FEMS Microbiology Ecology, 91, 2015

doi: 10.1093/femsec/fiu005
Advance Access Publication Date: 5 December 2014
Research Article

RESEARCH ARTICLE

Distinct composition signatures of archaeal and bacterial phylotypes in the Wanda Glacier forefield, Antarctic Peninsula

Igor S. Pessi[1,§], César Osorio-Forero[2], Eric J. C. Gálvez[2,¥], Felipe L. Simões[3], Jefferson C. Simões[4], Howard Junca[2,*,¶] and Alexandre J. Macedo[1,5]

[1]Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves, 9500 - Agronomia, Porto Alegre - RS, 91501-970, Brazil, [2]Research Group Microbial Ecology: Metabolism, Genomics and Evolution of Communities of Environmental Microorganisms, CorpoGen, Carrera 5 66A-34, 110231, Bogotá, DC, Colombia, [3]Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves, 9500 - Agronomia, Porto Alegre - RS, 91501-970, Brazil, [4]Centro Polar e Climático, Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves, 9500 - Agronomia, Porto Alegre - RS, 91501-970, Brazil and [5]Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Avenida Ipiranga, 2752 - Azenha, Porto Alegre - RS, 90610-000, Brazil

*Corresponding author. Research Group Microbial Ecology: Metabolism, Genomics and Evolution of Communities of Environmental Microorganisms, CorpoGen, Carrera 5 66A-34, 110231, Bogota, DC, Colombia; Tel: +57-1-805-0106; Fax: +57-1-348-4607; E-mail: info@howardjunca.com
§Present Address: Centre d'Ingénierie des Protéines, Université de Liège, Liège, Belgium.
¥Present Address: Research Group Microbial Immune Regulation, Helmholtz Centre for Infection Research, Braunschweig, Germany
¶Present Address: Faculty of Basic & Applied Sciences, UMNG, Campus Cajicá, Bogotá, DC, Colombia.

One Sentence Summary: Pyrosequencing analysis of 16S amplicons from soil microbiomes at the Wanda Glacier forefield, King George Island, Antarctic Peninsula, reveals distinct archaeal and bacterial phylotypes involved in early successional stages after deglaciation.

Editor. Tillmann Lueders

ABSTRACT

Several studies have shown that microbial communities in Antarctic environments are highly diverse. However, considering that the Antarctic Peninsula is among the regions with the fastest warming rates, and that regional climate change has been linked to an increase in the mean rate of glacier retreat, the microbial diversity in Antarctic soil is still poorly understood. In this study, we analysed more than 40 000 sequences of the V5-V6 hypervariable region of the 16S rRNA gene obtained by 454 pyrosequencing from four soil samples from the Wanda Glacier forefield, King George Island, Antarctic Peninsula. Phylotype diversity and richness were surprisingly high, and taxonomic assignment of sequences revealed that communities are dominated by *Proteobacteria*, *Bacteroidetes* and *Euryarchaeota*, with a high frequency of archaeal and bacterial phylotypes unclassified at the genus level and without cultured representative strains, representing a distinct microbial community signature. Several phylotypes were related to marine microorganisms, indicating the importance of the marine environment as a source of colonizers for this recently deglaciated environment. Finally, dominant phylotypes were related to different microorganisms possessing a large array of metabolic strategies, indicating that early successional communities in Antarctic glacier forefield can be also functionally diverse.

Key words: next-generation sequencing; microbial diversity; recently deglaciated environments

Received: 30 June 2014; **Accepted:** 27 October 2014

INTRODUCTION

Increasing attention has been given to the effects of climate change on Earth's biodiversity, since perceptible atmospheric changes have been observed in the past decades (Blois *et al.*, 2013). The Antarctic Peninsula is among the regions that showed the fastest warming rate in the last five decades, accompanied by significant changes in the precipitation regimes (Vaughan *et al.*, 2001; Turner *et al.*, 2005). Recent regional climate change has been linked to the increase in the mean rate of glacier retreat observed in this region since 1954 (Cook *et al.*, 2005; Davies *et al.*, 2012). As a consequence of glacier retreat, terrain that has been previously covered by ice becomes exposed, providing new habitats for the colonization by pioneering organisms.

Early successional communities in glacier forefields (i.e. communities that establish themselves relatively quickly after glacier retreat) are usually limited by the low availability of organic matter, but have key roles in soil development and nutrient cycling, facilitating the establishment of more complex communities (Nemergut *et al.*, 2007; Schütte *et al.*, 2010; Zumsteg *et al.*, 2012; Bajerski and Wagner 2013). In Antarctica, a region largely dominated by microbial ecosystems, microorganisms are the main drivers of biogeochemical cycling (Wynn-Williams 1996; Vincent 2002). Therefore, the assessment and monitoring of Antarctic microbial diversity is vital for predicting the impact of climate change on this continent, since the effects of climate change have a higher and more rapid impact in polar environments (Mayewski *et al.*, 2013). Significant efforts to assess the microbial diversity in Antarctic soils have been carried out to date, showing the presence of highly diverse microbial communities (Saul *et al.*, 2005; Aislabie *et al.*, 2006; Teixeira *et al.*, 2010; Ganzert *et al.*, 2011; Chong *et al.*, 2012; Kim *et al.*, 2012; Roesch *et al.*, 2012; Bajerski and Wagner 2013; Zdanowski *et al.*, 2013). However, still much more additional knowledge is needed to accomplish a full description of Antarctic microbial diversity. High-throughput sequencing technologies such as Roche's 454 pyrosequencing have been proven very valuable for the detection of rare and/or difficult to cultivate taxa (Sogin *et al.*, 2006; Roesch *et al.*, 2007), since culture-based methods or molecular methods such as DGGE and clone libraries are known to substantially underestimate the diversity found in natural environments (Schloss and Handelsman 2004).

Wanda Glacier, King George Island, Antarctic Peninsula, has retreated considerably in the last decades as a consequence of the atmospheric warming recorded in the northern Antarctic Peninsula in the last five decades (Rosa *et al.*, 2009). The Wanda Glacier forefield was formed in the late 1990s, when the glacier shifted from a tidewater terminus to a land terminus glacier. This environment is under constant reworking by melt-water flow from the glacier, gravitational and melting processes and through the action of waves and annual tide fluctuations. Therefore, this site presents itself as an interesting environment for microbial ecology studies, more specifically in regard to microbial distribution and processes in recently deglaciated environments. In a previous study, the functional diversity of microbial communities in the Wanda Glacier forefield was assessed by means of C source utilization using Biolog EcoPlates (Pessi *et al.*, 2012). Results showed that communities were highly diverse and capable of metabolizing a wide range of C compounds. Here, we report the microbial diversity found in this peculiar environment based on the taxonomic evaluation of partial 16S rRNA gene sequences obtained by 454 pyrosequencing, applying methodological and technical adaptations we have developed for the analysis of microbial communities in other contrasting extreme environments (Bohorquez *et al.*, 2012).

MATERIALS AND METHODS

Study site and soil sampling

King George Island is the largest island of the South Shetland Islands, Antarctic Peninsula, with an area of approximately 1250 km^2, of which 93% is ice covered (Braun *et al.*, 2001). Due to the location of the ice cap off the Antarctic Peninsula northeast, this region presents a relatively warm, maritime climate with low annual variability in mean monthly atmospheric temperature, which can rise above 0°C in summer. Wanda Glacier is a land terminus glacier located on the oriental coast of Admiralty Bay, King George Island (62°6′30′ S, 58°21′0′ W; Fig. 1), comprising an area of 1.56 km^2. This glacier has been retreating rapidly since 1956, having lost 0.64 km^2 in the last five decades. The exposure of the Wanda Glacier forefield is linked to the stabilization of retreats since the late 1990s, when the glacier became land-based terminus (Rosa *et al.*, 2009). This deglaciation environment, which now extends for approximately 200 m from the actual glacier front to the shoreline, comprises a small forefield area and a proglacial lagoon, the latter formed as a consequence of melt-water glacier outflow. Sediments are transported towards the ocean through a channel in the proglacial lagoon, and the glacier forefield gets flooded eventually due to the action of wave and tide fluctuations. Soil samples were collected in January 2010 from four sites in the Wanda Glacier forefield (Fig. 1) after discarding the top 1–2 cm layer, with each sample consisting of five sub-samples within a 1 m^2 quadrant. Samples S1 and S4, collected close to the shore, are characterized by a coarse grain size, with very low nutrient content (C_{org} = 0.09%; $N < 0.01\%$; see Pessi *et al.*, 2012). Nutrient content is higher in samples S2 and S3, collected close to the proglacial lagoon, which also present a finer granularity (C_{org} = 0.26–0.34%; N = 0.02%). Samples are characterized by alkaline pH values between 8.59 (sample S2) and 9.86 (sample S3), and water contents between 7.51% (sample S4) and 17.91% (sample S1). All sites are ice-free since 1995–2000 (Rosa *et al.*, 2009), and have therefore the same time of exposure since glacier retreat.

DNA extraction and amplification of the 16S rRNA gene

Environmental DNA was extracted from the soil samples using the PowerSoil DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA), according to manufacturer's instructions with some modifications: 5 g of soil was added to a sterile 15 mL falcon tube and the extraction was carried out using 10 times the amount of each reagent, resulting in a final volume of 500 μL. The primer set 807F (5′-GGATTAGATACCCBRGTAGTC-3′) and 1050R (5′-AGYTGDCGACRRCCRTGCA-3′) was used to amplify the V5-V6 hypervariable region of the 16S rRNA gene in a two-step PCR protocol as previously described by Bohorquez *et al.*, (2012). First, the target region was amplified using primers without the barcode and 454 adaptor sequences. PCR products were then used as templates on a second reaction, including primers to extend the barcode and 454 adaptor sequences to the amplicons. The first PCR reaction consisted of 2 ng μL^{-1} DNA, 1X PCR buffer, 0.2 mM of each dNTP, 0.2 μM of each primer, 1.5 mM $MgCl_2$, 1 U Platinum Taq DNA Polymerase (Invitrogen, Carlsbad, CA, USA) and sterile Milli-Q water to a final volume of 50 μL. Amplification was performed using an initial denaturation step at 94°C for 2 min, followed by 30 cycles at 94°C for 30 s, 58°C for 30 s

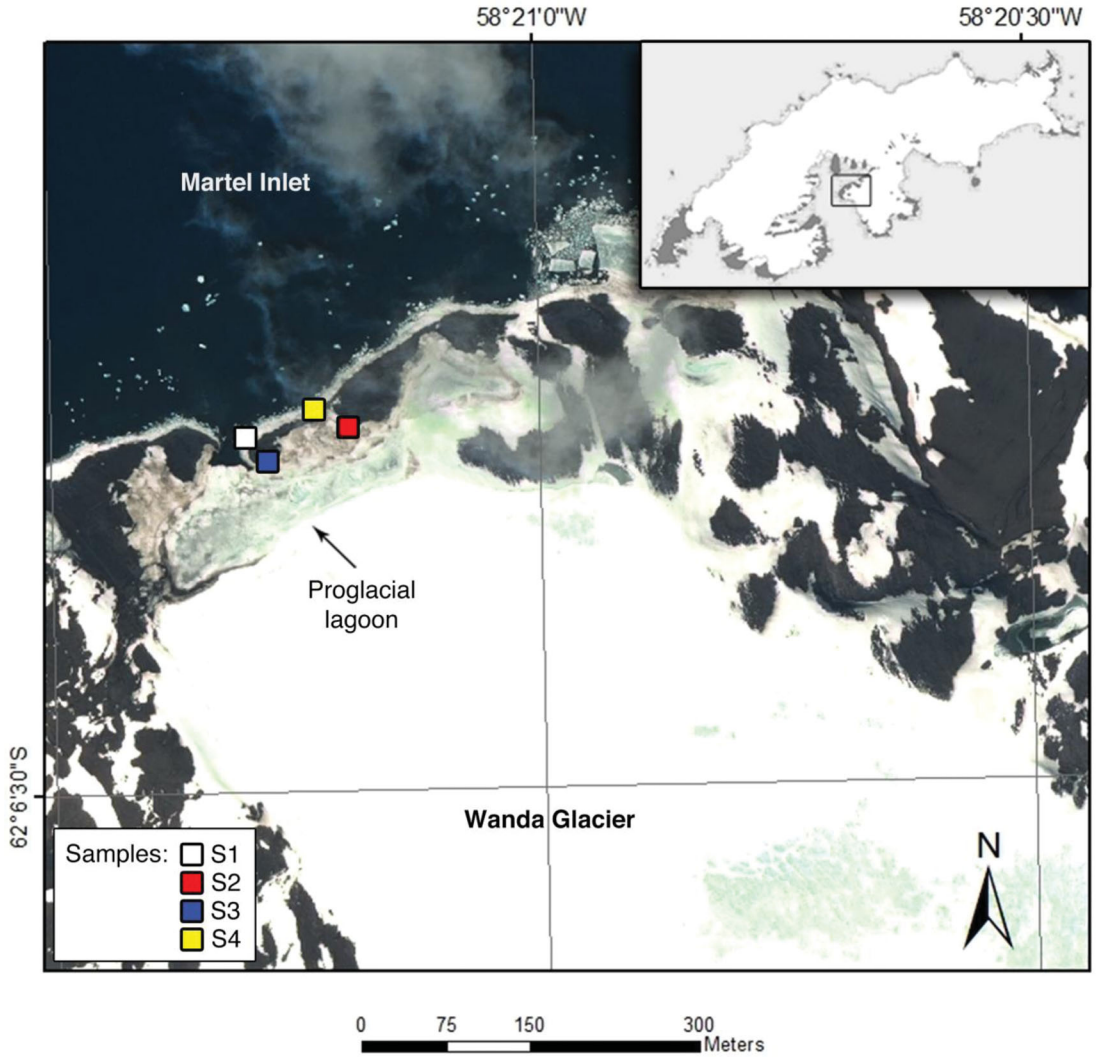

Figure 1. Satellite image of the Wanda Glacier showing the sampling sites used in this study (QuickBird image provided by Laboratório de Monitoramento da Criosfera, FURG, Rio Grande, Brazil). Inset: map of King George Island, Antarctic Peninsula, showing the location of the Wanda Glacier.

and 72°C for 1 min, and a final extension at 72°C for 5 min. Following amplification, PCR products were purified using the UltraClean PCR Clean-Up Kit (MO BIO Laboratories). The second PCR reaction was carried out at the same conditions as above but using reverse primers containing the barcode and 454 adaptor sequences and changing the number of cycles to five and the concentration of template DNA to 10 ng μL^{-1}. PCR products were purified using the Agencourt AMPure XP Kit (Beckman Coulter Inc., Brea, CA, USA) and pooled at equimolar concentrations. Sequences were obtained using the 454 GS FLX Titanium technology (454 Life Sciences, Branford, CT, USA).

Bioinformatic procedures

Quality filtering of reads, taxonomic assignment of sequences and operational taxonomic unit (OTU) picking were performed using the Quantitative Insights into Microbial Ecology (QIIME) package (Caporaso *et al.*, 2010). First, sequences were trimmed and demultiplexed using QIIME's default parameters (minimum quality score = 25, minimum length = 200, no ambiguous bases allowed and no mismatches allowed in the primer sequence). The resulting sequences were denoised using Denoiser (Reeder and Knight 2010), and chimeric sequences were removed using ChimeraSlayer (Haas *et al.*, 2011). Quality-filtered sequences were classified using the RDP Naïve Bayesian Classifier (Wang *et al.*, 2007) with a confidence threshold of 80%, and clustered into OTUs at 97% similarity using CD-HIT (Huang *et al.*, 2010). Rarefaction curves were generated using mothur (Schloss *et al.*, 2009). Chao1, ACE and Shannon indices were calculated using the phyloseq package in R (McMurdie and Holmes 2013), after rarefying datasets to 6792 sequences. The dominant phylotypes present in each dataset were selected by adding the highest relative abundances until the cumulative relative abundance reached 50%. One representative sequence was extracted for each dominant phylotype, classified using the RDP Classifier and searched for its nearest neighbour using the SeqMatch tool from the RDP. Phylogenetic reconstruction of archaeal dominant phylotypes was carried out along with the sequences from their five best SeqMatch isolate hits. Sequences were aligned using MUSCLE (Edgar 2004), and maximum likelihood trees based on the GTR model were inferred using MEGA5 (Tamura *et al.*, 2011). The sequences of the dominant phylotypes were deposited in GenBank under the accession numbers listed in Table S1.

Table 1. Microbial richness and diversity indices obtained for four soil samples from the Wanda Glacier forefield.

Sample	Number of OTUs	ACE index	Chao1 index	Shannon index
S1	782	2034	1697	5.15
S2	521	1159	1136	4.32
S3	903	2269	1930	5.41
S4	771	1938	1458	5.10

Multivariate analyses

To assess the relationships between environmental parameters and microbial community structures at the phylum and class levels, multivariate constrained correspondence analyses (CCA) were carried out using the vegan package in R (Oksanen *et al.*, 2013). Environmental variables included soil water, pH, C_{org}, N, P, K, Ca, Mg, Mn and Na contents (Pessi *et al.*, 2012). All variables except pH were ln+1 transformed. Similarly, a CCA was carried out to investigate possible relationships between taxonomic diversity and the previously observed patterns of C source utilization (Pessi *et al.*, 2012). In this case, the C source utilization data were used as dependent (response) variables and the taxonomic diversity data as independent (explanatory) variables.

RESULTS

Microbial community composition

Partial archaeal and bacterial 16S rRNA gene sequences were obtained from DNA extracted from four soil samples collected in the Wanda Glacier forefield (Fig. 1). A total of 49 003 reads were obtained by 454 pyrosequencing and submitted to stringent quality control, resulting in 40 802 high-quality sequences with an average length of 304 bp. Sequences clustered into 2426 OTUs at a 97% similarity cut-off, ranging from 649 to 1025 OTUs per sample. All datasets had a similar pattern of OTU accumulation; however, rarefaction curves did not reach an asymptote (Fig. S1, Supporting Information). To account for uneven sampling efforts, alpha diversity was calculated after rarefying datasets to 6792 sequences, the number of sequences in the smallest dataset (sample S3). Even after rarefying datasets, phylotype richness still differed almost 2-fold across samples, ranging from 521 to 903 (Table 1). Similarly, non-parametric estimates of phylotype richness ranged from 1159 to 2269 (ACE index) and from 1136 to 1930 (Chao1 index), following the same trend at every estimate. Diversity indices and rarefaction analyses pointed to a lower phylotype richness in sample S2 in comparison to the other samples.

Taxonomic assignment of sequences revealed the presence of 36 phyla from both prokaryotic domains across all four samples (Fig. 2a). *Proteobacteria* was by far the most predominant phylum (47.8% of the sequences), followed by *Bacteroidetes* (18.5%) and *Euryarchaeota* (10.0%). *Actinobacteria* (4.9%), *Planctomycetes* (3.1%), *Crenarchaeota* (2.8%), *Acidobacteria* (2.0%) and *Verrucomicrobia* (1.5%) were also found in lower abundance. Moreover, 4.9% of the sequences could not be assigned to any prokaryotic phyla. Overall, minor differences were observed between samples at the phylum level. Communities S2 and S3 had a higher proportion of *Proteobacteria* and *Bacteroidetes* and a lower relative abundance of *Euryarchaeota*. Moreover, *Crenarchaeota* was nearly absent in communities S3 and S4. At the class level, communities were composed mainly of taxa belonging to the classes

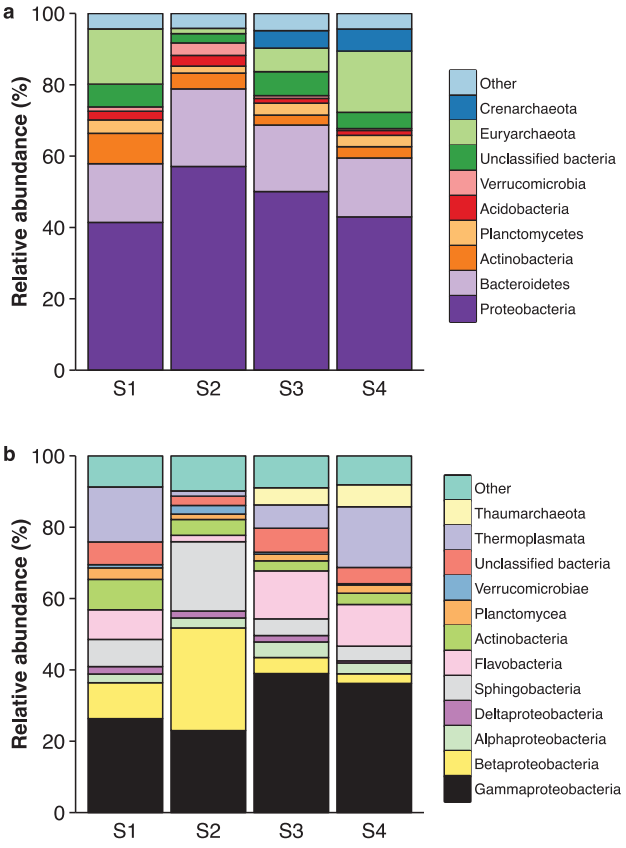

Figure 2. Taxonomic assignment at the (a) phylum and (b) class levels of partial 16S rRNA gene sequences obtained from the four soil samples from the Wanda Glacier forefield. Taxa with relative abundances lower than 1% are grouped in 'Other'.

Gammaproteobacteria (31.0%), *Betaproteobacteria* (11.6%), *Thermoplasmata* (10.0%), *Sphingobacteria* (9.1%) and *Flavobacteria* (9.0%) (Fig. 2b). Community dissimilarities were more evident at this deeper taxonomic level. Community S2 had a considerably distinct microbial community, with a higher proportion of *Betaproteobacteria* and *Sphingobacteria*. Finally, differences in community structure between the four samples were even more contrasting at the phylotype level. Only 3.1% of the OTUs were common to all four samples and 68.3% of the OTUs were not detected in more than one sample.

Analysis of dominant phylotypes

Due to the high OTU richness found in the samples, a more thorough analysis was undertaken with the dominant phylotypes of each dataset, determined by adding the highest relative abundances until the cumulative relative abundance reached 50%. This procedure selected 90 OTUs (between 9 and 33 OTUs per dataset), which were classified using the RDP Classifier and analysed using the SeqMatch tool of the RDP to find the nearest neighbour in the database.

Only 2 of the 13 dominant archaeal phylotypes were classified at the genus level and/or have cultured representatives, i.e. had at least 97% similarity with sequences from isolated microorganisms (Table S1). These represented 4.8 and 6.3% of the sequences in samples S3 and S4, respectively, and were assigned to the genus *Nitrosopumilus*, with 97.8% similarity to a strain of *Nitrosopumilus maritimus* isolated from a tropical

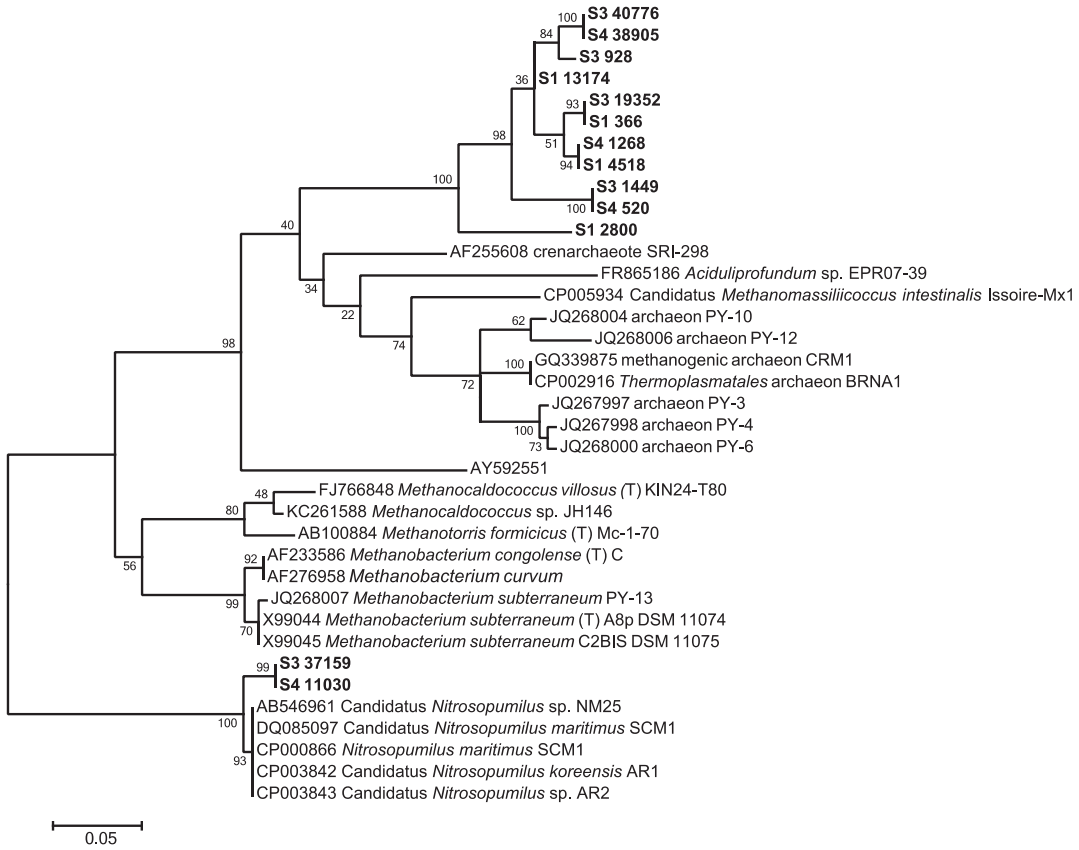

Figure 3. Phylogenetic analysis of the dominant archaeal phylotypes in the Wanda Glacier forefield. Maximum-likelihood tree based on the GTR model of phylotype sequences (in bold) and their five best SeqMatch isolate hits. Bootstrap values are shown as percentages of 1000 replicates.

marine tank in the USA (Könneke *et al.*, 2005). The other 11 dominant archaeal phylotypes, which were classified only up to the class level, individually represented up to 13.4% of the sequences in each dataset. They were only distantly related to a strain of *Methanobacterium subterraneum* (80.1% similarity) and to other uncharacterized archaea isolated from Chinese permafrost (79.5–82.1%), as well as to methanogenic rumen and gut archaea (78.8–82.0%). Moreover, phylogenetic analyses showed that they form a cluster within the class *Thermoplasmata*, separated from their nearest neighbours (Fig. 3).

The majority of the dominant bacterial phylotypes were classified at the genus level and/or have cultured representatives, indicating the presence of the genera *Ilumatobacter*, *Nocardioides*, *Flavobacterium*, *Maribacter*, *Ulvibacter*, *Winogradskyella*, *Antarcticimonas*, *Bizionia*, *Maritimimonas*, *Sediminicola*, *Nitrospira*, *Pelomonas*, *Acidovorax*, *Thiobacillus*, *Methylotenera*, *Glaciecola*, *Zhongshania*, *Neptunomonas*, *Pseudomonas*, *Arenicella* and *Granulosicoccus* (Table S1). Sequences were related to bacteria found in a wide range of different environments across the globe, such as seashore sand (Japan), hydrocarbon-contaminated soil (Antarctica), freshwater (Japan, Netherlands and Germany), marine sediment (Antarctica, North Sea and China), marine basalt (Iceland), seawater (Antarctica and Atlantic Ocean), rice field soil (Italy) and whale carcass (Japan). Moreover, several dominant phylotypes had over 99% similarity to bacteria isolated from Antarctic environments. Except for one phylotype related to a strain of *Nocardioides* sp. isolated from hydrocarbon-contaminated soil (Saul *et al.*, 2005), all were related to bacteria isolated from Antarctic seawater or marine sediment, including *Granulosicoccus* sp., *Flavobacterium frigidarium*, *Antarcticimonas flava*, *Zhongshania guokonii*, *Pseudomonas* sp. and an unclassified bacteria (classified as *Glaciecola* by the RDP Classifier).

Several highly abundant bacterial phylotypes were not classified at the genus level and had low similarity with sequences available in the RDP. One phylotype, representing 12.6% of the sequences in sample S2, had 95.0% similarity to an unclassified *Bacteroidetes* isolated from a hot spring in Svalbard. Another phylotype, present in all samples with relative abundances ranging from 0.9 to 5.0%, had 95.4% similarity to a strain of *Panacagrimonas perspica* isolated from a ginseng field in Korea. One phylotype, accounting for 1.0 and 4.3% of the sequences in samples S3 and S4, respectively, had 95.4% of similarity to an unclassified *Gammaproteobacteria* isolated from seawater. Finally, one phylotype present in samples S1, S3 and S4 had very low similarity (84.9%) to a strain of *Allocatelliglobosispora scoriae* isolated from volcanic ash in Korea.

Links between microbial community structure, environmental variables and C source utilization

The relationship between environmental parameters and microbial community composition at the phylum and class levels was assessed by CCA (Fig. S2, Supporting Information). However, multivariate analysis failed to detect significant correlations between microbial community structure and environmental variables ($P > 0.05$), and no clustering of samples was evident.

Putative relationships between previously observed patterns of C source utilization (Pessi *et al.*, 2012) and taxonomic

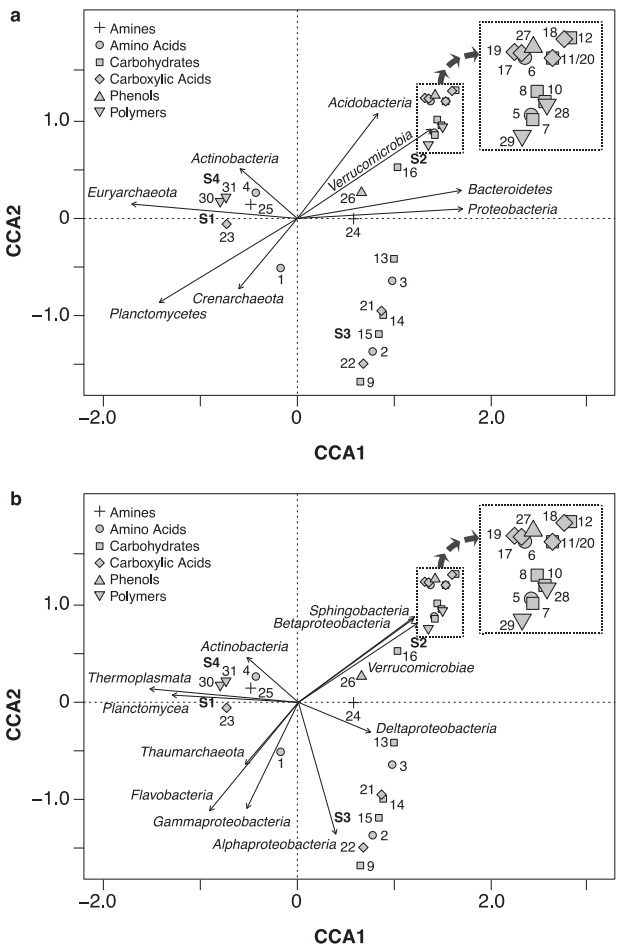

Figure 4. A CCA showing the relationship between C source utilization and microbial taxonomic diversity at the (a) phylum and (b) class levels in the Wanda Glacier forefield. Samples are in bold, taxa are italicized and C sources are represented by symbols followed by numbers. Only taxa with relative abundance higher than 1% are shown. 1: L-arginine; 2: L-asparagine; 3: glycyl-L-glutamic acid; 4: L-phenylalanine; 5: L-serine; 6: L-threonine; 7: D-cellobiose; 8: I-erythritol; 9: D-galactonic acid γ-lactone; 10: N-acetyl-D-glucosamine; 11: glucose-1-phosphate; 12: β-methyl-D-glucoside; 13: DL-α-glycerol phosphate; 14: α-D-lactose; 15: D-mannitol; 16: D-xylose; 17: γ-hydroxybutyric acid; 18: α-keto butyric acid; 19: D-galacturonic acid; 20: D- glucosaminic acid; 21: itaconic acid; 22: D-malic acid; 23: pyruvic acid methyl ester; 24: phenylethylamine; 25: putrescine; 26: 2-hydroxy benzoic acid; 27: 4-hydroxy benzoic acid; 28: α-cyclodextrin; 29: glycogen; 30: Tween 40; 31: Tween 80. Data on C source utilization were taken from Pessi *et al.*, (2012).

diversity at the phylum and class levels were also assessed by CCA (Fig. 4). The abundance of *Thermoplasmata* (phylum *Euryarchaeota*) was correlated with the utilization of Tween 40, Tween 80, L-phenylalanine, pyruvic acid methyl ester and putrescine ($P < 0.05$) by the communities. *Sphingobacteria* (phylum *Bacteroidetes*) and *Betaproteobacteria* were correlated with the intake of several C sources, including compounds of all chemical natures except amines ($P < 0.05$). Similarly, the abundance of *Alphaproteobacteria* was correlated with the utilization of different amino acids, carbohydrates and carboxylic acids ($P < 0.05$).

DISCUSSION

This exploratory study was carried out to describe the microbial diversity found in the Wanda Glacier forefield (King George Island, Antarctic Peninsula). The exposure of the Wanda Glacier forefield as a result of glacier retreat is a relatively recent event (<20 years) (Rosa *et al.*, 2009), which reflects in a low availability of organic matter for the development of microorganisms. However, a previous study has shown that this environment harbours microbial communities with surprisingly high metabolic potentials as assessed by culture-dependent techniques (Pessi *et al.*, 2012). Here, we extend this description by assessing the taxonomic diversity of these microbial communities based on pyrosequencing analysis of 16S rRNA gene amplicons.

Non-parametric estimates of phylotype richness calculated for the communities (Table 1) were surprisingly high given the oligotrophic characteristic of the soils. However, it was comparable to what has been recently observed in other similar Antarctic and Arctic environments (Schütte *et al.*, 2010; Teixeira *et al.*, 2010). Shannon diversity indices calculated for the communities were two to three times higher than what has been previously observed for these communities by means of C source utilization (Pessi *et al.*, 2012). However, no clear correspondence was observed between the diversity estimates obtained for each community in both studies. Sample S2, which had the lowest phylotype diversity (Table 1), was the most functionally diverse sample based on the number of utilized C compounds (Pessi *et al.*, 2012). At the same time, samples S1 and S4, which were the least functionally diverse, had intermediate values of phylotype diversity among the analysed samples. However, a lack of correspondence between phylotype and functional diversity indices can be due to limitations of both techniques. Biolog EcoPlates analyses are considered a culture-dependent approach, which may favour the growth of specific microorganisms and therefore underestimate the diversity and the extent of functions in the environment. Microbial diversity estimations based on PCR amplification of the 16S rRNA gene are known to suffer from technical limitations as well, either due to preferential DNA extraction or gene amplification, or differences in the number of copies of the target gene between taxa (Polz and Cavanaugh 1998), although recent studies have supported the general reliability of pyrosequencing results (Jumpstart Consortium Human Microbiome Project Data Generation Working Group 2012).

An almost 2-fold variation in phylotype richness was observed across the four samples (Table 1, Fig. S1, Supporting Information), showing a relative heterogeneity in the structure of the microbial communities at this small spatial scale (Fig. 1). Indeed, 68.3% of the OTUs were unique to one sample and multivariate analysis showed that communities are distinct from one another (Fig. S2, Supporting Information). On the other hand, taxonomic assignment of sequences revealed a somewhat more homogeneous community composition at the phylum and class levels (Fig. 2). This is consistent with recent observations that soil bacterial community structures seem to be stable at higher taxonomic levels, while at the genus and species level communities are more sensitive to spatial and environmental gradients (Chong *et al.*, 2012 and references therein). Despite differences in the microbial community structure observed between samples, no influence of environmental parameters was detected by multivariate analysis (Fig. S2, Supporting Information). This may be explained to some extent by the restricted amount of sampling points available and the relatively short environmental gradients they represent.

In general, *Proteobacteria* and *Bacteroidetes* were the dominant bacterial groups, with *Actinobacteria*, *Planctomycetes*, *Acidobacteria* and *Verrucomicrobia* also representing a significant fraction of the communities (Fig. 2a). These are among the nine bacterial phyla most frequently found in soils worldwide (Janssen 2006), and several studies have demonstrated the ubiquity of these phyla

in Antarctic habitats, including glacier forefield (Foong, Ling and González 2010; Bajerski and Wagner 2013; Zdanowski *et al.*, 2013), soil (Aislabie *et al.*, 2006; Foong, Ling and González 2010; Teixeira *et al.*, 2010; Ganzert *et al.*, 2011; Chong *et al.*, 2012; Kim *et al.*, 2012; Roesch *et al.*, 2012) and coastal waters (Ghiglione and Murray 2012; Zeng *et al.*, 2014), as well as in other cold habitats such as the Himalayan Mountain Ranges (Srinivas *et al.*, 2011) and the Andes (Nemergut *et al.*, 2007). Our results are also consistent with previous descriptions of microbial diversity in rhizosphere soil from Keller Peninsula, King George Island, Antarctic Peninsula (Teixeira *et al.*, 2010; Roesch *et al.*, 2012). Microbial assemblages in soil samples from different plant communities and soil types were dominated by *Proteobacteria*, *Bacteroidetes*, *Acidobacteria* and *Actinobacteria* (Roesch *et al.*, 2012). However, Teixeira *et al.*, (2010) found that *Firmicutes* also constituted an important fraction of the microbial communities in the rhizosphere of *Deschampsia antarctica* and *Colobanthus quitensis*. Moreover, the presence of *Archaea*, which constitutes an important fraction of the diversity found in the present study, was not reported in both works.

Euryarchaeota was the dominant archaeal phylum in the samples and the third most abundant group after *Proteobacteria* and *Bacteroidetes* (Fig. 2a). Several studies have shown that *Euryarchaeota* are ubiquitous in the Antarctic Ocean (DeLong *et al.*, 1994; Massana *et al.*, 1998; López-García *et al.*, 2001; Alonso-Sáez *et al.*, 2011). Moreover, Zumsteg *et al.*, (2012) showed that *Euryarchaeota* are abundant in the Damma glacier forefield (central Switzerland), with a higher predominance in younger (2 years old) than intermediate (62 years) and old soil (110 years). In our study, the relative abundance of *Euryarchaeota* was slightly higher in soils with lower organic matter (samples S1 and S4; Fig. 2). This is in agreement to what has been suggested in the aforementioned study that *Euryarchaeota* seem to be adapted to the oligotrophic conditions found in recently exposed soil, being outcompeted by other microorganisms in soil of later successional stages, richer in organic matter (Zumsteg *et al.*, 2012). However, the role of euryarchaeal populations in this environment remains yet to be elucidated.

Analysis of the most abundant phylotypes revealed that the communities are dominated by a small number of archaeal and bacterial OTUs, unclassified at the genus level and without cultured representatives (Table S1). The only identified archaeal phylotype was related to the chemolithoautotrophic nitrifier *N. maritimus*, which is able to grow using NH_3 and CO_2 as energy and carbon sources, respectively (Könneke *et al.*, 2005). This organism and other closely related uncultivated mesophilic archaea, which were previously affiliated with *Crenarchaeota*, represent in fact a deep phylogenetic branch, constituting the recently proposed phylum *Thaumarchaeota* (Brochier-Armanet *et al.*, 2008). All other archaeal phylotypes, with relative abundances as high as 13.4%, had extremely low similarity to sequences from both isolated and uncultured microorganisms (as low as 78.8%), and phylogenetic analyses showed that they form a divergent cluster within the class *Thermoplasmata* (Fig. 3). Unclassified bacterial phylotypes were also found among the dominant members of the communities, representing up to 12.6% of the sequences in a single sample (Table S1). In summary, unclassified phylotypes made up a significant fraction of the microbial communities, but due to the lack of information on the taxonomic identity of this core microbiome, their functional role remains to be elucidated. Moreover, given the low similarity with sequences available in the RDP, they likely represent novel taxa at the genus level, indicating that Antarctica is still a source of untapped microbial diversity.

On the other hand, identified bacterial phylotypes were also among the dominant members of the communities (Table S1), providing valuable information on the ecology of the microbial assemblages. The genus *Methylotenera* was the dominant phylotype in sample S2, representing 10.65% of the sequences in this dataset. Methylotrophic bacteria belonging to the family *Methylophilaceae* are ubiquitous in a wide range of terrestrial and aquatic ecosystems, and are able to use single C compounds such as methylamine and methanol as C source for growth (Kalyuzhnaya *et al.*, 2012). OTUs assigned to the genus *Neptunomonas* dominated sample S3, accounting for 7.9% of the sequences in this dataset, and were also present in sample S4 with a slightly lower relative abundance (3.2%). This genus of facultative anaerobic *Gammaproteobacteria*, originally isolated from polluted coastal marine sediments, is able to use several polycyclic aromatic hydrocarbons and long chain alkanes as C and electron source for growth (Kersters *et al.*, 2006). Sequences were identical to the strain *N. japonica* JAMM 0745, isolated from deep sediment adjacent to whale carcasses and described as an endosymbiont of the bone-eating polichaete *Osedax japonicus* (Miyazaki *et al.*, 2008). To the best of our knowledge, this is the first study reporting a coastal soil habitat dominated by a species of the genus *Neptunomonas*.

As outlined above, communities were dominated by phylotypes related to microorganisms with distinct metabolisms such as chemolithoautotrophs, methylotrophs, methanogens and heterotrophs, indicating that oligotrophic environments may select for functionally diverse communities that can efficiently exploit the limited resources. Indeed, in a previous assessment of the functional diversity, we found that these communities were able to oxidize a wide range of organic compounds, and a clear relationship between microbial functional diversity and soil C_{org} content was observed (Pessi *et al.*, 2012). Samples with lower C_{org} content (S1 and S4) harboured simple, highly specialized communities, able to metabolize just a few compounds but to a very high extent. On the other hand, microbial communities in samples with higher C_{org} content (S2 and S3) were functionally more diverse, able to metabolize a higher number of compounds but at much lower rates. In the present study, we extend this description by relating the observed patterns of C source utilization to microbial community structure at the phylum and class levels (Fig. 4).

Multivariate ordination analysis carried out using the C source utilization data, along with multiple regressions of taxa abundances with ordination axes, gave insights into which microbial groups may be driving the intake of determined compounds (Fig. 4). The higher abundance of *Thermoplasmata* (phylum *Euryarchaeota*) in samples S1 and S4 was associated with a higher utilization of Tween 40, Tween 80, L-phenylalanine, pyruvic acid methyl ester and putrescine by these communities (Fig. 4b). Although molecular studies have suggested that *Archaea* play important ecological roles in most ecosystems (including a significant contribution to the C cycle through heterotrophic transformation in the cold biosphere), very few archaea have been isolated and cultivated in the laboratory (Cavicchioli 2006). The lack of data on the physiological attributes of cultivated archaea precludes the validation of our observations, and, therefore, they should be seen as a suggestion of the metabolic potential of *Archaea* in this environment. These results suggest that further ecophysiological studies of heterotrophic archaea are thus needed to assess their contribution to the Antarctic food chain.

Proteobacteria, the most common phylum worldwide, is a phenotypically versatile phylum, including chemoorganotrophs and

phototrophs as well as chemolithoautotrophs such as sulphur-oxidizing and nitrifying bacteria (Kersters *et al.*, 2006). The ability to obtain energy from the oxidation of inorganic compounds and C from the fixation of CO_2 may be advantageous in recently deglaciated environments, where the availability of organic matter is usually limited. However, our results suggest an important role of heterotrophic *Proteobacteria* in this environment, given the correlation of *Alpha-* and *Betaproteobacteria* with the utilization of a wide range of C compounds (Fig. 4b).

Bacteroidetes, on the other hand, are well known for their ability to degrade a wide range of polymers, and several Antarctic isolates have been shown to produce extracellular enzymes such as lipases, proteases and phosphatases (Aislabie *et al.*, 2006). Moreover, Bajerski and Wagner (2013) have found that, in two glacier forefields of Larsemann Hills, East Antarctica, *Bacteroidetes* occurred at highest abundance near the glacier tongue. Hence, this phylum seems to be well adapted to cold and poorly developed habitats and may be playing an important role in soil development, by degrading polymers and producing extracellular enzymes. Our results agree with these observations, showing the putative relationship between *Sphingobacteria* and the intake of a series of C compounds (Fig. 4b).

A large number of the dominant phylotypes were related to microorganisms isolated from seawater and marine sediments from the Southern Ocean and also from the North Sea and the Atlantic and Pacific Oceans (Table S1). One remarkable finding was that three phylotypes, representing from 2.1 to 4.4% of the sequences in each dataset, were identical to the type strain of the recently described *Granulosicoccus* genus, isolated from surface water off King George Island, Antarctic Peninsula (Lee *et al.*, 2007). This indicates that the original isolation of this bacterium was not circumstantial, as our study shows its environmental importance (based on its high relative abundance) in nearby costal soil. A particular phenotypic feature of this bacterium is the production of poly-β-hydroxybutyrate, both as a response to environmental stress and as energy storage. In this context, it may indicate an additional important adaptive feature for survival in this oligotrophic environment. However, as mentioned above, the Wanda Glacier forefield is under the direct influence of the adjacent marine environment due to the action of waves and tide fluctuations (Rosa *et al.*, 2009). Therefore, it is possible that at least a fraction of the diversity found in the analysed soil may represent a somewhat transient population of marine origin, which may or may not be a relevant and metabolically active fraction of the microbial community. Nevertheless, it is likely that the marine environment plays an important role on the colonization of this recently deglaciated habitat, serving as a source of pioneering microorganisms and being responsible for the high diversity found in the Wanda Glacier forefield.

Microbial activity in glacier forefield is limited by the low availability of C and N, and in Antarctic environments microorganisms also have to cope with other extreme conditions such as low temperatures, constant freeze-thaw cycles and high ultraviolet radiation. The soil analysed in the present study is particularly oligotrophic, with very low C and almost negligible N contents (Pessi *et al.*, 2012), although it is known that allochtonous inputs are an important but ephemeral source of organic matter for early successional heterotrophic microbial communities (Stibal *et al.*, 2008). Nevertheless, combined results from the present study and a previous work on the functional diversity of these communities (Pessi *et al.*, 2012) showed that microbial assemblages in recently exposed Antarctic forefield soil (<20 years) can be taxonomically and functionally highly diverse, reinforcing the idea that Antarctic soil bacterial communities are indeed more diverse than what was previously thought. In addition, the present study highlights how small is our understanding of Antarctic microbial diversity, given the surprisingly high abundance of several unclassified archaeal and bacterial phylotypes in the Wanda Glacier forefield. Therefore, this study contributes to our collective knowledge about microbial diversity and distribution on Earth, particularly in understudied extreme environments.

ACKNOWLEDGEMENTS

We would like to acknowledge Maria Mercedes Zambrano (GeBiX network, Colombia) and Patricia Del Portillo (CorpoGen, Colombia) for providing infrastructure and support for experimental and bioinformatics analyses, Jose Ricardo Bustos for custom bioinformatics scripting and support and David Velázquez (UAM, Spain), Julia Kleinteich (ULg, Belgium) and the members of Microbial Ecology Group at CorpoGen for their valuable discussions and suggestions. We also thank the anonymous reviewers for their rigorous and helpful contributions for the improvement of the manuscript.

FUNDING

This research was supported by grants Universal/2009 (CNPq), NANOBIOTEC-BRASIL/2008 (CAPES), CNPq-Colciencias/2008, Colciencias-SENA (project 6570-392-19990), Colciencias (contracts 427-2009 and 718-2009), European Commission MAGIC-PAH research consortium (FP7-KBBE-2009–245226) and FAPERGS. Logistic support was provided by the Brazilian Antarctic Program (PROANTAR) during the 28th Brazilian Antarctic Expedition in the austral summer season.

Conflict of interest statement. None declared.

SUPPLEMENTARY DATA

Supplementary data is available at FEMSEC online.

REFERENCES

Aislabie JM, Chhour K-L, Saul DJ, *et al.* Dominant bacteria in soils of Marble Point and Wright Valley, Victoria Land, Antarctica. *Soil Biol Biochem* 2006;**38**:3041–56.

Alonso-Sáez L, Andersson A, Heinrich F, *et al.* High archaeal diversity in Antarctic circumpolar deep waters. *Environ Microbiol Rep* 2011;**3**:689–97.

Bajerski F, Wagner D. Bacterial succession in Antarctic soils of two glacier forefields on Larsemann Hills, East Antarctica. *FEMS Microbiol Ecol* 2013;**85**:128–42.

Blois JL, Zarnetske PL, Fitzpatrick MC, *et al.* Climate change and the past, present, and future of biotic interactions. *Science* 2013;**341**:499–504.

Bohorquez LC, Delgado-Serrano L, López G, *et al.* In-depth characterization via complementing culture-independent approaches of the microbial community in an acidic hot spring of the Colombian Andes. *Microb Ecol* 2012;**63**:103–15.

Braun M, Saurer H, Vogt S, *et al.* The influence of large-scale atmospheric circulation on the surface energy balance of the King George Island ice cap. *Int J Climatol* 2001;**36**:21–36.

Brochier-Armanet C, Boussau B, Gribaldo S, *et al.* Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol* 2008;**6**:245–52.

Caporaso JG, Kuczynski J, Stombaugh J, *et al*. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;**7**:335–6.

Cavicchioli R. Cold-adapted archaea. *Nat Rev Microbiol* 2006;**4**:331–43.

Chong CW, Pearce DA, Convey P, *et al*. Patterns in the distribution of soil bacterial 16S rRNA gene sequences from different regions of Antarctica. *Geoderma* 2012;**181–2**:45–55.

Cook AJ, Fox AJ, Vaughan DG, *et al*. Retreating glacier fronts on the Antarctic Peninsula over the past half-century. *Science* 2005;**308**:541–4.

Davies BJ, Carrivick JL, Glasser NF, *et al*. Variable glacier response to atmospheric warming, northern Antarctic Peninsula, 1988–2009. *Cryosphere* 2012;**6**:1031–48.

DeLong EF, Wu KY, Prézelin BB, *et al*. High abundance of Archaea in Antarctic marine picoplankton. *Nature* 1994;**371**:695–7.

Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;**32**:1792–7.

Foong CP, Ling CMWV, González M. Metagenomic analyses of the dominant bacterial community in the Fildes Peninsula, King George Island (South Shetland Islands). *Polar Sci* 2010;**4**:263–73.

Ganzert L, Lipski A, Hubberten H-W, *et al*. The impact of different soil parameters on the community structure of dominant bacteria from nine different soils located on Livingston Island, South Shetland Archipelago, Antarctica. *FEMS Microbiol Ecol* 2011;**76**:476–91.

Ghiglione JF, Murray AE. Pronounced summer to winter differences and higher wintertime richness in coastal Antarctic marine bacterioplankton. *Environ Microbiol* 2012;**14**:617–29.

Haas BJ, Gevers D, Earl AM, *et al*. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 2011;**21**:494–504.

Huang Y, Niu B, Gao Y, *et al*. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 2010;**26**:680–2.

Janssen PH. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl Environ Microb* 2006;**72**:1719–28.

Jumpstart Consortium Human Microbiome Project Data Generation Working Group. Evaluation of 16S rDNA-based community profiling for human microbiome research. *PLOS One* 2012;**7**:e39315.

Kalyuzhnaya MG, Beck DAC, Vorobev A, *et al*. Novel methylotrophic isolates from lake sediment, description of *Methylotenera versatilis* sp. nov. and emended description of the genus *Methylotenera*. *Int J Syst Evol Micr* 2012;**62**:106–11.

Kersters K, De Vos P, Gillis M, *et al*., Introduction to the proteobacteria. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E (eds). *The Prokaryotes, Vol*. 5. New York, NY: Springer, 2006, 3–37.

Kim O-S, Chae N, Lim HS, *et al*. Bacterial diversity in ornithogenic soils compared to mineral soils on King George Island, Antarctica. *J Microbiol* 2012;**50**:1081–5.

Könneke M, Bernhard AE, De la Torre JR, *et al*. Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 2005;**437**:543–6.

Lee K, Lee HK, Choi T-H, *et al*. *Granulosicoccaceae* fam. nov., to include *Granulosicoccus antarcticus* gen. nov., sp. nov., a nonphototrophic, obligately aerobic chemoheterotroph in the order *Chromatiales*, isolated from Antarctic seawater. *J Microbiol Biotechn* 2007;**17**:1483–90.

López-García P, López-López A, Moreira D, *et al*. Diversity of free-living prokaryotes from a deep-sea site at the Antarctic Polar Front. *FEMS Microbiol Ecol* 2001;**36**:193–202.

Massana R, Taylor LT, Murray AE, *et al*. Vertical distribution and temporal variation of marine planktonic archaea in the Gerlache Strait, Antarctica, during early spring. *Limnol Oceanogr* 1998;**43**:607–17.

Mayewski PA, Maasch KA, Dixon D, *et al*. West Antarctica's sensitivity to natural and human-forced climate change over the holocene. *J Quaternary Sci* 2013;**28**:40–8.

Miyazaki M, Nogi Y, Fujiwara Y, *et al*. *Neptunomonas japonica* sp. nov., an *Osedax japonicus* symbiont-like bacterium isolated from sediment adjacent to sperm whale carcasses off Kagoshima, Japan. *Int J Syst Evol Micr* 2008;**58**:866–71.

McMurdie PJ, Holmes S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS One* 2013;**8**:e61217.

Nemergut DR, Anderson SP, Cleveland CC, *et al*. Microbial community succession in an unvegetated, recently deglaciated soil. *Microb Ecol* 2007;**53**:110–22.

Oksanen J, Blanchet FG, Kindt R, *et al*. *vegan: Community Ecology Package. R Package Version 2.0-10*. http://CRAN.R-project.org/package=vegan, 2013 (Last accessed in 13 September 2014).

Pessi IS, Elias SO, Simões FL, *et al*. Functional diversity of microbial communities in soils in the vicinity of Wanda Glacier, Antarctic Peninsula. *Microbes Environ* 2012;**27**:200–3.

Polz MF, Cavanaugh CM. Bias in template-to-product ratios in multitemplate PCR. *Appl Environ Microb* 1998;**64**:3724–30.

Reeder J, Knight R. Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nat Methods* 2010;**7**:668–9.

Roesch LFW, Fulthorpe RR, Pereira AB, *et al*. Soil bacterial community abundance and diversity in ice-free areas of Keller Peninsula, Antarctica. *Appl Soil Ecol* 2012;**61**:7–15.

Roesch LFW, Fulthorpe RR, Riva A, *et al*. Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J* 2007;**1**:283–90.

Rosa KK, Vieira R, Ferrando FJ, *et al*. Feições sedimentológicas e geomorfológicas do ambiente de deglaciação das geleiras Wanda e Ecology, ilha Rei George, Antártica. *Pesquisa em Geociências (Porto Alegre, Brazil)* 2009;**36**:315–26. (in portuguese).

Saul DJ, Aislabie JM, Brown CE, *et al*. Hydrocarbon contamination changes the bacterial diversity of soil from around Scott Base, Antarctica. *FEMS Microbiol Ecol* 2005;**53**:141–55.

Schloss PD, Handelsman J. Status of the microbial census. *Microbiol Mol Biol R* 2004;**68**:686–91.

Schloss PD, Westcott SL, Ryabin T, *et al*. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microb* 2009;**75**:7537–41.

Schütte UME, Abdo Z, Foster J, *et al*. Bacterial diversity in a glacier foreland of the high Arctic. *Mol Ecol* 2010;**19**:54–66.

Sogin ML, Morrison HG, Huber JA, *et al*. Microbial diversity in the deep sea and the underexplored "rare biosphere". *PNAS* 2006;**103**:12115–20.

Srinivas TNR, Singh SM, Pradhan S, *et al*. Comparison of bacterial diversity in proglacial soil from Kafni Glacier, Himalayan Mountain ranges, India, with the bacterial diversity of other glaciers in the world. *Extremophiles* 2011;**15**:673–90.

Stibal M, Tranter M, Benning LG, *et al*. Microbial primary production on an Arctic glacier is insignificant in comparison

with allochthonous organic carbon input. *Environ Microbiol* 2008;**10**:2172–8.

Tamura K, Peterson D, Peterson N, *et al*. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011;**28**:2731–79.

Teixeira LCRS, Peixoto RS, Cury JC, *et al*. Bacterial diversity in rhizosphere soil from Antarctic vascular plants of Admiralty Bay, maritime Antarctica. *ISME J* 2010;**4**:989–1001.

Turner J, Colwell SR, Marshall GJ, *et al*. Antarctic climate change during the last 50 years. *Int J Climatol* 2005;**25**:279–94.

Vaughan DG, Marshall GJ, Connolley WM, *et al*. Devil in the detail. *Science* 2001;**293**:1777–9.

Vincent W. *Microbial Ecosystems in Antarctica*. Cambridge, UK: Cambridge University Press, 2002.

Wang Q, Garrity GM, Tiedje JM, *et al*. Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microb* 2007;**73**:5261–7.

Wynn-Williams DD. Antarctic microbial diversity: the basis of polar ecosystem processes. *Biodivers Conserv* 1996;**5**:1271–93.

Zdanowski MK, Żmuda-Baranowska MJ, Borsuk P, *et al*. Culturable bacteria community development in postglacial soils of Ecology Glacier, King George Island, Antarctica. *Polar Biol* 2013;**36**:511–27.

Zeng Y-X, Yu Y, Qiao Z-Y, *et al*. Diversity of bacterioplankton in coastal seawaters of Fildes Peninsula, King George Island, Antarctica. *Arch Microbiol* 2014;**196**:137–47.

Zumsteg A, Luster J, Göransson H, *et al*. Bacterial, archaeal and fungal succession in the forefield of a receding glacier. *Microb Ecol* 2012;**63**:552–64.