FEMS Microbiology Ecology, 92, 2016, fiv164

doi: 10.1093/femsec/fiv164
Advance Access Publication Date: 21 December 2015
Research Article

RESEARCH ARTICLE

Supraglacial bacterial community structures vary across the Greenland ice sheet

Karen A. Cameron[1,2,*], Marek Stibal[1,2,3], Jakub D. Zarsky[3], Erkin Gözdereliler[1,2], Morten Schostag[2,4] and Carsten S. Jacobsen[2,5]

[1]Department of Geochemistry, Geological Survey of Denmark and Greenland (GEUS), 1350 Copenhagen, Denmark, [2]Center for Permafrost (CENPERM), University of Copenhagen, 1350 Copenhagen, Denmark, [3]Department of Ecology, Charles University in Prague, 128 43 Praha 2, Czech Republic, [4]Department of Biology, University of Copenhagen, 2100 Copenhagen, Denmark and [5]Department of Environmental Science, Aarhus University, 4000 Roskilde, Denmark

*Corresponding author: Department of Geochemistry, Geological Survey of Denmark and Greenland (GEUS), Øster Voldgade 10, 1350 Copenhagen, Denmark. Tel: +45 91 33 35 94; E-mail: kac.geus@gmail.com

One sentence summary: Bacteria sampled from surface ice and debris across the Greenland ice sheet have noteworthy kilometer-scale variations in their community structures and compositions.

Editor: Rosa Margesin

ABSTRACT

The composition and spatial variability of microbial communities that reside within the extensive (>200 000 km^2) biologically active area encompassing the Greenland ice sheet (GrIS) is hypothesized to be variable. We examined bacterial communities from cryoconite debris and surface ice across the GrIS, using sequence analysis and quantitative PCR of 16S rRNA genes from co-extracted DNA and RNA. Communities were found to differ across the ice sheet, with 82.8% of the total calculated variation attributed to spatial distribution on a scale of tens of kilometers separation. Amplicons related to *Sphingobacteriaceae*, *Pseudanabaenaceae* and WPS-2 accounted for the greatest portion of calculated dissimilarities. The bacterial communities of ice and cryoconite were moderately similar (global $R = 0.360$, $P = 0.002$) and the sampled surface type (ice versus cryoconite) did not contribute heavily towards community dissimilarities (2.3% of total variability calculated). The majority of dissimilarities found between cryoconite 16S rRNA gene amplicons from DNA and RNA was calculated to be the result of changes in three taxa, *Pseudanabaenaceae*, *Sphingobacteriaceae* and WPS-2, which together contributed towards 80.8 ± 12.6% of dissimilarities between samples. Bacterial communities across the GrIS are spatially variable active communities that are likely influenced by localized biological inputs and physicochemical conditions.

Keywords: ice; cryoconite; Greenland; bacteria; diversity; biogeography

INTRODUCTION

The surface of the Greenland ice sheet (GrIS), the largest body of ice in the northern hemisphere, harbors a simple yet unique microbial ecosystem, comprising snow, ice, surface melt waters and surface debris called cryoconite. The biologically active area of the ice sheet covers more than 200 000 km^2 around the margins of the GrIS (Hodson *et al.* 2010), can be up to ~100 km wide, and has tended to increase in area each year within recent history (Hanna *et al.* 2008; Fettweis *et al.* 2011). This area supports microbial communities that are important for biogeochemical processes (Stibal, Sabacká and Zárský 2012a; Telling *et al.* 2012) and surface melt (Takeuchi, Kohshima and Seko 2001; Lutz *et al.*

Received: 19 September 2015; **Accepted:** 16 December 2015

2014), and that are likely important for the injection of biota and nutrients into downstream environments (Stibal, Sabacká and Zárský 2012a).

Several studies have targeted the diversity and activity of surface debris within the biologically active area of the GrIS (e.g. Cameron, Hodson and Osborn 2012a,b; Stibal *et al.* 2012b, 2015a; Telling *et al.* 2012; Musilova *et al.* 2015); however, these studies have so far been limited to a singular small area within the South West of the GrIS that encompasses the Russell and Leverett Glacier catchments. Previous research has shown that the microbial composition and activity within surface snow, ice and cryoconite on the GrIS varies spatially and correlations have been made to the distance from ice-free land as well as the sampling locality (Stibal *et al.* 2012b; Telling *et al.* 2012; Edwards *et al.* 2014; Cameron *et al.* 2015; Stibal *et al.* 2015b). Microbial communities in Greenland cryoconite feature bacteria from the classes Proteobacteria, Cyanobacteria, Bacteroidetes and Actinobacteria (Cameron, Hodson and Osborn 2012b; Edwards *et al.* 2014; Musilova *et al.* 2015; Stibal *et al.* 2015a). Cryoconite communities are biogeochemically active (Stibal, Sabacká and Zárský 2012a; Telling *et al.* 2012); however proportional differences between the active and total components of communities have been found at different sites, based on DNA/RNA co-analysis (Stibal *et al.* 2015a). In contrast, the diversity and community composition of microbes in surface ice is less known, and it is unclear whether both ice and cryoconite supraglacial niches harbor the same microbial community. Two recent reports have started to address surface ice ecosystems, detailing the cell abundance of surface ice (Stibal *et al.* 2015b), and the diversity, activity and temporal stability of marginal 'dirty ice' (Musilova *et al.* 2015).

This study aims to explore biogeographical variations in the diversity and community composition of cryoconite and surface ice microbial communities sampled from several locations across the GrIS (Fig. 1, Table 1). We test the hypotheses that the community composition is uniform across the GrIS and does not differ between cryoconite and surface ice. To do so, we collected surface ice and cryoconite, extracted DNA from all samples and RNA from cryoconite, and performed quantitative PCR and Illumina MiSeq sequencing of 16S rRNA genes from the DNA (termed 16S rDNA herein) and/or RNA (termed 16S rRNA herein) to determine the diversity and community structure. Multivariate statistics were then used to identify significant differences between surface ice and cryoconite communities, and between communities sampled from different regions and sites on the ice sheet.

MATERIALS AND METHODS

Sample collection

A total of 84 surface ice and cryoconite samples were collected from 14 sites on the GrIS and from a neighboring isolated ice cap between May and September 2013. These sites were grouped into six geographically distinct 'regions', each separated by >250 km (mean distance between regions = 1073 ± 481 km). Furthermore, within five of the regions, samples were taken from geographically distinct 'areas' along a linear transect, with >3.5 km between each area (mean distance between neighboring areas = 21.67 ± 27.32 km; Table 1; Fig. 1). These sites have been described previously (Stibal *et al.* 2015a).

Ten surface ice cores (~15 cm long) were extracted at each site using a small handheld drill and custom-built stainless steel corers (~20 cm^2 surface area). The corers were autoclaved and kept sterile in polypropylene bags prior to use. The ice cores were transferred to sterile 750 ml WhirlPak bags (Nasco, WI, USA). All samples were kept frozen in insulated boxes until transportation to Copenhagen where samples were stored at −20°C until analysis.

Cryoconite was collected from 2–18 holes at each site, depending on their availability, using sterile 50 ml syringes. Nitrile gloves were worn at all times to prevent contamination of the samples. Aliquots of the sediment (0.3–0.5 g) were immediately placed into 2 ml Eppendorf tubes filled with 1.2 ml of LifeGuard Soil Preservation Solution (MO BIO Laboratories, Carlsbad, CA, USA) in order to preserve the RNA. Five tubes were filled from each syringe-collected cryoconite sample. LifeGuard preserved samples were frozen to −20°C no later than 2 hours after collection, and were transported frozen to Copenhagen.

Sample analysis

Ice samples from each location were pooled together and placed in detergent, acid- and heat-treated glasswear, according to methods described in Vital *et al.* (2007), and allowed to melt at 4°C. A 300 ml aliquot of the sample was then filtered through Sterivex GP 0.22 μm polyethersulfone filters (Millipore, MA, USA). DNA was extracted from the Sterivex filters using the PowerWater Sterivex DNA Isolation Kit (MO BIO Laboratories), following the manufacturer's protocol. An unused Sterivex filter was extracted alongside the samples as a procedural control.

DNA and RNA were co-extracted from the LifeGuard-preserved cryoconite samples using the RNA PowerSoil Total RNA Isolation Kit with the RNA PowerSoil DNA Elution Accessory Kit (MO BIO Laboratories) according to the manufacturer's instructions. Between 0.79 and 2.42 g (wet weight) of sediment was used for each extraction, and a blank containing 2 mL of DEPC water (MO BIO Laboratories) was extracted in parallel. All RNA extracts were immediately subjected to DNase treatment using the RTS DNase Kit (MO BIO Laboratories) according to the manufacturer's instructions. A subsample of each DNase-treated RNA sample was then used as a template in cDNA production via RT-PCR using the RevertAid Premium RT kit and random hexamer primers (Fermentas, St Leon-Rot, Germany) according to the manufacturer's instructions.

Quantitative PCR of 16S rRNA genes was performed using a CFX96 Touch system (Bio-Rad, CA, USA). Reaction mixtures (20 μl total) consisted of 1 μl of template DNA or cDNA which was diluted 1:100 into molecular grade water (MO BIO Laboratories), 10 μl of SYBR Premix DimerEraser (TaKaRa, Japan) and 0.6 μl of forward and reverse primers (10 pmol μl^{-1}). The primers used were 341F (5′-CCTACGGGAGGCAGCAG-3′) and 518R (5′-ATTACCGCGGCTGCTGG-3′) (Muyzer, De Waal and Uitterlinden 1993). The cycle program was 95°C for 30 s followed by 50 cycles of 95°C for 30 s, 55°C for 30 s and 72°C for 30 s. The reaction was completed by a final 72°C elongation step for 6 min and followed by high-resolution melt curve analysis in 0.5°C increments from 72°C to 98°C. All qPCR reactions were performed in triplicate and were prepared under DNA-free conditions in a pressurized clean lab with a HEPA-filtered air inlet and nightly UV irradiation. Standards of bacterial 16S rRNA genes were prepared by extracting DNA from a serially diluted culture of *Escherichia coli*. The gene copy number of the highest standard was 1.12×10^7 μl^{-1}. The detection limit was 1.7×10^3 gene copies μl^{-1}. Due to the diluted nature of our samples, potential inhibition due to humic acid or other inhibitory compounds was considered unlikely and was not evaluated.

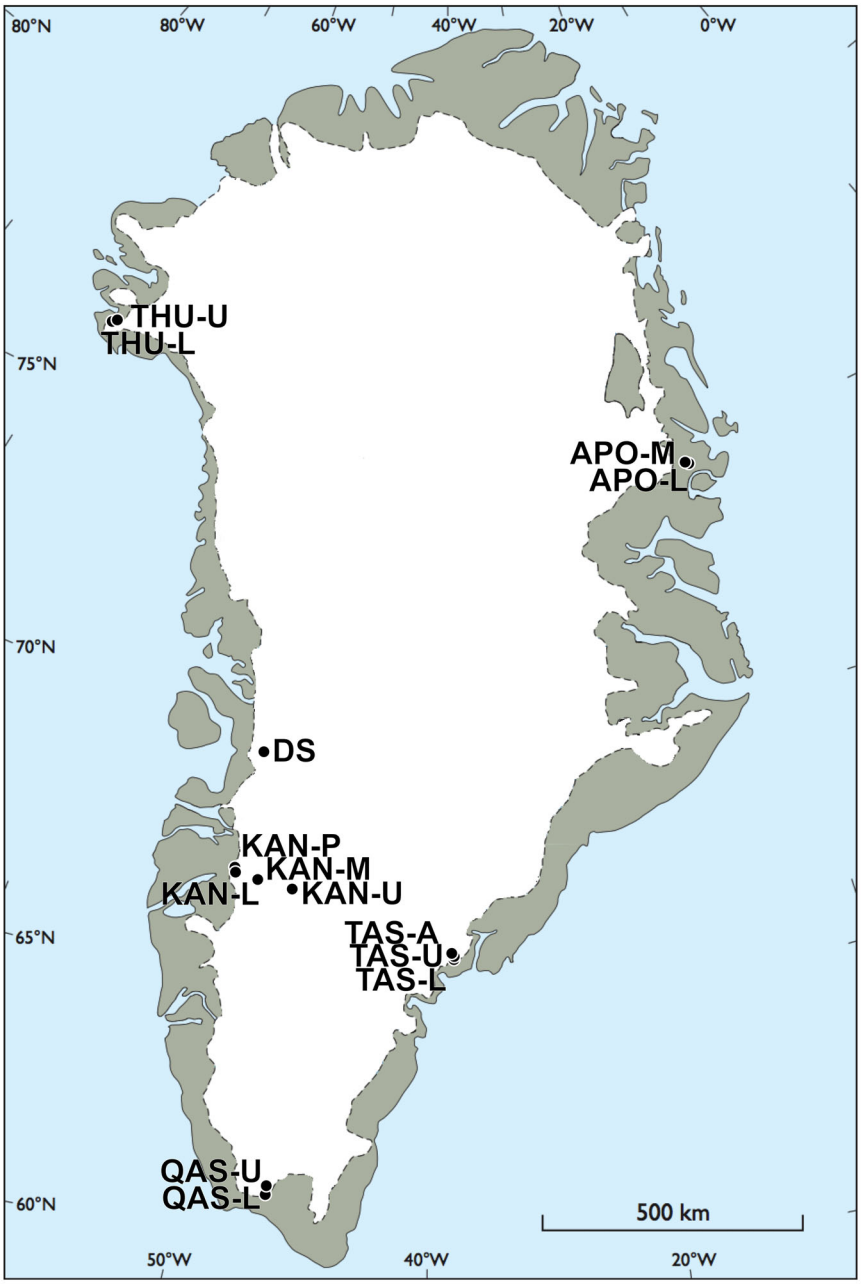

Figure 1. Map of Greenland depicting the 14 sample areas. Figure reproduced from van As *et al.* (2011).

Sequence preparation

DNA and RNA extracts were prepared for paired-end MiSeq sequencing (Illumina, CA, USA) using a two-step PCR method to amplify the V3–V4 hypervariable region of bacterial and archaeal 16S rRNA genes, and to attach MiSeq adapter and barcode sequences. Modified 341F (5′-CCTAYGGGRBGCASCAG-3′) and 806R (5′-GGACTACNNGGGTATCTAAT-3′) primers (Yu *et al.* 2005; Xu *et al.* 2012) tagged with Nextera transposase sequences (Illumina; 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3′ and 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3′, respectively) were used for the first PCR within a PCRBIO HiFi polymerase (PCR BioSystems, UK) master-mix, according to the manufacturer's instructions, using a cycle program of 95°C for 2 min followed by 25 cycles of 95°C for 15 s, 55°C for 15 s and 72°C for 40 s followed by a final 72°C elongation step for 4 min. Products were checked by running 5 μl on a 1.5% agarose gel. Amplicons were cleaned up using HighPrep PCR (MagBio, MD, USA) according to the manufacturer's instructions with the exception of using a 1:1.4 amplicon : bead ratio and resuspending the final elutant into 20 μl of molecular grade water. The elutant (2 μl) was used within a second PCR alongside Phusion High-Fidelity DNA Polymerase (Thermo Scientific, MA, USA) according to the manufacturer's instructions and using barcoding primers from Nextera DNA Library Preparation Kit v2 (Illumina) to obtain unique combinations of P5 and P7 primer barcodes on each sample. The second PCR cycle program was 98°C for

Table 1. Description of samples.

Site name (Area)	Region	Position	Distance from ice free land (km)	Altitude (m)	Date (DOY)	Sample type (number of samples after sequence processing)
APO-L	APO	74°37.471′N, 21°22.507′W	0.5	644	1 May (121)	ice (3)
APO-M	APO	74°38.634′N, 21°28.110′W	0.5	874	1 May (121)	ice (2)
DS	DS	69°28.560′N, 49°34.838′W	18	956	25 Jun (176)	ice (3), cryoconite (2)
KAN-L	KAN	67°05.798′N, 49°56.303′W	5	680	19 Sept (262)	ice (3)[a]
KAN-M	KAN	67°03.964′N, 48°49.356′W	42	1270	19 Sept (262)	ice (3)
KAN-P	KAN	67°09.139′N, 50°00.184′W	2	583	18 Aug (230)	cryoconite (18)
KAN-U	KAN	67°00.014′N, 47°01.162′W	112	1850	22 Sept (265)	ice (3)[a]
QAS-L	QAS	61°01.873′N, 46°50.910′W	1.5	310	20 Aug (232)	ice (3), cryoconite (2)
QAS-U	QAS	61°10.653′N, 46°49.042′W	12	890	20 Aug (232)	ice (3)
TAS-A	TAS	65°46.864′N, 38°54.193′W	10	891	27 Aug (239)	ice (3)
TAS-L	TAS	65°38.460′N, 38°53.895′W	1.5	270	27 Aug (239)	ice (3), cryoconite (18)
TAS-U	TAS	65°41.975′N, 38°51.995′W	5	580	29 Aug (241)	ice (3), cryoconite (4)
THU-L	THU	76°23.991′N, 68°15.921′W	1.5	570	12 Aug (224)	ice (3), cryoconite (2)
THU-U	THU	76°25.181′N, 68°08.706′W	3	770	13 Aug (225)	ice (3)[a]

DOY; day of year 2013.
[a]no sequencing output.

1 min followed by 13 cycles of 98°C for 10 s, 55°C for 20 s and 72°C for 20 s followed by a final 72°C elongation step for 5 min. Amplicons were cleaned up using HighPrep PCR according to the manufacturer's instructions, with the exception of using a 1:0.8 amplicon:bead ratio and resuspending the final elutant into 25 μl of molecular grade water. The cleaned amplicon products were checked by running 5 μl on a 1.5% agarose gel. Quantification of amplicons was performed on a QUBIT fluorometer (Life Technologies, CA, USA) using a QUBIT dsDNA HS assay kit (Life Technologies). Amplicons were normalized to 10 nM and pooled together. Amplicons that were lower than the QUBIT detection limit (<0.5 ng ml^{-1}) were omitted from further analysis.

Negative control PCRs, using molecular grade water instead of DNA or cDNA samples, were performed alongside these reactions. These negative controls followed processing steps until the point of pooling and normalization prior to sequencing. At this point, all negative controls were found to be below the detection limit of QUBIT quantification analysis (<0.50 ng ml^{-1}), and were therefore not included in the normalization step. Of the Sterivex procedural controls that were set up in parallel to the ice DNA extractions, one resulted in a final QUBIT reading of 2.34 ng ml^{-1} and was therefore included in the MiSeq sequencing library.

Sequences were quality filtered and processed using a QIIME processing platform (Caporaso *et al.* 2010b) and using the default quality filters unless otherwise stated. Paired-end sequences were joined and libraries were split using Phred scores of 25 or higher, with a maximum of five consecutive low-quality bases. Operational taxonomic units (OTUs) were defined as sequences that possessed ≥97% identity. OTUs were clustered using a reference based UCLUST algorithm (Edgar 2010) against a Greengenes (GG) reference library (August 2013 release; DeSantis *et al.* 2006) with a minimum cluster size of two. Sequences were aligned using PyNAST (Caporaso *et al.* 2010a). Chimeras were identified using CHIMERASLAYER (Haas *et al.* 2011) against the GG library. Chimeras and singletons were removed from the data set. Taxonomies were assigned from the GG reference library using a BLAST method (Altschul *et al.* 1990). Shared gaps were finally removed before an OTU table was made.

CatchAll (Bunge 2011) was used to calculate parametric alpha diversity. Redundancy analyses were performed in the multivariate data analysis software Canoco 5 (Microcomputer Power, NY, USA). Untransformed Bray–Curtis resemblance, CLUSTER analysis, analysis of similarity (ANOSIM) and contributions of variables to similarity (SIMPER) were calculated from OTU matrices using PRIMER-E version 6 (Plymouth, UK). Metabolic information was obtained from the KEGG database (Kanehisa and Goto 2000). Amplicon data sets are available at The European Bioinformatics Institute under study accession number PRJEB10526. (http://www.ebi.ac.uk).

RESULTS

Abundance analysis

The summarized results of qPCR analyses grouped by area are displayed in Table 2. Cryoconite 16S rDNA copy numbers per wet weight were highly variable between individual samples, regardless of sample area; with a range of almost five orders of magnitude between the highest and lowest copy numbers (1.78×10^{11} copies g^{-1}; sample taken from TAS-L, and 2.36×10^{6} copies g^{-1}; sample taken from KAN-P, respectively, data not shown). The mean cryoconite 16S rDNA copy number was $4.78 \times 10^{10} \pm 5.38 \times 10^{10}$ copies g^{-1}. Cryoconite 16S rRNA gene copy numbers were similarly variable between samples with almost seven orders of magnitude difference between the most and least abundant samples (1.01×10^{13} copies g^{-1}; sample taken from TAS-L, and 2.07×10^{6} copies g^{-1}; sample taken from KAN-P, respectively). The mean cryoconite 16S rRNA gene copy number was an order of magnitude higher than that of 16S rDNA, at $6.91 \times 10^{11} \pm 1.66 \times 10^{12}$ copies g^{-1}. Cryoconite 16S rDNA and rRNA gene copy numbers were found to be significantly different (two sample z-test; $z = 2.41$, $P = 0.016$). Ice 16S rDNA gene copy numbers per milliliter (Stibal *et al.* 2015b) were significantly and several orders of magnitude lower than the gram equivalent in cryoconite (mean copy number in ice; $7.74 \times 10^{6} \pm 9.79 \times 10^{6}$ copies ml^{-1}, two sample z-test; $z = 5.76$, $P < 0.0001$).

Table 2. Mean qPCR 16S rRNA gene copy numbers from DNA and RNA extracts.

Site	Cryoconite 16S rDNA qPCR (copies g^{-1})	Cryoconite 16S rRNA qPCR (copies g^{-1})	Ice 16S rDNA qPCR (copies ml^{-1})[a]
APO-L	–	–	$1.08 \times 10^7 \pm 2.38 \times 10^6$
APO-M	–	–	$2.39 \times 10^7 \pm 7.80 \times 10^6$
DS	$1.62 \times 10^{11} \pm 7.97 \times 10^8$	9.90×10^{11}	$1.97 \times 10^7 \pm 8.74 \times 10^5$
KAN-L	–	–	$2.43 \times 10^4 \pm 1.11 \times 10^4$
KAN-M	–	–	$8.81 \times 10^5 \pm 2.96 \times 10^5$
KAN-P	$7.27 \times 10^9 \pm 9.92 \times 10^9$	$1.22 \times 10^{11} \pm 1.71 \times 10^{11}$	–
KAN-U	–	–	$2.94 \times 10^4 \pm 1.35 \times 10^4$
QAS-L	–	–	$2.63 \times 10^7 \pm 1.17 \times 10^7$
QAS-U	–	–	$2.19 \times 10^6 \pm 6.25 \times 10^5$
TAS–A	–	–	$1.73 \times 10^6 \pm 2.24 \times 10^5$
TAS-L	$7.37 \times 10^{10} \pm 5.37 \times 10^{10}$	$1.28 \times 10^{12} \pm 2.33 \times 10^{12}$	$1.37 \times 10^7 \pm 5.30 \times 10^5$
TAS-U	$2.40 \times 10^{10} \pm 7.40 \times 10^9$	$4.97 \times 10^{10} \pm 1.13 \times 10^{10}$	$6.26 \times 10^6 \pm 4.32 \times 10^6$
THU-L	$7.24 \times 10^{10} \pm 7.19 \times 10^9$	8.76×10^{11}	$4.53 \times 10^5 \pm 1.28 \times 10^5$
THU-U	–	–	$4.58 \times 10^4 \pm 3.89 \times 10^4$

[a]From Stibal *et al.* (2015b).

Amplicon sequencing read output

After sequence processing, the total number of reads across all 118 16S rDNA and rRNA samples was 466 012, with an mean number of reads per samples at 3949 ± 3722, and a maximum number of reads of 15 219. Sequences were rarefied to 752 sequences per sample, which resulted in the omission of two DS 16S rRNA samples, two THU-L 16S rRNA samples, a TAS-L 16S rRNA sample and a TAS-U 16S rRNA sample. After rarefaction the mean OTU count of all samples was 35.62 ± 13.38, ranging from 16.33 ± 3.06 (TAS-U cryoconite16S rRNA) to 56.00 ± 3.61 (QAS-U ice 16S rDNA; Table S1, Supporting Information). Mean percentage coverage of the rarefied libraries, calculated against CatchAll calculated species richness of the rarefied libraries was 56.88 ± 9.70% (Table S1, Supporting Information).

Alpha diversity

Mean CatchAll calculated alpha diversity of amplicons generated from ice and cryoconite DNA extracts, grouped by area, ranged from 31.45 ± 6.16 (THU-L cryoconite) to 116.35 ± 58.05 (DS cryoconite; Table S1, Supporting Information). Mean CatchAll calculated alpha diversity of amplicons generated from cryoconite 16S rDNA and rRNA was calculated to be 72.92 ± 32.80, and 42.65 ± 19.15, respectively. Differences between CatchAll calculated diversities from cryoconite 16S rDNA and rRNA extracts were found to be statistically different when a two-tailed paired Z-test was performed ($z = 5.25$, $P < 0.0001$). Ice 16S rDNA generated amplicons had a CatchAll calculated alpha diversity of 81.54 ± 26.61. Differences between CatchAll calculated diversities from ice and cryoconite 16S rDNA were not found to be statistically different when a two-tailed paired Z-test was performed ($z = -1.248$, $P = 0.212$).

Factors contributing to community variability

To identify significant factors affecting bacterial community structure on the GrIS, a redundancy analysis (RDA) was performed with sample type (ice versus cryoconite) and site characteristics (region, area, position along the N-S transect expressed as the N coordinate, altitude, distance from the margin; Table 1) as the explanatory variables. This analysis explained 81.6% of the total variation in the data. Area was the most significant variable, contributing 82.8% to the explained variation (*pseudoF* = 32.4; $P = 0.001$), followed by sample type (2.3%, *pseudoF* = 8, $P = 0.001$; Fig. 2). Position along the N-S transect, altitude and distance from the margin were all found to be covariables of area. To remove any potential bias due to the different times of sampling at each locality, a partial RDA was conducted with sampling day as a covariate. This analysis explained 68.5% of the total variation in the data. Region contributed 69.4% (*pseudoF* = 35.9; $P = 0.001$) and area 17.8% (*pseudoF* = 10.7; $P = 0.001$) to the explained variation.

Spatial variability of communities

To test for dissimilarities between samples collected from different areas, a global two-way crossed ANOSIM analysis of Bray–Curtis similarities, grouped by area and accounting for differences in surface type, was performed. This analysis suggested that there are large and significant differences between groups (global $R = 0.926$, $P = 0.001$). However, due to small sample numbers within each group, only 14 of the 45 pairwise analyses were found to be statistically significant at the $P = 0.05$ level, and of these only 11 at the $P = 0.01$ level. A global two-way crossed ANOSIM analysis performed with region groupings and accounting for differences in surface type suggested that samples from different regions are also dissimilar (global $R = 0.874$, $P = 0.001$). Pairwise two-way crossed ANOSIM results from this analysis are reported in Table 3. Of the pairwise correlations that were found to be statistically significant at the $P = 0.05$ level, all were calculated to be moderately to strongly dissimilar ($R = 0.568 \geq 1$; Table 3). When only cryoconite samples were grouped by region they were found to be strongly dissimilar (global $R = 0.906$, $P = 0.001$), calculated using a one-way ANOSIM analysis, while cryoconite from TAS-L and TAS-U areas were found to be only moderately dissimilar ($R = 0.792$, $P = 0.001$) when calculated using a pairwise ANOSIM test. Ice samples grouped by region were found to be moderately dissimilar (global $R = 0.658$, $P = 0.001$), calculated using a one-way ANOSIM analysis, however statistically significant ANOSIM analyses of ice from different areas within the same region (i.e. APO-L and APO-M, QAS-L and QAS-U, and TAS-L and TAS-U) could not be calculated due to small sample sizes.

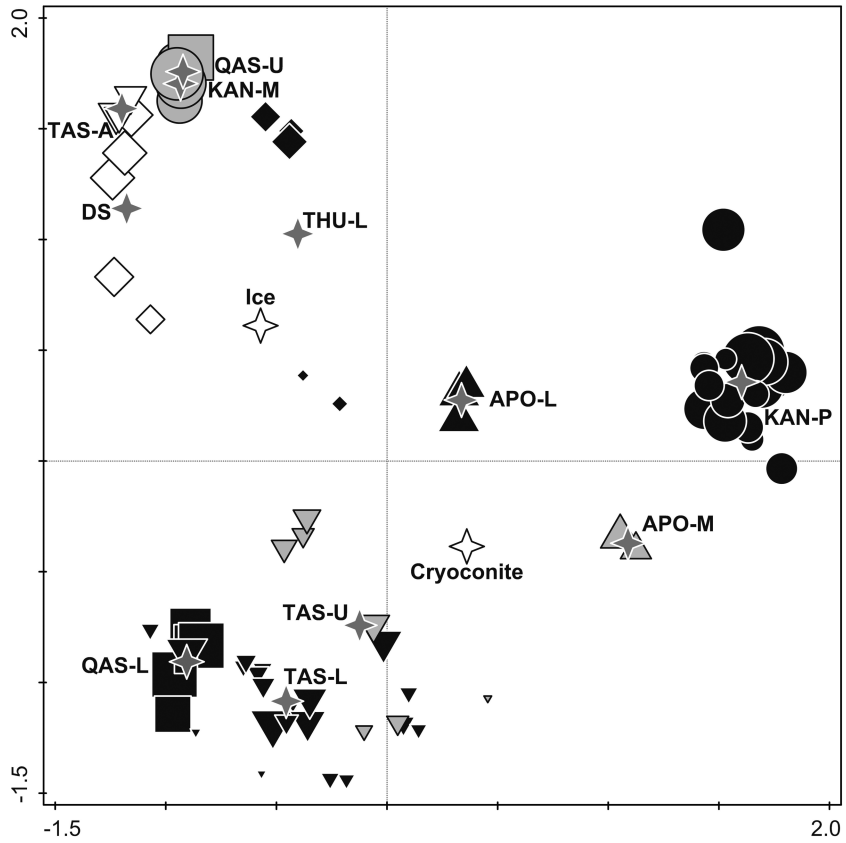

Figure 2. RDA biplot describing the variation of species diversity in relation to area (grey star symbol) and sample type (white star symbol) variables when samples were rarefied and randomly resampled to the lowest number of sequences per sample (752 reads) and average values from resamplings were used for further analysis. The size of symbols is proportional to sample richness. Symbol shapes denote the sample region while the color differentiates between different sampling areas; black circles: KAN-P, grey circles: KAN-M, black square: QAS-L, grey square: QAS-U, black down triangle: TAS-L, grey down triangle: TAS-U, white down triangle: TAS-A, white diamond: DS, black diamond: THU-L, black up triangle: APO-L, grey up triangle: APO-M.

Table 3. Two-way crossed ANOSIM analysis of Bray–Curtis similarities using both surface type and region as factors, and displaying the R values and (p values) for region grouped analyses.

Region	APO	DS	KAN	QAS	TAS
DS	1 (0.018)[a]				
KAN	1 (0.018)[a]	1 (0.001)			
QAS	0.568 (0.002)	0.077 (0.325)[b]	0.886 (0.002)		
TAS	0.844 (0.002)	0.648 (0.001)	0.988 (0.001)	0.103 (0.251)[b]	
THU	1 (0.018)[a]	1 (0.033)[a]	1 (0.002)	0.635 (0.016)[a]	0.947 (0.001)

[a]Not significant at the 0.01 level.
[b]Not significant at the 0.05 level.

Sample type variability of communities

To test for dissimilarity between communities grouped by surface types, accounting for differences that occur between regions, a two-way crossed ANOSIM analysis of Bray–Curtis similarities was performed, and revealed that the bacterial communities of ice and cryoconite were moderately similar (global $R = 0.360$, $P = 0.002$). Statistically significant comparisons of pairwise cryoconite and ice samples from the same region or area could not be calculated using ANOSIM analysis due to the small sample sizes.

OTU contributions towards area grouped community structures

OTUs with the greatest contribution towards the uniqueness of area-grouped samples, separated by sample type, were tested by one-way SIMPER analysis, using a Bray–Curtis dissimilarity matrix. Bray–Curtis dissimilarity of area grouped cryoconite samples averaged at 0.79 ± 0.08 (Table 4). OTUs most closely related to *Sphingobacteriaceae* contributed heavily towards the composition of all samples of cryoconite grouped by area, with the exception of KAN-P (mean group contribution excluding KAN-P;

Table 4. Heat map of one-way SIMPER analysis generated from cryoconite and ice OTU contributions that were found to be >1% in at least one sample area group. Amplicons related to family (f) level, or class (c) or phyla (p) level where necessary, are combined together, with number of OTUs involved shown in brackets. Mean calculated within-area and within-sample type Bray–Curtis similarity is shown in brackets within the column headers. Shading is on a scale of light-dark; 0%–100% contribution.

	Cryoconite (0.70)						Ice (0.70)									
	DS (0.85)	KAN P (0.73)	QAS L (0.91)	TAS L (0.74)	TAS U (0.72)	THU L (0.82)	APO L (0.87)	APO M (0.88)	DS (0.85)	KAN M (0.88)	QAS L (0.90)	QAS U (0.86)	TAS A (0.89)	TAS L (0.92)	TAS U (0.91)	THU L (0.84)
f: *Acetobacteraceae* (2)	0.78	0.22	0.44	0.31	0.09	0.49	2.56	0.61	3.14	5.86	0.93	8.48	1.95	0.73	1.37	17.57
f: *Acidobacteriaceae* (3)	7.06	0.17	0.88	0.91	1.2	1.3	1.28	0	14.41	5.3	1.92	1.23	11.97	1.98	5.33	5.71
f: *ACK-M1* (1)	0.63	11.54	0.15	0.47	1.88	0	0.05	0.76	0.63	5.06	0	0.36	0	0.05	0.15	0.05
f: *Chitinophagaceae* (1)	0	0.35	0.44	0	0	0	1.07	1.06	0	0.66	0.93	0.98	0	0	0	0
f: *Clostridiaceae* (2)	0	0	2.04	0.1	0.03	0	0	0	0.1	0.2	0.84	1.9	0.05	0.39	0.15	0
f: *Comamonadaceae* (2)	0	0	0	0	0	0	0.31	0	0	0.05	0	0	0	0	0	0
f: *Cytophagaceae* (8)	1.89	1.11	0.74	0.36	0.15	2.1	13	2.12	7.65	3.39	9.39	18.44	13.78	3.24	5.23	24.28
f: *Deinococcaceae* (1)	0	0.37	0.15	0.1	0.12	0.32	1.38	0.91	0	0	0.34	0	0	0.19	0	0.74
f: *Microbacteriaceae* (4)	8.01	0.79	0.87	1.47	2.34	0.97	4.09	1.66	15.78	32.56	1.47	39.92	22.56	2.17	3.86	7.84
f: *Nostocaceae* (2)	0	0.58	0	0	0	3.72	9.62	3.48	0	0	0	0	0	0	0	0.95
f: *Pseudanabaenaceae* (3)	0.16	62.63	0.73	10.91	32.08	48.06	26.97	70.95	0	0	1.57	0	0	11.41	26.73	26.08
f: *Solibacteraceae* (1)	0	0.94	1.02	0.57	0.12	0	0.26	0.45	0	0.05	0.49	0	0	0	0.05	0.05
f: *Sphingobacteriaceae* (2)	73.31	12.31	58.48	57.69	28.37	42.07	28.1	7.41	51.57	37.01	33.92	15.11	40.42	50.44	44.53	12.59
f: *Sphingomonadaceae* (3)	0.63	1.1	0.3	0.06	0	0	0.56	0.15	1.15	1.67	0.39	2.73	3.75	0.24	0.15	0
f: *Sporichthyaceae* (1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
f: *Thermogemmatisporaceae* (1)	0	2.62	0	2.85	0	0.49	2	0.61	0	0	0	0	0	2.47	0	0.32
f: *Xanthomonadaceae* (1)	0	0	0	0	0	0	0	0	0.1	0	0.1	0.31	1	0	0	0
c: Deltaproteobatceria (1)	0	0	0	0	0	0	0.05	0	0	0	0	0	0	0	0	0
p: TM7 (1)	6.12	0.06	0	0	0.03	0	0	0	1.52	1.87	0	1.64	0.35	0	0	0
p: WPS-2 (4)	0.16	0.81	31.43	23.32	32.29	0.16	3.02	6.5	0	1.47	44.19	2.52	0	26.07	11.59	0.63

51.98 ± 17.22%), with Greengene OTU identity (GG OTU id) 825696 being the sole contributing OTU at each site. At KAN-P OTUs most closely related to *Pseudanabaenaceae* predominated in contributing towards group identity (62.63%), due to the presence of GG OTU id 174204 (*Leptolyngbya*; 35.42%) and GG OTU id 241071 (*Pseudanabaena*; 23.23%), while at THU-L 48.06% of the group identity was contributed to by GG OTU id 3274 (*Leptolyngbya*). At QAS-L, TAS-L and TAS-U, GG OTU id 831263, which is most closely related to the candidate phylum WPS-2, contributed on average 29.01 ± 4.95% towards group identity.

Bray–Curtis similarity of area grouped ice samples averaged at 0.88 ± 0.03 (Table 4). Similar to cryoconite samples, OTUs most closely related to *Sphingobacteriaceae* and *Pseudanabaenaceae* featured in their contribution towards group identity (32.11 ± 15.85% and 16.73 ± 22.63% mean contribution, respectively), and WPS-2 related OTUs contributed to group identity at QAS-L, TAS-L and TAS-U (44.19%, 26.07% and 11.59%, respectively; Table 4). Amplicons from only one *Sphingobacteriaceae*-related OTU (GG OTU id 825696) were generated from all areas, with the exception of the TAS-L group which lacked this OTU and instead had 50.44% contribution to group identity from another *Sphingobacteriaceae*-related OTU (GG OTU id 836522). In addition, four OTUs related to *Microbacteriaceae* and seven OTUs related to *Cytophagaceae* featured as contributing factors to the group identity (13.19 ± 14.04% and 10.05 ± 7.29% mean contribution, respectively), while *Acetobacteraceae* contributed towards 17.57% of the THU-L community identity, and *Acidobacteriaceae* was responsible for 14.41% and 11.97% of community identity at DS and TAS-A, respectively.

Comparisons of cryoconite 16S rDNA and rRNA amplicon libraries

To identify the significant factors affecting active and total community structures of cryoconite sampled from the GrIS, an RDA was performed with molecule type (DNA versus RNA) and site characteristics (region, area, position along the N-S transect expressed as the N coordinate, altitude, distance from the margin; Table 1) as the explanatory variables. This analysis explained 74.2% of the total variation in the data. Region was still the most significant variable, contributing 49.9% to the explained variation ($pseudoF = 0.001$; $P = 0.003$), followed by molecule type (44.8% explained, $pseudoF = 0.001$, $P = 0.003$). However, when a partial RDA was conducted with sampling day as a covariate, in order to remove a potential bias due to the different times of sampling at each locality, 60.1% of the total variation was explained, with molecule type contributing 69.8% of the explained variation ($pseudoF = 0.001$; $P = 0.003$) and region only explaining 25.2% ($pseudoF = 0.001$; $P = 0.002$).

A two-way ANOSIM analysis of 16S rDNA and rRNA amplicon libraries from all areas showed that these communities are statistically dissimilar (Global R = 0.867, $P = 0.001$). To investigate OTUs responsible for the differences that exist between amplicon communities that were generated from DNA and RNA extracts, two-way SIMPER analyses, using Bray–Curtis dissimilarity matrices of DNA- and RNA-generated amplicons were performed using area groupings and assuming that differences may exist as a result of localized meter-scale differences in community composition. Cryoconite from DS and THU-U were excluded from the analysis as 16S rRNA amplicon libraries were below the 752 sequences per sample rarefaction cut-off point selected. The mean Bray–Curtis dissimilarity between all 16S rDNA and rRNA samples was calculated to be moderate at 0.56. The results of these area grouped SIMPER analyses of dissimilarities are summarized in Table 5. The majority of differences found between 16S rDNA and rRNA amplicon libraries were the result of the changing abundance of three taxa; *Pseudanabaenaceae*, *Sphingobacteriaceae* and WPS-2 (80.8 ± 12.6% mean combined contribution; Table 5). *Pseudanabaenaceae* had a strong presence within the 16S rDNA amplicon libraries of KAN-P (57.34%) and TAS-U (33.70%) grouped communities, with amplicons most closely related to GG OTU id 174204 and 241071 predominating at KAN-P and 3274 and 174204 predominating at TAS-U (data not shown). These OTUs were found to have a greater relative abundance within the 16S rRNA amplicon libraries (87.47% and 65.43%, respectively) suggesting that *Pseudanabaenaceae* are active within these communities. In contrast, QAS-L and TAS-L 16S rRNA amplicon libraries were found to have lower relative abundances of *Pseudanabaenaceae* (0.20% and 13.44%, respectively) and were instead dominated by WPS-2 related OTUs (94.68% and 80.36%, respectively), predominately GG OTU id 831263 (data not shown), which had a lesser presence within KAN-P and TAS-U grouped 16S rRNA amplicon libraries (3.33% and 30.77%, respectively). *Sphingobacteriaceae*, which had a noticeable abundance in 16S rDNA amplicon libraries from all area groups analyzed (mean abundance: 11.49%–53.89%), were markedly less abundant within the 16S rRNA libraries (mean abundance: 0.29%–2.21%), suggesting that they are predominantly inactive within cryoconite holes. Similarly, *ACK-M1* had a mean abundance of 10.81% within KAN-P grouped 16S rDNA amplicon libraries; however, this family only constituted 0.21% of the KAN-P grouped 16S rRNA amplicon libraries.

Procedural control analysis

A single procedural control Sterivex sample was found to contain amplifiable DNA. Once sequenced it was revealed that 73.1% of the amplicon reads clustering to a single *Herbaspirillum* OTU (GG OTU id 181996). Triplicate ice samples from KAN-L, KAN-U and THU-U, all of which had been problematic to amplify and were already known to be in low abundance (Stibal *et al.* 2015b), similarly had a high abundance of this OTU (68.7 ± 24.8%) and were found to cluster with up to 82% similarity to the procedural control using a Bray–Curtis similarity matrix. These nine samples were therefore excluded from further analysis. In addition, the two most abundant OTUs from the procedural control extraction (GG OTU ids 181996 and 127549), together constituting 81.0% of the total abundance within these samples, were removed from all sample sequence profiles.

DISCUSSION

Bacterial communities sampled from surface ice within the biologically active zone around the GrIS and grouped by area were found to be markedly variable, with >80% of the variation found being explainable by the area (10 km scale) from which samples were taken. When communities were grouped by region (100 km scale) they were found to be less dissimilar, suggesting that dissimilarities between communities exist on a localized 'area' scale. Previous biogeographical studies of cryoconite reported communities to be distinct when sampled between landmasses (Mueller and Pollard 2004; Cameron, Hodson and Osborn 2012b; Edwards *et al.* 2014) as well as on more localized scales (Edwards *et al.* 2011, 2013; Cameron, Hodson and Osborn 2012b; Grzesiak *et al.* 2015; Stibal *et al.* 2015a). This suggests that in addition to environmental factors associated with regional differences, such as the region's climate and daylight conditions, more localized environmental variables such as altitude, distance from

Table 5. Heat map of two-way SIMPER analysis of OTUs generated from cryoconite 16S rDNA and rRNA and grouped by area, that were found to have >1% contribution towards the dissimilarity between groups in at least one group. Amplicons related to family (f) level, or order (o) or phyla (p) level where necessary, taxonomies are combined together, with number of OTUs involved shown in brackets. Mean calculated within-area Bray–Curtis similarity is shown in brackets within the column headers. 16S rDNA and rRNA mean abundance in percentage (m.Ab.) and percentage contribution to dissimilarities (Contrib%) of each taxa are shown for each area. Shading is on a percentage scale of light-dark; 0%–100%.

	KAN P (0.37)			QAS L (0.65)			TAS L (0.64)			TAS U (0.38)		
	DNA m.Ab.	cDNA m.Ab.	Contrib%	DNA m.Ab.	cDNA m.Ab.	Contrib%	DNA m.Ab.	cDNA m.Ab.	Contrib%	DNA m.Ab.	cDNA m.Ab.	Contrib%
f: *Acidobacteriaceae* (1)	0.1	0	0.15	0.2	0.07	0.1	0.64	0.05	0.44	1	0	1.39
f: *ACK-M1* (1)	10.81	0.21	14.18	0.13	0.07	0.05	0.9	0.02	0.75	3.26	0	3.06
f: *Sphingobacteriaceae* (1)	11.49	0.29	15.78	53.86	1.66	40.05	50.49	2.21	37.78	22.17	0.27	30.31
f: *Thermogemmatisporaceae* (1)	2.82	0.31	3.67	0	0	0	3.07	0.29	2.19	0.03	0	0.06
f: *Nostocaceae* (1)	1.34	0.8	1.13	0	0	0	0.01	0	0.01	0	0	0
f: *Clostridiaceae* (1)	0.01	0	0.01	2.26	0.13	1.63	0.23	0	0.18	0	0	0
f: *Acetobacteraceae* (1)	0.18	0.63	0.61	0.13	0.2	0.05	0.09	0.52	0.36	0.37	1.24	1.04
f: *Sphingomonadaceae* (1)	1.14	0.14	1.14	0.27	0.07	0.15	0.13	0.04	0.09	0	0	0
f: *Microbacteriaceae* (2)	0.15	0.08	0.28	0.67	0	0.51	1.58	0.22	1.11	2.86	0	3.29
f: *Pseudanabaenaceae* (4)	57.34	87.47	43.58	0.73	0.2	0.51	13.51	13.44	4.38	33.7	65.43	37.18
o: Ellin6067 (1)	1.3	0.04	1.6	0	0	0	0	0	0	0	0	0
p: WPS-2 (3)	1.03	3.33	3.18	31.38	94.68	49.18	21.86	80.36	46.36	30.49	30.77	14.9

the ice margin, periods of snow free exposure, physical topography and neighboring ecological niches are important factors in shaping the microbial communities that reside on ice surfaces. Indeed, RDA of the GrIS surface communities within this study identified N coordinate, altitude and distance from the margin as significant variables, which further suggests that localized factors have an influence on the establishment of surface ice communities.

Surface type was calculated to account for only ~2% of the variation found between samples, and ice and cryoconite communities that were grouped by region were found to be moderately similar. In addition, no statistically significant differences were found between the alpha diversities of ice and cryoconite 16S rDNA amplicon communities, despite several orders higher abundance of copy numbers within cryoconite in comparison to ice. These results resonate with the findings of a similar analysis performed on Leverett Glacier (SW Greenland, 10 km from KAN-P) where no significant differences were found between cryoconite and marginal surface ice samples (Musilova *et al.* 2015). Based on these results, we consider that there are three potential pathways for the replenishment and/or maintenance of supraglacial communities; first, that local or long-range transported biota, likely distributed through the atmosphere, may simultaneously inoculate both surface ice and cryoconite niches, thereafter minor community shifts occur in response to the differing physico-chemical conditions. Several studies have suggested that localized sources, transported in the atmosphere, are responsible for seeding supraglacial environments (e.g. Edwards *et al.* 2013; Cameron *et al.* 2015; Musilova *et al.* 2015; Stibal *et al.* 2015a); however, more focused and detailed studies are essential to fully substantiate and quantify this suggestion. Top BLAST hits of cultured organisms that were most closely related to the most abundant OTUs within this current study often similarly originated from polar environments such as glacial ice (i.e. GG OTU id 825696; Zeng *et al.* 2013, 4310803; Segawa *et al.* 2011, 279325; García-Echauri *et al.* 2011 and 827674; Klassen and Foght 2011), snow (i.e. GG OTU IDS 174204, 3274 and 760221; Harding *et al.* 2011), glacial forefields (i.e. GG OTU IDs 4310803; Bajerski and Wagner 2013, and 279325; Zdanowski *et al.* 2013), sea water (i.e. GG OTU id 4310803; Lo Giudice *et al.* 2010), microbial mats (GG OTU IDs 828231; Peeters, Ertz and Willems 2011 and 760221; Kleinteich *et al.* 2012), clouds (i.e. GG OTU id 828231; Vaïtilingom *et al.* 2012) and Antarctic lakes (i.e. GG OTU IDS 174204, 3274 and 760221; Taton *et al.* 2006, and 241071; Nadeau, Milbrandt and Castenholz 2001). This trend is suggestive of the endemic nature of communities from polar environments, and additionally it is further suggestive of the dispersal of neighboring communities onto the supraglacial environment. The second potential pathway involves the homogenization of surface ice and cryoconite communities, where surface ice communities are transported through hydrological flow-paths into surface ice depressions, which harbor cryoconite material, and where cryoconite is washed onto glacial surfaces as a consequence of surface melt. Third, the surface ice itself may act as a self-contained reservoir and incubator of biota; constituting a frozen storage unit for microorganisms that become locked into the ice during colder periods, and a chemically and physically suitable environment for the growth of these same stored biota during warmer periods. Given the dynamic nature of glacial surfaces (Irvine-Fynn *et al.* 2011), their seasonal physical variability, and the similarities found between ice and cryoconite amplicon communities, it is likely that all three scenarios occur to some degree in unison.

Of the SIMPER calculated identities of area-grouped cryoconite and surface ice samples, OTUs with the largest contributions towards within group identity were most closely related to *Sphingobacteriaceae*, *Pseudanabaenaceae*, including *Leptolyngbya* and *Pseudanabaena*, and WPS-2. It is interesting to note the vastly changing relative abundance of photosynthetic *Pseudanabaenaceae* between all 16S rDNA and rRNA amplicon communities, which ranged from 0% to 87%. Within areas where these primary producers were absent or low in abundance, it is likely that an alternative photosynthetic community of eukaryotic organisms are responsible for pulling bioavailable carbon into the system (Cameron, Hodson and Osborn 2012b; Lutz *et al.* 2014). In contrast, the low relative abundance of *Sphingobacteriaceae* related OTUs within the 16S rRNA amplicon libraries (mean relative abundance: 0.29%–2.21%) in comparison to their notable presence within 16S rDNA libraries (mean relative abundance: 11.49%–53.89%) is suggestive of their inactivity within cryoconite holes at the time of sampling. Interestingly, the mean total abundance of 16S rRNA copies per wet weight was significantly different and an order of magnitude higher than those of 16S rDNA, which suggests that a large proportion of the community was active at the time of sampling (Stibal *et al.* 2015a). However, in contrast to studies within the KAN region by Stibal *et al.* (2015b), comparisons of 16S rDNA : rRNA amplicon ratios between transect sites TAS-U and TAS-L, revealed a higher proportion of active cells at the more marginal TAS-L sample area, which is once again suggestive of the biogeographical variability across the GrIS.

The identification of contaminants within three triplicate ice samples during this study stands as a reminder of the care and awareness that is required during the processing and analysis of challenging low biomass samples. The presence of OTUs most closely related to *Herbaspirillum*, which was found in abundance within some of the samples, has previously been noted within several studies of laboratory contamination issues, including those from commercial kits (Salter *et al.* 2014). While *Herbaspirillum* is among several taxa that have previously been sequenced and cultured from glacial environments (e.g. Schütte *et al.* 2010; Bajerski *et al.* 2013; Shtarkman *et al.* 2013) the presence of *Herbaspirillum* related OTUs within a control and samples within this current study is too prevalent to ignore.

In conclusion, we do not find a core, uniform bacterial community within samples taken across the biologically active area of the GrIS. Instead, we surmise that noteworthy kilometer-scale variations between the microbial compositions of sampled communities are the result of localized biological, chemical or physical influencing factors. The communities found within surface debris and on surface ice can be treated as a whole, with few differences existing between them. Differences between the total and potentially active community compositions of these supraglacial environments were found to be spatially variable. It is therefore of paramount importance to address spatial variability the when overviewing or up-scaling the diversity and functionality of biota within the biologically active area of the ice sheet.

SUPPLEMENTARY DATA

Supplementary data are available at FEMSEC online.

ACKNOWLEDGEMENTS

We thank Pernille Stockmarr and Karin Vestberg for technical assistance, Jason Box, Michele Citterio and Martin Veicherts for field assistance and Tue Kjærgaard Nielsen for assistance with primer selection. Figure 1 was reproduced and amended with kind permission from Dirk van As.

FUNDING

This research was supported by a Villum Foundation Young Investigator Programme grant [grant number VKR 023121 to MS] and sequencing costs were additionally co-funded by a Danish Geocenter Grant [grant number 5298507 to CSJ]. In addition, financial support was received from the Danish National Research Foundation [grant number CENPERM DNRF100].

Conflict of interest. None declared.

REFERENCES

Altschul SF, Gish W, Miller W *et al.* Basic local alignment search tool. *J Mol Biol* 1990;**215**:403–10.

Bajerski F, Ganzert L, Mangelsdorf K *et al.* *Herbaspirillum psychrotolerans* sp. nov., a member of the family Oxalobacteraceae from a glacier forefield. *Int J Syst Evol Micr* 2013;**63**:3197–203.

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

Bunge J. Estimating the number of species with CatchAll. *Biocomputing* 2011:121–30.

Cameron KA, Hagedorn B, Dieser M *et al.* Diversity and potential sources of microbiota associated with snow on western portions of the Greenland Ice Sheet. *Environ Microbiol* 2015;**17**:594–609.

Cameron KA, Hodson AJ, Osborn AM. Carbon and nitrogen biogeochemical cycling potentials of supraglacial cryoconite communities. *Polar Biol* 2012a;**35**:1375–93.

Cameron KA, Hodson AJ, Osborn AM. Structure and diversity of bacterial, eukaryotic and archaeal communities in glacial cryoconite holes from the Arctic and the Antarctic. *FEMS Microbiol Ecol* 2012b;**82**:254–67.

Caporaso JG, Bittinger K, Bushman FD *et al.* PyNAST: A flexible tool for aligning sequences to a template alignment. *Bioinformatics* 2010a;**26**:266–7.

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

DeSantis TZ, Hugenholtz P, Larsen N *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with arb. *Appl Environ Microb* 2006;**72**:5069–72.

Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;**26**:2460–1.

Edwards A, Anesio AM, Rassner SM *et al.* Possible interactions between bacterial diversity, microbial activity and supraglacial hydrology of cryoconite holes in Svalbard. *ISME J* 2011;**5**:150–60.

Edwards A, Mur LA, Girdwood SE *et al.* Coupled cryoconite ecosystem structure-function relationships are revealed by comparing bacterial communities in alpine and Arctic glaciers. *FEMS Microbiol Ecol* 2014;**89**:222–37.

Edwards A, Rassner SM, Anesio AM *et al.* Contrasts between the cryoconite and ice-marginal bacterial communities of Svalbard glaciers. *Polar Res* 2013;**32**:19468.

Fettweis X, Tedesco M, van den Broeke M *et al.* Melting trends over the Greenland ice sheet (1958–2009) from spaceborne microwave data and regional climate models. *The Cryosphere* 2011;**5**:359–75.

García-Echauri SA, Gidekel M, Gutiérrez-Moraga A *et al.* Isolation and phylogenetic classification of culturable psychrophilic prokaryotes from the Collins glacier in the Antarctica. *Folia Microbiol* 2011;**56**:209–14.

Grzesiak J, Górniak D, Świątecki A *et al.* Microbial community development on the surface of Hans and Werenskiold Glaciers (Svalbard, Arctic): a comparison. *Extremophiles* 2015;**5**:885–97.

Hanna E, Huybrechts P, Steffen K *et al.* Increased runoff from melt from the Greenland ice sheet: a response to global warming. *J Clim* 2008;**21**:331–41.

Harding T, Jungblut AD, Lovejoy C *et al.* Microbes in high Arctic snow and implications for the cold biosphere. *Appl Environ Microb* 2011;**77**:3234–43.

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

Hodson A, Bøggild C, Hanna E *et al.* The cryoconite ecosystem on the Greenland ice sheet. *Ann Glaciol* 2010;**51**:123–9.

Irvine-Fynn TDL, Hodson AJ, Moorman BJ *et al.* Polythermal glacier hydrology: a review. *Rev Geophys* 2011;**49**:RG4002.

Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000;**28**:27–30.

Klassen J, Foght J. Characterization of *Hymenobacter* isolates from Victoria Upper Glacier, Antarctica reveals five new species and substantial non-vertical evolution within this genus. *Extremophiles* 2011;**15**:45–57.

Kleinteich J, Wood SA, Küpper FC *et al.* Temperature-related changes in polar cyanobacterial mat diversity and toxin production. *Nature Climate Change* 2012;**2**:356–60.

Lo Giudice A, Casella P, Caruso C *et al.* Occurrence and characterization of psychrotolerant hydrocarbon-oxidizing bacteria from surface seawater along the Victoria Land coast (Antarctica). *Polar Biol* 2010;**33**:929–43.

Lutz S, Anesio AM, Jorge Villar SE *et al.* Variations of algal communities cause darkening of a Greenland glacier. *FEMS Microbiol Ecol* 2014;**89**:402–14.

Mueller DR, Pollard WH. Gradient analysis of cryoconite ecosystems from two polar glaciers. *Polar Biol* 2004;**27**:66–74.

Musilova M, Tranter M, Bennett SA *et al.* Stable microbial community composition on the Greenland ice sheet. *Front Microbiol* 2015;**6**:193.

Muyzer G, De Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microb* 1993;**59**:695–700.

Nadeau T-L, Milbrandt EC, Castenholz RW. Evolutionary relationships of cultivated Antarctic Oscillatorians (Cyanobacteria). *J Phycol* 2001;**37**:650–4.

Peeters K, Ertz D, Willems A. Culturable bacterial diversity at the Princess Elisabeth Station (Utsteinen, Sør Rondane Mountains, East Antarctica) harbours many new taxa. *Syst Appl Microbiol* 2011;**34**:360–7.

Salter SJ, Cox MJ, Turek EM *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 2014;**12**:87.

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

Segawa T, Yoshimura Y, Watanabe K *et al.* Community structure of culturable bacteria on surface of Gulkana Glacier, Alaska. *Polar Sci* 2011;**5**:41–51.

Shtarkman YM, Koçer ZA, Edgar R *et al.* Subglacial Lake Vostok (Antarctica) accretion ice contains a diverse set of sequences from aquatic, marine and sediment-inhabiting bacteria and eukarya. *PLoS One* 2013;**8**:e67221.

Stibal M, Gözdereliler E, Cameron KA *et al.* Microbial abundance in surface ice on the Greenland Ice Sheet. *Front Microbiol* 2015b;**6**:225.

Stibal M, Sabacká M, Zárský J. Biological processes on glacier and ice sheet surfaces. *Nat Geosci* 2012a;**5**:771–4.

Stibal M, Schostag M, Cameron KA *et al.* Different bulk and active bacterial communities in cryoconite from the margin and interior of the Greenland ice sheet. *Environ Microbiol Rep* 2015a;**7**:293–300.

Stibal M, Telling J, Cook J *et al.* Environmental controls on microbial abundance and activity on the Greenland ice sheet: a multivariate analysis approach. *Microb Ecol* 2012b;**63**:74–84.

Takeuchi N, Kohshima S, Seko K. Structure, formation, and darkening process of albedo-reducing material (cryoconite) on a Himalayan glacier: a granular algal mat growing on the glacier. *Arct Antarct Alp Res* 2001;**33**:115–22.

Taton A, Grubisic S, Ertz D *et al.* Polyphasic study of Antarctic cyanobacterial strains. *J Phycol* 2006;**42**:1257–70.

Telling J, Stibal M, Anesio AM *et al.* Microbial nitrogen cycling on the Greenland ice sheet. *Biogeosciences* 2012;**9**:2431–42.

Vaïtilingom M, Attard E, Gaiani N *et al.* Long-term features of cloud microbiology at the Puy de Dôme (France). *Atmos Environ* 2012;**56**:88–100.

van As D, Fausto RS, Ahlstrøm A *et al.* Programme for Monitoring of the Greenland Ice Sheet (PROMICE): first temperature and ablation record. *Geol Surv Den Greenl Bull* 2011;**23**:73–6.

Vital M, Füchslin HP, Hammes F *et al.* Growth of *Vibrio cholerae* O1 ogawa eltor in freshwater. *Microbiol* 2007;**153**:1993–2001.

Xu Y, Moser C, Al-Soud WA *et al.* Culture-dependent and -independent investigations of microbial diversity on urinary catheters. *J Clin Microbiol* 2012;**50**:3901–8.

Yu Y, Lee C, Kim J *et al.* Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol Bioeng* 2005 **89**:670–9.

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

Zeng Y-X, Yan M, Yu Y *et al.* Diversity of bacteria in surface ice of Austre Lovénbreen glacier, Svalbard. *Arch Microbiol* 2013;**195**:313–22.