environmental microbiology
environmental microbiology reports

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

Journal:	*Environmental Microbiology and Environmental Microbiology Reports*
Manuscript ID	EMI-2021-1632
Journal:	Environmental Microbiology
Manuscript Type:	EMI - Research article
Date Submitted by the Author:	22-Sep-2021
Complete List of Authors:	Quiroga, María; Instituto Tecnológico de Chascomús, Valverde, Angel; IRNASA, ; Mataloni, Gabriela; Instituto de Investigación e Ingeniería Ambiental (IIIA, UNSAM-CONICET) Casa, Valeria; Instituto de Investigación e Ingeniería Ambiental (IIIA, UNSAM-CONICET) Stegen, James; Pacific Northwest National Laboratory, Biological Sciences Division Cowan, Don
Keywords:	bacterial community assembly, null models, phylogenetic resolution, Antarctic Peninsula, Cierva Point, wetlands

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]; 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, USA

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

[†] These authors contributed equally to this work.

Corresponding author: María V Quiroga

INTECH, Av. Intendente Marino Km 8.200, Chascomús (7130), Buenos Aires, Argentina.

Email: mvquiroga@iib.unsam.edu.ar

Running title: Community assembly across phylogenetic resolutions

Keywords: bacterial community assembly, null models, phylogenetic resolution, Antarctic Peninsula, Cierva Point, wetlands

Originality-Significance Statement: The relative influence of the bacterial community assembly processes varies across levels of phylogenetic resolution in an Antarctic wetland complex.

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.

Introduction

Understanding the processes governing community assembly is a key topic in microbial ecology (Vellend, 2016). It is now recognised 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). Yet, only a single phylogenetic resolution (mainly ASVs or 97% similarity 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 was 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 it has been demonstrated that taxon phylogenetic resolution may determine the relative influence of ecological drivers of the diversity patterns of arbuscular mycorrhizal fungi (Roy *et al.*, 2019) and vascular plants (Swenson *et al.*, 2006).

Bacterial phylogenetic diversity studies have typically involved the clustering of sequences into operational taxonomic units (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

104 *et al.*, 2019). This approach provides an opportunity to evaluate the influences of ecological
105 processes on bacterial community assembly at different phylogenetic resolutions.

106 Cierva Point Wetland Complex (CPWC) is a macro-biodiversity hotspot on the north-
107 west coast of the Antarctic Peninsula (Agraz *et al.*, 1994; Antarctic Treaty Secretariat, 2013;
108 Wilhelm *et al.*, 2016). Its complex, fragmented topography defines a mosaic of distinct
109 environmental units characterized by different combinations of land cover, slope and
110 orientation, most hosting a large number of different environments (Agraz *et al.*, 1994), which
111 in this study have been grouped into lentic, lotic and terrestrial environments for simplicity. The
112 complex is completely covered by snow from April to December but is mostly snow-free during
113 the austral summer (Wilhelm *et al.*, 2016). It has been shown that slope angle determines the
114 extent and direction of hydrological connectivity during snow melt or rain events in this system
115 (Mataloni *et al.*, 2005, 2010). In addition, this protected area hosts an increasingly large colony
116 of gentoo penguin (*Pygoscelis papua*) (González-Zevallos *et al.*, 2013) that contribute to
117 nutrient input and may contribute to the dispersal of microorganisms across the wetland
118 complex.

119 Here, we aimed to study the relative influence of the ecological processes shaping the
120 CPWC bacterial metacommunity and how the assessment of these processes may vary across
121 different levels of phylogenetic resolution. We used high-throughput sequencing of 16S rRNA
122 genes and applied the null models proposed by Stegen *et al.* (2013, 2015) to the phylogenetic
123 data, implementing the approach of García-García *et al.* (2019) by clustering ASVs into OTUs
124 with decreasing sequence identity thresholds (i.e., 99, 98, 97 and 94% similarity OTUs). We
125 hypothesize that the mosaic of different environments are subject to spatially varying
126 environmental conditions, which will impose high variable selection in the different CPWC
127 microbiomes; that is, each environment will select for different taxa. In addition, we expect
128 dispersal to be enhanced across the different local communities during the austral summer
129 period by snow melt, rain events and penguin movements, resulting in the homogenization of
130 the metacommunity (i.e., homogenizing dispersal). We therefore expect to observe simultaneous
131 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. 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 Cierva Point Wetland Complex (CPWC, ca. area: 1 square kilometre) 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) (Supporting Information Table S1 and Fig. 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. S2). 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. S3); 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. S4).

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. S5), and this resolution was therefore excluded from subsequent downstream analyses. The PERMANOVA analyses based on phylogenetic dissimilarity (βMNTD) showed that the bacterial assemblages from lentic and lotic environments were significantly different from those of terrestrial environments (all $P < 0.01$) at each phylogenetic resolution, although this was not clearly visualized in the principal coordinate plots (Fig. 1). 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. S6), reflecting the general low phylogenetic variability between the local communities. There were no significant differences in community heterogeneity between environments at each phylogenetic resolution (betadispersion, 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. S5), 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 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. 2). The opposite trend was

observed for RC_{bray} values (global Kruskal–Wallis test: $P < 0.001$; significant pairwise Wilcoxon Tests: $P < 0.004$, Fig. 2).

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 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. 3). 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. Using βNTI distances, the contribution of the measured environmental factors in shaping the βNTI values increased towards 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

212 explained by pH ($R^2 = 0.20$, $P = 0.001$), conductivity ($R^2 = 0.09$, $P = 0.001$) and penguin impact
213 ($R^2 = 0.05$, $P = 0.006$).

214 Using RC_{bray} distances, we found that there was dispersal limitation between the
215 different environments across all phylogenetic resolutions (Fig. 5). PERMANOVA analyses
216 confirmed that RC_{bray} variation was significantly (though weakly) explained by the type of
217 environment ($R^2 = 0.10$, $R^2 = 0.10$, $R^2 = 0.10$, $R^2 = 0.12$ (all $P = 0.001$) for OTU97, OTU98,
218 OTU99 and ASV data respectively). Penguin impact ($R^2 = 0.06$, $R^2 = 0.04$, $R^2 = 0.05$ (all $P <$
219 0.002) for OTU97, OTU98 and OTU99 levels respectively) and conductivity ($R^2 = 0.04$, $P =$
220 0.003 for OTU97 data) significantly (though weakly) explained RC_{bray} variation.

221

222 **Discussion**

223 **Community assembly processes vary across levels of phylogenetic resolution**

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

233 The strong influence of selection processes at the ASV level appears to reflect
234 environmental filters/constraints acting at this finer sub-species level, which is supported by the
235 high phylogenetic signal of ASVs (Supporting Information Fig. S5). As ASVs can represent
236 ecotypes within species (García-García *et al.*, 2019) and display niche partitioning (Martiny *et*
237 *al.*, 2006), we propose that the CPWC ASV patterns may reflect a significant influence of
238 microevolutionary processes in the assembly of this bacterial metacommunity. As the 16S

239 rRNA genes are too conserved to detect recent evolutionary changes (see Chase *et al*., 2021 and
240 references therein) further studies are needed to corroborate this.

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

253

254 **Environmental variables related to selection processes**

255 We hypothesized that the mosaic of different environments from the CPWC would lead to
256 spatially heterogeneous environmental conditions, and impose high variable selection, with each
257 environment selecting for different taxa. However, we did not detect variable selection, and
258 therefore rejected our hypothesis. Instead, we observed homogeneous selection at all
259 phylogenetic resolutions in concordance with overlapping environmental conditions between
260 the environments (Supporting Information Fig. S7).

261 Across all phylogenetic resolutions, pH was identified as a variable imposing selection,
262 which agrees with previous observations showing that pH preference is a trait that is relatively
263 deeply conserved (Martiny *et al.*, 2015). At the ASV level, conductivity appeared to have a
264 selective role influencing bacterial phylogenetic turnover. The presence of a colony of gentoo
265 penguins in the vicinity also appeared to affect bacterial turnover. Penguin guano modifies the
266 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 (Figure 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. 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.

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. We may therefore expect a diverse range of environmental variables to play important roles in governing bacterial metacommunities across Antarctica. As more studies accumulate it will become possible to evaluate any dominant patterns in key variables, types of influential assembly processes, and how patterns change across levels of phylogenetic resolution.

Dispersal within the Cierva Point Wetland Complex

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, that can reach down to ca. −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).

A conceptual model

The observed differences in community assembly processes at different phylogenetic resolutions have led us to propose and discuss a conceptual model describing these differences. 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 of separation among environments based on βNTI (Fig. 4a), the clear separation of environments based on RC_{bray} (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 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 and elevation of sampling sites were established using a global positioning systems equipment (GPS eTrex, Garmin International Inc., 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 (*Pygoscellis 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) was 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 Co., 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 ca. 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 were scraped using one sterile toothbrush per site. The detached biofilm was suspended in ca. 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 Inc. (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,

405 Quast *et al.*, 2013). Unassigned ASVs or classified as chloroplasts or mitochondria were
406 removed. The resulting count table with no singletons was normalized to an equal sampling
407 depth of 6284 reads per sample using *rarefy_even_depth* function from *phyloseq* package
408 (McMurdie and Holmes, 2013). A total of 402176 total reads and 3960 ASVs were retained for
409 further analysis. These ASVs were clustered into OTUs with decreasing sequence identity
410 thresholds (i.e., 99, 98, 97 and 94% similarity OTUs) using the Opticlust algorithm in Mothur
411 software following García-García *et al.* (2019), which resulted in OTU tables with a) 723
412 OTU99, b) 438 OTU98, c) 312 OTU97 and d) 160 OTU94, and 402176 reads each.
413 Phylogenetic trees for ASVs and OTUs were constructed using qiime2 (Bolyen *et al.*, 2019)
414 with the q2-phylogeny plugin (align-to-tree-mafft-fasttree pipeline). The sequence data obtained
415 in this work were deposited at NCBI BioProject database (ID PRJNA719989).
416
417 **Statistical Analyses**
418 ***Community structure.*** Weighted βMNTD distance matrices were calculated with *comdistnt*
419 function from *picante* package (Kembel *et al.*, 2010) in R (R Core Team, 2018). Differences in
420 phylogenetic compositions between samples were visualized with Principal Coordinates
421 Analysis (PCoA) using *vegan* package (Oksanen *et al.*, 2019). Differences between
422 environments were tested with permutational analysis of variance (PERMANOVA; Anderson,
423 2001) using *adonis_pairwise* function with FDR correction for multiple comparisons
424 (*metagMisc* package; Mikryukov, 2020). Homogeneity of multivariate dispersion
425 (betadispersion; Anderson, 2006) was evaluated with the former function. The relationship
426 among geographic distance and phylogenetic dissimilarity (βMNTD) was studied with Mantel
427 test.
428
429 ***Phylogenetic signal***. A Mantel correlogram (*mantel.correlog* function from *vegan* package) was
430 used to test for phylogenetic signal, based on Pearson correlation coefficients between taxa
431 differences in environmental optima and phylogenetic distances. Significance tests for each of
432 30 phylogenetic distance classes were based on 999 permutations, no distance class cut-off and

433 a progressive Bonferroni correction (Legendre and Legendre, 1998). Environmental optimum
434 for abundant taxa (i.e., relative abundance > 1% in any sample) were estimated by means of
435 canonical correspondence analysis (CCA) with explanatory pH, log-transformed conductivity,
436 slope and penguin impact values, and type of environment as a dummy variable. Permutation
437 tests of the overall analysis and the first two canonical axes showed significant canonical
438 relationships ($P < 0.05$). Taxa scores on the first two canonical axes were used as synthetic
439 descriptors of their ecological optima (Borcard *et al.*, 2018), and used for calculating Euclidian
440 distances in order to estimate between-taxa environmental optimum differences following
441 Llames *et al.* (2017). Between-taxa cophenetic distances were calculated using *cophenetic.phylo*
442 function from *ape* package (Paradis and Schliep, 2019). These analyses were performed for
443 each taxonomic resolution independently.
444
445 ***Assembly processes.*** The null model approach proposed by Stegen *et al.* (2013, 2015) was
446 applied to investigate bacterial community assembly processes across phylogenetic resolutions.
447 βNTI and RC_{bray} indices were calculated based on entire-community null model analysis with
448 the *qpen* function from *iCAMP* package (Ning *et al.*, 2020). Differences in mean βNTI or RC_{bray}
449 values between taxonomic resolutions were evaluated with global Kruskal–Wallis test and
450 Mann–Whitney post-hoc pairwise comparisons applying Bonferroni correction.
451
452 ***Environmental features related with assembly processes.*** The quantitative environmental
453 features (pH, log-transformed conductivity, slope and penguin impact) and the qualitative
454 environmental feature (type of environment) were tested as explanatory variables in a distance-
455 based redundancy analysis (dbRDA) model selection procedure with either βNTI or RC_{bray}
456 matrices as the response variables, independently for each taxonomic resolution. Both βNTI and
457 RC_{bray} distance matrices were normalized to vary between 0 and 1 according to Stegen *et al.*
458 (2013) before stepwise model selection (*ordistep* function, argument direction = "both", $P <$
459 0.05). The features that significantly explained variation in βNTI were considered as
460 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 for their help to PH 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. AV 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'). JCS was supported by the U.S. Department of Energy-BER program, as part of an Early Career Award to JCS at Pacific Northwest National Laboratory, a multi-program national laboratory operated by Battelle for the US Department of Energy under Contract DE-AC05-76RL01830.

Conflict of Interest

The authors declare no conflict of interest.

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

489 Antarctica). *Polar Biol.* **36**: 629–644.
490 Amato, P., Doyle, S.M., Battista, J.R., and Christner, B.C. (2010) Implications of subzero
491 metabolic activity on long-term microbial survival in terrestrial and extraterrestrial
492 permafrost. *Astrobiology* **10**: 789–798.
493 Anderson, M.J. (2001) A new method for non-parametric multivariate analysis of variance.
494 *Austral Ecol.* **26**: 32–46.
495 Anderson, M.J. (2006) Distance-based tests for homogeneity of multivariate dispersions.
496 *Biometrics* **62**: 245–53.
497 Antarctic Treaty Secretariat (2013) Management Plan for Antarctic Specially Protected Area
498 No. 134. Cierva Point and offshore Islands, Danco Coast, Antarctic Peninsula. ATCM
499 XXXVI Final Report. Measure 5 Annex., pp. 63–76.
500 Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., et al.
501 (2019) Reproducible, interactive, scalable and extensible microbiome data science using
502 QIIME 2. *Nat. Biotechnol.* **37**: 852–857.
503 Borcard, D., Gillet, F., and Legendre, P. (2018) Numerical Ecology with R Springer
504 International Publishing, Cham.
505 Brooks, P.D., Williams, M.W., and Schmidt, S.K. (1998) Inorganic nitrogen and microbial
506 biomass dynamics before and during spring snowmelt. *Biogeochemistry* **43**: 1–15.
507 Callahan, B.J., McMurdie, P.J., and Holmes, S.P. (2017) Exact sequence variants should replace
508 operational taxonomic units in marker-gene data analysis. *ISME J.* **11**: 2639–2643.
509 Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P.
510 (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nat.*
511 *Methods* **13**: 581–583.
512 Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-lyons, D., Huntley, J., Fierer, N., et al.
513 (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and
514 MiSeq platforms. *ISME J.* 1621–1624.
515 Cavicchioli, R. (2015) Microbial ecology of Antarctic aquatic systems. *Nat. Rev. Microbiol.* **13**:
516 691–706.

517 Chase, A.B., Weihe, C., and Martiny, J.B.H. (2021) Adaptive differentiation and rapid evolution
518 of a soil bacterium along a climate gradient. *Proc. Natl. Acad. Sci.* **118**: e2101254118.
519 Danczak, R.E., Daly, R.A., Borton, M.A., Stegen, J.C., Roux, S., Wrighton, K.C., and Wilkins,
520 M.J. (2020) Ecological Assembly Processes Are Coordinated between Bacterial and Viral
521 Communities in Fractured Shale Ecosystems. *mSystems* **5**: 1–13.
522 Dini-Andreote, F., Stegen, J.C., Van Elsas, J.D., and Salles, J.F. (2015) Disentangling
523 mechanisms that mediate the balance between stochastic and deterministic processes in
524 microbial succession. *Proc. Natl. Acad. Sci. U. S. A.* **112**: E1326–E1332.
525 Edgar, R.C. (2018) Updating the 97% identity threshold for 16S ribosomal RNA OTUs.
526 *Bioinformatics* **34**: 2371–2375.
527 Fodelianakis, S., Washburne, A.D., Bourquin, M., Pramateftaki, P., Kohler, T.J., Styllas, M., et
528 al. (2021) Microdiversity characterizes prevalent phylogenetic clades in the glacier-fed
529 stream microbiome. *ISME J.* https://doi.org/10.1038/s41396-021-01106-6
530 García-García, N., Tamames, J., Linz, A., Pedrós-Alió, C., and Puente-Sánchez, F. (2019)
531 Microdiversity ensures the maintenance of functional microbial communities under
532 changing environmental conditions. *ISME J.* **13**: 2969–2983.
533 Georgiades, K. and Raoult, D. (2011) Defining Pathogenic Bacterial Species in the Genomic
534 Era. *Front. Microbiol.* **1**: 151.
535 González-Zevallos, D., Santos, M.M., Rombolá, E.F., Juáres, M.A., and Coria, N.R. (2013)
536 Abundance and breeding distribution of seabirds in the northern part of the Danco Coast,
537 Antarctic Peninsula. *Polar Res.* **32**: 11133.
538 Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., and Martiny, J.B.H. (2012) Beyond
539 biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.*
540 **10**: 497–506.
541 Huber, P., Metz, S., Unrein, F., Mayora, G., Sarmento, H., and Devercelli, M. (2020)
542 Environmental heterogeneity determines the ecological processes that govern bacterial
543 metacommunity assembly in a floodplain river system. *ISME J.* **14**: 2951–2966.
544 Ji, M., Kong, W., Stegen, J., Yue, L., Wang, F., Dong, X., et al. (2020) Distinct assembly

545 mechanisms underlie similar biogeographical patterns of rare and abundant bacteria in
546 Tibetan Plateau grassland soils. *Environ. Microbiol.* **22**: 2261–2272.
547 Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., et al.
548 (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**: 1463–
549 1464.
550 Kirby, B.M., Easton, S., Marla Tuffin, I., and Cowan, D.A. (2014) Bacterial Diversity in Polar
551 Habitats. In, Miller,R.V. and Whyte,L.G. (eds), *Polar Microbiology*. ASM Press,
552 Washington, DC, USA, pp. 1–31.
553 Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., and Glöckner, F.O.
554 (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and
555 next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**: 1–11.
556 Kraemer, S.A., Barbosa da Costa, N., Shapiro, B.J., Fradette, M., Huot, Y., and Walsh, D.A.
557 (2020) A large-scale assessment of lakes reveals a pervasive signal of land use on bacterial
558 communities. *ISME J.* **14**: 3011–3023.
559 Langenheder, S., Wang, J., Karjalainen, S.M., Laamanen, T.M., Tolonen, K.T., Vilmi, A., and
560 Heino, J. (2017) Bacterial metacommunity organization in a highly connected aquatic
561 system. *FEMS Microbiol. Ecol.* **93**: 1–9.
562 Lavrinienko, A., Jernfors, T., Koskimäki, J.J., Pirttilä, A.M., and Watts, P.C. (2020) Does
563 Intraspecific Variation in rDNA Copy Number Affect Analysis of Microbial
564 Communities? *Trends Microbiol.* **29**: 19–27.
565 Legendre, P. and Legendre, L. (1998) Numerical Ecology Second Eng. Elsevier Science,
566 Amsterdam, The Netherlands.
567 Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., et al.
568 (2004) The metacommunity concept: a framework for multi-scale community ecology.
569 *Ecol. Lett.* **7**: 601–613.
570 Lindström, E.S. and Langenheder, S. (2012) Local and regional factors influencing bacterial
571 community assembly. *Environ. Microbiol. Rep.* **4**: 1–9.
572 Llames, M.E., Huber, P., Metz, S., and Unrein, F. (2017) Interplay between stochastic and

573 deterministic processes in the maintenance of alternative community states in
574 Verrucomicrobia-dominated shallow lakes. *FEMS Microbiol. Ecol.* **93**: 1–10.
575 Logares, R., Deutschmann, I.M., Junger, P.C., Giner, C.R., Krabberød, A.K., Schmidt, T.S.B.,
576 et al. (2020) Disentangling the mechanisms shaping the surface ocean microbiota.
577 *Microbiome* **8**: 1–17.
578 Logares, R., Tesson, S.V.M., Canbäck, B., Pontarp, M., Hedlund, K., and Rengefors, K. (2018)
579 Contrasting prevalence of selection and drift in the community structuring of bacteria and
580 microbial eukaryotes. *Environ. Microbiol.* **20**: 2231–2240.
581 Lu, H.P., Yeh, Y.C., Sastri, A.R., Shiah, F.K., Gong, G.C., and Hsieh, C.H. (2016) Evaluating
582 community-environment relationships along fine to broad taxonomic resolutions reveals
583 evolutionary forces underlying community assembly. *ISME J.* **10**: 2867–2878.
584 Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads.
585 *EMBnet.journal* **17**: 10–12.
586 Martiny, A.C., Coleman, M.L., and Chisholm, S.W. (2006) Phosphate acquisition genes in
587 Prochlorococcus ecotypes: Evidence for genome-wide adaptation. *Proc. Natl. Acad. Sci.*
588 **103**: 12552–12557.
589 Martiny, J.B.H., Jones, S.E., Lennon, J.T., and Martiny, A.C. (2015) Microbiomes in light of
590 traits: A phylogenetic perspective. *Science (80-.).* **350**: aac9323.
591 Mataloni, G., Garraza, G.G., Bölter, M., Convey, P., and Fermani, P. (2010) What shapes
592 edaphic communities in mineral and ornithogenic soils of Cierva Point, Antarctic
593 Peninsula? *Polar Sci.* **4**: 405–419.
594 Mataloni, G., Vinocur, A., and De Tezanos Pinto, P. (2005) Abiotic characterization and
595 epilithic communities of a naturally enriched stream at Cierva Point, Antarctic Peninsula.
596 *Antarct. Sci.* **17**: 163–170.
597 McMurdie, P.J. and Holmes, S. (2013) phyloseq: An R Package for Reproducible Interactive
598 Analysis and Graphics of Microbiome Census Data. *PLoS One* **8**: e61217.
599 Mikryukov, V. (2020) metagMisc: Miscellaneous functions for metagenomic analysis. R
600 package version 0.0.4.

601 Mortola, N., Romaniuk, R., Cosentino, V., Eiza, M., Carfagno, P., Rizzo, P., et al. (2019)
602 Potential Use of a Poultry Manure Digestate as a Biofertiliser: Evaluation of Soil
603 Properties and Lactuca sativa Growth. *Pedosphere* **29**: 60–69.
604 Mysara, M., Vandamme, P., Props, R., Kerckhof, F.-M., Leys, N., Boon, N., et al. (2017)
605 Reconciliation between operational taxonomic units and species boundaries. *FEMS*
606 *Microbiol. Ecol.* **93**: 1–12.
607 Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F., et al.
608 (2013) Patterns and Processes of Microbial Community Assembly. *Microbiol. Mol. Biol.*
609 *Rev.* **77**: 342–356.
610 Ning, D., Yuan, M., Wu, L., Zhang, Y., Guo, X., Zhou, X., et al. (2020) A quantitative
611 framework reveals ecological drivers of grassland microbial community assembly in
612 response to warming. *Nat. Commun.* **11**: 4717.
613 Obu, J., Westermann, S., Vieira, G., Abramov, A., Balks, M.R., Bartsch, A., et al. (2020) Pan-
614 Antarctic map of near-surface permafrost temperatures at 1 km2 scale. *Cryosph.* **14**: 497–
615 519.
616 Okazaki, Y., Fujinaga, S., Salcher, M.M., Callieri, C., Tanaka, A., Kohzu, A., et al. (2021)
617 Microdiversity and phylogeographic diversification of bacterioplankton in pelagic
618 freshwater systems revealed through long-read amplicon sequencing. *Microbiome* **9**: 24.
619 Oksanen, J., Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019)
620 Vegan: Community Ecology Package. R package version 2.5-6.
621 Oloo, F., Valverde, A., Quiroga, M.V., Vikram, S., Cowan, D., and Mataloni, G. (2016) Habitat
622 heterogeneity and connectivity shape microbial communities in South American peatlands.
623 *Sci. Rep.* **6**: 25712.
624 Paradis, E. and Schliep, K. (2019) Ape 5.0: An environment for modern phylogenetics and
625 evolutionary analyses in R. *Bioinformatics* **35**: 526–528.
626 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013) The SILVA
627 ribosomal RNA gene database project: improved data processing and web-based tools.
628 *Nucleic Acids Res.* **41**: D590–D596.

629 R Core Team (2018) R: A Language and Environment for Statistical Computing. Vienna: R
630 Foundation for Statistical Computing. Available online at: https://www.R-project.org/.
631 Ramos Marín, S. (2018) Spatial modelling of the Temperature at the Top of Permafrost in
632 Cierva Point (Antarctic Peninsula). *Master Thesis, Univ. Lisboa*.
633 Rosselló-Móra, R. and Amann, R. (2015) Past and future species definitions for Bacteria and
634 Archaea. *Syst. Appl. Microbiol.* **38**: 209–216.
635 Roy, J., Mazel, F., Sosa-Hernández, M.A., Dueñas, J.F., Hempel, S., Zinger, L., and Rillig,
636 M.C. (2019) The relative importance of ecological drivers of arbuscular mycorrhizal
637 fungal distribution varies with taxon phylogenetic resolution. *New Phytol.* **224**: 936–948.
638 Schloss, P.D. and Handelsman, J. (2004) Status of the microbial census. *Microbiol Mol Biol Rev*
639 **68**: 686–691.
640 Stegen, J.C., Lin, X., Fredrickson, J.K., Chen, X., Kennedy, D.W., Murray, C.J., et al. (2013)
641 Quantifying community assembly processes and identifying features that impose them.
642 *ISME J.* **7**: 2069–2079.
643 Stegen, J.C., Lin, X., Fredrickson, J.K., and Konopka, A.E. (2015) Estimating and mapping
644 ecological processes influencing microbial community assembly. *Front. Microbiol.* **6**: 1–
645 15.
646 Stegen, J.C., Lin, X., Konopka, A.E., and Fredrickson, J.K. (2012) Stochastic and deterministic
647 assembly processes in subsurface microbial communities. *ISME J.* **6**: 1653–1664.
648 Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J., and Zimmerman, J.K. (2006) The
649 problem and promise of scale dependency in community phylogenetics. *Ecology* **87**:
650 2418–2424.
651 Tuorto, S.J., Darias, P., McGuinness, L.R., Panikov, N., Zhang, T., Häggblom, M.M., and
652 Kerkhof, L.J. (2014) Bacterial genome replication at subzero temperatures in permafrost.
653 *ISME J.* **8**: 139–149.
654 Vellend, M. (2010) Conceptual Synthesis in Community Ecology. *Q. Rev. Biol.* **85**: 183–206.
655 Vellend, M. (2016) The theory of ecological communities. 1st ed. Princeton University Press,
656 Woodstock, UK.

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 web-site.

Table S1 Bacterial community sampled.

Fig. S1 Location of sampled sites within the Cierva Point Wetland Complex.

Fig. S2 Rarefaction curves for the 64 samples sequenced.

Fig. S3 Diversity measures for the different environments.

Fig. S4 Relative abundance of the top 10 most abundant phyla.

Fig. S5 Phylogenetic Mantel correlograms for each phylogenetic resolution.

Fig. S6 Violin plots of βMNTD indices across phylogenetic resolutions.

Fig. S7 Principal component analysis (PCA) ordination plot of abiotic data.

683 **Figure legends**

684 **Fig. 1** Principal coordinates analyses (PCoA) based on phylogenetic dissimilarities (βMNTD)
685 between bacterial communities at ASV (a), OTU99 (b), OTU98 (c) and OTU97 (d)
686 phylogenetic resolutions in the Cierva Point Wetland Complex. Different superscript letters in
687 the legends of the ordinations indicate significant differences between environments
688 (PERMANOVA, $P < 0.01$). Env., environments.

689 **Fig. 2** Violin plots of βNTI (a) and RC_{bray} (b) indices across phylogenetic resolutions. Black
690 points indicate the mean value; bars represent ± 1 standard deviation. Differences between
691 taxonomic resolutions were evaluated with global Kruskal–Wallis test and Mann–Whitney *post-*
692 *hoc* pairwise comparisons applying Bonferroni correction. Different letters indicate significant
693 differences in mean between phylogenetic resolutions ($P < 0.05$).

694 **Fig. 3** Contribution of the ecological processes shaping bacterial community structure across
695 phylogenetic resolutions.

696 **Fig. 4** Distance based redundancy analyses (dbRDA) based on ßNTI matrices for ASV (a),
697 OTU99 (b), OTU98 (c) and OTU97 (d) phylogenetic resolutions. Different superscript letters in
698 the legends of the ordinations indicate significant differences between environments
699 (PERMANOVA, $P < 0.01$). Env., environments.

700 **Fig. 5** Distance based redundancy analyses (dbRDA) based on RC_{bray} matrices for ASV (a),
701 OTU99 (b), OTU98 (c) and OTU97 (d) phylogenetic resolutions. Different superscript letters in
702 the legends of the ordinations indicate significant differences between environments
703 (PERMANOVA, $P < 0.01$). Env., environments.

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

2

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

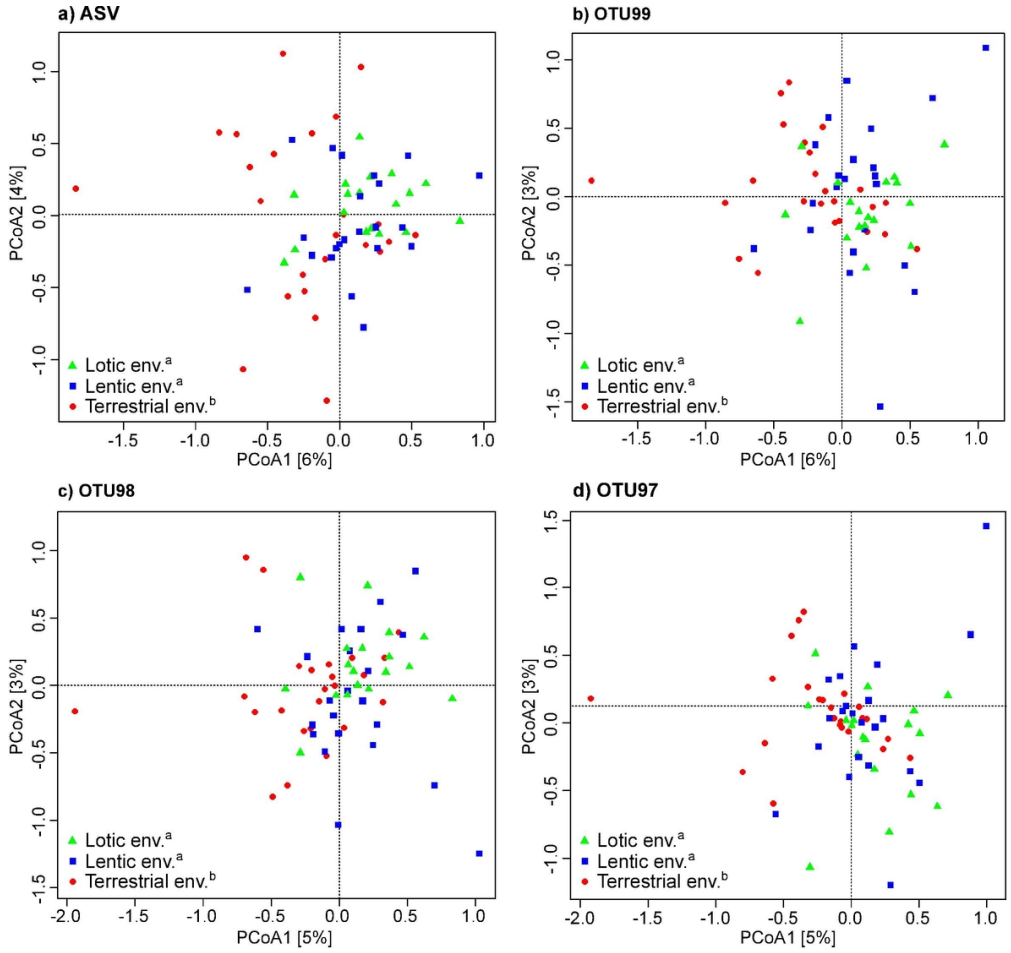

Fig. 1. Principal coordinates analyses (PCoA) based on phylogenetic dissimilarities (βMNTD) between bacterial communities at ASV (a), OTU99 (b), OTU98 (c) and OTU97 (d) phylogenetic resolutions in the Cierva Point Wetland Complex. Different superscript letters in the legends of the ordinations indicate significant differences between environments (PERMANOVA, $P < 0.01$). Env., environments.

100x94mm (300 x 300 DPI)

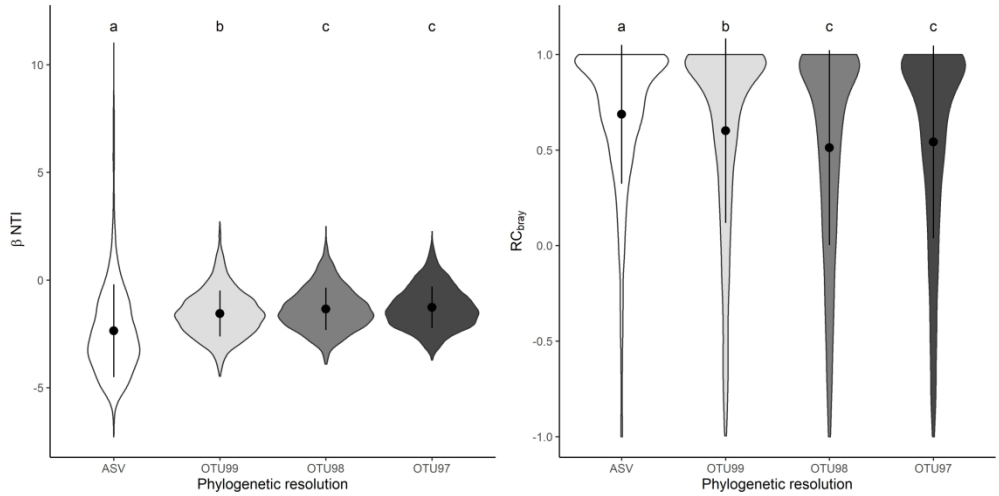

Fig. 2. Violin plots of βNTI (a) and RC_{bray} (b) indices across phylogenetic resolutions. Black points indicate the mean value; bars represent ± 1 standard deviation. Differences between taxonomic resolutions were evaluated with global Kruskal–Wallis test and Mann–Whitney post-hoc pairwise comparisons applying Bonferroni correction. Different letters indicate significant differences in mean between phylogenetic resolutions ($P < 0.05$).

266x133mm (300 x 300 DPI)

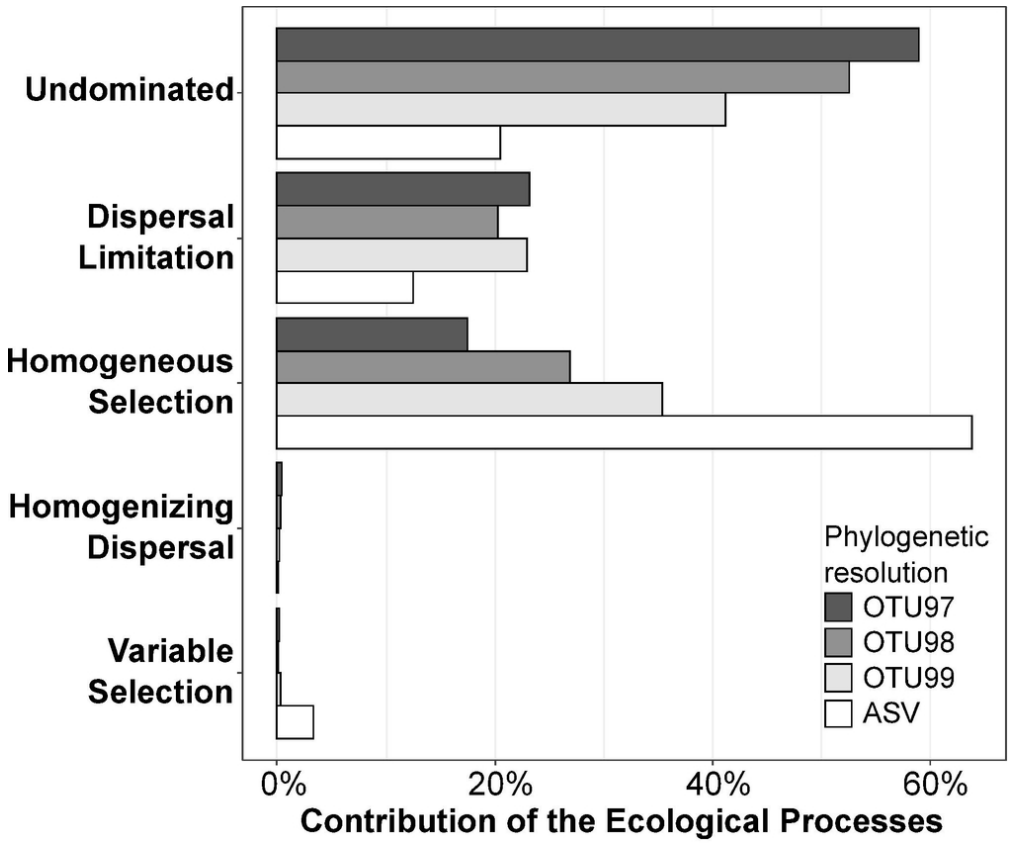

Fig. 3. Contribution of the ecological processes shaping bacterial community structure across phylogenetic resolutions.

85x71mm (300 x 300 DPI)

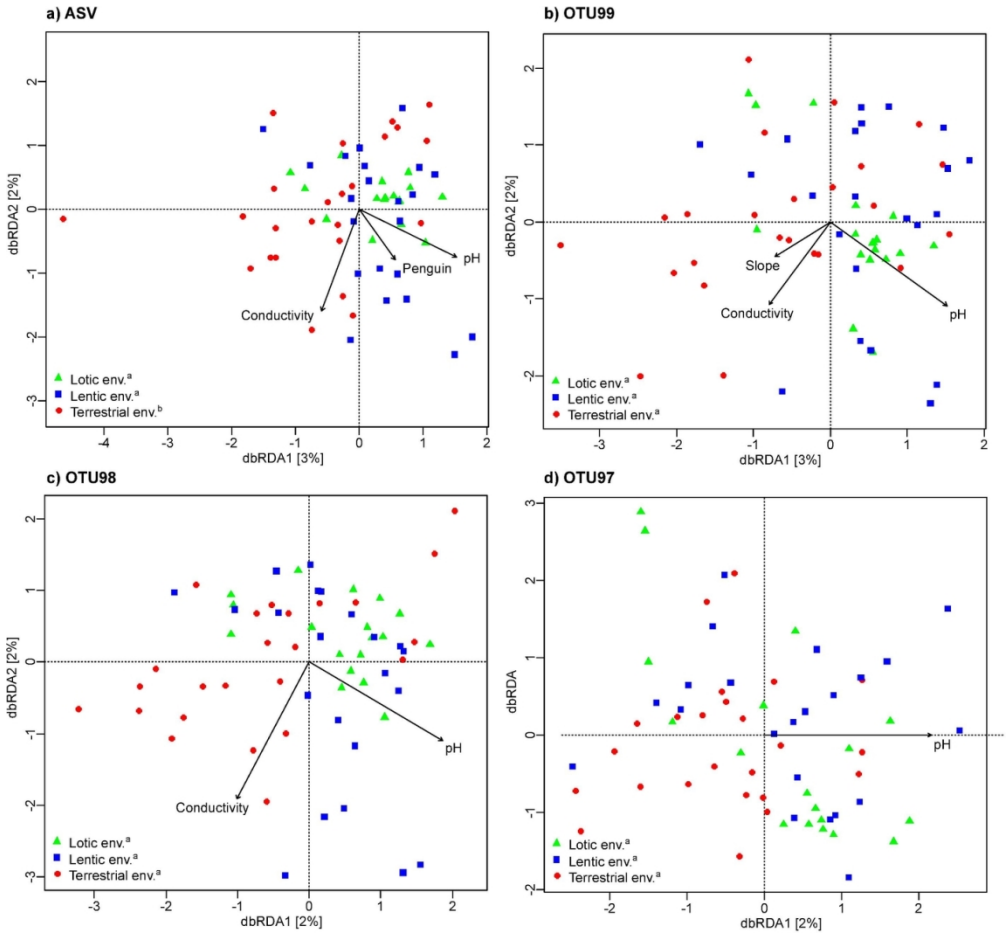

Fig. 4. Distance based redundancy analyses (dbRDA) based on βNTI matrices for ASV (a), OTU99 (b), OTU98 (c) and OTU97 (d) phylogenetic resolutions. Different superscript letters in the legends of the ordinations indicate significant differences between environments (PERMANOVA, $P < 0.01$). Env., environments.

119x111mm (300 x 300 DPI)

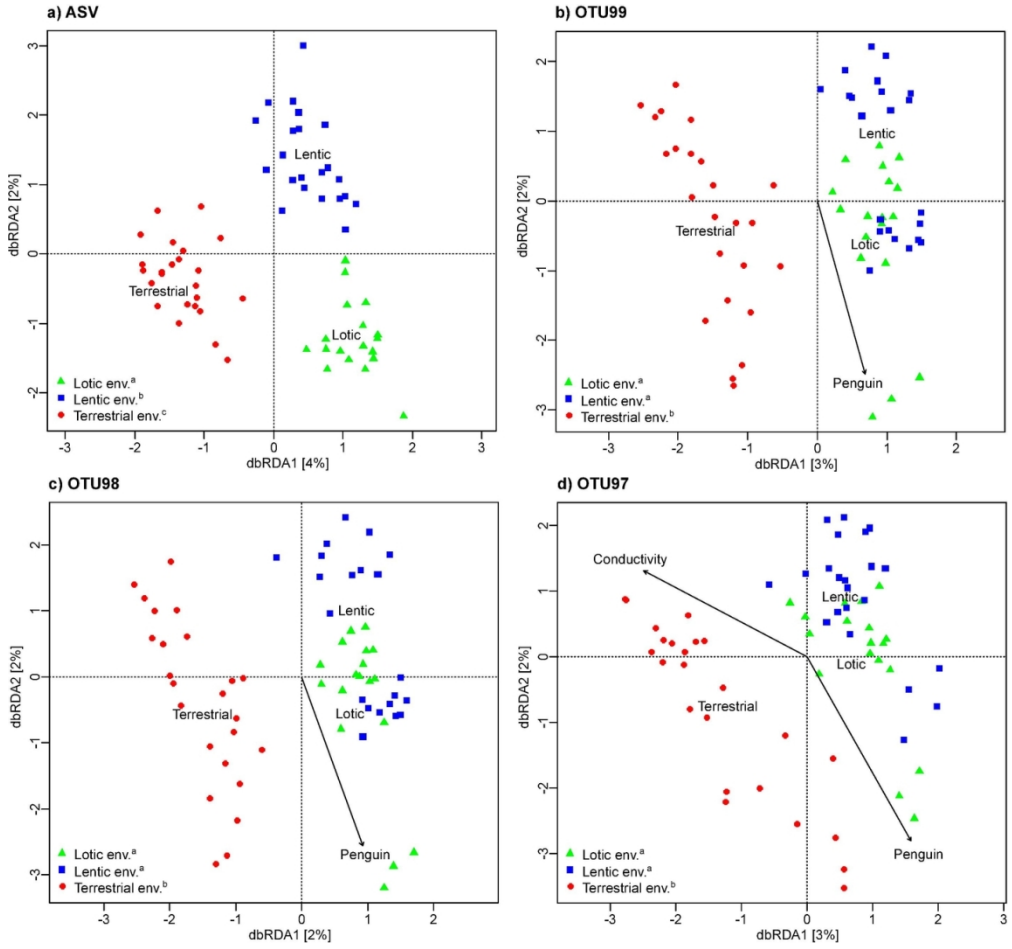

Fig. 5. Distance based redundancy analyses (dbRDA) based on RC_{bray} matrices for ASV (a), OTU99 (b), OTU98 (c) and OTU97 (d) phylogenetic resolutions. Different superscript letters in the legends of the ordinations indicate significant differences between environments (PERMANOVA, $P < 0.01$). Env., environments.

120x112mm (300 x 300 DPI)