ORIGINAL ARTICLE

Molecular evidence for an active endogenous microbiome beneath glacial ice

Trinity L Hamilton[1], John W Peters[1], Mark L Skidmore[2] and Eric S Boyd[1]

[1]*Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT, USA and* [2]*Department of Earth Sciences, Montana State University, Bozeman, MT, USA*

Geologic, chemical and isotopic evidence indicate that Earth has experienced numerous intervals of widespread glaciation throughout its history, with roughly 11% of present day Earth's land surface covered in ice. Despite the pervasive nature of glacial ice both today and in Earth's past and the potential contribution of these systems to global biogeochemical cycles, the composition and phylogenetic structure of an active microbial community in subglacial systems has yet to be described. Here, using RNA-based approaches, we demonstrate the presence of active and endogenous archaeal, bacterial and eukaryal assemblages in cold (0–1 °C) subglacial sediments sampled from Robertson Glacier, Alberta, Canada. Patterns in the phylogenetic structure and composition of subglacial sediment small subunit (SSU) ribosomal RNA (rRNA) assemblages indicate greater diversity and evenness than in glacial surface environments, possibly due to facilitative or competitive interactions among populations in the subglacial environment. The combination of phylogenetically more even and more diverse assemblages in the subglacial environment suggests minimal niche overlap and optimization to capture a wider spectrum of the limited nutrients and chemical energy made available from weathering of bedrock minerals. The prevalence of SSU rRNA affiliated with lithoautotrophic bacteria, autotrophic methane producing archaea and heterotrophic eukarya in the subglacial environment is consistent with this hypothesis and suggests an active contribution to the global carbon cycle. Collectively, our findings demonstrate that subglacial environments harbor endogenous active ecosystems that have the potential to impact global biogeochemical cycles over extended periods of time.
The ISME Journal (2013) **7,** 1402–1412; doi:10.1038/ismej.2013.31; published online 14 March 2013
Subject Category: geomicrobiology and microbial contributions to geochemical cycles
Keywords: archaea; cold; eukarya; methane; RNA; subsurface

Introduction

Glacial ice covers ∼11% of Earth's land mass today and a much larger fraction at times in Earth's past (for example, in the late Neoproterozoic, so called 'Snowball Earth' ∼650 Mya; Kirschvink *et al.*, 2000). The advance and retreat of ice during glacial–interglacial cycles is thought to have driven large scale shifts in the distribution of taxa to habitats with more stable climatic conditions (refugia; Hodson *et al.*, 2008) enabling biodiversity to persist. Recently, it was suggested that glacial beds, which tend to be comparably more stable than other surface environments (Skidmore *et al.*, 2005), may have served as refugia for biodiversity during periods of inclement climatic, atmospheric or geologic conditions (Hodson *et al.*, 2008). Here, the continual exposure of fresh mineral surfaces at the bed of the glacier would serve as a steady source of nutrients and energy capable of supporting a functionally diverse assemblage of microorganisms over extended periods of time (Sharp *et al.*, 1999; Skidmore *et al.*, 2000, 2005). One-dimensional models that integrate metabolic potentials of putative microbial assemblages in sub-ice sheet sediments during glacial periods indicate that such ecosystems could contribute significantly to global biogeochemical cycles, specifically the carbon cycle (Wadham *et al.*, 2008, 2012). However, our understanding of life in subglacial sediment environments is incomplete and evidence for active microbiomes in such systems has yet to be firmly demonstrated.

Molecular phylogenetic analysis of RNA affords the unique opportunity to examine both the composition and structure of the active fraction of microbial assemblages sampled from natural environments (Felske *et al.*, 1998). Single-stranded RNA is synthesized only by active cells and it degrades relatively rapidly once produced when compared to DNA (Hirsch *et al.*, 2010). Thus, RNA-based approaches have been employed to identify the functioning members of natural microbial communities. Despite

the clear advantage of RNA-based approaches to microbial ecology research, such approaches have received far less attention in subsurface research. This is due primarily to the low biomass associated with many subsurface environments (Whitman et al., 1998) including subglacial systems (Kastovska et al., 2007), a feature that is often confounded by the presence of complex organic and mineral matrices in these systems (Hodson et al., 2008) that interfere with the successful extraction and purification of intact RNA. However, recent technological advances now make it possible to characterize assemblages from biomass-limited eco-systems using RNA-based approaches, including the deep subsurface (Ogram et al., 1995). These advances facilitate not only examination of the composition of communities, but also their phylogenetic structure. Patterns in the phylogenetic structure of assemblages can provide unique insight into the relative roles of stochastic (species neutral) and deterministic (competitive or facilitative inter-species interactions or environmental filtering) processes in dictating the assembly of macrobiological communities (Webb et al., 2002; Cavender-Bares et al., 2004; Horner-Devine and Bohannan, 2006; Bryant et al., 2008; Cavender-Bares et al., 2009; Meuser et al., 2013). Such approaches hold tremendous promise in the study of microbial ecosystems where they can provide critical insight into the mechanisms that generate and maintain diversity, such as speciation, extinction, dispersal and species interactions (Martiny et al., 2006). Nonetheless, they have only recently been applied to microbial assemblages (Horner-Devine and Bohannan, 2006; Bryant et al., 2008; Meuser et al., 2013).

Robertson Glacier (RG), Alberta, Canada is an alpine glacier in the Canadian Rockies that, for the past decade, has served as a model system to understand the contribution of glacial systems to local and global biogeochemical cycles (Sharp et al., 2002; Boyd et al., 2010, 2011). Several indirect lines of evidence suggest that the subsurface environment beneath RG may host an active microbiome. First, geochemical data indicate the production of significant amounts of sulfate in the cold (0–1 °C) subglacial environment, likely through the weathering of pyrite present in bedrock (Sharp et al., 2002). Second, genetic and microcosm studies reveal that RG subglacial sediments host a diversity of organisms that are capable of performing several redox transformations ex situ, including methanogenesis (Boyd et al., 2010) and nitrate reduction and nitrification (Boyd et al., 2011). Third, the detection of coenzyme M, a biomarker of methanogens and archaeal methanotrophs (Elias et al., 1999; Hallam et al., 2004), was detected in subglacial sediments at a concentration corresponding to ~ 3000 cells gram dry weight sediment^{-1} (Boyd et al., 2010). However, it is not known if the populations were active or in a state of quiescence because of an incomplete understanding of the stability of coenzyme M under these environmental conditions. In this study, we evaluated the potential for RG subglacial sediments to host active microbial communities by quantifying and sequencing archaeal, bacterial, and eukaryal small subunit (SSU) ribosomal RNA (rRNA) genes and their transcripts in genomic DNA and RNA extracts, respectively. In order to test for endogenicity in the subglacial ecosystem, these data were then compared with data from representative surficial glacial environments (snow debris and cryoconite sediments) that when melted or disturbed, could introduce exogenous populations to the subglacial system via moulins and crevasses. Patterns in the phylogenetic structure and composition identified in the respective communities are discussed in terms of the role of subglacial systems in (i) global biogeochemical cycles and (ii) supporting biodiversity during extended periods of inclement climatic, atmospheric or geological conditions (Hodson et al., 2008).

Materials and methods

Sample collection
Samples for nucleic acid extraction were collected on 14 October 2010 from RG (115°20′W, 50°44′N) in Peter Lougheed Provincial Park, Kananaskis County, Alberta, Canada. Fine-grained basal sediments were collected aseptically from within an ice cave that formed at the terminus of the glacier due to subglacial discharge at 1200 hours. Subsamples of cryoconite sediment and snow, the latter of which contained a noticeable amount of debris, were collected from the surface of the glacier at 1300 hours. A detailed description of sample sites and sample collection strategy is provided in the supporting online material (SOM), SOM Figure 1
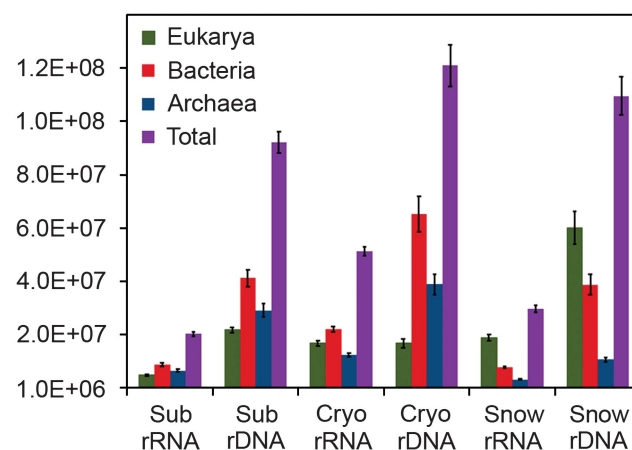

Figure 1 qRT-PCR of SSU rRNA transcripts (rRNA) and genes (rDNA) from subglacial sediment, cryoconite sediment and snow debris. Results are presented as the mean of triplicate qPCR assays; error bars represent the s.d. of replicates. Copy number is normalized to per gram dry weight sediment. Cryo, cryoconite sediment; Sub, subglacial sediment.

Table 1 Sequencing and diversity statistics for assemblages sampled at Robertson Glacier. SSU cDNA and SSU rDNA assemblages for each sample site are separated by a slash

Diversity statistics	Mean phylogenetic diversity[a]			Predicted coverage of diversity[b]		
	Subglacial	Cryoconite	Snow	Subglacial	Cryoconite	Snow
Archaea	0.165/0.151	0.164/0.068	**0.062/0.024**	0.985/0.985	0.956/0.993	0.985/0.985
Bacteria	0.517/0.527	0.509/0.520	0.434/0.450	0.962/0.937	0.974/0.871	0.992/0.943
Eukarya	0.730/0.599	0.549/0.569	0.429/0.482	0.942/0.955	0.905/0.788	0.970/0.900

Evenness statistics	NRI[c]			NTI[c]		
	Subglacial	Cryoconite	Snow	Subglacial	Cryoconite	Snow
Archaea	0.732/1.149	− 0.061/0.811	**1.713/0.976**	0.333/0.760	0.877/0.413	0.913/**1.034**
Bacteria	− 0.823/ − 1.015	− 0.714/ − 0.845	− 0.369/ − 0.158	**− 1.574**/ − **2.320**	− 1.067/ − **3.269**	0.494/ − 0.590
Eukarya	− **1.426**/ − 0.231	0.408/0.440	0.885/1.060	− **1.697**/ − **1.146**	− 0.234/0.033	0.949/**1.546**

Abbreviations: cDNA, complementary DNA; MPD, mean phylogenetic distance; NRI, net relatedness index; NTI, nearest taxon index; OTU, operational taxonomic unit; rRNA, ribosomal RNA; SSU, small subunit.
[a]MPD, calculated using representative OTUs defined at 97.0% identities. Higher MPD indices are indicative of a greater phylogenetic diversity. Boldface values are statistically significant from randomized samples (null model 3) at $P<0.10$.
[b]The percentage of the predicted diversity (defined at 97.0% identities) that was sampled from each assemblage.
[c]The NRI and NTI as an indicator of tree-wide and branch tip phylogenetic evenness, respectively, in each assemblage. Increasingly negative NRI and NTI are indicative of phylogenetic overdispersion/evenness whereas increasingly positive values are indicative of phylogenetic clustering. Boldface values are statistically significant from randomized samples (null model 3) at $P<0.10$.

and SOM Table 1. Samples for DNA-based analyses (∼1 g) were collected in sterile 1.5-ml microcentrifuge tubes with flame-sterilized spatulas and immediately flash-frozen in a dry ice–ethanol slurry. Sediment aliquots (∼ 1 g) for RNA were collected in sterile 2-ml tubes containing 0.5 ml RNALater (Qiagen, Valencia, CA, USA) and flash-frozen in a dry ice–ethanol slurry. Samples were stored on dry ice during transport to the field station and during transport back to Montana State University where they were stored at − 80 °C until further processed. With the exception of snow samples, all samples for RNA extraction were stored in RNALater (Qiagen) prior to RNA extraction. Snow samples were kept at − 80 °C until they were melted (4 °C) and concentrated via centrifugation (14 000 × g, 5 min., 4 °C) immediately prior to molecular-based analyses (additional details of snow processing is provided in SOM).

Nucleic acid extraction, quantification, and generation of complementary DNA (cDNA)
DNA extraction and purification was carried with a Fast DNA Spin Kit for Soil (MP Biomedicals, Solon, OH, USA) as modified from the manufacturer's instructions as previously described (Boyd et al., 2007b). DNA was extracted in triplicate from three independent ∼250 mg subsamples of wet weight sediment in the case of subglacial and cryoconite samples or ∼250 mg of debris concentrated from melted snow. Equal volumes of each triplicate extraction were pooled for further analyses. RNA extraction and initial purification was carried out using a FastRNA Spin kit for Soil (MP Biomedicals). RNA was extracted in triplicate from

approximately ∼400 mg of wet weight sediment (subglacial and cryoconite samples) or ∼400 mg of wet weight debris (snow sample). Following initial purification, RNA was subjected to DNAse I digestion (Roche, Indianapolis, IN, USA) for 1 h at room temperature (∼22 °C). Following digestion, RNA was further purified using a High Pure RNA Isolation Kit (Roche) and was stored at − 80 °C in a solution of 100% ethanol and 0.3 M sodium acetate until further processed. The dry solid content of the subglacial sediments, cryoconite sediments and snow debris used in molecular analyses was determined by drying the residual material remaining after nucleic acid extraction at 80 °C for 24 h.

The concentration of DNA or RNA in each pooled volume was determined using a Qubit dsDNA HS Assay kit or a Qubit RNA Assay kit (Molecular Probes, Eugene, OR, USA) and a Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA), respectively. As an additional precaution, RNA extracts were screened for the presence of contaminating genomic DNA by performing a PCR using ∼1 ng of RNA as template and archaeal and bacterial 16S rRNA gene primers, described below. Equal volumes of each RNA extract were pooled and subjected to cDNA synthesis. cDNA was synthesized from 20 ng of purified RNA extracted from the subglacial and cryoconite sediments and the snow debris using the iScript cDNA Synthesis Kit (Bio-Rad Laboratories, Hercules, CA, USA) and the following reaction cycling conditions: 5 min at 25 °C, 30 min at 42 °C and 5 min at 85 °C. Following synthesis of cDNA, samples were purified by ethanol precipitation and re-suspended in nuclease-free water for use in pyrotag sequencing.

PCR amplification of SSU rRNA genes from genomic and complementary DNA

Purified genomic DNA extracts or cDNA were subjected to amplification of archaeal and bacterial 16S ribosomal DNA (rDNA) using primers 344F (5′-ACGGGGYGCAGCAGGCGCGA-3′) and 915R (5′-GTGCTCCCCGCCAATTCCT-3′) at an annealing temperature of 61 °C or primers 1100F (5′-YAACGA GCGCAACCC-3′) and 1492R (5′-GGTTACCTTGTTA CGACTT-3′) and an annealing temperature of 55 °C, respectively, (Boyd *et al.*, 2007a). For amplification of eukaryotic 18S rDNA, primers euk-A7F (5′-AACC TGGTTGATCCTGCCAGT-3′) (Medlin *et al.*, 1988) and Euk-570R (5′-GCTATTGGAGCTGGAATTAC-3′) (Weekers *et al.*, 1994) were used at an annealing temperature of 42 °C. For each set of primers, ∼1 ng of purified genomic DNA or cDNA was subjected to PCR in triplicate using the following reaction conditions: initial denaturation at 94 °C (4 min), followed by 35 cycles of denaturation at 94 °C (1 min), annealing at the optimal temperature for each primer pair (1 min), primer extension at 72 °C (1.5 min), followed by a final extension step at 72 °C for 20 min. Reactions contained 2 mM MgCl$_2$ (Invitrogen), 200 μM each deoxynucleotide triphosphate (Eppendorf, Hamburg, Germany), 0.5 μM forward and reverse primer (Integrated DNA Technologies, Coralville, IA, USA), 0.4 mg ml^{-1} molecular-grade bovine serum albumin (Roche), and 0.25 units Taq DNA polymerase (Invitrogen) in a final reaction volume of 50 μl. Positive control reactions were performed using genomic DNA from *Azotobacter vinelandii* DJ, *Roseiflexus castenholzii*, *Sulfolobus solfataricus* P2 and *Chlamydomonas reinhardtii*. Negative control reactions were performed in the absence of added genomic DNA or cDNA.

Quantitative PCR (qPCR)

qPCR was used to estimate the number of SSU rDNA templates in genomic DNA extracted from subglacial sediments, cryoconite sediment and snow debris. qPCR followed protocols described previously (Boyd *et al.*, 2011). Briefly, SSU rDNA genes amplified from the archaeal, bacterial and eukaryal positive controls (as listed in previous section) were used to generate standard curves for use in relating template copy number to threshold qPCR amplification signal. Triplicate PCR products were purified using a QIAquick PCR Purification Kit (Qiagen) and quantified as described above. The concentrations of each replicate PCR product varied by less than a factor of 1.3 and thus were averaged for use in generating standard curves. An archaeal 16S rDNA standard curve was generated over 5 orders of magnitude from 2.6×10^1 to 2.2×10^6 copies of template per assay ($R^2 = 0.998$). A bacterial 16S rDNA standard curve was generated over 5 orders of magnitude from 3.1×10^1 to 5.7×10^6 copies of template per assay ($R^2 = 0.998$). An eukaryotic 18S rDNA standard curve was generated over 5 orders of magnitude from 9.1×10^1 to 3.7×10^6 copies of template per assay ($R^2 = 0.997$). The detection limit for all assays was between 10 and 15 copies of template per assay. qPCR assays were performed in optically clear PCR tubes (Qiagen) in a Rotor-Gene 300 quantitative real-time PCR machine (Qiagen) using a SsoFast EvaGreen Supermix qPCR Kit (Bio-Rad Laboratories). Assay reactions were amended with molecular-grade bovine serum albumin to a final concentration of 0.4 mg ml^{-1} (Roche). qPCR cycling conditions were as follows: initial denaturation (95 °C for 10 min) followed by 40 cycles of denaturation (95 °C for 10 s), annealing (at the optimal annealing temperature for each primer as described above for 10 s) and extension (72 °C for 20 s). Specificity of the qPCR assays was verified by melt curve analysis. The reported template abundances are the average and standard deviation of qPCR assays performed in triplicate.

Quantitative reverse transcription-PCR (qRT-PCR)

qRT-PCR was used to quantify the abundance of 16S rRNA transcripts in RNA extracted from subglacial sediment, cryoconite sediment and snow debris. qRT-PCR was performed with the Power SYBR Green RNA-to-C$_T$ one-step kit from Invitrogen according to the manufacturer's protocol and were assayed on a RotorGene-Q real-time PCR detection system from Qiagen. Reactions were performed in triplicate with each reaction containing ∼10 ng of total RNA, 500 nM forward and reverse primer, in a final reaction volume of 20 μl. The following cycling conditions were employed: reverse transcription at 48 °C (30 min) followed by initial activation of the DNA polymerase at 95 °C (10 min) followed by 40 cycles of denaturation at 95 °C (15 s), annealing and extension (at the optimal annealing temperature for each primer as described above for 1 min). Specificity of the qRT-PCR assays was verified by melt curve analysis. Control reactions to ensure that the amplified product was not due to genomic DNA contamination were performed in the absence of reverse transcriptase or in the absence of added template RNA. An archaeal 16S rDNA standard curve was generated over 5 orders of magnitude from 6.5×10^2 to 3.5×10^7 copies of template per assay ($R^2 = 0.997$). A bacterial 16S rDNA standard curve was generated over 5 orders of magnitude from 9.7×10^1 to 2.3×10^6 copies of template per assay ($R^2 = 0.998$). A eukaryotic 18S rDNA standard curve was generated over 5 orders of magnitude from 4.4×10^2 to 5.1×10^7 copies of template per assay ($R^2 = 0.997$). The detection limit for all assays was between 14 and 18 copies of template per assay.

Sequence analysis

SSU rDNA and SSU cDNA were sequenced by the Research and Testing Laboratory (Lubbock, TX, USA). SSU rDNA and cDNA from each fraction

were barcoded as previously described (Dowd et al., 2008) using the primers described above for SSU rDNA analysis and were sequenced using a 454 Genome Sequencer FLX System (Roche, Nutley, NJ, USA). Each sample was sequenced once. Post sequence processing was performed using the Mothur (ver. 1.24.1) sequence analysis platform (Schloss et al., 2009). Raw libraries were trimmed to a minimum length of 200 bases and were subjected to a filtering step using the quality scores file to remove sequences with anomalous base calls. Unique sequences were aligned using SILVA databases specific for SSU rDNA from each domain and sequences were trimmed using a defined start and end site based on inclusion of 75% of the total sequences; sequences that started before or after these defined positions were removed without further consideration. The resulting unique sequences were pre-clustered to remove amplification and sequencing errors and chimeras were identified and removed using UCHIME (Edgar et al., 2011). Operational taxonomic units (OTUs) were assigned at a sequence similarity of 97.0% using the furthest-neighbor method. The remaining sequences were randomly sub-sampled in order to normalize the total number of sequences in each library. Collectively, these steps resulted in a normalized size of 1136, 1278 and 938 sequences for each archaeal, bacterial and eukaryal library, respectively. Rarefaction curves were used to compute the percent coverage of the predicted taxonomic richness for each library, which indicating that $>87\%$ of the predicted sequence richness was sampled for each library (Table 1). Sequences were classified using the Bayesian classifier (Wang et al., 2007) and the RDP database, with manual verification using BLASTn (Supplementary Tables 2–4). Sequences representing each OTU have been deposited in the NCBI SRA database under accession numbers SAMN01729343 to SAMN01730044 and KC572493 (Supplementary Tables 2–4).

Sequences representing each unique OTU (defined at 97.0% sequence identities) were compiled for each domain. ClustalX (ver. 2.0.9; Larkin et al., 2007) was used to align nucleic acid sequences using default parameters. Archaeal, bacterial and eukaryal SSU rDNA and SSU cDNA alignments were subjected to evolutionary model prediction via jModeltest (ver. 2.1.1; Darriba et al., 2012), Maximum-likelihood phylogenetic reconstruction via PhyML (version 3.0; Guindon and Gascuel, 2003) specifying the general time reversible model and gamma distributed rate variation with a proportion of invariable sites, and rate smoothing using the multidimensional version of Rambaut's parameterization as implemented in PAUP (ver. 4.0; Swofford, 2001) as previously described (Meuser et al., 2013).

Community ecology analysis
Rate smoothed ultrameric trees were used to calculate the mean phylogenetic distance (MPD), net relatedness index (NRI) and nearest taxon index

(NTI) for each assemblage while specifying among individual abundances (-a parameter) over 999 iterations with the program Phylocom (ver. 4.0.1; Webb et al., 2008). MPD is an abundance weighted metric that describes the pairwise phylogenetic distance between sequences in a community, when compared with the total sequence pool. Assemblages with higher MPD indices exhibit a greater phylogenetic diversity relative to the total sequence pool. The NRI is a measure of tree-wide phylogenetic clustering of sequences, whereas NTI is more sensitive to clustering at the terminals of the tree (Webb et al., 2002). Increasingly positive NRI and NTI metrics indicate that co-occurring species are more phylogenetically related than expected by chance (phylogenetic clustering). In contrast, increasingly negative NRI and NTI metrics indicate that co-occurring species are less phylogenetically related than expected by chance (phylogenetic overdispersion). We tested whether these values differed significantly from that of a randomly assembled community by comparison with a null model generated by the independent swap method (null model 3 within Phylocom). The independent swap method was selected because of the results of a previous empirical study (Kembel, 2009) that demonstrated that of the models available in Phylocom, this method is the most robust for detecting niche-based assembly processes and is less prone to type I errors. One thousand swaps were performed over 999 permutations, specifying sequence abundance weights (-a) in the analysis. A two-tailed significance test was used to evaluate the rank of observed values. When >900 permutations supported the observed values rather than the random or null model ($P<0.10$), the observed rank was deemed to be significant.

Phylocom was also used to calculate Rao's community phylogenetic relatedness for archaeal, bacterial and eukaryal assemblages using relative sequence abundance weights and rate-smoothed ultrameric trees. PAST (ver. 2.17b; Hammer et al., 2001) was used to generate cluster dendograms specifying the single linkage method and Euclidean distances. Bootstrap values correspond to the frequency that each node was observed in a given position out of 1000 replicates.

Results and discussion

Abundance of SSU rRNA and rDNA
The detection of archaeal, bacterial and eukaryal SSU rRNA in subglacial sediments suitable for reverse transcription to cDNA (Figure 1) provided the first direct evidence that the RG subglacial system harbors active microbial populations. Surprisingly, subglacial sediments harbored only a slightly lower total (archaea + bacteria + eukarya) SSU rDNA gene abundance than surface snow or cryoconite (Figure 1), although the total SSU rRNA

abundance was lower in the subglacial sediments than surface snow or cryoconite. As such, the ratio of total SSU rRNA to SSU rDNA genes was significantly lower ($P < 0.03$ in all two-tailed pairwise T-tests) in the subglacial sediments (0.22 ± 0.01) than in cryoconite debris and surface snow (0.42 ± 0.03 and 0.27 ± 0.02, respectively; Figure 1). The lower abundance of SSU rRNA and lower ratio of SSU rRNA to SSU rDNA in the subglacial sediments could be attributable to the dependence on limited chemical energy and oligotrophic conditions in the RG subglacial environment (Boyd et al., 2011) when compared with the surface environment, which has direct access to light energy capable of supporting photosynthesis.

Phylogenetic structure and diversity of assemblages
Unlike traditional alpha diversity metrics (for example, pairwise comparison of OTU overlap among communities), beta diversity metrics such as NRI, NTI and MPD use a phylogenetic framework for determining the extent of evolutionary divergence between taxa that comprise a community, which can vary considerably (Lozupone and Knight, 2008). Given that many metabolic traits are conserved during the evolution of a lineage (Blomberg et al., 2003), positive relationships between the phylogenetic relatedness of species and their overall ecological similarity are often observed (Webb et al., 2002; Wiens and Graham, 2005; Cadotte et al., 2009). For this reason, beta metrics of diversity (for example, MPD) can provide insight into both the taxonomic and ecological variation among communities. Similarly, beta metrics of phylogenetic structure, which describe the extent to which phylotypes comprising a community are clustered on ultrameric phylogenetic trees (for example, NRI/NTI), provide insight into the role of stochastic (species neutral) and deterministic mechanisms, either biological (competitive interactions) or environmental (habitat filtering), in the assembly of the community (Webb et al., 2002; Cavender-Bares et al., 2004; Horner-Devine and Bohannan, 2006). This information can provide a basis for predicting a number of properties of an ecosystem, including overall productivity given finite resources (Maherali and Klironomos, 2007), the extent of niche overlap (Cavender-Bares et al., 2004), as well as the susceptibility of a community when confronted with environmental change (Knapp et al., 2008).

Archaeal assemblages at RG tended to be phylogenetically clustered as evinced by positive NRI and/or NTI metrics regardless of whether they were derived from genomic DNA or cDNA fractions or if they were derived from surface cryoconite, surface snow or subglacial sediment. This suggests that archaea dispersing into each of the respective environments are excluded or included based on their ecologies, which is dictated largely by the physiological traits harbored by an organism (Webb et al., 2002; Horner-Devine and Bohannan, 2006; Martiny et al., 2006). In other words, the conditions of the environment, which are the combined result of abiotic and biotic factors, are likely to select for a particular set of traits that are compatible with the characteristics of an environment and which permit colonization. The inferred physiologies of archaeal populations (described in detail below) in the subglacial and cryoconite assemblages is that of strict anaerobes, whereas those in the surface snow assemblages are strict aerobes. Thus, aerobic/anaerobic boundaries may select for specific traits related to the utilization or avoidance of oxygen in the respective ecosystems. Such a situation is consistent with data that suggest anoxic conditions prevail at the bed of an alpine glacier in Switzerland (Tranter et al., 2002).

In contrast, bacterial and eukaryal SSU cDNA assemblages associated with subglacial sediments and cryoconite exhibited significant and negative NRI and/or NTI, with those associated with the subglacial sediments typically exhibiting more statistically significant and negative values (Table 1). Increasingly negative NRI/NTI indices suggest that assemblages contain phylotypes that are more phylogenetically divergent or that are affiliated with higher taxonomic ranks, given the species richness of the system (that is, phylogenetically overdispersed or phylogenetically even) (Webb et al., 2002). Evidence for evenly structured communities has been interpreted to be the result of (i) competitive interactions among closely related species that share a fundamental niche thereby limiting their co-existence over the long term (competitive exclusion) (Losos 1994; Webb et al., 2002; Cavender-Bares et al., 2004; Maherali and Klironomos, 2007) or (ii) facilitative interactions where facilitator species create microhabitats that support distantly related specie(s) with non-overlapping niches (Lortie, 2007; Valiente-Banuet and Verdu, 2007). Evenly structured soil communities are associated with the utilization of a greater spectrum of the resources when compared with phylogenetically clustered assemblages (Maherali and Klironomos, 2007; Fornara and Tilman, 2008). Thus, phylogenetically even assemblages in the RG subglacial environment may reflect the oligotrophic conditions present in the RG subsurface environment (Boyd et al., 2010, 2011) and selection to optimize the utilization of limited resources.

In this study, the MPD of archaeal, bacterial and eukaryal SSU cDNA was higher in the subglacial ecosystem than surface snow or cryoconite debris ecosystems (Table 1), consistent with conclusions drawn from analyses of NRI/NTI metrics. As closely related species tend to have similar physiological traits and often have overlapping ecological niches (Webb et al., 2002; Wiens and Graham, 2005), it follows that distantly related species should have divergent physiological traits, a feature that would help maintain minimal niche overlap and

potentially minimize competitive interactions (Maherali and Klironomos, 2007). Alternatively, the tendency for overdispersed subglacial assemblages to exhibit elevated phylogenetic diversity may point to an important role for facilitation among divergent, co-occurring species. Previous studies have shown that facilitative or mutualistic interactions among plant populations increase the availability of resources and alter substrate characteristics (Callaway, 1995), features that may lead to increased phylogenetic diversity (Valiente-Banuet and Verdu, 2007) and which may stabilize ecosystems (Ringel et al., 1996). Mutualistic relationships are common in nature and may be more pronounced in nutrient-limited ecosystems such as RG, which has been shown to be oligotrophic with respect to bioavailable nitrogen (Boyd et al., 2011) and phosphorus (ES Boyd, unpublished).

The greater phylogenetic diversity of active (RNA-based SSU cDNA analysis) and total (DNA-based SSU rDNA analysis) archaea, bacteria and eukarya associated with the subglacial sediments when compared with surface assemblages was unanticipated. In particular, the detection of active eukarya in the hypoxic subglacial sediments was surprising given the lack of previous evidence for this domain in cold subglacial environments. Recent evidence, however, suggests that eukaryote populations are active in a number of deep sea marine sediment environments characterized by cold, hypoxic and oligotrophic conditions (Edgcomb et al., 2010; Schippers et al., 2012). In deep marine sediments, eukaryotes are likely to be bacterivores that graze bacterial cells or heterotrophs dependent on buried organics. Phylogenetic inference of the dominant active eukaryotes in RG subglacial sediments (see below) indicate that they are heterotrophic and dependent on organic carbon. This finding is consistent with the abundance of particulate organic carbon (25 ± 14 mg g sediment^{-1}; Boyd et al., 2010) and pore water dissolved organic carbon (60μmol l^{-1}; Boyd et al., 2011) in RG sediments.

Comparison of SSU rDNA and cDNA composition

Archaeal, bacterial and eukaryal SSU cDNA and rDNA assemblages from each environment tended to form sister groups or branch closely together in cluster analysis (Figure 2), providing evidence that the SSU cDNA in each subglacial assemblage is the result of endogenous synthesis. In the case of each domain, the subglacial and cryoconite assemblages clustered closely together and distantly from the snow assemblage, suggesting that the subglacial sediment and cryoconite ecosystems are more similar ecologically. Intriguingly, the depth (or the extent to which the assemblages differ) of the clusters that comprise the SSU cDNA and rDNA sister groups for the subglacial and cryoconite archaeal, bacterial and eukaryal microbiomes was generally shallower than that from the snow
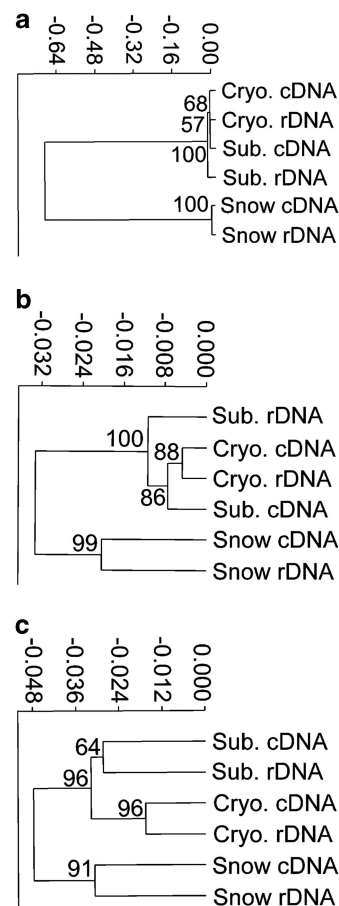

Figure 2 Cluster dendograms depicting the Rao phylogenetic relatedness of archaeal (a), bacterial (b) and eukaryal (c) SSU rDNA and SSU cDNA assemblages. The cophenetic correlation coefficient indicating the extent to which the pairwise distances between the original unmodeled dissimilarity data points were preserved in each dendogram was 0.9997, 0.9486 and 0.8791, respectively. Cryo, cryoconite sediment; Sub, subglacial sediment.

microbiome, suggesting that the majority of the SSU rDNA genes present in the subglacial and cryoconite microbiomes are being transcribed, with a much lower fraction of the SSU rDNA genes being transcribed in the snow microbiome at the time of sampling (midday).

Composition of archaeal, bacterial and eukaryal assemblages

The composition of SSU rDNA and cDNA communities was similar in each respective environment especially in the subglacial ecosystem (Figure 3), consistent with the results of the cluster analysis. In general, the inferred physiology of the dominant populations associated with subglacial sediments suggests the presence of a complex ecosystem that contains both aerobic and anaerobic taxa likely harbored in numerous microhabitats (Supplementary Tables 2–4). Archaeal SSU cDNA was dominated (70% of total) by sequences
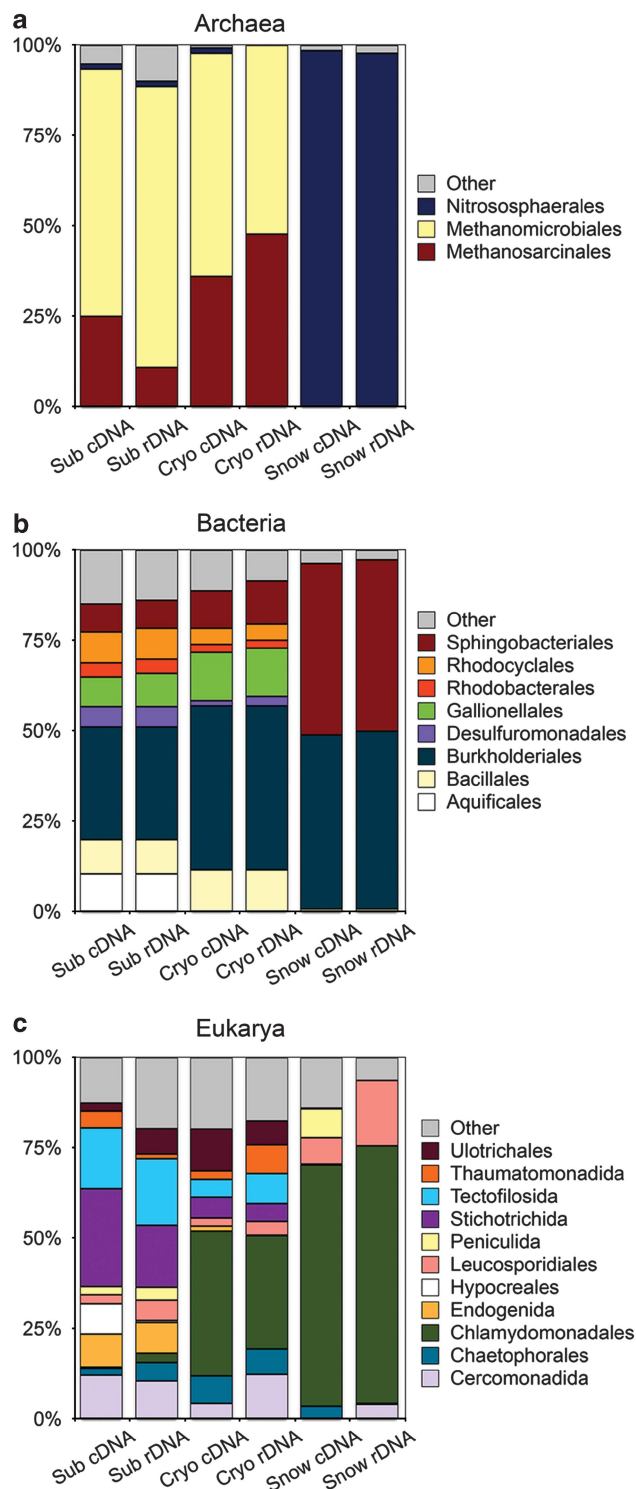
Figure 3 Composition of SSU rDNA and SSU cDNA in nucleic acids recovered from subglacial sediment, cryoconite sediment and snow debris. Representative OTUs for each library were binned at the order level for eukarya (**a**) and bacteria (**b**) and at the family level for archaea (**c**). Taxonomic bins that represented <5.0% of the total sequences from each assemblage were pooled and depicted as 'Other'. Cryo, cryoconite sediment; Sub, subglacial sediment.

affiliated with the strictly anaerobic methanogen order Methanomicrobiales with a smaller fraction (25% of total) affiliated with the Methanosarcinales (Supplementary Table 2). In contrast, bacterial SSU cDNA was dominated by sequences affiliated with aerobic autotrophs *Sideroxydans* sp. (12% of total), *Sulfurihydrogenibium* sp. (12% of total) and *Planococcus* sp. (10% of total) (Supplementary Table 3). Sequences affiliated with *Sulfurihydrogenibium* sp., which were only identified in cold (0–1 °C) subglacial SSU cDNA and rDNA assemblages, are of particular interest as sequences affiliated with this lineage have yet to be identified in any environment with a temperature of <50 °C (Reysenbach *et al.*, 2005). The eukaryal SSU cDNA assemblage in the subglacial ecosystem was dominated by sequences affiliated with two unicellular orders, including the ciliate order Strichotrichida (*Amphisiella* sp., 29% of total) and the amoeba order Tectofilosida (*Capsellina* sp., 17% of sequences), both of which are heterotrophic (Supplementary Table 4). The eukaryal assemblage associated with the cDNA fraction exhibited differences from that associated with the DNA fraction, including a greater abundance of sequences affiliated with the amoeba *Stichotrichida* and the flagellate *Thaumatomonadida* and, importantly, an absence of sequences affiliated with the green alga *Ulotrichales* (Figure 3; Supplementary Table 4). The recovery of active and dominant archaea, bacteria and eukarya in subglacial sediments that are putatively involved in autotrophic methane production, lithoautotrophic carbon dioxide fixation and heterotrophy, respectively, suggests that these ecosystems have significant potential to impact global biogeochemical cycles, in particular the carbon cycle. These results confirm previous data that were suggestive of the presence of an active methanogen assemblage in this environment (Boyd *et al.*, 2010) and corroborate recent reports suggesting that these ecosystems may contribute to the global methane cycle over recent glacial–interglacial cycles (Wadham *et al.*, 2008, 2012).

Comparison of the taxonomic affiliation of SSU cDNA derived from subglacial sediment, surface snow and cryoconite communities reveal distinct differences among the communities (Figure 3; Supplementary Tables 1–3). In general, the cryoconite and subglacial sediment SSU cDNA communities were most similar, with pronounced differences noted between these assemblages and the assemblages associated with snow. In particular, the abundance of eukaryal sequences affiliated with *Chlamydomonas* sp., an alga, in the surface snow (67% of total) and cryoconite (40% of total) but not the subglacial sediment is consistent with previous hypotheses suggesting that subglacial assemblages are supported by chemical energy (Skidmore *et al.*, 2000, 2005; Gaidos *et al.*, 2008) while surface communities are supported primarily by light energy (Telling *et al.*, 2010). Sequences affiliated with the green alga *Ulotrichales* were detected in

1410

both the SSU cDNA and rDNA fraction of the cryoconite assemblage, but only in the rDNA fraction of the subglacial sediment assemblage. This may indicate the subglacial environment receives exogenous material from the surface, such has been demonstrated in a separate Canadian Arctic glacial system using DNA-based fingerprinting methods (Bhatia *et al.*, 2006). However, the physical characteristics of the environment (for example, lack of available light) are likely to preclude activity of this phototrophic population, which may explain the lack of detectable sequences affiliated with this taxon in the cDNA fraction. Such observations further support endogenous synthesis of RNA in the subglacial environment.

Subglacial sediments and cryoconite harbored similar archaeal SSU cDNA assemblages, both of which were dominated by sequences affiliated with methanogens within the Methanomicrobiales and to a lesser extent Methanosarcinales (Figure 3). In contrast, the surface snow archaeal SSU cDNA assemblage was dominated (98% of total) by sequences affiliated with nitrifying archaea (Nitrososphaerales). Bacterial SSU cDNA assemblages also exhibited distinct differences, most notably the detection of sequences affiliated with Aquificales only in the subglacial community, and the detection of sequences affiliated with the iron utilizing Gallionellales (*Sideroxydans* sp.) only in the subglacial sediment and cryoconite debris communities.

Conclusions

The data presented here provide the first RNA-based evidence for an active, diverse and endogenous microbial community adapted to the dark and oligotrophic conditions that characterize subglacial environments. At RG, the subglacial sediments appear to host phylogenetically more diverse assemblages than cryoconite and surface snow assemblages. This observation, coupled with phylogenetic evidence that is suggestive of competitive or facilitative interactions among members of the subglacial microbiome, may indicate that the assembly of the community reflects the limited nutrient and energy condition of the subglacial system. We propose that the continual exposure of fresh mineral surfaces due to bedrock comminution by flowing glacial ice provides ample mineral-based energy and nutrients capable of sustaining an active and stable subglacial ecosystem during extended periods of widespread glaciation (Snowball Earth, for example; Hoffman *et al*, 1998; Kirschvink *et al.*, 2000).

Acknowledgements

References

Bhatia M, Sharp M, Foght J. (2006). Distinct bacterial communities exist beneath a high arctic polythermal glacier. *Appl Environ Microbiol* **72**: 5838–5845.

Blomberg SP, Garland Jr T, Ives AR. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* **57**: 717–745.

Boyd ES, Cummings DE, Geesey GG. (2007a). Mineralogy influences structure and diversity of bacterial communities associated with geological substrata in a pristine aquifer. *Microb Ecol* **54**: 170–182.

Boyd ES, Jackson RA, Encarnacion G, Zahn JA, Beard T, Leavitt WD *et al.* (2007b). Isolation, characterization, and ecology of sulfur-respiring Crenarchaea inhabiting acid-sulfate-chloride geothermal springs in Yellowstone National Park. *Appl Environ Microbiol* **73**: 6669–6677.

Boyd ES, Skidmore M, Mitchell AC, Bakermans C, Peters JW. (2010). Methanogenesis in subglacial sediments. *Environ Microbiol Rep* **2**: 685–692.

Boyd ES, Lange RK, Mitchell AC, Havig JR, Hamilton TL, Lafrenière MJ *et al.* (2011). Diversity, abundance, and potential activity of nitrifying and nitrate-reducing microbial assemblages in a subglacial ecosystem. *Appl Environ Microbiol* **77**: 4778–4787.

Bryant JA, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ, Green JL. (2008). Colloquium paper: microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. *Proc Natl Acad Sci USA* **105**: 11505–11511.

Cadotte MW, Cavender-Bares J, Tilman D, Oakley TH. (2009). Using phylogenetic, functional and trait diversity to understand patterns of plant community productivity. *PLoS ONE* **4**: e5695.

Callaway RM. (1995). Positive interactions among plants. *Bot Rev* **61**: 306–349.

Cavender-Bares J, Ackerly DD, Baum DA, Bazzaz FA. (2004). Phylogenetic overdispersion in Floridian oak communities. *Am Nat* **163**: 823–843.

Cavender-Bares J, Kozak KH, Fine PVA, Kembel SW. (2009). The merging of community ecology and phylogenetic biology. *Ecol Lett* **12**: 693–715.

Darriba D, Taboada G, Doallo R, Posada D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat Meth* **9**: 772.

Dowd SE, Wolcott RD, Sun Y, McKeehan T, Smith E, Rhoads D. (2008). Polymicrobial nature of chronic diabetic foot ulcer biofilm infections determined using bacterial tag encoded FLX amplicon pyrosequencing (bTEFAP). *PLoS ONE* **3**: e3326.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194–2200.

Edgcomb VP, Beaudoin D, Gast R, Biddle JF, Teske AT. (2010). Marine subsurface eukaryotes: the fungal majority. *Environ Microbiol* **13**: 172–183.

Elias DA, Krumholz LR, Tanner RS, Suflita JM. (1999). Estimation of methanogen biomass by quantitation of coenzyme M. *Appl Environ Microbiol* **79**: 5541–5545.

Felske A, Wolterink A, Van Lis R, Akkermans ADL. (1998). Phylogeny of the main bacterial 16 S rRNA sequences in Drentse A grassland soils (The Netherlands). *Appl Environ Microbiol* **64**: 871–879.

Fornara DA, Tilman D. (2008). Plant functional composition influences rates of soil carbon and nitrogen accumulation. *J Ecol* **96**: 314–322.

Gaidos E, Marteinsson V, Thorsteinsson T, Jóhannesson T, Rúnarsson ÁR, Stefansson A et al. (2008). An oligarchic microbial assemblage in the anoxic bottom waters of a volcanic subglacial lake. ISME J 3: 486–497.

Guindon S, Gascuel O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 696–704.

Hallam SJ, Putnam N, Preston CM, Detter JC, Rokshar D, Richardson PM et al. (2004). Reverse methanogenesis: testing the hypothesis with environmental genomics. Science 305: 1457–1462.

Hammer O, Harper DAT, Ryan PD. (2001). PAST: paleontological statistics software package for education and data analysis. Palaeontologia Electronica 4: 1–9.

Hirsch PR, Mauchline TH, Clark IM. (2010). Culture-independent molecular techniques for soil microbial ecology. Soil Biol Biochem 42: 878–887.

Hodson A, Anesio AM, Tranter M, Fountain A, Osborn M, Priscu J et al. (2008). Glacial ecosystems. Ecol Monogr 78: 41–67.

Hoffman PF, Kaufman AJ, Halverson GP, Schrag DP. (1998). A Neoproterozoic snowball earth. Science 281: 1342–1346.

Horner-Devine MC, Bohannan BJM. (2006). Phylogenetic clustering and overdispersion in bacterial communities. Ecology 87: S100–S108.

Kastovska K, Stibal M, Sabacka M, Cerna B, Santruckova H, Elster J. (2007). Microbial community structure and ecology of subglacial sediments in two polythermal Svalbard glaciers characterized by epifluorescence microscopy and PLFA. Polar Biol 30: 277–287.

Kembel SW. (2009). Disentangling niche and neutral influences on community assembly: assessing the performance of community phylogenetic structure tests. Ecol Lett 12: 949–960.

Kirschvink JL, Gaidos EJ, Bertani LE, Beukes NJ, Gutzmer J, Maepa LN et al. (2000). Paleoproterozoic snowball Earth: extreme climatic and geochemical global change and its biological consequences. Proc Natl Acad Sci USA 97: 1400–1405.

Knapp S, Kühn I, Schweiger O, Klotz S. (2008). Challenging urban species diversity: contrasting phylogenetic patterns across plant functional groups in Germany. Ecol Lett 11: 1054–1064.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947–2948.

Lortie CJ. (2007). An ecological tardis: the implications of facilitation through evolutionary time. Trends Ecol Evol 22: 627–630.

Losos JB. (1994). Integrative approaches to evolutionary ecology: anolis lizards as model systems. Annu Rev Ecol Syst 25: 467–493.

Lozupone CA, Knight R. (2008). Species divergence and the measurement of microbial diversity. FEMS Microbiol Rev 32: 557–578.

Maherali H, Klironomos JN. (2007). Influence of phylogeny on fungal community assembly and ecosystem functioning. Science 316: 1746–1748.

Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL et al. (2006). Microbial biogeography: putting microorganisms on the map. Nat Rev Microbiol 4: 102–112.

Medlin L, Elwood HJ, Stickel S, Sogin ML. (1988). The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene 71: 491–499.

Meuser JE, Baxter BK, Spear JR, Peters JW, Posewitz MCBoyd ES. (2013). Contrasting patterns of community assembly in the stratified water column of Great Salt Lake, Utah. Microb Ecol; doi:10.1007/s00248-013-0180-9.

Ogram A, Sun W, Brockman FJ, Fredrickson JK. (1995). Isolation and characterization of RNA from low-biomass deep-subsurface sediments. Appl Environ Microbiol 61: 763–768.

Reysenbach A-L, Banta A, Civello S, Daly J, Mitchel K, Lalonde S et al. (2005). Aquificales in Yellowstone National Park. Montana State University: Bozeman, MT, USA.

Ringel MS, Hu HH, Anderson G. (1996). The stability and persistence of mutualisms embedded in community interactions. Theoretical Pop Biol 50: 281–297.

Schippers A, Kock D, Höft C, Köweker G, Siegert M. (2012). Quantification of microbial communities in subsurface marine sediments of the Black Sea and off Namibia. Front Microbiol 3: 16.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75: 7537–7541.

Sharp M, Parkes J, Cragg B, Fairchild IJ, Lamb H, Tranter M. (1999). Widespread bacterial populations at glacier beds and their relationship to rock weathering and carbon cycling. Geology 27: 107–110.

Sharp M, Creaser RA, Skidmore M. (2002). Strontium isotope composition of runoff from a glaciated carbonate terrain. Geochim Cosmochim Acta 66: 595–614.

Skidmore M, Anderson SP, Sharp M, Foght J, Lanoil BD. (2005). Comparison of microbial community compositions of two subglacial environments reveals a possible role for microbes in chemical weathering processes. Appl Environ Microbiol 71: 6986–6997.

Skidmore ML, Foght JM, Sharp MJ. (2000). Microbial life beneath a high arctic glacier. Appl Environ Microbiol 66: 3214–3220.

Swofford DL. (2001). Paup: Phylogenetic Analysis Using Parsimony (and Other Methods), 4.0b10 edn. Sinauer Associate: Sunderland, MA, USA.

Telling J, Anesio AM, Hawkings JR, Tranter M, Wadham JL, Hodson A et al. (2010). Measuring rates of gross photosynthesis and net community production in cryoconite holes: a comparison of field methods. Ann Glaciol 51: 153–162.

Tranter M, Sharp MJ, Lamb HR, Brown GH, Hubbard BP, Willis IC. (2002). Geochemical weathering at the bed of Haut Glacier d'Arolla, Switzerland-a new model. Hydrol Process 16: 959–993.

Valiente-Banuet A, Verdu M. (2007). Facilitation can increase the phylogenetic diversity of plant communities. Ecol Lett 10: 1029–1036.

Wadham JL, Tranter M, Tulaczyk S, Sharp M. (2008). Subglacial methanogenesis: A potential climatic amplifier? Global Biogeochem Cycles 22: GB2021.

Wadham JL, Arndt S, Tulaczyk S, Stibal M, Tranter M, Telling J et al. (2012). Potential methane reservoirs beneath Antarctica. Nature 488: 633–637.

Wang Q, Garrity GM, Tiedje JM, Cole JR. (2007). Naive Bayesian classifier for rapid assignment of rRNA

1412

sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**: 5261–5267.

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ. (2002). Phylogenies and community ecology. *Annu Rev of Ecol Syst* **33**: 475–505.

Webb CO, Ackerly DD, Kembel SW. (2008). Phylocom: software for the analysis of phylogenetic community structure and character evolution. *Bioinformatics* **24**: 2098–2100.

Weekers PHH, Gast RJ, Fuerst PA, Byers TJ. (1994). Sequence variations in small-subunit ribosomal RNAs of *Hartmannella vermiformis* and their phylogenetic implications. *Mol Biol Evol* **11**: 684–690.

Whitman WB, Coleman DC, Wiebe WJ. (1998). Prokaryotes: the unseen majority. *Proc Natl Acad Sci USA* **95**: 6578–6583.

Wiens JJ, Graham CH. (2005). Niche conservatism: integrating evolution, ecology and conservation biology. *Annu Rev Ecol Evol Syst* **36**: 519–539.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)