

**Original citation:**

Pearce, David A., Newsham, Kevin K., Thorne, Michael A. S., Calvo-Bado, Leo A., Krsek, Martin, Laskaris, Paris, Hodson, Andy and Wellington, Elizabeth M. H.. (2012) Metagenomic analysis of a southern maritime antarctic soil. *Frontiers in Microbiology*, Volume 3 . Article number 403.

**Permanent WRAP url:**

<http://wrap.warwick.ac.uk/56284>

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0) license and may be reused according to the conditions of the license. For more details see: <http://creativecommons.org/licenses/by-nc-nd/3.0/>

Copyright: © 2012 Pearce, Newsham, Thorne, Calvo-Bado, Krsek, Laskaris, Hodson and Wellington. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

**A note on versions:**

The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: [publications@warwick.ac.uk](mailto:publications@warwick.ac.uk)



# Metagenomic analysis of a southern maritime Antarctic soil

David A. Pearce<sup>1\*</sup>, Kevin K. Newsham<sup>1</sup>, Michael A. S. Thorne<sup>1</sup>, Leo Calvo-Bado<sup>2</sup>, Martin Krsek<sup>2†</sup>, Paris Laskaris<sup>2†</sup>, Andy Hodson<sup>3</sup> and Elizabeth M. Wellington<sup>2</sup>

<sup>1</sup> Ecosystems Programme, Natural Environment Research Council, British Antarctic Survey, Cambridge, UK

<sup>2</sup> School of Life Sciences, University of Warwick, Coventry, UK

<sup>3</sup> Department of Geography, University of Sheffield, Sheffield, UK

## Edited by:

Diana R. Nemerut, University of Colorado Boulder, USA

## Reviewed by:

Rebecca Gast, Woods Hole Oceanographic Institution, USA  
Cristina Takacs-Vesbach, University of New Mexico, USA

## \*Correspondence:

David A. Pearce, Ecosystems Programme, British Antarctic Survey, Natural Environment Research Council, Madingley Road, Cambridge CB3 0ET, UK.  
e-mail: dpearce@bas.ac.uk

## † Present address:

Martin Krsek, Masaryk University, Brno, Czech Republic;  
Paris Laskaris, Department of Plant Pathology, University of Florida, Gainesville, FL 32611, USA.

Our current understanding of Antarctic soils is derived from direct culture on selective media, biodiversity studies based on clone library construction and analysis, quantitative PCR amplification of specific gene sequences and the application of generic microarrays for microbial community analysis. Here, we investigated the biodiversity and functional potential of a soil community at Mars Oasis on Alexander Island in the southern Maritime Antarctic, by applying 454 pyrosequencing technology to a metagenomic library constructed from soil genomic DNA. The results suggest that the commonly cited range of phylotypes used in clone library construction and analysis of 78–730 OTUs (de-replicated to 30–140) provides low coverage of the major groups present (~5%). The vast majority of functional genes (>77%) were for structure, carbohydrate metabolism, and DNA/RNA processing and modification. This study suggests that prokaryotic diversity in Antarctic terrestrial environments appears to be limited at the generic level, with *Proteobacteria*, *Actinobacteria* being common. Cyanobacteria were surprisingly under-represented at 3.4% of sequences, although ~1% of the genes identified were involved in CO<sub>2</sub> fixation. At the sequence level there appeared to be much greater heterogeneity, and this might be due to high divergence within the relatively restricted lineages which have successfully colonized Antarctic terrestrial environments.

**Keywords:** Antarctica, bacteria, metagenomics, polar, soil, 454 pyrosequencing

## INTRODUCTION

Antarctic soils have fascinated microbiologists throughout the last century. Expeditions to the continent as far back as 1901 collected soils specifically for bacteriological study. The results of these studies, and later twentieth century research, indicated that Antarctic soils typically harbor low numbers of bacterial taxa (Wynn-Williams, 1996). It is thought that this low taxonomic diversity is associated with increasingly severe environmental conditions, such as restricted water and nutrient availability, and frequent freeze-thaw cycling in summer. These environmental factors also increase in severity with progression from the island archipelagos of the Scotia arc and northern Antarctic Peninsula southwards to the Antarctic continent (Convey, 2001). However, the evidence for this restricted taxonomic diversity has been largely based on data from culture-based studies (e.g., Wynn-Williams, 1983) and, more recently, on data from clone libraries (Table 1). These studies have typically only provided levels of coverage of 0.50–0.80 (but see Aislabie et al., 2009), and all have recorded similar levels of biodiversity (at ~30–140 OTUs), all of which could be considered to be relatively low when compared with soils sampled from temperate or tropical environments (Fierer et al., 2003, 2007; Fierer and Jackson, 2006). Many of the dominant bacterial taxa identified by previous studies of Antarctic soil fall into similar taxonomic categories (Table 2). However, in recent years, with the advent of

metagenomic technology (Handelsman et al., 1998; Hugenholtz et al., 1998; Eisen, 2007; Delmont et al., 2011) it is now possible to make a more comprehensive assessment of the scope of the microbial biodiversity present in these soils, and even to determine some of the potential geochemical functions of these microbial communities.

Antarctic soils are of particular interest because chemical analyses has shown that they are relatively low in nutrient content (Lawley et al., 2004; Newsham et al., 2010). This can lead to strong gradients in physicochemical parameters at a wide range of spatial scales; of the order of meters (Chong et al., 2010), kilometers (Chong et al., 2011), or hundreds of kilometers (Yergeau et al., 2007). Antarctic soils also provide extremely good early indications of the potential effects of environmental change. The Antarctic Peninsula, for example, is warming three times faster than the global average (Turner et al., 2005).

Unsurprisingly, an increasing number of studies are beginning to show that the microbial biodiversity associated with these environmental gradients could be much larger than was once thought. Mars Oasis was chosen for this study as preliminary data already exist for this site. It has been suggested to be a potential biodiversity hotspot (Yergeau et al., 2007) and it has unique soil chemistry when compared to the surrounding area (Chong et al., 2011). It is also geographically isolated, being situated 1,000 km from South America on the south-eastern coast of Alexander Island in

**Table 1 | Summary data from a selection of previous studies of Antarctic soil microbial diversity.**

Reference	Latitude and longitude (place name)	T	S	R	% D	Shannon–Weaver DI	Coverage	Richness
Aislabie et al. (2006)	S 77° 25' E 163° 41' (Marble Point)	728	33	52–85	44–56	2.65–3.95	0.50–0.52	n/a
	S 77° 31' E 161° 52' (Bull Pass)	n/a	n/a	29–47	82–85	2.53–3.19	0.81–0.83	n/a
	S 77° 31' E 161° 40' (Lake Vanda)	n/a	n/a	47–61	67–69	3.27–3.32	0.70–0.70	n/a
Saul et al. (2005)	S 77° 50' E 166° 45' (Scott Base)	522	62	56	n/a	3.70–3.76	n/a	n/a
Aislabie et al. (2008)	S 77° 55' E 166° 45' (Scott Base)	155	n/a	45–51	11–18	n/a	0.52–0.57	46–182
	S 77° 25' E 163° 41' (Marble Point)	131	n/a	47–85	4–12	n/a	0.50	n/a
	S 77° 31' E 161° 52' (Bull Pass)	236	n/a	29–47	23–24	n/a	0.78–0.83	n/a
	S 77° 31' E 161° 40' (Lake Vanda)	211	n/a	47–49	16–17	n/a	0.63–0.70	n/a
	S 77° 19' E 170° 13' (Cape Hallett)	173	n/a	26	19–27	n/a	0.77–0.80	n/a
Aislabie et al. (2009)	S 77° 19' E 170° 13' (Cape Hallett)	580	52	27–57	29–76	n/a	0.45–0.78	63–256
	S 77° 13' E 166° 26' (Cape Bird)	168	11	4–19	78–99	n/a	0.85–0.99	5–36
Smith et al. (2006)	S 78° 05' E 165° 53' (PENP)	181	61	n/a	n/a	1.598	0.73	n/a
	S 78° 06' E 165° 49' (MVG)	n/a	n/a	n/a	n/a	1.331	0.64	n/a
	S 78° 01' E 165° 33' (BIS)	n/a	n/a	n/a	n/a	1.238	0.56	n/a
Taton et al. (2003)	S 77° 37' E 163° 07' (Lake Fryxell)	78	16	15	n/a	2.88	0.79	n/a
Niederberger et al. (2008)	S 72° 22' E 169° 53' (Luther Vale)	323	323	n/a	n/a	3.32–4.04	n/a	n/a
Yergeau et al. (2007)	54° 15' S, 36° 30' W (South Georgia)	178	2111	130	n/a	n/a	n/a	470
	60° 43' S, 45° 38' W (Signy Island)	174	n/a	128	n/a	n/a	n/a	420
	67° 34' S, 68° 08' W (Anchorage Island)	154	n/a	100	n/a	n/a	n/a	430
	71° 19' S, 68° 18' W (Fossil Bluff)	183	n/a	60	n/a	n/a	n/a	180
	71° 53' S, 68° 15' W (Mars Oasis)	168	n/a	138	n/a	n/a	n/a	460
	72° 03' S, 68° 31' W (Coal Nunatak)	187	n/a	40	n/a	n/a	n/a	100
	78° 26' S, 85° 60' W (Ellsworth Mountains)	170	n/a	98	n/a	n/a	n/a	270
Yergeau et al. (2009)	60° 43' S, 45° 38' W (Signy Island)	320	n/a	n/a	n/a	n/a	n/a	420
	67° 34' S, 68° 08' W (Anchorage Island)	367	n/a	n/a	n/a	n/a	n/a	430
	71° 19' S, 68° 18' W (Fossil Bluff)	107	n/a	n/a	n/a	n/a	n/a	180
	72° 03' S, 68° 31' W (Coal Nunatak)	160	n/a	n/a	n/a	n/a	n/a	100

T, total number of clones; S, number of sequences deposited; R, number of ribotype or phylotype patterns; % D, clones assigned to dominant ribotype (percentage); DI, diversity Index; n/a, not available.

**Table 2 | Many of the dominant bacterial taxa identified by previous studies on Antarctic soil fall into similar taxonomic categories.**

*Rubrobacter*, *Arthrobacter*, *Acidobacteria*, *Oscillatoria*, *Phormidium*, *Deinococcus*, *Sphingomonas*, *Bacteroides*, *Brevundimonas*, *Chloroflexus*, *Hymenobacter*, *Leptolyngbya*, *Nostoc*, *Pseudonocardia*, *Psychrobacter*, *Rhodococcus*, *Synechococcus*, *Actinobacteria*, *Anabaena*, *Cytophaga*, *Fervidobacterium*, *Friedmanniella*, *Microcoleus*, *Microcystis*, *Nitrosospora*, *Pseudomonas*, *Sphingobacterium*, *Sporosarcosina*, and *Xanthomonas*

the southern Maritime Antarctic, and is isolated by the Antarctic Circumpolar current and prevailing wind direction from the continental interior. Through studies of aerobiological transfer at Rothera (Hughes et al., 2004) and Halley (Pearce et al., 2010) we have some idea of the type of colonist arriving via aerial transfer, and there is relatively little wildlife or human impact at the site. One such study (Newsham et al., 2010) showed no difference between microbial biodiversity across two different parts of the same site at the 97% sequence homology level. However, the

effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene can all influence biodiversity estimates (Schloss, 2010). A re-analysis of data from this study, showed that the biodiversity could be different at each of the two study sites examined depending upon the specific criteria used for sequence differentiation. Hence the site may contain a higher diversity than that shown by clone library analysis alone.

Here, we report the biodiversity and functional potential of the soil community at Mars Oasis, based on the application of 454 pyrosequencing technology to a metagenomic library. It is important to recognize that all techniques in molecular biology impose some degree of bias or selection, and indeed numerous studies have investigated new methods to improve extraction, purification, amplification, and quantification of DNA from soils. In addition, comparative studies have been performed to analyze the efficiency of methods for extraction and purification of soil DNA recovered, and there are a number of excellent reviews in the literature which consider this topic in some detail (for example,

Wintzingerode et al., 1997; Frostegård et al., 1999; Courtois et al., 2001; Martin-Laurent et al., 2001; Feinstein et al., 2009; Delmont et al., 2011; Mahmoudi et al., 2011). For this reason, we do not attempt to provide a systematic analysis of the differentiation based on 16S rDNA. We rather highlight that a polyphasic approach can significantly increase the apparent diversity present and to focus on the relative magnitude and direction of the difference rather than absolute values. This is particularly important for Antarctic soils now, where the total biodiversity was believed to be limited. This view is changing. So the aim of this study was to gain a more comprehensive understanding of the taxonomic diversity of bacteria present in the soil and to determine an initial frequency distribution of potential functional genes. By combining the latter data with analyses of the chemistry of runoff and lake water, we also aimed to try to gain some preliminary insight into the main elements being utilized and cycled by the soil microbial community.

## MATERIALS AND METHODS

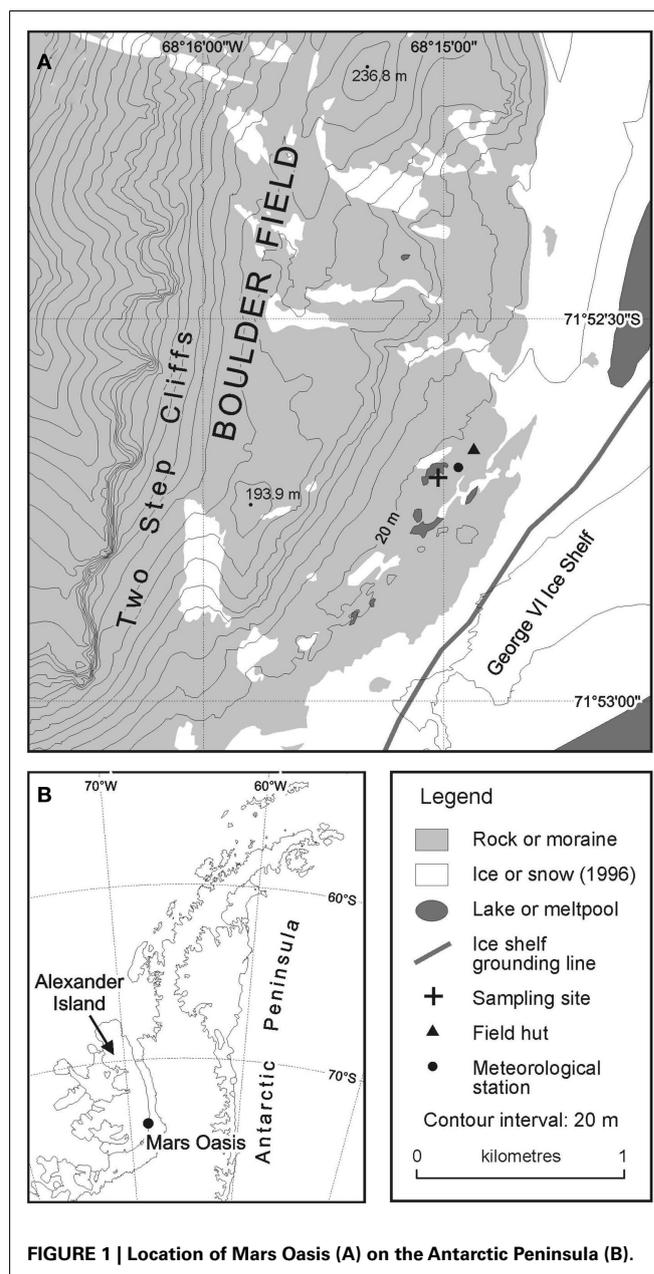
### SITE DESCRIPTION AND SAMPLING

Mars Oasis consists of an upper and lower terrace formed from shales. The soil used in this study was collected from the lower terrace, which is situated on a moraine ridge formed by contact between the George VI ice-shelf and Alexander Island. The lower terrace consists of areas of till, fluvial, and lacustrine sediments, with streams and ponds forming during the austral spring and summer. Restricted stands of bryophytes occur on the lower terrace close to meltwater streams and ponds. Lichens are sparse at the site, from which higher animals, including seals and nesting birds, are absent. Mean monthly air and soil (20 mm depth) temperatures at Mars Oasis vary between 1 and 6°C in January and –20 and –15°C in June, respectively. Snow depth at the site is typically ~2 m in winter, but snow ablates rapidly, usually in November, coinciding with a marked rise in soil water content close to the meltwater ponds at the lower site (H. J. Peat, personal communication).

The site was accessed from Rothera Research Station on Adelaide Island by fixed-wing aircraft fitted with skis in December 2004. Samples of moraine soil were collected from an area of level ground at the south-eastern margin of a permanent meltwater pool (71°52.6960' S; 68°14.9879' W; **Figure 1**). The soil was collected by inserting four sterile Vacutainers, with their lids removed, to a depth of 5 cm into the soil. Bryophytes were absent from this soil. The lids of the Vacutainers were replaced and the samples placed into re-sealable polythene bags, which were packed in ice in an insulated box. The soils were returned the following day to Rothera Research Station and were frozen at –20°C, prior to their return to the UK at the same temperature.

### METAGENOMIC LIBRARY CONSTRUCTION

The four soil samples (75 g each) were each suspended in 1% SDS solution (25 ml), to which 0.05 g of glucanex and glucanase had been added. The suspensions were vortexed for a few seconds and then incubated at 37°C for 4 h. They were cooled and filtered (1 mm) and 250 µg of RNAase was added to each solution. The solutions were centrifuged six times at 7,500 rpm for 15 min. After each centrifuge run, the supernatant (12 ml) was decanted. Three



molar sodium acetate at pH 7.0 (1.2 ml) and ethanol (26.4 ml) was added to each aliquot of the supernatant, which was centrifuged at 10,000 g for 10 min. The pellet of DNA was drained and dried for several minutes. TE buffer (0.1 ml) was added to each pellet, which was then incubated at 4°C for 16 h. All of the TE buffer solutions were combined (~240 µl), mixed with an equal volume of loading buffer, and ran out in a large-welled 1% low melting point agarose gel (Sigma-Aldrich) at 20 V overnight. The gel ran for 48 h, after which a block of agarose containing the target DNA, which had advanced 14–17 mm, was excised from the gel with a sterile scalpel. A size standard was used to select the region of the gel containing 35–45 kbp fragments. The agarose containing the target nucleic acids was then kept at 4°C for 60 h.

The agarose was digested in Gelase according to the manufacturers' instructions and enzymes were then denatured by heating to 60°C for 30 s. DNA was precipitated in three separate aliquots and was re-suspended in TE buffer (4.5 µl). The concentration of DNA ( $3 \times 20 \text{ ng } \mu\text{l}^{-1}$  aliquots) was determined by running against known standards in 1.5% agarose gels (1 h at 120 V). The DNA was then end-repaired by combining aliquots (12 µl) on ice with  $10 \times$  end-repair buffer (1.85 µl), 2.5 mM dNTP mix (1.85 µl), 10 mM ATP (1.85 µl), and end-repair enzyme mix (0.92 µl). The mixture was incubated at room temperature for 45 min and then at 70°C for 10 min. The DNA, consisting of a solution containing 240 ng of DNA, was then ligated into the pEpiFOS-5 fosmid vector (EpiCentre, Madison, WI, USA) by combining it with sterile water (2 µl),  $10 \times$  fast link ligation buffer (3 µl), 10 mM ATP solution (3 µl), fastlink ligase (3 µl), and vector (1 µl). The ligation reaction was then incubated at 4°C for 7 days.

The ligation mix was heated to 70°C for 10 min and the fosmid clones were packaged into lambda phages using Max-Plax lambda packaging extracts according to the manufacturer's instructions (Epicentre, Madison, WI, USA). This process yielded three aliquots (1.025 ml) of cloned cells. The packaged library was transduced into *E. coli* EPI-100, and *E. coli* transformants were selected on LB agar supplemented with  $12.5 \mu\text{g ml}^{-1}$  chloramphenicol. After determining the number of viable cells present, aliquots (200 µl) were spread onto dry Luria broth (100 ml) containing  $12.5 \mu\text{g ml}^{-1}$  chloramphenicol in 47 Petri dishes (150 mm diameter). Chloroformed phage (0.5 ml) was added to EPIFOS cells (5 ml, OD 1.0) and were incubated at 37°C for 20 min. Aliquots (100 µl) were added to each plate, spread, and grown at 37°C for 17 h. Colonies were picked into individual wells of 96 well plates containing Luria broth with  $12.5 \mu\text{g ml}^{-1}$  chloramphenicol (40 µl). The plates were incubated at 37°C for 17 h before sterile glycerol (10% v/v) was added to each well and the plates covered with plastic seals prior to storage at -80°C. All procedures described above took place under a sterile hood.

Quality control was established by end sequencing 20 random fosmids using pEpiFOS<sup>TM</sup>-5 forward and reverse end sequencing primers to ensure environmental DNA had been successfully incorporated, from microorganisms that one might expect to find in this extreme environment. One full fosmid was also sequenced. Primers were used to identify specific sequences from the fosmid library. The metagenomic library was screened using a range of primers for viral (Cyanophage CPS4GC, CPS5 Fuller et al., 1998; Wilson et al., 1999 and Phycodnaviridae AVS1, AVS2 Chen and Suttle, 1995), fungal (ITS1F/ITS4F; White et al., 1990; Gardes and Bruns, 1993), phosphonate (Gilbert et al., 2009), and N cycling (nosZ-F/nosZ-R, nirS1F/nirS6R, and nifHF/nifHRb; Thröback et al., 2004; Rösche and Bothe, 2005) genes. A selection of *E. coli* cells containing fosmids were screened for antibiotic production.

Cells from 25 plates (10% of the total) were combined (to favor depth of sequencing rather than coverage) and cultured in Luria broth with  $12.5 \mu\text{g ml}^{-1}$  chloramphenicol overnight in a shaking incubator at 37°C until an OD of 0.8 was obtained. The cells were centrifuge-concentrated and used to construct a

10,000 Gbp metagenomic library for 454 pyrosequencing. Fosmids were extracted from *E. coli* cells using the QIAGEN Plasmid Midi Kit (QIAGEN Plasmid Midi Kit, Cat. No. 12145, QIAGEN) and then treated with ATP-dependent Exo-nuclease (Plasmid-Safe<sup>TM</sup> ATP-Dependent Dnase,  $10 \text{ U } \mu\text{l}^{-1}$  10,000 U, Cat. No. E3110K, Epicentre). Extracts were sequenced by Macrogen (South Korea) according to the emPCR Method Manual – Lib-L MV (Anon, 2009a) and the Sequencing Method Manual (Anon, 2009b).

#### MARS OASIS CLONE LIBRARY RE-ANALYSIS

In clone library based studies, it is common to de-replicate samples through RFLP, or to assign sequences to groups with a pre-determined sequence similarity (commonly 97%), for the purposes of comparison across different samples, studies, or environmental gradients. The consequence of this approach is a potential underestimate of the total sequence diversity present in any given sample. To estimate the magnitude of this uncertainty, we selected 21 groupings of OTUs derived from a Mars Oasis clone library study (Newsham et al., 2010) and independently aligned them in CLUSTALW, to determine the levels of variation or similarity within each designated group. In the original study, PCR products were aligned in ClustalW and vector sequences removed. Initially these sequences were grouped according to gross similarity by aligning all sequences in Clustal and generating a single average distance tree based on percentage identities. Groups of sequences and any ungrouped sequences were then analyzed as separate data sets. In this re-analysis, this step was modified so that sequences within each group were only retained in that group if they had  $\geq 97\%$  sequence similarity to other members of the same grouping.

#### CHEMISTRY

Samples of runoff, lake water, sediment pore water from lake margins and snow were collected in December 2007. The samples were filtered immediately in the field ( $1 \mu\text{m}$ ) and then frozen. Sub-samples were stored at  $\sim 1^\circ\text{C}$  in the dark for  $\sim 10$  days until pH and  $\text{HCO}_3^-$  analyses (alkalinity titration using 1 mM HCl) could be made. Thawed samples were analyzed for major ions ( $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Cl}^-$ ,  $\text{NO}_3^-$ ,  $\text{SO}_4^{2-}$ ) in the UK using Dionex DX90 ion chromatography units, and for  $\text{NH}_4^+$  using a Skalar Autoanalyzer. Precision errors were  $< 5\%$  according to mid-range standards for all tests.

#### DATA ANALYSES

After stringent removal of technical replicates (Gomez-Alvarez et al., 2009) with cd-hit (at 99.5%; Li and Godzik, 2006), vector screening with Lucy (Chou and Holmes, 2001) and MG-RAST (Meyer et al., 2008), and eliminating shorter ( $< 100$  bp) reads, 262,086 reads (average length 441 bp) were then analyzed using MG-RAST (Meyer et al., 2008). Searches with a minimum cut-off of  $1e^{-05}$ , were made against the RDP (Cole et al., 2009), GreenGenes (DeSantis et al., 2006), and SEED (Overbeek et al., 2005) databases. An alternate independent OTU analysis was carried out by screening the original set of reads for 16S sequence using both RDP and GenBank (Benson et al., 2005), eliminating redundancy, and selecting those above 90% identity. Sequences were deposited in GenBank accession number SRA060370.

## RESULTS

### METAGENOMIC LIBRARY

#### Preliminary analysis

End sequencing gave matches to phototrophs and halotolerant organisms such as *Nocardiooides* sp., *Actinobacteria*, *Chlamydomonas reinhardtii*, *Halobacterium* sp., Halophilic archaeon, *Chromohalobacter salexigens*, *Phytophthora sojae*, and a plant pathogen, confirming that genomic DNA incorporated into the metagenomic library was predominantly derived from typical soil micro-organisms. Specific gene probing for phosphonate genes, fungal genes, and phage genes all generated successful amplifications.

#### Phylogeny from 454 pyrosequencing data

The total number of sequences containing some taxonomic information was 261,840. Blasting these sequences against the SEED database produced 322 Phyla hits (including bacteria, eukarya, viruses, and archaea). The full phylogeny of these sequences (>1,000 hits) is shown in **Table 3**. The numbers of genera in each class in the 454 phylogeny were (with percentage abundances in parentheses): Proteobacteria 320 genera (48.7%) as: alpha 100 genera (12.2%), beta 46 genera (4.2%), delta 27 genera (4.1%), epsilon 18 genera (0.1%), gamma 129 genera (7.5%), unclassified (0.1%), then Actinobacteria 78 genera (10.6%), Firmicutes 85 genera (7.9%), Bacteroidetes 21 genera (5.9%), Planctomycetes 3 genera (4.7%), Acidobacteria 2 genera (4.0%), Cyanobacteria 26 genera (3.4%), Verrucomicrobia 2 genera (3.1%), and Chloroflexi 9 genera (2.3%).

A rarefaction analysis of the final 454 data matches from the 261,840 sequences yielded 1,160 genera identified (**Figure 2**). An analysis of frequency distribution allowed an assessment of rare diversity, with the most common individual sequence match occurring 5,652 times (**Figure 3**). One hundred sixty-eight sequences occurred only once and 58 only twice. These data gave a Chao estimated sequence number of 1,400 (82.9% coverage) and a coverage estimate (Good, 1953) of 85.52%.

Of the 1,160 genera identified, the top 10 matches were to *Candidatus Solibacter usitatus* (5,652), *Burkholderia* (5,405), *Streptomyces* (5,348), *Xanthomonas* (3,685), *Pseudomonas* (3,554), *Sphingomonas* (3,432), *Planctomyces* (3,417), *Bradyrhizobium* (3,308), *Rhodopseudomonas* (3,295), and *Bacillus* (3,279). Other important groups were also present, for example, the *Methylobacteria* (2,562). Only 228 sequences had a single hit and 10,236 sequences in 73 groups were unclassified genera (3.88%).

Of the 3,316 species identified, the top 10 matches were to *Candidatus Solibacter usitatus* (5,652), *Rhodopseudomonas palustris* (3,293), *Rhodopirellula baltica* (2,764), *Gemmatimonas aurantiaca* (2,735), *Sphingomonas wittichii* (2,507), *Candidatus Koribacter versatilis* (2,343), *Salmonella enterica* (2,299), *Sphingopyxis alaskensis* (2,272), *Novosphingobium aromaticivorans* (2,239), and *Chthoniobacter flavus* (2,100). Other important groups were also identified, e.g., bacterium Ellin514 (1,341), which is commonly recorded in polar studies. Only 2,326 sequences in 328 groups were unclassified species (0.82%).

Screening the sequences using the SEED database and selecting for virus sequences gave 494 phage sequences in 28 genera (shown in parentheses). The top 10 phage type occurrences in order of

**Table 3 | Phylogeny of bacterial sequences derived from 454 pyrosequencing data, with number of sequences in each group (where frequency >1,000).**

<i>Candidatus Solibacter usitatus</i>	5,652
<i>Rhodopseudomonas palustris</i>	3,293
<i>Rhodopirellula baltica</i>	2,764
<i>Gemmatimonas aurantiaca</i>	2,735
<i>Sphingomonas wittichii</i>	2,507
<i>Candidatus Koribacter versatilis</i>	2,343
<i>Salmonella enterica</i>	2,299
<i>Sphingopyxis alaskensis</i>	2,272
<i>Novosphingobium aromaticivorans</i>	2,239
<i>Chthoniobacter flavus</i>	2,100
<i>Pirellula staleyi</i>	1,931
<i>Ruminococcus albus</i>	1,889
<i>Gemmata obscuriglobus</i>	1,811
<i>Planctomyces limnophilus</i>	1,692
<i>Myxococcus xanthus</i>	1,624
<i>Xylella fastidiosa</i>	1,601
<i>Bradyrhizobium japonicum</i>	1,571
<i>Verrucomicrobium spinosum</i>	1,519
<i>Opitutus terrae</i>	1,379
<i>Frankia</i> sp.	1,352
Bacterium Ellin514	1,341
<i>Bradyrhizobium</i> sp.	1,340
<i>Xanthomonas campestris</i>	1,332
<i>Erythrobacter</i> sp.	1,302
<i>Spirosoma linguale</i>	1,263
<i>Chitinophaga pinensis</i>	1,182
<i>Blastopirellula marina</i>	1,161
<i>Planctomyces maris</i>	1,128
<i>Escherichia coli</i>	1,127
<i>Synechococcus</i> sp.	1,120
<i>Erythrobacter litoralis</i>	1,105
<i>Sinorhizobium meliloti</i>	1,074
<i>Sorangium cellulosum</i>	1,055
<i>Sphingobium japonicum</i>	1,037
<i>Roseiflexus</i> sp. RS-1	1,019

frequency were *Mycobacterium* 107 (10 types), *Burkholderia* 104 (8 types), *Bordetella* 59 (3 types), *Pseudomonas* 51 (7 types), *Enterobacteria* 39 (10 types), *Flavobacterium* 22 (1 type), *Myxococcus* 14 (1 type), *Synechococcus* 11 (2 types), *Prochlorococcus* 9 (3 types), and *Sinorhizobium* 9 (1 type).

Screening sequences using the SEED database and selecting for Archaea sequences gave 32 Euryarchaeota, 16 Crenarchaeota, and 1 Korarchaeota. The top 10 Archaeal species that were recorded were *Methanosarcina acetivorans* 296 (3 types), *Methanospirillum hungatei* 77 (1 type), *Pyrococcus abyssi* 75 (3 types), *Sulfolobus acidocaldarius* 64 (3 types), *Haloarcula marismortui* 60 (1 type), *Methanococcus maripaludis* 59 (1 type), *Pyrobaculum aerophilum* 55 (4 types), *Methanoculleus marisnigri* 50 (1 type), *Archaeoglobus fulgidus* 48 (1 type), and *Methanosphaerula palustris* 48 (1 type).

Screening sequences using the SEED database and selecting for eukaryotic sequences generated few matches. These included the nematode *Caenorhabditis* sp., the liverwort *Marchantia* sp., the

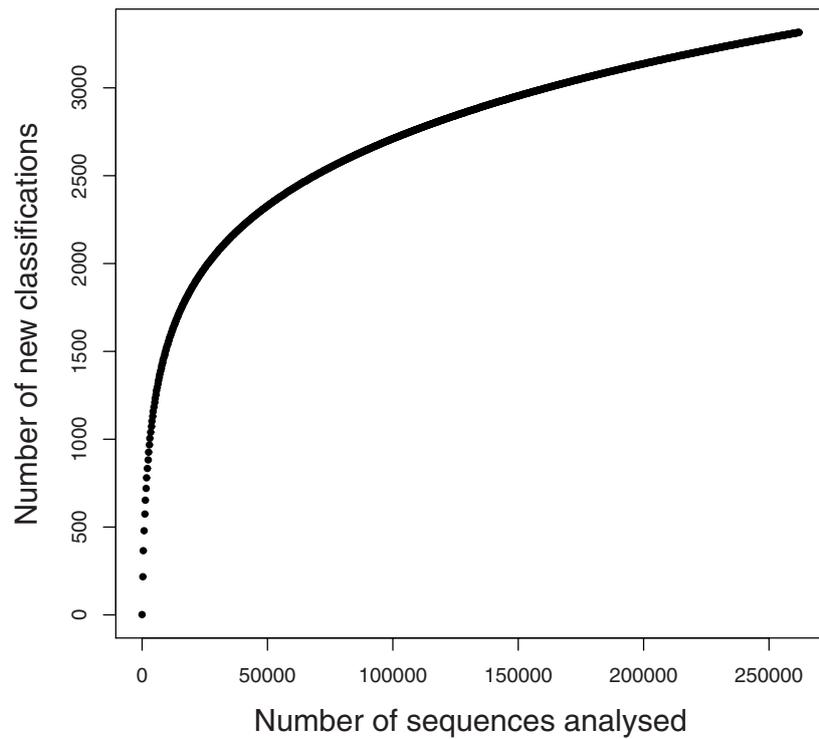


FIGURE 2 | Rarefaction graph from SEED data.

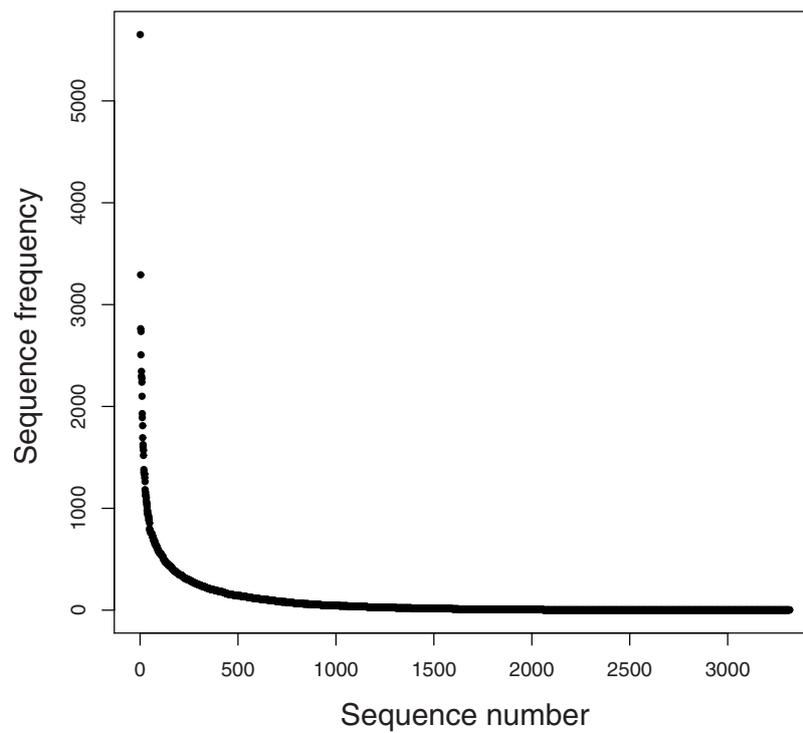


FIGURE 3 | Rare diversity graph.

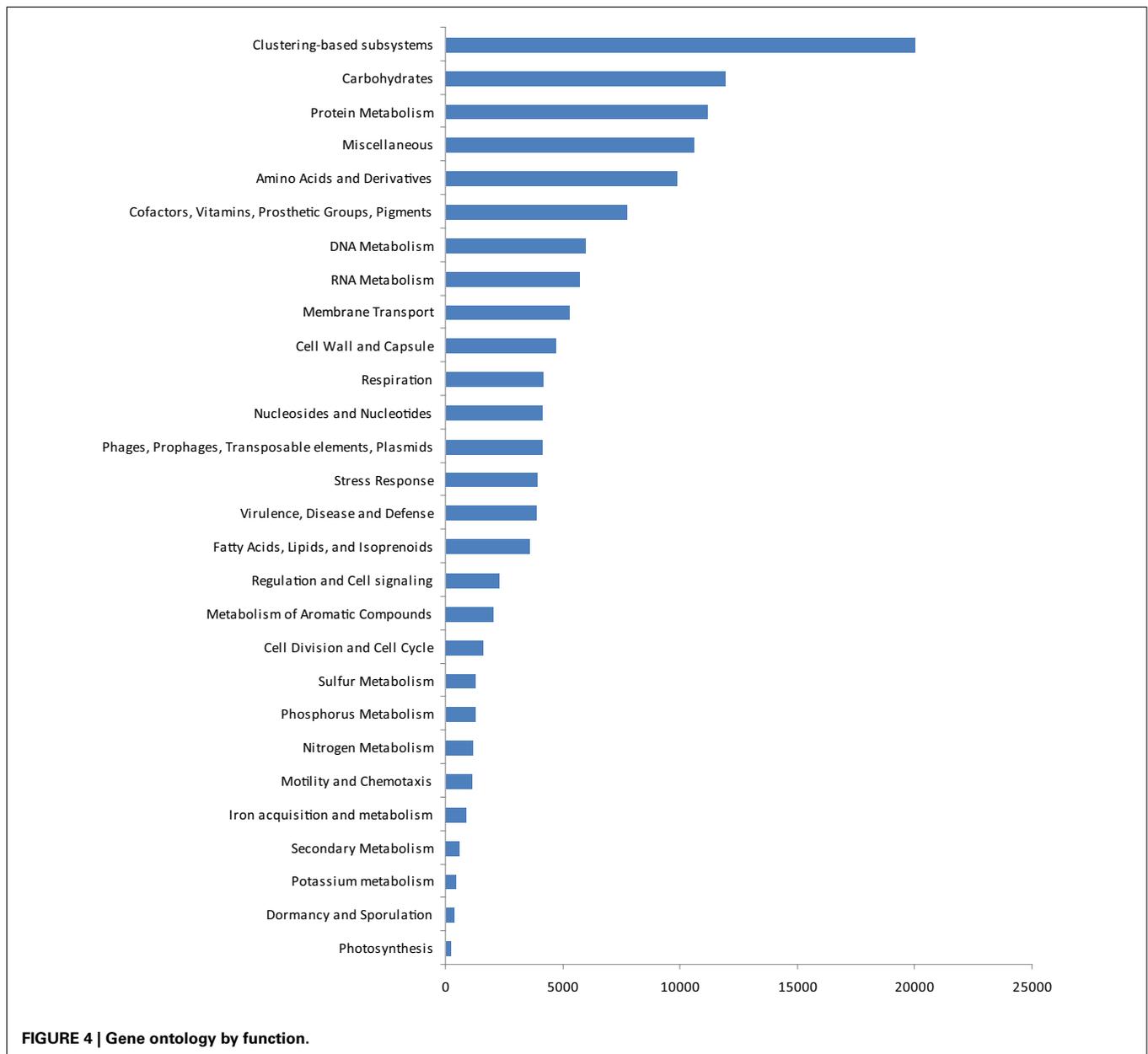
marine diatom *Odontella* sp., the fungi (*Gibberella* sp., *Neurospora* sp., *Magnaporthe* sp., *Schizosaccharomyces* sp., *Saccharomyces* sp., and *Eremothecium* sp.), protozoa (*Mesostigma* sp., *Naegleria* sp., and *Paramecium* sp.), and algae (*Cyanidium* sp., *Cyanidioschyzon* sp., and *Laminaria* sp.).

#### Gene ontology from 454 data

The gene ontologies derived from the metagenomic data are shown in **Figure 4**. Based on activity class, the most frequently encountered genes were those for clustering-based sub-systems (the precise functions of which are currently unknown), with the next most frequent genes being for carbohydrate metabolism, protein metabolism, amino acids and their derivatives, and cofactors, vitamins, prosthetic groups, and pigments. After these, the next most common genes were for DNA and RNA metabolism, membrane

transport, and the cell wall and capsule. Genes for respiration, nucleosides and nucleotides, three classes of potentially ecologically important genes phages, prophages, transposable elements, plasmids, stress response genes, virulence, disease, and defense followed these. After these, four further classes of housekeeping genes (fatty acids, lipids, and isoprenoids; regulation and cell signaling, metabolism of aromatic compounds and cell division and cell cycle) occurred. Finally, ecologically important classes of genes for sulfur, phosphorus, and nitrogen metabolism, motility and chemotaxis, iron acquisition and metabolism, secondary metabolism, potassium metabolism, dormancy and sporulation, and photosynthesis were recorded.

Removing all structural, carbohydrate metabolism, and DNA/RNA processing and modification genes (77% of matches) resulted in 31 classes of protein with >500 matches (10% of



matches). The most common (<6% of all matches, ~10,000 matches) was to a protein (DUF1446) of unknown function. The next most frequent category (>1%, ~3,000 matches) contained genes for prophage associated DNA combinational repair protein (2.4%), resistance to antibiotics and toxic compounds (2.4%), one-carbon metabolism (1.5%), and oxidative stress (1.5%). The next category, with 1,500–3,000 matches (0.5–1.0%) contained genes for CO<sub>2</sub> fixation (0.86%), flagellar motility (0.7%), phosphate metabolism (0.6%), and phospholipids (0.6%). After these, 1,000–1,500 matches (0.3–0.5%) were for osmotic stress (0.45%), heat shock (0.42%), quinone cofactors (0.38%), Ton and Tol transport systems (0.38%), and ammonia assimilation (0.36%). In the next category, several groups of potentially ecologically important proteins for Antarctic soils were then recorded, 1,000–600 (0.2–0.3%) containing genes for siderophores (0.3%), high affinity phosphate transporter and control of PHO regulon (0.3%), iron acquisition (0.29%), potassium homeostasis (0.29%), organic sulfur assimilation (0.28%), nitrate and nitrite ammonification (0.28%), nucleotidyl-phosphate metabolic clusters (0.28%), inorganic sulfur assimilation (0.27%), bacterial chemotaxis (0.25%), regulation of virulence (0.24%), P uptake in cyanobacteria (0.21%), and pathogenicity islands (0.21%). Finally, <600 hits (0.15–0.2%) occurred to quorum sensing and biofilm formation (0.19%), periplasmic stress (0.19%), and detoxification genes (0.17%). A further 43,168 categories had <500 matches, with 17 classes having a single sequence associated with them.

#### Clone library re-analysis

Of 43 sequences that had earlier been grouped into 21 sequence types based on a ≥97% cut-off level, pair-wise comparison of sequences within the originally assigned groups showed that only four of these independently the criteria within the group itself (Table 4), suggesting that 14 of the original groupings underestimated total diversity.

#### Chemistry

Chemical data are shown in Table 5. Runoff and soil pore water from the margin of the lake at Mars Oasis had pHs of 7.4–7.6. Sulfate and calcium were the dominant ions present in the water (2,700–4,100 μ equivalents L<sup>-1</sup>). Carbonate and magnesium ions were less frequent (450–1,400 μ equivalents L<sup>-1</sup>), followed by sodium, chloride, and potassium (13–333 μ equivalents L<sup>-1</sup>). Finally, ammonium-N and nitrate-N were the least frequent ions in runoff, and ammonium was infrequent in pore water (each 3–6 μ equivalents L<sup>-1</sup>), but nitrate was relatively frequent in the latter (200 μ equivalents L<sup>-1</sup>).

#### DISCUSSION

Studies of bacterial communities from around the world suggest a wide spectrum of taxonomic diversity, from the Amazonian soils, where every sequence sampled could be different (Fierer et al., 2007) to the highly selective Rio Tinto river in Europe, with a relatively restricted biodiversity (Palacios et al., 2008). It is apparent from the current study that the bacterial community diversity in the soil at Mars Oasis lies somewhere between the extremes, with a total of 1,160 genera from 3,318 phylotypes detected in the 454 library. This is an order of magnitude greater than data from clone library studies alone, which have to date recorded 78–730 (de-replicated to 30–140) phylotypes present in Antarctic soils sampled from the sub-Antarctic Islands, the Antarctic Peninsula, and the continent itself (see references in Table 1). Of the 1,160 genera recorded in our study, 71 (6.12%) have also been identified by other studies of Antarctic soil biodiversity.

#### DIVERSITY AT THE GENUS LEVEL

The most frequent genera in Antarctic soils (identified in >3 studies) are *Rubrobacter*, *Arthrobacter*, *Oscillatoria*, *Sphingomonas*, *Chloroflexus*, *Anabaena*, *Actinobacteria*, *Microcoleus*, *Microcystis*, *Nitrosospira*, *Pseudomonas*, *Fervidobacterium*, *Xanthomonas*,

Table 4 | Mars Oasis clone library re-analysis.

Nominal identification given	Original (≥97% BLAST i.d.)	Sequence number	Sequence pair-wise comparisons (mismatch/sequence length similarity%)		
Uncultured bact clone MeCl 62	12/657 (98.2%)	3	27/841 (96.8%)	38/721 (94.7%)	33/721 (95.4%)
Uncultured eukaryote clone	36/730 (99%)	3	12/769 (98.4%)	74/802 (90.8%)	51/760 (93.3%)
Uncultured bact clone FRCH17502	3/745 (99.6%)	2	20/767 (97.4%)		
Uncultured bact clone LVH3-G7	4/715 (99%)	2	4/715 (99.0%)		
Uncultured bacteroidetes AS28	209/728 (71.3%)	2	92/725 (87.3%)		
Uncultured bacteroidetes clone	34/780 (95.6%)	3	317/826 (97.9%)	25/781 (96.8%)	29/782 (96.3%)
Uncultured cyanobacterium clone	20/699 (97.14%)	2	20/708 (97.2%)		
Uncultured Micrococccineae	64/715 (91.1%)	3	23/715 (96.8%)	48/708 (93.2%)	53/757 (93.0%)
Soil bacterial clone U8	24/795 (97.0%)	2	26/801 (96.8%)		
Actinomycetes clone FB-2 A11	28/724 (96.1%)	2	24/725 (96.7%)		
Uncultured bacterium clone CM131	32/414 (92.3%)	2	33/414 (92.0%)		
Bacterial clone KuyT-IWBP 17	26/744 (96.5%)	2	29/744 (96.1%)		
<i>Solirubrobacter</i> Gsoil 921	33/709 (95.3%)	3	7/710 (99.0%)	33/701 (95.3%)	30/767 (96.1%)
Uncultured bact 071021-ONK-KR1-12	69/812 (91.5%)	2	62/799 (92.2%)		
Uncultured bacterium FACH1766	16/805 (98.0%)	2	20/802 (97.5%)		
Uncultured bacterial clone F1-2F-F12	163/775 (79.0%)	2	80/774 (89.7%)		
Uncultured Caldilineaceae bacterium	59/734 (92.0%)	3	30/744 (96.0%)	48/761 (93.7%)	48/744 (93.5%)
<i>Anabaena</i> sp.	37/756 (92.0%)	3	18/772 (97.7%)	29/770 (96.2%)	39/780 (95.0%)

**Table 5 | pH and ion concentrations of lake runoff, lake water, pore water, and snow at Mars Oasis.**

Sample	pH	NH <sub>4</sub> <sup>+</sup>	Na <sup>+</sup>	K <sup>+</sup>	Ca <sup>2+</sup>	Mg <sup>2+</sup>	HCO <sub>3</sub> <sup>-</sup>	Cl <sup>-</sup>	SO <sub>4</sub> <sup>2-</sup>	NO <sub>3</sub> <sup>-</sup>	O <sub>2</sub>	DOC
Runoff	7.6	5.61	118	12.8	2,981	784	1,392	32.6	4,082	3.33	13.0	0.35
Lake water (sample 1)	7.3	6.25	101	14.2	1,238	259	1,771	141	956	2.38	12.2	6.53
Lake water (sample 2)	7.4	4.09	129	20.0	1,540	233	1,573	108	1,310	1.80	12.9	2.43
Lake water (sample 3)	7.5	6.14	85.7	16.6	849	115	768	103	781	0.86	12.7	1.15
Lake water (sample 4)	7.7	3.11	36.8	8.37	510	64.0	1,088	26.1	165	5.92	12.6	m.d.
Pore water	7.4	2.91	124	16.2	2,693	449	652	333	2,464	199	7.31	5.59
Snow	5.7	2.03	16.5	12.0	46.1	3.91	m.d.	19.2	46.2	2.38	m.d.	3.64

DOC, dissolved organic carbon; m.d., missing data.

Data are expressed in  $\mu$  equivalents L<sup>-1</sup> except for dissolved oxygen and DOC (both mg L<sup>-1</sup>) and pH. Data are means of 7, 12, 3, and 4 measurements for runoff, lake water, pore water, and snow, respectively. PO<sub>4</sub><sup>3-</sup> was below detection limits of 1  $\mu$  equivalents L<sup>-1</sup>.

and *Acidobacteria*, *Phormidium*, *Deinococcus*, *Bacteroides*, *Brevundimonas*, *Hymenobacter*, *Leptolyngbya*, *Nostoc*, *Pseudonocardia*, *Psychrobacter*, *Rhodococcus*, *Synechococcus*, *Cytophaga*, *Friedmanniella*, and *Sphingobacterium*. The former group was identified in both clone library based and 454 based studies, indicating that there is some broad agreement with previous studies, whilst the latter group was specifically identified in this study.

#### DIVERSITY AT THE SEQUENCE LEVEL

The 10 Antarctic soil biodiversity studies examined (based upon PCR amplification, cloning, and sequencing of 16S rRNA genes) each reported relatively low prokaryotic biodiversity in Antarctic terrestrial ecosystems. However, estimated coverage in each of these studies ranged from 30–70%, only two of the collectors' curves approached anywhere near saturation and all were based upon the assumption that taxonomic differentiation occurs with <97% sequence similarity over 200–800 bp. Despite 97% sequence similarity being adopted for species differentiation in many biodiversity studies, it has already been demonstrated that 100% sequence similarity in the 16S rRNA gene can be found between isolates with different ecological phenotypes (Peña et al., 2010), and conversely, <97% sequence similarity has been found between isolates which have an identical ecological function. We therefore conducted a detailed re-analysis of 300 sequences (Newsham et al., 2010), and show that at a sequence similarity of 97%, only 7 of the 300 sequences were duplicated in the clone library. This re-analysis showed that the diversity of closely related sequences may have been underestimated by up to 40%. If we add to this an approximate doubling of the total number of genera identified through the addition of 454 data to the combined clone library summary, then there could be a minimum of four times greater diversity than was previously described. So although prokaryotic diversity may indeed be restricted at the generic level, there appears to be high sequence diversity present in the soil at Mars Oasis.

#### BIODIVERSITY

The range of taxa identified extended (indeed increased by five times), rather than contradicted the taxa identified in earlier clone library studies (140 Genera were present in both; 56 from clone library studies only and 1,026 from this 454 study only). A number of key taxa were identified, that have been indicated to be important in other studies of Antarctic soils – these include the

Actinobacteria and Cyanobacteria. The taxa identified were also similar to those recovered in airborne samples of the Peninsula region. Hughes et al. (2004) sampled air for a 2-week period above Rothera Station on Adelaide Island, and found a range of microorganisms, including cyanobacteria, actinomycetes, diatom plastids, and other uncultivated bacterial groups. Elsewhere, on the Brunt Ice-Shelf over an isolated scientific research station, on an ice-shelf in continental Antarctica Pearce et al. (2010) found Bacilli, Pseudomonads, and Sphingomonads. In common with other environmental studies, particularly in the polar regions, many of the sequences obtained were from as yet uncultivated organisms. The detected aerial microorganisms were different from those obtained over the Antarctic Peninsula. In both aerobiological studies, a low microbial biodiversity was detected, which included many sequence replicates. In this study, there were also important genera identified that had not been described in a range of selected clone library studies (Table 6).

#### GENE ONTOLOGY

The gene ontology data in the present study, when expressed by activity class, gave some potential insights into the presence of functional genes in the soil at Mars Oasis. Although many genes for clustering-based sub-systems were encountered in the 454 library, suggesting the functional coupling of genes whose present purpose is unknown, the data clearly indicated that the microbial community was active, with the presence of many genes for cell division and the cell cycle, cell wall formation, nucleotides and nucleosides, and RNA metabolism. This indicates that there is potential for the soil community at the oasis to express these genes, at least during the austral summer, when temperatures are typically above freezing point during the daytime and liquid water is freely available. There is also the potential for active competition between microbes in the soil, with the presence of antibiotic and toxic compound resistance genes, quorum sensing, and biofilm formation genes and many genes relating to virulence. Given the abundance of *Actinobacteria* in the soil, and particularly genera such as *Streptomyces*, which are active synthesizers of antibiotics, it is unsurprising that many virulence genes were encountered in the soil. Genes found at lower frequencies than expected were those for stress responses (including oxidative stress, osmotic stress, periplasmic stress, cold shock, and detoxification genes), perhaps reflecting the not unfavorable environmental conditions for soil microbial growth at

**Table 6 | Diversity of bacteria (and other groups) identified to the Genus level from Antarctic soil from by 454 sequencing.**

*Abiotrophia, Acanthamoeba, Acaryochloris, Acetitomaculum, Acetivibrio, Acetobacter, Acetohalobium, Acholeplasma, Achromobacter, Acidaminococcus, Acidilobus, Acidiphilium, Acidithiobacillus, Acidithiomicrobium, Acidobacterium, Aciduliprofundum, Acorus, Actinocorallia, Actinomadura, Actinoplanes, Actinosynnema, Acuclavella, Acyrthosiphon, Adineta, Aedes, Aerococcus, Aeromonas, Aeropyrum, Afipia, Aggregatibacter, Agromyces, Ahrensia, Ailuropoda, Ajellomyces, Akkermansia, Albidiferax, Alcanivorax, Algoriphagus, Alicyclophilus, Alicyclobacillus, Aliivibrio, Alistipes, Alkalilimnicola, Alkaliphilus, Allium, Allochromatium, Alphabaculovirus, Alphapapillomavirus, Alternaria, Alteromonas, Amaranthus, Amatitlania, Aminobacterium, Aminomonas, Ammonifex, Amplypterus, Amycolatopsis, Anaerobaculum, Anaerococcus, Anaerofustis, Anaerolinea, Anaeromyxobacter, Anaerostipes, Anaerotruncus, Anaplasma, Ancylobacter, Aneurinibacillus, Anopheles, Anoplopoma, Anoxybacillus, Anthopleura, Antonospora, Aphanizomenon, Aphanomyces, Apis, Apteryx, Aquimonas, Arabidopsis, Arcanobacterium, Archaeoglobus, Arcobacter, Aromatoleum, Arsenophonus, Artemia, Arthroderma, Arthrospira, Ascidia, Asfivirus, Aspergillus, Asticcacaulis, Ateles, Atopobium, Aulacoseira, Aurantimonas, Aureobasidium, Aureococcus, Azorhizobium, Azospirillum, Babesia, Bacteroides, Basfia, Bathymodiolus, Batrachovirus, Beggatoa, Beijerinckia, Bermanella, Beryx, Beutenbergia, Bilophila, Blastocystis, Blastomonas, Blastopirellula, Blattabacterium, Blautia, Bombyx, Bordetella, Borrelia, Bos, Botryotinia, Bpp-1-like viruses, Brachybacterium, Brachymonas, Brachypodium, Brachyspira, Branchiostoma, Brassica, Brevibacillus, Brevundimonas, Bromus, Brucella, Brugia, Bryonia, Bulleidia, Butyrivibrio, Caenorhabditis, Caldanaerobacter, Calditerrivibrio, Caldivirga, Caligus, Callithrix, Camelus, Caminibacter, Camponotus, Campylobacter, Canavalia, Candida, Candidatus, Accumulibacter, Candidatus, Amoebophilus, Candidatus, Azobacteroides, Candidatus, Blochmannia, Candidatus, Chloracidobacterium, Candidatus, Cloacamonas, Candidatus, Desulfurudis, Candidatus, Hamiltonella, Candidatus, Korarchaeum, Candidatus, Koribacter, Candidatus, Kuenenia, Candidatus, Liberibacter, Candidatus, Magnetobacterium, Candidatus, Micrarchaeum, Candidatus, Nitrososphaera, Candidatus, Odysseella, Candidatus, Parvarchaeum, Candidatus, Pelagibacter, Candidatus, Phytoplasma, Candidatus, Protochlamydia, Candidatus, Puniceispirillum, Candidatus, Regiella, Candidatus, Rhodoluna, Candidatus, Solibacter, Candidatus, Sulcia, Canis, Capnocytophaga, Capsaspora, Capsicum, Carassius, Cardiobacterium, Carnobacterium, Catabena, Catenibacterium, Catenulispora, Catharanthus, Catonella, Cavia, Cellulomonas, Cellulosilyticum, Cellulosimicrobium, Cellvibrio, Cenarchaeum, Chaetoceros, Chaetomium, Chara, Chattonella, Chelativorans, Chitinophaga, Chlamydia, Chlamydomonas, Chlamydomonada, Chlorella, Chlorobaculum, Chlorobium, Chloroherpeton, Chlorokybus, Chlorovirus, Chondrus, Chrysopathes, Chthoniobacter, Cicer, Ciona, Citreicella, Citrobacter, Citromicrobium, Clavibacter, Clavisporea, Coccidioides, Cochliobolus, Collimonas, Collinsella, Colossendeis, Comamonas, Compsopogon, Congregibacter, Coprinopsis, Coprobacillus, Coprococcus, Coprothermobacter, Coptotermes, Coraliomargarita, Corallina, Corynebacterium, Coxiella, Crassostrea, Cricetulus, Croceibacter, Crocosphaera, Cronobacter, Crustomastix, Cryptocodium, Cryptobacterium, Cryptosporidium, Cucumis, Cucurbita, Culex, Culicoides, Cupriavidus, Curvibacter, Curvularia, Cyanidioschyzon, Cyanidium, Cyanobium, Cyanophora, Cylindrospermopsis, Cystobacter, Cytophaga, Dactylosporangium, Danio, Daphnia, Dasypus, Daucus, Debaryomyces, Deferribacter, Dehalogenimonas, Deinococcus, Denitrovibrio, Dermacentor, Dermacoccus, Desulfarculus, Desulfatibacillum, Desulfobacterium, Desulfobulbus, Desulfocella, Desulfococcus, Desulfohalobium, Desulfomicrobium, Desulfomonile, Desulfonatronospora, Desulfotomaculum, Desulfurispirillum, Desulfurivibrio, Desulfurococcus, Desulfuromonas, Dethiobacter, Dethiosulfovibrio, Dialister, Dicathais, Dichelobacter, Dickeya, Dictyoglomus, Dictyostelium, Dietzia, Discophora, Dokdonia, Dolichospermum, Dorea, Drosophila, Durinskia, Dyadobacter, Echinops, Ectocarpus, Edwardsiella, Eggerthella, Ehrlichia, Eikenella, Elaeis, Eleocharis, Eleotris, Elizabethkingia, Elusimicrobium, Elymus, Emericella, Encephalitozoon, Enchytraeus, Endoriftia, Enhydrobacter, Enhygromyxa, Ensifer, Entamoeba, Epiphyas, Epulopiscium, Equus, Eremococcus, Eremothecium, Erinaceus, Erwinia, Erysipelothrix, Ethanologenens, Eubacterium, Euglena, Faecalibacterium, Felis, Ferrimonas, Ferroglobus, Ferroplasma, Filifactor, Filobasidiella, Finegoldia, Flammeovirga, Flexithrix, Fragilariopsis, Francisella, Fructobacillus, Fulvimarina, Fusarium, Fusobacterium, Gallionella, Gallus, Gardnerella, Gasterosteus, Gemella, Gemmatimonas, Gentiana, Geobacillus, Giardia, Gibberella, Glacielecola, Glomerella, Glossina, Gluconacetobacter, Glycine, Gordonia, Gordonibacter, Gorilla, Gracilaria, Gramella, Granulicatella, Grosmannia, Guillardia, Haemophilus, Hafnia, Hahella, Haladaptatus, Halalkalicoccus, Halanaerobium, Haliangium, Haliotis, Haloarcula, Halobacterium, Haloferax, Halogeometricum, Halomicrobium, Halomonas, Haloquadratum, Halorhabdus, Halorubrum, Haloterrigena, Halothermothrix, Halothiobacillus, Harpegnathos, Hartmannella, Haslea, Helianthus, Heliobacillus, Heliobacterium, Heliothis, Herbaspirillum, Herminiimonas, Hirschia, Histophilus, Hoefflea, Holdemania, Homo, Hordeum, Hydra, Hydrogenivirga, Hydrogenobacter, Hydrogenobaculum, Hydrogenophaga, Hyles, Hyperthermus, Hyphomicrobium, Hyphomonas, Ictalurus, Ignicoccus, Ignisphaera, Ilyobacter, Ipomoea, Isosphaera, Ixodes, Jackiella, Jannaschia, Jonesia, Jonquetella, Kalidium, Kangiella, Karenia, Ketogulonicigenium, Kineococcus, Kingella, Kitasatospora, Kluveromyces, Kocuria, Kordia, Kosmotoga, Ktedonobacter, Kutzneria, Kytococcus, L5-like viruses, Labrenzia, Laccaria, Lachancea, Lactobacillus, Lactococcus, Lambda-like viruses, Laminaria, Lamprocystis, Laribacter, Larimichthys, Lautropia, Lawsonia, Leadbetterella, Leclercia, Leeuwenhoekella, Legionella, Leishmania, Lentisphaera, Lentivirus, Lepeophtheirus, Lepidium, Leptolyngbya, Leptosphaeria, Leptospira, Leptospirillum, Leptothrix, Leptotrichia, Limnobacter, Limnoria, Listeria, Listonella, Loa, Lodderomyces, Lolium, Loxodonta, Lunularia, Lutiella, LUZ24-like viruses, Lymnaea, Lysinibacillus, Lysobacter, Lytechinus, Macaca, Macrocooccus, Magnaporthe, Magnetococcus, Magnetospirillum, Malassezia, Malus, Mannheimia, Marchantia, Maribacter, Maricaulis, Marinitoga, Marinobacter, Mariprofundus, Maritimibacter, Marivirga, Medicago, Megamonas, Megasphaera, Mesembryanthemum, Mesorhizobium, Mesostigma, Metajapyx, Metallosphaera, Metarhizium, Methanobacterium, Methanobrevibacter, Methanocaldococcus, Methanocella, Methanococcoides, Methanococcus, Methanocorpusculum, Methanoculleus, Methanohalobium, Methanohalophilus, Methanoplanus, Methanopyrus, Methanoregula, Methanosaeta, Methanosarcina, Methanosphaera, Methanosphaerula, Methanospirillum, Methanothermobacter, Methanothermococcus, Methanothermus, Methylacidiphilum, Methylbium, Methylobacter, Methylobacterium, Methylocapsa, Methylocella, Methylocystis, Methylomonas, Methylophaga, Methylophilus, Methylosinus, Methylothermus, Methyloversatilis, Methylovorus, Metridium, Meyerozyma, Microbacterium, Microchaete, Micromonas, Micromonospora, Microtus, Mimivirus, Mitsuoella, Mnemiopsis, Mobiluncus, Monilophthora, Monocercomonoides, Monodelphis, Monosiga, Moorella, Moraxella, Mucilaginibacter,*

(Continued)

Table 6 | Continued

*Mus*, *Mussa*, *Mycobacterium*, *Myotis*, *Myrothecium*, *Myxococcus*, *Myzus*, N15-like viruses, N4-like viruses, *Naegleria*, *Nakamurella*, *Nakaseomyces*, *Nannocystis*, *Nanoarchaeum*, *Nasonia*, *Natranaerobius*, *Natrialba*, *Natronomonas*, *Nautilia*, *Nectria*, *Neisseria*, *Nematostella*, *Neosartorya*, *Nephroselmis*, *Neptuniibacter*, *Neurospora*, *Nicotiana*, *Nitratifactor*, *Nitratiruptor*, *Nitrobacter*, *Nitrococcus*, *Nitrosomonas*, *Nitrosopumilus*, *Nitrospira*, *Nocardioopsis*, *Nonomuraea*, *Nostoc*, *Notechis*, *Novosphingobium*, *Oceanibulbus*, *Oceanicaulis*, *Oceanithermus*, *Oceanobacillus*, *Ochrobactrum*, *Octadecabacter*, *Odontella*, *Oenococcus*, *Oenothera*, *Oikopleura*, *Okibacterium*, *Oligotropha*, *Olsenella*, *Oncorhynchus*, *Opsanus*, *Oribacterium*, *Orientia*, *Ornithobacterium*, *Ornithorhynchus*, *Oryctolagus*, *Oryza*, *Oryzias*, *Oscillochloris*, *Osmerus*, *Ostreococcus*, *Otolemur*, *Ovis*, P1-like viruses, P22-like viruses, P2-like viruses, *Paenibacillus*, *Pagrus*, *Paludibacter*, *Pan*, *Pantoea*, *Parabacteroides*, *Paracoccidioides*, *Paramecium*, *Parascardovia*, *Parvibaculum*, *Parvularcula*, *Pasteurella*, *Paucimonas*, *Paulinella*, *Pectobacterium*, *Pediculus*, *Pediococcus*, *Pelagibacter*, *Pelobacter*, *Pelodictyon*, *Pelotomaculum*, *Penicillium*, *Peperomia*, *Peptoniphilus*, *Peptostreptococcus*, *Perilla*, *Perittia*, *Perkinsus*, *Persephonella*, *Persicobacter*, *Pervagor*, *Petroselinum*, *Petrotoga*, *Pfiesteria*, *Phaeobacter*, *Phaeodactylum*, *Phaeosphaeria*, *Phascolarctobacterium*, *Phenyllobacterium*, Phi29-like viruses, PhiC31-like viruses, PhiC32-like viruses, phiKMV-like viruses, phiKZ-like viruses, *Phormidium*, *Phoronis*, *Photobacterium*, *Photorhabdus*, *Physarum*, *Physcomitrella*, *Phytophthora*, *Picea*, *Pichia*, *Picrophilus*, *Picrorhiza*, *Pimelobacter*, *Pinctada*, *Pinus*, *Pisum*, *Planktothrix*, *Planobispora*, *Plasmodium*, *Plesiocystis*, *Pleurotus*, *Ploceus*, *Pneumocystis*, *Podospora*, *Poecilia*, *Polaromonas*, *Polynucleobacter*, *Polysphondylium*, *Pongo*, *Populus*, *Porphyra*, *Porphyrobacter*, *Porphyromonas*, *Postia*, *Potorous*, *Prasinovirus*, *Prauserella*, *Prionoxystus*, *Propionibacterium*, *Prostheobacter*, *Prostheochloris*, *Prototheca*, *Providencia*, *Pseudechis*, *Pseudendozonium*, *Pseudonocardia*, *Pseudoramibacter*, *Pseudovibrio*, *Pseudoxanthomonas*, *Psilotum*, *Psychrobacter*, *Psychroflexus*, *Psychromonas*, *Pteris*, *Puccinia*, *Pyramidobacter*, *Pyrenophora*, *Pyrobaculum*, *Pyrococcus*, *Rahnella*, *Ralstonia*, *Rana*, *Raphidiopsis*, *Rattus*, *Reclinomonas*, *Reinekea*, *Renibacterium*, *Rhadinovirus*, *Rhinoceros*, *Rhodobacter*, *Rhodococcus*, *Rhodomonas*, *Rhodopirellula*, *Rhodopseudomonas*, *Rhodothermus*, *Ricinus*, *Rickettsia*, *Rickettsiella*, *Ricordea*, *Riemerella*, *Robiginitalea*, *Roseburia*, *Roseibium*, *Roseomonas*, *Roseovarius*, *Rothia*, *Rubritalea*, *Ruegeria*, *Ruminococcus*, *Rupicapra*, *Saccharomonospora*, *Saccharomyces*, *Saccharophagus*, *Saccharopolyspora*, *Saccharum*, *Saccoglossus*, *Sagittula*, *Salinibacter*, *Salinispora*, *Salmo*, *Sanguibacter*, *Scardovia*, *Scenedesmus*, *Scheffersomyces*, *Schistocerca*, *Schistosoma*, *Schizophyllum*, *Schizosaccharomyces*, *Sclerotinia*, *Scutigera*, *Scytonema*, *Sebaldella*, *Segniliparus*, *Selaginella*, *Selenomonas*, *Serratia*, *Shewanella*, *Shigella*, *Shuttleworthia*, *Sideroxydans*, *Simonsiella*, *Simplexvirus*, *Sinorhizobium*, *Slackia*, *Sodalis*, *Solanum*, *Solobacterium*, *Sordaria*, *Sorex*, *Sorghum*, SP6-like viruses, SPbeta-like viruses, *Spermophilus*, *Sphaerotilus*, *Sphingobacterium*, *Sphingobium*, *Sphingopyxis*, *Spinacia*, *Spirochaeta*, *Spirogyra*, *Spirosoma*, *Spisula*, SPO1-like viruses, *Squalus*, *Stackebrandtia*, *Staphylothermus*, *Starkeya*, *Staurastrum*, *Stephos*, *Stigmatella*, *Streptoalloteichus*, *Streptobacillus*, *Streptosporangium*, *Strongylocentrotus*, *Suaeda*, *Subdoligranulum*, *Succinatimonas*, *Sulfitobacter*, *Sulfolobus*, *Sulfuricurvum*, *Sulfurihydrogenibium*, *Sulfurimonas*, *Sulfurospirillum*, *Sulfurovum*, *Sutterella*, *Synechococcus*, *Syntrophobacter*, *Syntrophomonas*, *Syntrophothermus*, *Syntrophus*, T1-like viruses, T4-like viruses, T7-like viruses, *Taenia*, *Taeniopygia*, *Takifugu*, *Talaromyces*, *Teredinibacter*, *Terracoccus*, *Terriglobus*, *Terrimonas*, *Tetrahymena*, *Tetraodon*, *Thalassiosira*, *Thalassobium*, *Thauera*, *Theileria*, *Thermaerobacter*, *Thermanaerovibrio*, *Thermincola*, *Thermoactinomyces*, *Thermoanaerobacter*, *Thermoanaerobacterium*, *Thermobaculum*, *Thermobispora*, *Thermococcus*, *Thermocrinis*, *Thermodesulfobacterium*, *Thermofilum*, *Thermomicrobium*, *Thermomonospora*, *Thermoplasma*, *Thermoproteus*, *Thermosediminibacter*, *Thermosinus*, *Thermosphaera*, *Thermosynechococcus*, *Thermovibrio*, *Thioalkalivibrio*, *Thiobacillus*, *Thiocapsa*, *Thiococcus*, *Thiorhodovibrio*, *Thylacodes*, *Tolomonas*, *Tolypothrix*, *Toxoplasma*, *Toxoptera*, *Trabulsiella*, *Tribolium*, *Trichinella*, *Trichocolea*, *Trichomonas*, *Trichophyton*, *Trichoplax*, *Triglochin*, *Triticum*, *Truepera*, *Trypanosoma*, *Tsukamurella*, *Tuber*, *Tubularia*, *Tupaia*, *Turbo*, *Turicibacter*, *Uncinocarpus*, *Ustilago*, *Vanderwaltozyma*, *Varicellovirus*, *Variovarax*, *Veillonella*, *Verrucomicrobium*, *Verticillium*, *Vicia*, *Victivallis*, *Vigna*, *Vitis*, *Volvox*, VP2-like phages, *Vulcanisaeta*, *Waddlia*, *Weeksella*, *Weissella*, *Wigglesworthia*, *Wolbachia*, *Wolinella*, *Xanthobacter*, *Xenopus*, *Xenorhabdus*, *Xylanimonas*, *Xylaria*, *Yarrowia*, *Yersinia*, *Zea*, *Zingiber*, *Zoophthora*, *Zunongwangia*, *Zygnema*, *Zygosaccharomyces*, *Zymomonas*

Mars Oasis during the summer, and those for photosynthesis. Given the abundance of Cyanobacteria in the lake margin at Mars Oasis (Wynn-Williams, 1996), it was surprising that only ~1% of the genes in the library encoded for CO<sub>2</sub> fixation. Although genes for nitrogen cycling expressed by phyla such as *Acidobacteria* were not found in the library, the use of probes indicated the presence of using *nifH*, *nosZ*, and *nirS* genes in the soil (data not shown). Genes for sulfur, phosphorus, and nitrogen metabolism were all present at about 1%, whilst those for iron acquisition and metabolism were 0.7% and potassium metabolism 0.3%.

The gene ontology data, if expressed as actual function, corroborate the view that the community at Mars Oasis is active during summer, with the potential expression of many genes for cytoskeleton and ribosome formation. The presence of genes for the utilization of lactose and galactose indicate that the microbes in the soil most probably utilize relatively simple sugars for growth: there are few plants at the oasis, and those that are present are bryophytes, which typically do not form complex aromatic molecules such as

lignin. Nevertheless, some capacity within the microbial community was found for the assimilation of aromatic compounds, with the presence of genes for the assimilation of peptides, which are known to be of importance to the nitrogen cycle in soils of the northern Maritime Antarctic (Hill et al., 2011).

#### SOIL CHEMISTRY

The dominant ion in runoff and soil pore water was sulfate, which is almost certainly derived from the oxidation of sulfide minerals in the local shales. Like the carbonate sources, there were also clear signs of secondary minerals (gypsum and/or anhydrite) contributing to the high SO<sub>4</sub><sup>2-</sup> concentrations. Furthermore, carbonate precipitates were visible around the base of all larger clasts in the soils (Andre and Hall, 2004), and so dissolution of secondary carbonates will have contributed to the high concentrations of Ca<sup>2+</sup> and HCO<sub>3</sub><sup>-</sup> also present in waters. The major ion geochemistry of surface waters at Mars Oasis therefore seems to be controlled by reactive carbonate and sulfide

mineral phases and the precipitation of secondary salts following the evaporation of sediment pore waters, with coupled pyrite oxidation and carbonate dissolution, and carbonate and anhydrite dissolution and precipitation dominating the rock weathering reactions.

## SUMMARY

Studies during the twentieth century suggested that Antarctic soils are of comparatively low microbial biodiversity (Wynn-Williams, 1996). This is certainly true when most Antarctic soils are compared to temperate or tropical soils (Fierer and Jackson, 2006; Fierer et al., 2007). It appears from the current study that prokaryotic diversity in soil at Mars Oasis is limited at the generic level, with the frequent occurrence of *Actinobacteria* and *Cyanobacteria*. However, at the sequence level, there appears to be much greater heterogeneity than was previously thought, perhaps owing to high divergence within the relatively restricted lineages that have successfully colonized Antarctic terrestrial environments. However, the process of grouping sequences can have an impact.

## REFERENCES

- Aislabie, J., Broady, P., and Saul, D. (2006). Culturable heterotrophic bacteria from high altitude, high latitude soil of La Gorce Mountains (86°30'S, 147°0'W), Antarctica. *Antarct. Sci.* 18, 313–321.
- Aislabie, J., Jordan, S., Ayton, J., Klassen, J. L., Barker, G. M., and Turner, S. (2009). Bacterial diversity associated with ornithogenic soil of the Ross Sea region, Antarctica. *Can. J. Microbiol.* 55, 21–36.
- Aislabie, J., Jordan, S., and Barker, G. M. (2008). Relation between soil classification and bacterial diversity in soils of the Ross Sea region, Antarctica. *Geoderma* 144, 9–20.
- Andre, M.-F., and Hall, K. (2004). Honeycomb development on Alexander island, glacial history of George VI sound and palaeoclimatic implications (two step cliffs/Mars Oasis, W Antarctica). *Geomorphology* 65, 117–138.
- Anon. (2009a). *emPCR Method Manual – Lib-L MV*. Mannheim: Roche Diagnostics, 12.
- Anon. (2009b). *Sequencing Method Manual*. Mannheim: Roche Diagnostics, 22.
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Wheeler, D. L. (2005). GenBank. *Nucleic Acids Res.* 33, D34–D38.
- Chen, F., and Suttle, C. A. (1995). Amplification of DNA-polymerase gene fragments from viruses infecting microalgae. *Appl. Environ. Microbiol.* 61, 1274–1278.
- Chong, C. W., Pearce, D. A., Convey, P., Tan, G. Y. A., Wong, R. C. S., and Tan, I. K. P. (2010). High levels of spatial heterogeneity in the biodiversity of soil prokaryotes on Signy Island, Antarctica. *Soil Biol. Biochem.* 42, 601–610.
- Chong, C. W., Pearce, D. A., Convey, P., and Tan, I. K. P. (2011). Assessment of soil bacterial communities on Alexander island (in the maritime and continental Antarctic transitional zone). *Polar Biol.* 35, 387–399.
- Chou, H.-H., and Holmes, M. H. (2001). DNA sequence quality trimming and vector removal. *Bioinformatics* 17, 1093–1104.
- Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., et al. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* 37, D141–145.
- Convey, P. (2001). “Antarctic ecosystems,” in *Encyclopedia of Biodiversity*, Vol. 1, ed. S. A. Levin (San Diego: Academic), 171–184.
- Courtois, S., Frostegård, A., Goransson, P., Depret, G., Jeannin, P., and Simonet, P. (2001). Quantification of bacterial subgroups in soil: comparison of DNA extracted directly from soil or from cells previously released by density gradient centrifugation. *Environ. Microbiol.* 3, 431–439.
- Delmont, T. O., Robe, P. R., Clark, I., Simonet, P., and Vogel, T. M. (2011). Metagenomic comparison of direct and indirect soil DNA extraction approaches. *J. Microbiol. Methods* 86, 397–400.
- DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006). Greengenes, a chimeric-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5072.
- Eisen, J. A. (2007). Environmental shotgun sequencing: its potential and challenges for studying the hidden world of microbes. *PLoS Biol.* 5:e82. doi:10.1371/journal.pbio.0050082
- Feinstein, L. M., Sul, W. J., and Blackwood, C. B. (2009). Assessment of bias associated with incomplete extraction of microbial DNA from soil. *Appl. Environ. Microbiol.* 75, 5428.
- Fierer, N., Breitbart, M., Nulton, J., Salamon, P., Lozupone, C., Jones, R., et al. (2007). Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of bacteria, archaea, fungi, and viruses in soil. *Appl. Environ. Microbiol.* 73, 7059–7066.
- Fierer, N., and Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U.S.A.* 103, 626–631.
- Fierer, N., Schimela, J. P., and Holden, P. A. (2003). Variations in microbial community composition through two soil depth profiles. *Soil Biol. Biochem.* 35, 167–176.
- Frostegård, A., Courtois, S., Ramisse, V., Clerc, S., Bernillon, D., Le Gall, F., et al. (1999). Quantification of bias related to the extraction of DNA directly from soils. *Appl. Environ. Microbiol.* 65, 5409–5420.
- Fuller, N., Wilson, W. H., Joint, I. R., and Mann, N. H. (1998). Occurrence of T4 gp20 homologues in marine cyanophages and their application to PCR-based detection and quantification techniques. *Appl. Environ. Microbiol.* 64, 2051–2060.
- Gardes, M., and Bruns, T. D. (1993). ITS primers with enhanced specificity for basidiomycetes – application to the identification of mycorrhizae and rusts. *Mol. Ecol.* 2, 113–118.
- Gilbert, J. A., Thomas, S., Cooley, N. A., Kulakova, A., Field, D., Booth, T., et al. (2009). Potential for phosphate utilisation by marine bacteria in temperate coastal waters. *Environ. Microbiol.* 11, 111–125.
- Gomez-Alvarez, V., Teal, T. K., and Schmidt, T. M. (2009). Systematic artifacts in metagenomes from complex microbial communities. *ISME J.* 3, 1314–1317.
- Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika* 40, 237–264.
- Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J., and Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* 5, 245–249.
- Hill, P. W., Farrar, J., Roberts, P., Farrell, M., Grant, H., Newsham, K. K., et al. (2011). Vascular plant success in a warming Antarctic may be due to efficient nitrogen acquisition. *Nat. Clim. Chang.* 1, 50–53.
- Hugenholtz, P., Goebel, B. M., and Pace, N. R. (1998). Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J. Bacteriol.* 180, 4765–4774.
- Hughes, K. A., McCartney, H. A., Lachlan-Cope, T. A., and Pearce, D. A. (2004). A preliminary study of airborne microbial biodiversity over Peninsular Antarctica. *Cell. Mol. Biol.* 50, 537–542.

- Lawley, B., Ripley, S., Bridge, P., and Convey, P. (2004). Molecular analysis of geographic patterns of eukaryotic diversity in Antarctic soils. *Appl. Environ. Microbiol.* 70, 5963–5972.
- Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659.
- Mahmoudi, N., Slater, G. F., and Fulthorpe, R. R. (2011). Comparison of commercial DNA extraction kits for isolation and purification of bacterial and eukaryotic DNA from PAH-contaminated soils. *Can. J. Microbiol.* 57, 623–628.
- Martin-Laurent, F., Philippot, L., Hallet, S., Chaussod, R., Germon, J. C., Soulas, G., et al. (2001). DNA extraction from soils: old bias for new microbial diversity analysis methods. *Appl. Environ. Microbiol.* 67, 2354–2359.
- Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., et al. (2008). The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9:386. doi:10.1186/1471-2105-9-386
- Newsham, K. K., Pearce, D. A., and Bridge, P. (2010). Minimal influence of water and nutrient content on the bacterial community composition of a maritime Antarctic soil. *Microbiol. Res.* 165, 523–530.
- Niederberger, T. D., McDonald, I. R., Hacker, A. L., Soo, R. M., Barrett, J. E., Wall, D. H., et al. (2008). Microbial community composition in soils of Northern Victoria Land, Antarctica. *Environ. Microbiol.* 10, 1713–1724.
- Overbeek, R., Begley, T., Butler, R., Choudhuri, J., Chuang, H., Cohoon, M., et al. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.* 33, 5691–5702.
- Palacios, C., Zettler, E., Amils, R., and Amaral-Zettler, L. (2008). Contrasting microbial community assembly hypotheses: a reconciling tale from the Río Tinto. *PLoS ONE* 3:e3853. doi:10.1371/journal.pone.0003853
- Pearce, D. A., Hughes, K. A., Harangozo, S. A., Lachlan-Cope, T. A., and Jones, A. E. (2010). Biodiversity of airborne microorganisms at Halley station, Antarctica. *Extremophiles* 14, 145–159.
- Peña, A., Teeling, H., Huerta-Cepas, J., Santos, F., Yarza, P., Brito-Echeverria, J., et al. (2010). Fine-scale evolution: genomic, phenotypic and ecological differentiation in two coexisting *Salinibacter ruber* strains. *ISME J.* 4, 882–895.
- Rösche, C., and Bothe, H. (2005). Improved assessment of denitrifying, N<sub>2</sub>-fixing, and total-community bacteria by terminal restriction fragment length polymorphism analysis using multiple restriction enzymes. *Appl. Environ. Microbiol.* 71, 2026–2035.
- Saul, D. J., Aislabie, J. M., Brown, C. E., Harris, L., and Foght, J. M. (2005). Hydrocarbon contamination changes the bacterial diversity of soil from around Scott Base, Antarctica. *FEMS Microbiol. Ecol.* 53, 141–155.
- Schloss, P. D. (2010). The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies. *PLoS Comput. Biol.* 6:e1000844. doi:10.1371/journal.pcbi.1000844
- Smith, J. J., Tow, L. A., Stafford, W., Cary, C., and Cowan, D. A. (2006). Bacterial diversity of three different Antarctic cold desert mineral soils. *Microb. Ecol.* 51, 413–421.
- Taton, A., Grubisic, S., Brambilla, E., De Wit, R., and Willemotte, A. (2003). Cyanobacterial diversity in natural and artificial microbial mats of Lake Fryxell (McMurdo Dry Valleys, Antarctica): a morphological and molecular approach. *Appl. Environ. Microbiol.* 69, 5157–5169.
- Throckmold, I. N., Enwall, K., Jarvis, A., and Hallin, S. (2004). Reassessing PCR primers targeting nirS, nirK and nosZ genes for community surveys of denitrifying bacteria with DGGE. *FEMS Microbiol. Ecol.* 49, 401–417.
- Turner, J., Colwell, S. R., Marshall, G. J., Lachlan-Cope, T. A., Carleton, A. M., Jones, P. D., et al. (2005). Antarctic climate change during the last 50 years. *Int. J. Climatol.* 25, 279–294.
- White, T. J., Bruns, T., Lee, S., and Taylor, J. (1990). “Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics,” in *PCR Protocols: A Guide to Methods and Applications*, eds M. A. Innis, H. Gelfand, J. S. Sninsky, and T. J. White (New York: Academic Press), 315–322.
- Wilson, W. H., Fuller, N. J., Joint, I. R., and Mann, N. H. (1999). Analysis of cyanophage diversity and population structure in a south-north transect of the Atlantic ocean. *Bull. Inst. Oceanogr.* 19, 209–216.
- Wintzingerode, F. V., Göbel, U. B., and Stackebrandt, E. (1997). Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol. Rev.* 21, 213–229.
- Wynn-Williams, D. D. (1983). Distribution and characteristics of chromobacterium in the maritime and sub-Antarctic. *Polar Biol.* 2, 101–108.
- Wynn-Williams, D. D. (1996). Antarctic microbial diversity: the basis of polar ecosystem processes. *Biodivers. Conserv.* 5, 1271–1293.
- Yergeau, E., Newsham, K. K., Pearce, D. A., and Kowalchuk, G. (2007). Patterns of bacterial diversity across a range of Antarctic terrestrial habitats. *Environ. Microbiol.* 9, 2670–2682.
- Yergeau, E., Schoondermark-Stolk, S. A., Brodie, E. L., Déjean, S., DeSantis, T. Z., Gonçalves, O., et al. (2009). Environmental microarray analyses of Antarctic soil microbial communities. *ISME J.* 3, 340–351.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 June 2012; accepted: 02 November 2012; published online: 05 December 2012.

Citation: Pearce DA, Newsham KK, Thorne MAS, Calvo-Bado L, Krsek M, Laskaris P, Hodson A and Wellington EM (2012) Metagenomic analysis of a southern maritime Antarctic soil. *Front. Microbiol.* 3:403. doi: 10.3389/fmicb.2012.00403

This article was submitted to *Frontiers in Extreme Microbiology*, a specialty of *Frontiers in Microbiology*.

Copyright © 2012 Pearce, Newsham, Thorne, Calvo-Bado, Krsek, Laskaris, Hodson and Wellington. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.