selection imposed by hypersalinity and low temperatures controlled the evolutionary development of the microbial communities from the Deep Lake, Vestfold Hills, Eastern Antarctica.

Caveats

Ecological processes occur along a continuum of space and time (Hanson *et al*., 2012), yet our sampling represents a snapshot in time of this isolated system in the Antarctic continent. Despite CPWC being accessible only during the austral summer, we may expect temporal changes in the assembly processes related to changes in the system hydrology over this season. A sampling design encompassing both spatial and temporal scales would provide more insights into the mechanisms of community assembly in this Antarctic wetland complex. Also, we should be aware that ASV data could be overestimating diversity, and therefore detecting a strong (but not necessarily real) selection effect, as not enough literature and genomic data to date allow us to fully understand intragenomic rDNA sequence polymorphisms (Lavrinienko *et al*., 2020; Okazaki *et al*., 2021).

Conclusions

Here, we investigated the community assembly processes applying null models (Stegen *et al*., 2013, 2015) at different levels of phylogenetic resolution. We found that, as suggested by Hanson *et al*. (2012), the relative influence of the processes that shape bacterial communities change with phylogenetic resolution. More specifically, we observed that selection processes seem to be more important at finer (i.e. ASV level) than at coarser (i.e. OTU99, OTU98 and OTU97 levels) resolution, which may suggest that microevolutionary processes are shaping the bacterial metacommunity from CPWC. Indeed, a recent study has demonstrated that both ecological and evolutionary processes can alter the diversity of a soil microbiome on annual timescales (Chase *et al*., 2021). To further quantify the relative contribution of evolutionary processes to microbial community assembly, the path forward involves using emerging sequencing and bioinformatic tools combined with simulation modelling to test and update refined hypotheses. In all, this study shows that assessing community assembly processes at different phylogenetic resolutions is key to improve our understanding of microbial ecology.

Experimental procedures

Study site, sampling and environmental data

Cierva Point (64°09′ S, 60° 57′ W) encompasses the ASPA (Antarctic Specially Protected Area) No. 134 (Agraz *et al*., 1994; Antarctic Treaty Secretariat, 2013). The area shows a mild, humid climate (mean annual air temperature ca. −3.2°C; Wilhelm *et al*., 2016). Remarkably, its mean annual ground temperature (ca. –0.95°C) is within the highest range of the continent (Obu *et al*., 2020), with an annual precipitation ranging from 400 to 1100 mm (Wilhelm *et al*., 2016).

Location of sampling sites was established using a global positioning systems equipment (GPS eTrex, Garmin International, Olathe, KS, USA). The slope of each sampling site was calculated with a field laser clinometer (Scout DX 1000 ARC, Bushnell, Overland Park, KS, USA). To assess the degree of penguin impact, a scale of use intensity with six nominal levels was established according to the abundance and permanence of gentoo penguins (*Pygoscelis papua*) or signs thereof (e.g. feathers or faeces), where 0 corresponds to the absence of penguins or signs and 5 to nesting areas with high abundance and permanence of penguins. At soil sites, composite samples were collected in sterile Whirl-Pak bags and frozen at −20°C for transport and further analysis. Soil pH and conductivity (both 1:2.5 water suspension) were analyzed at the Soil Institute, National Institute of Agricultural Technology (INTA, Hurlingham, Buenos Aires, Argentina), following standard protocols described in Mortola *et al*. (2019). For lentic and lotic environments, water pH and conductivity were measured *in situ* using a pHmeter (HI98108, Hanna Instruments, Woonsocket, RI, USA) and a multiparametric probe (Sension 156, Hach, Loveland, CO, USA). At moss sites, interstitial water was obtained by aseptically squeezing the mosses *in situ* (Oloo *et al*., 2016), followed by water pH and conductivity measure. Snow was collected in 500 ml sterile pots, retained frozen and transported to the laboratory, where the parameters were measured on freshly melted snow.

Composite soil samples were transferred to sterile cryovials (ExtraGene, Taichung City, Taiwan) and preserved with 1 ml LifeGuard soil preservation solution (Qiagen, Hilden, Germany) at 4°C until further processing. Aliquots of approximately 200 ml of water samples from lentic water bodies and mosses were sequentially filtered through a 55 μm mesh size net, and 3 and 0.22 μm sterile nitrocellulose membranes (Nalgene, Rochester, NY, USA). The surface of rocks from lentic and lotic sampling sites was scraped using one sterile toothbrush per site. The detached biofilm was suspended in approximately 30 ml of 0.22 μm-filtered distilled water and sequentially filtered as described above.

Approximately 250 ml of coloured snow were allowed to melt and also sequentially filtered. The 0.22 μm membranes were preserved in sterile cryovials with 3.5 ml RNA*later* stabilization solution (Sigma-Aldrich, St. Louis, MO, USA) at 4°C until further processing.

DNA extraction and amplicon sequencing

DNA was extracted from 0.5 g of soil samples or half 0.22 μm membranes using the PowerSoil DNA isolation kit (Qiagen). A two-Step PCR was performed with primers 337F and 805R (Klindworth *et al*., 2013) for the 16S rRNA gene (V3–V4 regions). Amplicons were sequenced using Illumina MiSeq 2 x 300 paired-end reads approach (Caporaso *et al*., 2012) at Applied Biological Materials (BC, Canada).

Sequence data processing

Primer sequences were removed with Cutadapt 1.18 (Martin, 2011). ASVs were determined using *DADA2* v1.16.0 (Callahan *et al*., 2016) with default parameters, unless specified otherwise. Briefly, forward reads were quality-filtered and trimmed using the *DADA2* function *filterAndTrim* (options maxEE = 2, minLen = 175, truncLen = 250). Error rate models were fitted using the function *learnErrors*. ASVs were then inferred for each sample using the functions *derepFastq* and *dada*. An ASV table was created using *makeSequenceTable*. Chimeric sequences were removed using *removeBimeraDenovo*, which resulted in a table with 5336 ASVs. ASVs were classified using *assignTaxonomy* with the SILVA database (version 138, Quast *et al*., 2013). Unassigned ASVs or classified as chloroplasts or mitochondria were removed. The resulting count table with no singletons was normalized to an equal sampling depth of 6284 reads per sample using *rarefy_even_depth* function from *phyloseq* package (McMurdie and Holmes, 2013). A total of 402 176 total reads and 3960 ASVs were retained for further analysis. These ASVs were clustered into OTUs with decreasing sequence identity thresholds (i.e. 99%, 98%, 97% and 94% similarity OTUs) using the Opticlust algorithm in Mothur software following García-García *et al*. (2019), which resulted in OTU tables with (a) 723 OTU99, (b) 438 OTU98, (c) 312 OTU97 and (d) 160 OTU94, and 402 176 reads each. Phylogenetic trees for ASVs and OTUs were constructed using qiime2 (Bolyen *et al*., 2019) with the q2-phylogeny plugin (align-to-tree-mafft-fasttree pipeline). The sequence data obtained in this work were deposited at NCBI BioProject database (ID PRJNA719989, 64 sequence data links, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA719989).

Statistical analyses

Community structure. Weighted βMNTD distance matrices were calculated with *comdistnt* function from *picante* package (Kembel *et al*., 2010) in R (R Core Team, 2018). Differences in phylogenetic compositions between samples were visualized with principal coordinates analysis (PCoA) using *vegan* package (Oksanen *et al*., 2019). Differences between environments were tested with PERMANOVA (Anderson, 2001) using *adonis_pairwise* function with FDR correction for multiple comparisons (*metagMisc* package; Mikryukov, 2020). Homogeneity of multivariate dispersion (beta dispersion; Anderson, 2006) was evaluated with the former function. The relationship between geographic distance and phylogenetic dissimilarity (βMNTD) was studied with the Mantel test.

Phylogenetic signal. A Mantel correlogram (*mantel.correlog* function from *vegan* package) was used to test for phylogenetic signal, based on Pearson correlation coefficients between taxa differences in environmental optima and phylogenetic distances. Significance tests for each of 30 phylogenetic distance classes were based on 999 permutations, no distance class cut-off and a progressive Bonferroni correction (Legendre and Legendre, 1998). Environmental optimum for abundant taxa (i.e. relative abundance >1% in any sample) were estimated by means of canonical correspondence analysis with explanatory pH, log-transformed conductivity, slope and penguin impact values, and type of environment as a dummy variable. Permutation tests of the overall analysis and the first two canonical axes showed significant canonical relationships ($P < 0.05$). Taxa scores on the first two canonical axes were used as synthetic descriptors of their ecological optima (Borcard *et al*., 2018), and used for calculating Euclidean distances in order to estimate between-taxa environmental optimum differences following Llames *et al*. (2017). Between-taxa cophenetic distances were calculated using *cophenetic.phylo* function from *ape* package (Paradis and Schliep, 2019). These analyses were performed for each taxonomic resolution independently.

Assembly processes. The null model approach proposed by Stegen *et al*. (2013, 2015) was applied to investigate bacterial community assembly processes across phylogenetic resolutions. βNTI and RC_{bray} indices were calculated based on entire-community null model analysis with the *qpen* function from *iCAMP* package (Ning *et al*., 2020). Differences in mean βNTI or RC_{bray} values between taxonomic resolutions were evaluated with global Kruskal–Wallis test and Mann–Whitney post hoc pairwise comparisons applying Bonferroni correction.

Environmental features related with assembly processes. The quantitative environmental features (pH, log-transformed conductivity, slope and penguin impact) and the qualitative environmental feature (type of environment) were tested as explanatory variables in a dbRDA model selection procedure with either βNTI or RC_{bray} matrices as the response variables, independently for each taxonomic resolution. Both βNTI and RC_{bray} distance matrices were normalized to vary between 0 and 1 according to Stegen *et al*. (2013) before stepwise model selection (*ordistep* function, argument direction = 'both', $P < 0.05$). The features that significantly explained variation in βNTI were considered as environmental variables imposing selection. The features not related to βNTI that explained variation in RC_{bray} represented environmental variables that impose dispersal limitation. The contribution of each significant feature in shaping the βNTI or RC_{bray} values were quantified with PERMANOVA, as implemented in the *adonis* function (*vegan* package).

Acknowledgements

This research contributes to the Ant-ICON (Integrated Science to Inform Antarctic and Southern Ocean Conservation) SCAR Scientific Research Programme and was jointly supported by ANPCyT – Argentina (grants PICT 2016-2517, 2016-1554) and NRF – South Africa. We thank two anonymous reviewers for their valuable comments on the manuscript. We thank for their help to P. H. Lebre, S. Metz, Y. Sica, P. Fermani, D. González, S. Ramos Marín, M. Libertelli, the crew of Base Primavera and the Instituto Antártico Argentino – Dirección Nacional del Antártico. A.V. was supported by the project 'CLU-2019–05 – IRNASA/CSIC Unit of Excellence', funded by the Junta de Castilla y León and co-financed by the European Union (ERDF 'Europe drives our growth'). J.C.S. was supported by the U.S. Department of Energy-BER program, as part of an Early Career Award to J.C.S. at Pacific Northwest National Laboratory, a multi-program national laboratory operated by Battelle for the US Department of Energy under Contract DE-AC05-76RL01830.

References

Agraz, J.L., Quintana, R.D., and Acero, J.M. (1994) Ecología de los ambientes terrestres en Punta Cierva (Costa de Danco, Península Antártica). *Contrib Inst Ant Arg* **439**: 1–32.

Allen, R., Hoffmann, L.J., Larcombe, M.J., Louisson, Z., and Summerfield, T.C. (2020) Homogeneous environmental selection dominates microbial community assembly in the oligotrophic South Pacific Gyre. *Mol Ecol* **29**: 4680–4691.

Allende, L., and Mataloni, G. (2013) Short-term analysis of the phytoplankton structure and dynamics in two ponds with distinct trophic states from Cierva Point (maritime Antarctica). *Polar Biol* **36**: 629–644.

Amato, P., Doyle, S.M., Battista, J.R., and Christner, B.C. (2010) Implications of subzero metabolic activity on long-term microbial survival in terrestrial and extraterrestrial permafrost. *Astrobiology* **10**: 789–798.

Anderson, M.J. (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecol* **26**: 32–46.

Anderson, M.J. (2006) Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* **62**: 245–253.

Antarctic Treaty Secretariat. (2013) Management Plan for Antarctic Specially Protected Area No. 134. Cierva Point and offshore Islands, Danco Coast, Antarctic Peninsula. ATCM XXXVI Final Report. Measure 5 Annex 63–76.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., *et al*. (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**: 852–857.

Borcard, D., Gillet, F., and Legendre, P. (2018) *Numerical Ecology with R*. Cham: Springer International Publishing.

Brooks, P.D., Williams, M.W., and Schmidt, S.K. (1998) Inorganic nitrogen and microbial biomass dynamics before and during spring snowmelt. *Biogeochemistry* **43**: 1–15.

Callahan, B.J., McMurdie, P.J., and Holmes, S.P. (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J* **11**: 2639–2643.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016) DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**: 581–583.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-lyons, D., Huntley, J., Fierer, N., *et al*. (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* **6**: 1621–1624.

Cavicchioli, R. (2015) Microbial ecology of Antarctic aquatic systems. *Nat Rev Microbiol* **13**: 691–706.

Chase, A.B., Weihe, C., and Martiny, J.B.H. (2021) Adaptive differentiation and rapid evolution of a soil bacterium along a climate gradient. *Proc Natl Acad Sci U S A* **118**: e2101254118.

Cohan, F.M. (2001) Bacterial species and speciation. *Syst Biol* **50**: 513–524.

Danczak, R.E., Daly, R.A., Borton, M.A., Stegen, J.C., Roux, S., Wrighton, K.C., and Wilkins, M.J. (2020) Ecological assembly processes are coordinated between bacterial and viral communities in fractured shale ecosystems. *mSystems* **5**: 1–13.

Dini-Andreote, F., Stegen, J.C., Van Elsas, J.D., and Salles, J.F. (2015) Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc Natl Acad Sci U S A* **112**: E1326–E1332.

Edgar, R.C. (2018) Updating the 97% identity threshold for 16S ribosomal RNA OTUs. *Bioinformatics* **34**: 2371–2375.

Fodelianakis, S., Washburne, A.D., Bourquin, M., Pramateftaki, P., Kohler, T.J., Styllas, M., *et al*. (2021) Microdiversity characterizes prevalent phylogenetic clades in the glacier-fed stream microbiome. *ISME J*. https://doi.org/10.1038/s41396-021-01106-6

García-García, N., Tamames, J., Linz, A., Pedrós-Alió, C., and Puente-Sánchez, F. (2019) Microdiversity ensures the maintenance of functional microbial communities under

changing environmental conditions. *ISME J* **13**: 2969–2983.

Georgiades, K., and Raoult, D. (2011) Defining pathogenic bacterial species in the genomic era. *Front Microbiol* **1**: 151.

González-Zevallos, D., Santos, M.M., Rombolá, E.F., Juáres, M.A., and Coria, N.R. (2013) Abundance and breeding distribution of seabirds in the northern part of the Danco Coast, Antarctic peninsula. *Polar Res* **32**: 11133.

Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., and Martiny, J.B.H. (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol* **10**: 497–506.

Huber, P., Metz, S., Unrein, F., Mayora, G., Sarmento, H., and Devercelli, M. (2020) Environmental heterogeneity determines the ecological processes that govern bacterial metacommunity assembly in a floodplain river system. *ISME J* **14**: 2951–2966.

Hughes, K.A., Cowan, D.A., and Wilmotte, A. (2015) Protection of Antarctic microbial communities – out of sight, out of mind. *Front Microbiol* **6**: 151.

Ji, M., Kong, W., Stegen, J., Yue, L., Wang, F., Dong, X., *et al*. (2020) Distinct assembly mechanisms underlie similar biogeographical patterns of rare and abundant bacteria in Tibetan Plateau grassland soils. *Environ Microbiol* **22**: 2261–2272.

Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., *et al*. (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**: 1463–1464.

Kirby, B.M., Easton, S., Marla Tuffin, I., and Cowan, D.A. (2014) Bacterial diversity in polar habitats. In *Polar Microbiology*, Miller, R.V., and Whyte, L.G. (eds). Washington, DC, USA: ASM Press, pp. 1–31.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., and Glöckner, F.O. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* **41**: 1–11.

Kraemer, S.A., Barbosa da Costa, N., Shapiro, B.J., Fradette, M., Huot, Y., and Walsh, D.A. (2020) A large-scale assessment of lakes reveals a pervasive signal of land use on bacterial communities. *ISME J* **14**: 3011–3023.

Langenheder, S., Wang, J., Karjalainen, S.M., Laamanen, T. M., Tolonen, K.T., Vilmi, A., and Heino, J. (2017) Bacterial metacommunity organization in a highly connected aquatic system. *FEMS Microbiol Ecol* **93**: 1–9.

Lavrinienko, A., Jernfors, T., Koskimäki, J.J., Pirttilä, A.M., and Watts, P.C. (2020) Does intraspecific variation in rDNA copy number affect analysis of microbial communities? *Trends Microbiol* **29**: 19–27.

Legendre, P., and Legendre, L. (1998) *Numerical Ecology Second Eng*. Amsterdam, The Netherlands: Elsevier Science.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., *et al*. (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett* **7**: 601–613.

Lindström, E.S., and Langenheder, S. (2012) Local and regional factors influencing bacterial community assembly. *Environ Microbiol Rep* **4**: 1–9.

Llames, M.E., Huber, P., Metz, S., and Unrein, F. (2017) Interplay between stochastic and deterministic processes in the maintenance of alternative community states in Verrucomicrobia-dominated shallow lakes. *FEMS Microbiol Ecol* **93**: 1–10.

Logares, R., Deutschmann, I.M., Junger, P.C., Giner, C.R., Krabberød, A.K., Schmidt, T.S.B., *et al*. (2020) Disentangling the mechanisms shaping the surface ocean microbiota. *Microbiome* **8**: 1–17.

Logares, R., Tesson, S.V.M., Canbäck, B., Pontarp, M., Hedlund, K., and Rengefors, K. (2018) Contrasting prevalence of selection and drift in the community structuring of bacteria and microbial eukaryotes. *Environ Microbiol* **20**: 2231–2240.

Lu, H.P., Yeh, Y.C., Sastri, A.R., Shiah, F.K., Gong, G.C., and Hsieh, C.H. (2016) Evaluating community-environment relationships along fine to broad taxonomic resolutions reveals evolutionary forces underlying community assembly. *ISME J* **10**: 2867–2878.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetJ* **17**: 10–12.

Martiny, A.C., Coleman, M.L., and Chisholm, S.W. (2006) Phosphate acquisition genes in Prochlorococcus ecotypes: evidence for genome-wide adaptation. *Proc Natl Acad Sci U S A* **103**: 12552–12557.

Martiny, J.B.H., Jones, S.E., Lennon, J.T., and Martiny, A.C. (2015) Microbiomes in light of traits: a phylogenetic perspective. *Science* **350**: aac9323.

Mataloni, G., Garraza, G.G., Bölter, M., Convey, P., and Fermani, P. (2010) What shapes edaphic communities in mineral and ornithogenic soils of Cierva Point, Antarctic Peninsula? *Polar Sci* **4**: 405–419.

Mataloni, G., Vinocur, A., and De Tezanos Pinto, P. (2005) Abiotic characterization and epilithic communities of a naturally enriched stream at Cierva Point, Antarctic Peninsula. *Antarct Sci* **17**: 163–170.

McMurdie, P.J., and Holmes, S. (2013) Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**: e61217.

Mikryukov, V. (2020) metagMisc: Miscellaneous functions for metagenomic analysis. R package version 0.0.4.

Mortola, N., Romaniuk, R., Cosentino, V., Eiza, M., Carfagno, P., Rizzo, P., *et al*. (2019) Potential use of a poultry manure digestate as a biofertiliser: evaluation of soil properties and Lactuca sativa growth. *Pedosphere* **29**: 60–69.

Mysara, M., Vandamme, P., Props, R., Kerckhof, F.-M., Leys, N., Boon, N., *et al*. (2017) Reconciliation between operational taxonomic units and species boundaries. *FEMS Microbiol Ecol* **93**: 1–12.

Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F., *et al*. (2013) Patterns and processes of microbial community assembly. *Microbiol Mol Biol Rev* **77**: 342–356.

Ning, D., Yuan, M., Wu, L., Zhang, Y., Guo, X., Zhou, X., *et al*. (2020) A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. *Nat Commun* **11**: 4717.

Obu, J., Westermann, S., Vieira, G., Abramov, A., Balks, M. R., Bartsch, A., *et al*. (2020) Pan-Antarctic map of near-

surface permafrost temperatures at 1 km2 scale. *Cryosphere* **14**: 497–519.
Okazaki, Y., Fujinaga, S., Salcher, M.M., Callieri, C., Tanaka, A., Kohzu, A., *et al*. (2021) Microdiversity and phylogeographic diversification of bacterioplankton in pelagic freshwater systems revealed through long-read amplicon sequencing. *Microbiome* **9**: 24.
Oksanen, J., Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., *et al*. (2019) Vegan: Community Ecology Package. R package version 2.5-6.
Oloo, F., Valverde, A., Quiroga, M.V., Vikram, S., Cowan, D., and Mataloni, G. (2016) Habitat heterogeneity and connectivity shape microbial communities in south American peatlands. *Sci Rep* **6**: 25712.
Paradis, E., and Schliep, K. (2019) Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**: 526–528.
Paver, S.F., Muratore, D., Newton, R.J., and Coleman, M.L. (2018) Reevaluating the salty divide: phylogenetic specificity of transitions between marine and freshwater systems. *mSystems* **3**: e00232-18. https://doi.org/10.1128/mSystems.00232-18
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., *et al*. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590–D596.
R Core Team. (2018) *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.
Ramos Marín, S. (2018) Spatial modelling of the temperature at the top of permafrost in Cierva Point (Antarctic Peninsula). Master Thesis. Universidade de Lisboa.
Rosselló-Móra, R., and Amann, R. (2015) Past and future species definitions for bacteria and archaea. *Syst Appl Microbiol* **38**: 209–216.
Roy, J., Mazel, F., Sosa-Hernández, M.A., Dueñas, J.F., Hempel, S., Zinger, L., and Rillig, M.C. (2019) The relative importance of ecological drivers of arbuscular mycorrhizal fungal distribution varies with taxon phylogenetic resolution. *New Phytol* **224**: 936–948.
Schloss, P.D., and Handelsman, J. (2004) Status of the microbial census. *Microbiol Mol Biol Rev* **68**: 686–691.
Stegen, J.C., Lin, X., Fredrickson, J.K., Chen, X., Kennedy, D.W., Murray, C.J., *et al*. (2013) Quantifying community assembly processes and identifying features that impose them. *ISME J* **7**: 2069–2079.
Stegen, J.C., Lin, X., Fredrickson, J.K., and Konopka, A.E. (2015) Estimating and mapping ecological processes influencing microbial community assembly. *Front Microbiol* **6**: 1–15.
Stegen, J.C., Lin, X., Konopka, A.E., and Fredrickson, J.K. (2012) Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J* **6**: 1653–1664.
Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J., and Zimmerman, J.K. (2006) The problem and promise of scale dependency in community phylogenetics. *Ecology* **87**: 2418–2424.
Tuorto, S.J., Darias, P., McGuinness, L.R., Panikov, N., Zhang, T., Häggblom, M.M., and Kerkhof, L.J. (2014) Bacterial genome replication at subzero temperatures in permafrost. *ISME J* **8**: 139–149.
Vellend, M. (2010) Conceptual synthesis in community ecology. *Q Rev Biol* **85**: 183–206.
Vellend, M. (2016) *The Theory of Ecological Communities*, 1st ed. Woodstock, UK: Princeton University Press.
Vellend, M., Srivastava, D.S., Anderson, K.M., Brown, C.D., Jankowski, J.E., Kleynhans, E.J., *et al*. (2014) Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos* **123**: 1420–1430.
Wilhelm, K.R., Bockheim, J.G., and Haus, N.W. (2016) Properties and processes of recently established soils from deglaciation of Cierva Point, Western Antarctic Peninsula. *Geoderma* **277**: 10–22.
Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.-H., *et al*. (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* **12**: 635–645.
Zhou, J., and Ning, D. (2017) Stochastic community assembly: does it matter in microbial ecology? *Microbiol Mol Biol Rev* **81**: e00002-17.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Bacterial community sampled.

Table S2. PERMANOVA results for pairwise comparisons among type of environment based on βMNTD, βNTI and RC_{bray}.

Table S3. Environmental data.

Fig. S1. Rarefaction curves for the 64 samples sequenced.

Fig. S2. Diversity measures for the different environments.

Fig. S3. Relative abundance of the top 10 most abundant phyla.

Fig. S4. Phylogenetic Mantel correlograms for each phylogenetic resolution.

Fig. S5. Violin plots of βMNTD indices across phylogenetic resolutions.

Fig. S6. Principal component analysis (PCA) ordination plot of abiotic data.