View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Aberystwyth Research Portal

# Aberystwyth University

## *Diversity and potential sources of microbiota associated with snow on western portions of the Greenland ice sheet*

Cameron, Karen; Hagedorn, Birgit; Dieser, Markus; Christner, Brent C.; Choquette, Kyla; Sletten, Ronald; Crump, Byron; Kellogg, Colleen; Junge, Karen

*Published in:*
Environmental Microbiology

*DOI:*
10.1111/1462-2920.12446

*Publication date:*
2014

*Citation for published version (APA):*
Cameron, K., Hagedorn, B., Dieser, M., Christner, B. C., Choquette, K., Sletten, R., Crump, B., Kellogg, C., & Junge, K. (2014). Diversity and potential sources of microbiota associated with snow on western portions of the Greenland ice sheet. *Environmental Microbiology*, *17*(3), 594-609. https://doi.org/10.1111/1462-2920.12446

**General rights**
Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

**Take down policy**
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400
email: is@aber.ac.uk

Download date: 09. Jul. 2020

# Diversity and potential sources of microbiota associated with snow on western portions of the Greenland Ice Sheet

Karen A. Cameron[1*], Birgit Hagedorn[2], Markus Dieser[3**], Brent C. Christner[3], Kyla Choquette[2], Ronald Sletten[4], Byron Crump[5], Colleen Kellogg[5], Karen Junge[1]

[1] Applied Physics Laboratory, University of Washington, Seattle, WA

[2] Environment and Natural Resources Institute, University of Alaska Anchorage, Anchorage, AK

[3] Department of Biological Sciences, Louisiana State University, Baton Rouge, LA

[4] Earth and Space Sciences, University of Washington, Seattle, WA

[5] Earth Ocean and Atmospheric Sciences, Oregon State University, Corvallis, OR

Author for correspondence: Karen A. Cameron, Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Box 355640, Seattle, WA 98105-6698. +45 91 33 35 94. cameronk@uw.edu

Running title: Greenland Ice Sheet snow microbial diversity

Current addresses:

* Geological Survey of Denmark and Greenland, Department of Geochemistry, Øster Voldgade 10, 1350 Copenhagen K, Denmark

** Center for Biofilm Engineering, Montana State University, Bozeman, MT

26 *Summary*

27 Snow overlays the majority of the Greenland Ice Sheet (GrIS). However, there is very little information available on the
28 microbiological assemblages that are associated with this vast and climate-sensitive landscape. In this study, the
29 structure and diversity of snow microbial assemblages from two regions of the western GrIS ice-margin were
30 investigated through the sequencing of small subunit rRNA genes. The origins of the microbiota were investigated by
31 examining correlations to molecular data obtained from marine, soil, freshwater and atmospheric environments and
32 geochemical analytes measured in the snow. Snow was found to contain a diverse assemblage of bacteria
33 (*Alphaproteobacteria*, *Betaproteobacteria* and *Gammaproteobacteria*) and eukarya (*Alveolata*, *Fungi*, *Stramenopiles* and
34 *Chloroplastida*). Phylotypes related to archaeal *Thaumarchaeota* and *Euryarchaeota* phyla were also identified. The
35 snow microbial assemblages were more similar to communities characterized in soil than to those documented in
36 marine ecosystems. Despite this, the chemical composition of snow samples was consistent with a marine contribution,
37 and strong correlations existed between bacterial beta diversity and the concentration of $Na^+$ and $Cl^-$. These results
38 suggest that surface snow from western regions of Greenland contains exogenous microbiota that were likely
39 aerosolized from more distant soil sources, transported in the atmosphere, and co-precipitated with the snow.

40

41 *Introduction*

42 Snows' transient nature, seasonal distribution and physical properties drive its widespread influences on climate,
43 hydrology and ecosystem functioning (Jones, 1999, Serreze *et al.*, 2006, Vavrus, 2007). Arctic, Antarctic and Alpine snow
44 has been found to harbor microbial communities (for example; Felip *et al.*, 1995, Thomas and Duval, 1995, Carpenter *et*
45 *al.*, 2000, Segawa *et al.*, 2005, Bachy *et al.*, 2011, Harding *et al.*, 2011, Hell *et al.*, 2013, Møller *et al.*, 2013), with snow
46 from Svalbard glaciers containing microbial abundances of 2 – 8 x $10^4$ cells $ml^{-1}$ (Amato *et al.*, 2007, Irvine-Fynn *et al.*,
47 2012). Arctic snow diversity studies have revealed microbial assemblages dominated by bacterial *Proteobacteria,*
48 *Bacteroidetes, Actinobacteria, Firmicutes, Acidobacteria* and cyanobacteria, and eukaryotic *Stramenopiles*, *Dikarya* and
49 *Alveolata*, using small subunit ribosomal RNA (rRNA) gene sequencing techniques (Amato *et al.*, 2007, Larose *et al.*, 2010,
50 Bachy *et al.*, 2011, Harding *et al.*, 2011, Hell *et al.*, 2013, Møller *et al.*, 2013). Snow-associated microbial communities
51 have been found to be metabolically active under *in situ* conditions in Alpine, Antarctic and Sub-Arctic locations (Felip *et*

52 *al.*, 1995, Carpenter *et al.*, 2000, Larsen *et al.*, 2007, Lopatina *et al.*, 2013), and snow microbial community production
53 has been suggested to be nutritionally important to supraglacial, subglacial and ice-marginal environments (Hodson *et*
54 *al.*, 2005, Wynn *et al.*, 2007, Hodson *et al.*, 2008, Schutte *et al.*, 2009, Telling *et al.*, 2011). Nevertheless, data on the
55 composition, biogeography, origins and metabolic activity of snow communities is sparse. Furthermore, there have been
56 no molecular-based studies to date describing the microbial diversity in snow from the vast (1.7 million $km^2$; Weidick
57 1995) and climate-sensitive landscape of the Greenland Ice Sheet (GrIS).
58
59 Snowpack ecosystems are believed to be seeded through the aeolian transportation of biota from local (Larose *et al.*,
60 2010, Harding *et al.*, 2011) as well as more distant sources (Harding *et al.*, 2011). The mobilization of soil particulates
61 from terrestrial surfaces via wind, and the production of marine aerosols during bubble dispersal at water-air interfaces,
62 are important mechanisms for the transportation of microorganisms across local, regional and intercontinental
63 distances (Finlay and Fenchel, 2004, Aller *et al.*, 2005, Rousseau *et al.*, 2005, Smith *et al.*, 2013). Once aerosolized, biotic
64 particles may influence meteorological processes by acting as nuclei for ice crystal formation within clouds (Lohmann
65 and Feichter, 2005, Möhler *et al.*, 2007, Pratt *et al.*, 2009). Over time frames of days to weeks (Burrows *et al.*, 2009),
66 bioaersols are transported in the atmosphere and are eventually returned to the Earth's surface through gravitational
67 deposition or by precipitation, where they may immigrate and integrate into ecological processes in the receiving
68 environment (Sattler *et al.*, 2001).
69
70 The GrIS possesses a largely ubiquitous snow cover that is free of major geographical features beyond the ice marginal
71 regions. This expanse of snow has a dynamic turnover, with an estimated snowfall of ~ 600 $km^3$ $yr^{-1}$ (data from 1979 –
72 2005; Fettweis, 2007) and an estimated surface run-off from snow and ice melt of 300 $km^3$ $yr^{-1}$ (data from 1979 – 2006;
73 Fettweis, 2007) to 400 $km^3$ $yr^{-1}$ (data from 2010; Bamber *et al.*, 2012). Snow accumulation is highest within central
74 regions of the ice sheet and melt water generation is larger within the ablation zones (Fettweis, 2007). The GrIS is
75 sensitive to the climate, and consequently, freshwater fluxes from the GrIS have been increasing annually for over two
76 decades, due to warmer surface air temperatures (Hanna *et al.*, 2008, Bamber *et al.*, 2012), and since 1985, the area of
77 snow melt has tended to expand each year (Comiso, 2006). Several extreme melt events have occurred within the last

78 decade (Comiso, 2006, Tedesco, 2007), the most notable of which occurred on July 12th 2012, where 98.6 % of the GrIS
79 was reported to experience surface melt, as opposed to 43.7 % four days earlier (Nghiem *et al.*, 2012). Due to
80 Greenland's topography and geographical location, the microbiology and ecology of GrIS surface snow environments
81 may differ to those of previously studied surface snows from glacial and sea ice environments (for example Hell *et al.*,
82 2013, Møller *et al.*, 2013). Snow accumulating on the surface of the GrIS is eventually transformed into melt water,
83 sublimated, compacted into glacial ice or redistributed through wind transportation (reviewed in; Hodson *et al.*, 2008,
84 Larose *et al.*, 2013b). Therefore, an understanding of the microorganisms that are co-deposited with snow provides
85 valuable information about the pool of species that serve as inoculum to all icy habitats associated with the GrIS
86 (Hodson *et al.*, 2005, Wynn *et al.*, 2007, Hodson *et al.*, 2008, Schutte *et al.*, 2009).
87
88 In this study, the structure, diversity and abundance of bacterial, archaeal and eukaryotic assemblages from North-West
89 (NW) and South-West (SW) regions of the GrIS (Figure 1) were investigated through Illumina sequencing of the small
90 subunit (16S or 18S) rRNA gene. The specific aims of this study were to characterize the microbial composition of the
91 snow, examine spatial bacterial variability in two regions of the GrIS, and infer the origins of these transient biota
92 through comparison with data from likely source environments (i.e., marine water, sea ice, soil, freshwater and air)
93 using correlation analyses of snow bacterial assemblages and chemical profiles. Together, these studies aimed to
94 improve our understanding of the origins, diversity and potential downstream significance of western GrIS surface snow
95 microbial assemblages.
96
97 *Experimental Procedures*
98 *Study sites*
99 To test the spatial variability of snow biota, snow was sampled along two, three point transects on the surface of the
100 GrIS, located ~ 1200 km apart in the NW (early June 2011; NW.1, NW.2, NW.3) and SW (late May 2012; SW.1, SW.2,
101 SW.3) regions of the GrIS (Figure 1; Table 1). NW transect samples were obtained within the Northern part of the Pituffik
102 Peninsula. Snow was sampled 9.5 km (NW.1), 3.5 km (NW.2) and 1.6 km (NW.3) from the ice margin, and the most
103 westerly point (NW.3) was located 9 km from open fjord water, 17 km from the marine waters of Baffin Bay, and 14 km

104 from the settlement of Thule. Weather observations recorded within the Thule region since 1971 (Global Historical
105 Climatology Network Database (GHCND) ID: GLW00017605; data available through National Oceanic & Atmospheric
106 Administration) show a mean air temperature of - 11.7 ± 3 °C, with the warmest temperatures recorded in July (5.3 ±
107 3 °C) and the lowest temperatures recorded in February (- 26.0 ± 5 °C). Annual trends show an increase in air
108 temperature since 1978 (0.04 °C $yr^{-1}$). The mean precipitation is ~ 124 mm $yr^{-1}$ water equivalent, of which ~ 40 % falls
109 during summer (June, July, August). Points within the SW transect were sampled 53 km (SW.1), 28 km (SW.2) and 2.5 km
110 (SW.1) from the ice margin. The most westerly point of the SW transect (SW.3) lies 50 km from open fjord water, 172 km
111 from the marine waters of the Davis Strait, and 43km from the settlement of Kangerlussuaq. Between 1948 and 2003,
112 the mean annual air temperature within this region was -5.1 °C, with a mean June to August temperature of 9.8 °C
113 (Aelby and Fritz, 2009). The mean annual precipitation was 158 mm $yr^{-1}$, and the mean calculated evaporation was 300
114 mm $yr^{-1}$ (Aelby and Fritz, 2009), indicating that this area has a negative water balance. Transect samples were obtained
115 within the ablation zone, however, the most easterly points (NW.1 and SW.1) were sampled within a region that is
116 believed to retain multiyear snow. Two additional snow samples were opportunistically obtained from the NW GrIS
117 region; one from the edge of the ice sheet (NW.IM), and one within the snow transect area in August (NW.4).
118 Furthermore, two freshwater supraglacial lakes, which were at the SW.2 and SW.3 snow transect sites (SW.2.lake,
119 SW.3.lake respectively), were sampled for comparison.
120
121 *Sampling*
122 All sampling sites were accessed by helicopter, with the exception of site NW.IM, which was accessed from the margin
123 on foot. Snow and supraglacial lake water were sampled for chemical analysis by collecting samples in 4.5 L Whirl-Pak
124 bags at sites that were approximately 10 m from microbiological sampling points (see Table 1). Snow samples were
125 thawed at ~ 25°C and stored at 4 °C in closed bags until analysis at the University of Anchorage Alaska. Snow samples for
126 microbial abundance analysis were collected into sterile 50 ml centrifuge tubes and melted at room temperature.
127 Samples were fixed by adding a 0.22 µm filtered formaldehyde solution to a final concentration of 2 % v/v. The samples
128 were stored at -20 °C until processed at the University of Washington. For the extraction of microbial DNA,
129 approximately 15 kg of snow from the top 30 cm of the snowpack was placed into autoclaved bags using a snow shovel

130 cleaned with 10 % v/v HCl and wiped with 70 % v/v ethanol. Approximately 15 L of supraglacial lake water was collected
131 into HCl (10 % v/v) cleaned carboys that were rinsed with 0.22 µm filtered deionized water. Snow samples were melted
132 at room temperature within the sampling bags for up to 36 h, after which they were processed immediately. Each
133 sample was filtered through a 0.22 µm Sterivex filter (Millipore, MA, USA), which was then filled with DNA storage buffer
134 (40 mM EDTA, 50 mM Tris HCl, 0.73M sucrose), and stored at -20 °C until DNA extractions were performed at the
135 University of Washington.
136
137 *Water chemistry*
138 The electrical conductivity and pH of melted snow was measured using a WTW Multi 3430 multiparameter meter (WTW,
139 Weilheim, Germany). Snow and supraglacial lake samples were passed through 0.2 µm filters prior to chemical analysis.
140 Samples collected for cation analysis were acidified to 1 % v/v using $HNO_3$ (Ultrex-grade) before analysis on an
141 inductively coupled plasma mass spectrometer (ICP MS 7500c; Agilent Technologies, CA, USA). Anion analysis was
142 performed using an ion chromatograph (Dionex IC 500, Dionex, and ICS 5000, Thermo Scientific, MA, USA). An
143 international NIST standard SRM 1643e was used to verify external calibration standards. The limit of detection was
144 calculated with 3.3 times the standard deviation of the calibration regression y-intercept, and the limit of quantitation
145 was defined as 3 times the limit of detection. Accuracy of analysis was < 5 % for concentrations above the reporting limit,
146 and up to 20 % for concentrations below the limit of quantification.
147
148 *Microbial abundance analysis*
149 Bacterial enumeration was performed by epifluorescent microscopy using a DAPI nucleic acid stain (Invitrogen, NY, USA)
150 alongside an acridine orange counter stain (Invitrogen) on a Zeiss AxioScope A1 microscope, as described by Collins *et al.*
151 (2008).
152
153 *DNA extraction, amplification and amplicon sequencing*
154 All DNA extractions and reagent preparations for PCR amplification were performed in a laminar flow hood, wiped with
155 ethanol and irradiated with germicidal UV. Filter pipet tips and DNA- and RNAase free certified plasticware was used

156 throughout. Sterivex filter units were opened and internal filter membranes were removed using a flame-heated razor
157 blade. DNA extractions were performed using PowerWater DNA Isolation Kits (Mo Bio Laboratories Inc., CA, USA).
158 Amplification of the bacterial 16S rRNA gene V4 hypervariable region was achieved using primers detailed by Caporaso
159 *et al.* (F515/R806; 2011). To obtain an overview of archaeal and eukaryotic diversity, PCRs were performed using pooled
160 DNA extracts from NW.1, NW.2 and NW.3 sites (named NW.1-3), and SW.1, SW.2 and SW.3 sites (named SW.1-3). The
161 V6 hypervariable region of archaeal 16S rRNA genes was targeted using V6-major and V6-minor primer sets, as
162 described by Huber *et al.* (958arcF/1048arcR-major/1048arcR-minor; 2007). The eukaryotic 18S rRNA V9 hypervariable
163 region was targeted using euk1 and euk2 primer sets, designed by Amaral-Zettler *et al.* (F1380/F1389/R1510; 2009). All
164 forward and reverse primers were modified to include a unique eight-nucleotide barcode. PCR reaction mixtures
165 contained 1X PCR Gold buffer (Applied Biosystems, CA, USA), 2.5 mM $MgCl_2$, 200 µM of each deoxyribonucleotide
166 triphosphates (Invitrogen, NY, USA), 0.3 µM of each primer, 2.5 U of AmpliTaq Gold - LD Polymerase (Applied
167 Biosystems) and up to 10 pg of the extracted DNA. The reaction volume was adjusted to a total of 50 µl with ultrapure
168 DNase/RNase free water. Thermal cycles consisted of an initial denaturation of 9 min at 95 °C, followed by 43 cycles of
169 94 °C for 30 sec, 55 °C for 60 sec and 72 °C for 60 sec, and a final extension of 7 min at 72 °C.
170
171 To prepare amplicons for sequencing, PCR products were cleaned using QIAquick PCR Purification Kits (Qiagen, Hilden,
172 Germany). Cleaned amplicons were combined into three pools, based on their bacterial, eukaryotic or archaeal origin.
173 The pooled samples contained 250 ng of DNA and were processed with the TruSeq RNA and DNA Sample Preparation Kit
174 (Illumina Inc., CA, USA). Sequencing was conducted on an Illumina MiSeq system (running MCS v.1.2.3 and RTA 1.14.23
175 software; Illumina, CA, USA).
176
177 *Sequence read analysis*
178 Forward orientated demultiplexed sequences were quality-filtered and processed using a QIIME processing platform
179 (Caporaso *et al.*, 2010b). QIIME standard operating procedure quality filters were used throughout. V-REVCOMP was
180 used to identify bacterial and archaeal forward orientated sequences (Hartmann *et al.*, 2011). Bacterial and archaeal
181 operational taxonomic units (OTUs) were defined as sequences that possessed ≥ 97% identity, which were clustered

182 using a reference based UCLUST algorithm against a Greengenes (GG) reference library (Desantis *et al.*, 2006, Edgar,
183 2010). Bacterial and archaeal sequences were aligned using PyNAST, and taxonomic classifications were assigned by
184 training the Ribosomal Database Project (RDP) classifier (Wang *et al.*, 2007, Caporaso *et al.*, 2010a) to use the February
185 2011 GG taxonomic dataset. Secondary taxonomic classifications were performed using NCBI BLASTn (Zheng *et al.,*
186 2000). Eukaryotic OTUs with ≥ 97 % identity were selected using a *de novo* UCLUST OTU clustering method, which allows
187 for the generation of OTU clusters based on sequence identities within the dataset, rather than generating OTU clusters
188 against a database of reference sequences. Eukaryotic taxonomic assignments were made using the Silva 104 reference
189 database (Quast *et al.*, 2013). OTUs that were not taxonomically classified under the Silva 104 reference database were
190 discarded. ChimeraSlayer was used to identify chimeric sequences within all sequence profiles (Haas *et al.*, 2011) and
191 QIIME was used to remove chimeric sequences and singletons. Sequence profiles were rarefied to the number of
192 sequences of the smallest sample sequence output within each pool. CatchAll was used to calculate parametric alpha
193 diversity (Bunge, 2011). Bray-Curtis resemblance, cluster analysis, non-metric multi-dimensional scaling (MDS), analysis
194 of similarity (ANOSIM), contributions of variables to similarity (SIMPER) and multi-variant environmental correlation
195 analysis (BIO-ENV) were calculated from OTU matrices using PRIMER-E version 6 (Plymouth, UK). Amplicon datasets are
196 available at The European Bioinformatics Institute under study accession number PRJEB4904 (www.ebi.ac.uk).
197
198 *Biogeographical analysis*
199 All snow sequences were compiled alongside 16S rRNA gene sequence profiles from 11 previously reported studies of
200 snow, slush, supraglacial ice, marine surface water, sea ice, periglacial soil, periglacial lake and air environments (Table
201 2), as well as to the supraglacial lake assemblages investigated within this current study (SW.2.lake, SW.3.lake).
202 Sequences were processed and analyzed using the same methodologies detailed for bacterial sequence read analysis.
203
204 *Results*
205 *Major ion chemical analysis*
206 Chemical analysis of GrIS snow and supraglacial lake samples revealed that NW GrIS samples had higher concentrations
207 of $Na^+$, $Cl^-$ and $SO_4^{2-}$ ions in comparison to samples obtained from the SW GrIS region (NW GrIS mean $Na^+$ concentration

208 was 6.0 fold higher than SW GrIS mean $Na^+$ concentration; NW GrIS mean $Cl^-$ concentration was 3.8 fold higher than SW
209 GrIS mean $Cl^-$ concentration; NW GrIS mean $SO_4^{2-}$ concentration was 1.6 fold higher than SW GrIS mean $SO_4^{2-}$
210 concentration; Table 1, Supplemental Figure 1). Within the NW GrIS snow transect, concentrations of $Na^+$ and $Cl^-$ were
211 found to be highest closest to the ice sheet margin and decreased in concentration further inland (Table 1). The
212 elemental composition of major cations and anions indicate that NW GrIS snow samples had the greatest similarity to
213 seawater (Supplemental Figure 1). The pH at all sites ranged between 5.1 and 5.6, and EC values ranged between 1.7 to
214 3.7 µS $cm^{-1}$ (Table 1).
215
216 *Microbial abundance*
217 Snow microbial abundance from the three NW snow transect sites, NW.1, NW.2 and NW.3, were calculated as $4.7 \times 10^2$
218 $\pm 2.8 \times 10^3$ cells $ml^{-1}$, $2.5 \times 10^2 \pm 2.2 \times 10^3$ cells $ml^{-1}$ and $4.1 \times 10^2 \pm 1.8 \times 10^3$ cells $ml^{-1}$ respectively. However, due to the
219 low microbial abundance, it was necessary to filter large volumes of snow, and cells may have been masked by sediment
220 particles. Hence, our cell abundance data probably underestimate actual cell concentrations. The mean microbial
221 abundance of NW GrIS sampled snow was $3.8 \times 10^2 \pm 1.1 \times 10^4$ cells $ml^{-1}$. With the exception of the SW.2 snow sample,
222 which had an estimated abundance of $2.6 \times 10^4 \pm 3.6 \times 10^4$ cells $ml^{-1}$, microbial abundance from SW region snow samples
223 could not be measured due to combined complications from low abundance and the masking effects of sediments.
224
225 *Bacterial structure, diversity and biogeochemical correlations*
226 PCR amplifications of 16S rRNA genes were successful for all snow and supraglacial lake samples, however, 43 thermal
227 cycles were used to achieve this (as guided by the AmpliTaq Gold manual), which will have likely increased PCR
228 associated artifacts (V. Wintzingerode *et al*., 1997). On average, 50.0 ± 26.9 % of sequences per sample were lost in
229 downstream quality filtering steps (Supplemental Table 1). Quality filtered sequence profiles ranged from 2404 reads
230 per sample (SW.3.lake) to 56803 reads per sample (NW.3) (Supplemental Table 1). Following the CatchAll calculation of
231 parametric alpha diversity, species richness estimates of rarefied snow bacterial profiles were found to range between
232 357 to 1756 OTUs per sample, with the lowest richness calculated at SW.2 and the highest richness calculated at the
233 NW.1 site (Supplemental Table 1). The mean taxon richness of snow sampled from the NW region (1037 ± 532 OTUs per

234 sample) was estimated to be more than twice as rich as snow sampled from the SW region (423 ± 94 OTUs per sample).
235 Both NW and SW snow sequence profiles has the highest diversity calculated at transect points that were furthest from
236 the ice margin (NW.1; 1756 OTUs per sample, and SW.1; 531 OTUs per sample). No significant trends were observed
237 between diversity estimates and distance from the ice margin (data not shown). Simprof tested cluster analyses of non-
238 transformed OTU profiles revealed that SW.1 and SW.3 profiles were 78 % similar, and NW.1, NW.2, NW.3 and NW.IM
239 samples were more than 28 % similar to each other. Sequences generated from SW.2 had only 9 % similarity to other
240 samples, and the August NW.4 sample had only 11 % similarity to other samples. Overall, NW and SW OTU profiles
241 shared less than 9 % similarity to each other. Pairwise ANOSIM tests of all non-transformed NW and SW GrIS snow OTU
242 profiles revealed that these assemblages were moderately similar ($R = 0.66$, $P = 0.018$, NW GrIS $n = 5$, SW GrIS $n = 3$).
243 However, on exclusion of the outlying SW.2 and NW.IM OTU profiles from the pairwise ANOSIM analysis, the NW and
244 SW OTU profiles were less similar ($R = 0.86$, $P = 0.067$, NW GrIS $n = 4$, SW GrIS $n = 2$), but were not significant at an $\alpha =$
245 $0.05$.
246
247 Sequences classifying within *Proteobacteria* dominated the bacterial phylotypes in all GrIS snow samples, with the
248 exception of SW.2 (Figure 2). The mean abundance of *Proteobacteria* was 60.4 ± 25.4 %, which predominantly consisted
249 of sequences related to *Beta*- (50.0 %), *Gamma*- (31.7 %) and *Alphaproteobacteria* (17.9 %). When all western GrIS snow
250 sequences were considered as a single compiled sample, *Sphingobacteria* (9.7 %), *Actinobacteria* (8.1 %), *Acidobacteria*
251 (3.9 %), *Bacilli* (3.7 %), *Clostridia* (3.6 %) and *Flavobacteria* (1.0 %) taxa were also major contributors of species richness.
252 Additionally, 52 classes of bacteria were identified that each represented less than 0.5 % of the total abundance
253 (compressed to the < 0.5 % fraction in Figure 2). OTUs that represented ≥ 10 % of the NW snow assemblages, sampled in
254 June, included GG OTU IDs 102382 (most closely related to *Massilia timonae;* NW.3; 10.6 % of total abundance), 165313
255 (most closely related to *Acinetobacter johnsonii;* NW.2; 10.3 % of total abundance) and 109056 (most closely related to
256 *Pseudomonas sp.;* NW.1; 22.8 %, NW.3; 16.0 % of total abundance). The August NW snow sample, NW.4, was not
257 heavily represented by any of these OTU sequences, however, over half of the sequence profile was represented by
258 three OTUs that were most closely related to GG OTU IDs 39819 (*Acinetobacter ursingii*; 17.0 %), 143746 (*Candidatus*
259 *Odyssella*; 19.3 %) and 354510 (*Collimonas arenae*; 20.2 %). No single OTU represented more than 10 % of the sequence

260 profile generated from the NW ice margin sample (NW.IM). Both SW.1 and SW.3 sequence profiles were dominated by
261 GG OTU IDs 20151 (most closely related to *Arcicella aquatica*; SW.1; 35.0 %, SW.3; 19.3 %) and 80104 (most closely
262 related to *Actimicrobium antarcticum*; SW.1; 40.5 %, SW.3; 47.5 %). Amplicons from SW.2 DNA extractions were most
263 heavily represented by sequences that were related to cyanobacteria (52.3 %, of which 48.6 % were most closely related
264 to the *Phormidium pristleyi* GG OTU ID 184822) and *Acidobacteria* (27.9 %, of which 25.9 % were most closely related to
265 the *Granulicella arctica* GG OTU ID 546864). Cyanobacteria related OTUs constituted < 1 % of all other sequence profiles.
266 OTUs (17) related to ice nucleating species of the *Enterobacteriaceae, Xanthomonadaceae and Pseudomonadaceae*
267 (reviewed in Christner *et al*., 2008 and reported in Joly *et al*., 2013) included *Pseudomonas fluorescens* (5 OTUs),
268 *Pseudomonas syringae* (1 OTU), *Pseudoxanthomonas sp.* (4 OTUs), *Xanthomonas sp.* (2 OTUs), *Pantoea agglomerans* (1
269 OTU) and *Erwinia* (4 OTUs). These targeted OTUs compiled between 0.2 % (SW.1) and 25.5 % (NW.1) of the total rarefied
270 assemblage population (2404 sequences per sample), with a mean abundance of 5.9 ± 10.0 %. OTUs that were found to
271 be most closely related to *Pseudomonas fluorescens* represented 5.7 ± 9.9 % of the total assemblage population. When
272 SIMPER tests of square root transformed Bray-Curtis OTU similarities were used to reveal order level taxonomies that
273 were responsible for the greatest levels of identity between NW and SW snow profiles, over half of the similarities found
274 (54 %) were accounted for by OTUs that were most closely related to *Burkholderiales*, *Pseudomonadales*,
275 *Sphingobacteriales, Actinomycetales* and *Sphingomonadales* orders.
276
277 SW GrIS supraglacial lake biota were sequenced and analyzed to generate OTU profiles from an adjacent yet distinct
278 environmental biome. CatchAll species richness calculations of the rarefied sequence profiles were found to be more
279 than twice as rich as the SW snow samples taken within the vicinity (SW.2.lake; 852 OTUs per sample, SW.3.lake; 843
280 OTUs per sample). Taxonomies that represented more that 10 % of the SW.2.lake OTU profile included
281 *Gammaproteobacteria* (32.4 %), *Alphaproteobacteria* (14.0 %), *Betaproteobacteria* (13.2 %) and *Actinobacteria* (10.4 %),
282 and sequences from the SW.3.lake were strongly represented by *Betaproteobacteria* (35.7 %), *Gammaproteobacteria*
283 (13.5 %) and *Bacilli* (11.5 %; Figure 2). No single OTU from these profiles dominated assemblages by more than 10 %.
284 CLUSTER analysis of SW GrIS supraglacial lake OTU profiles with GrIS snow profiles revealed that there were minimal

285 resemblances to profiles generated from the adjacent SW GrIS snow (both SW.2.lake and SW.3.lake had < 10 % similarity
286 to SW snow samples), however, stronger similarities were found to the NW snow samples (> 20 % similarity).
287
288 To identify correlations between chemical compositions (including $Na^+$, $Cl^-$, $Mg^{2+}$, $K^+$, $Ca^{2+}$, $F^-$, $SO_4^{2-}$, $NO_3^-$; see Table 1) and
289 GrIS snow and supraglacial lake bacterial assemblages, a BIO-ENV statistical analysis was performed. Analyses of square-
290 root transformed OTU profiles were found to correlate most strongly to a combined presence of $Na^+$ and $Mg^{2+}$
291 concentrations ($\rho = 0.856$, $P = 0.01$). Within this analysis, correlations to singular chemical compositions were highest for
292 $Na^+$ and $Cl^-$ ($\rho = 0.852$, $\rho = 0.718$ respectively), followed by $Mg^{2+}$ ($\rho = 0.662$), and were poor for $SO_4^{2-}$, $K^+$, $NO_3^-$ and $Ca^{2+}$ ($\rho$
293 $< 0.2$ for all analyses). Within the 50 strongest diversity correlations to multiple chemical combinations, 41 combinations
294 included $Na^+$ as a factor influencing the strength of the correlation ($\rho$ ranged from 0.613 to 0.856). Correlations where $\rho$
295 $> 0.8$ included influences from combinations of $Na^+$, $Cl^-$, $Mg^{2+}$, $K^+$, and $Ca^{2+}$, where up to four ions contributed towards
296 the correlations. BIO-ENV analyses of square-root transformed OTU profiles that were segregated into phyla taxonomies
297 were found to correlate with less statistical significance ($P = 0.02 – 0.75$), however, OTUs in the *Actinobacteria* were
298 found to strongly correlate with $Na^+$ ($\rho = 0.790$, $P = 0.02$) while *Betaproteobacteria* related OTUs correlated with $Na^+$ and
299 $NO_3^-$ ($\rho = 0.729$, $P = 0.04$).
300
301 *Archaeal diversity*
302 Archaeal diversity was investigated using two archaeal specific primers sets (V6-major and V6-minor) to amplify 16S
303 rRNA genes from pooled DNA extracts of NW and SW snow traverses. After OTUs were quality filtered, V6-major and V6-
304 minor sequences from pooled NW samples had low sequence outputs (70 and 327 sequences per sample respectively;
305 Supplemental Table 2), therefore, a low rarefaction cut-off point of 70 sequences per sample was selected for all
306 samples. CatchAll calculated alpha diversity of each rarefied sample was 32 ± 10 OTUs per sample post-rarefaction, and
307 up to 365 OTUs per sample pre-rarefaction (Supplemental Table 2).
308
309 Archaeal assemblages were comprised of *Thaumarchaeota* and *Euryarchaeota* phyla (Figure 3). *Thaumarchaeota*
310 dominated the archaeal sequences obtained (mean composition of all four profiles; 74.3 ± 9.3 %), with the genus

311 *Candidatus Nitrososphaera* representing 55.7 ± 6.8 % of the mean composition for all four profiles. Both NW (NW.1-3)
312 and SW (SW.1-3) assemblages contained sequences that were related to the *Halobacteria* class (mean abundance of V6-
313 major and V6-minor amplicons; NW.1-3; 5.7 ± 4.0 %, SW.1-3; 25.7 ± 14.1 %). V6-major and V6-minor amplicons that
314 were most related to methanogenic organisms comprised an average of 17.1 % of the archaeal NW.1-3 sequence profile,
315 however, amplicons that were related to methanogenic organisms were absent from the SW.1-3 V6-minor amplicon
316 profile and were only minimally identified within the SW.1-3 V6-major sequence profile (1.4 % composition).
317
318 *Eukaryotic diversity*
319 Eukaryotic sequence profiles were generated through the identification of 18S rRNA genes from pooled NW (NW.1-3)
320 and SW (SW.1-3) snow transect DNA extractions, using two primer sets to target the V9 hypervariable region (total
321 number of quality filtered sequences generated per sequenced amplicon ranged from 38,674 to 83,224). After
322 rarefaction of all profiles to 38,674 sequences per sample, the CatchAll calculated alpha diversity of each pooled transect
323 sample, and from both primer sets (euk1 and euk2) was high (NW.1-3; 1134 and 1252 OTUs estimated, SW.1-3; 836 and
324 862 OTUs estimated respectively).
325
326 Both NW and SW snow transect amplicon assemblages were strongly represented by *Alveolata*, *Fungi*, *Stramenopiles*
327 and *Chloroplastida* (Figure 4). *Fungi* related sequences made up 63.4 % and 65.8 % of the NW euk1 and euk2 primer
328 generated amplicons (respectively). Of these euk1 and euk2 primer generated amplicons, 76.6 % and 73.5 % were most
329 closely related to *Basidiomycota* (respectively; data not shown). *Alveolata* related sequences dominated the SW snow
330 profiles (mean OTU representation of euk1 and euk2 profiles; 62.4 %) and were most strongly represented by sequences
331 that were related to species of the order *Gymnodiniphycidae* (mean OTU representation of euk1 and euk2 profiles;
332 98.1 %; data not shown). Sequences that clustered taxonomically at the phylum level but represented less than 0.5 % of
333 the total sequence profile included taxonomies from 18 different phylum groups within *Alveolata*, *Amoebozoa*,
334 *Centrohelida*, *Discoba*, *Holozoa*, *Metamonada*, *Rhizaria*, *Stramenopiles* and *Chloroplastida*.
335
336 *Bacterial biogeographical analysis*

337 The bacterial sequence data from this study were compared with archived data from a selection of environmental types
338 representing possible sources of atmospheric bioaerosols (see Table 2). Archaeal and eukaryotic biogeographical
339 analyses were not performed due to a lack of available comparative data. Cluster analysis of square-root transformed
340 Bray-Curtis indices of OTU profiles grouped each environmental type into distinct clades. Exceptions to this included
341 bacterial assemblages sampled from frozen freshwater (including snow, slush, supraglacial surface ice and supraglacial
342 lake environments) and marine environments (marine surface water and sea ice samples, termed marine environments
343 herein) which each clustered into their own separate group (Figure 5; cluster analysis results not shown). In addition,
344 bacterial assemblages sequenced from Colorado snow (Bowers *et al.*, 2009) clustered separately from other frozen
345 freshwater assemblages, with one Bowers *et al.* sample clustering alongside sequences obtained from air samples within
346 the same study, and one sample clustering separately (Figure 5; cluster analysis results not shown). Cluster analysis of
347 each environmental clade revealed that OTU profiles obtained within the same study had higher percentages of
348 similarity (up to 75 % similarity) than those obtained from different studies (< 12 % similarity).
349
350 To test for correlations between OTU assemblages sampled from different types of environmental origin, pairwise
351 ANOSIM tests of square root transformed Bray-Curtis similarities were calculated. NW and SW GrIS snow assemblages
352 from this current study were calculated to be most similar to soil sampled assemblages ($R = 0.475$, $P = 0.001$ and $R =$
353 $0.720$, $P = 0.001$ respectively; Supplemental Table 3). SIMPER tests of non-transformed Bray-Curtis similarity indices
354 were used to reveal OTUs that were responsible for the greatest levels of identity between environmental groups.
355 Bacterial order taxonomies that contributed over 5 % of similarities between NW GrIS snow and soil bacterial taxonomic
356 profiles included the *Burkholderiales* (18.7 %), *Actinomycetales* (14.8 %), *Sphingomonadales* (11.8 %), *Clostridiales*
357 (6.4 %) and *Lactobacillales* (6.2%). Similarities between SW GrIS snow and soil bacterial order taxonomies were greatest
358 among *Burkholderiales* (49.9 %), *Sphingobacteriales* (23.7 %) and *Acidobacteriales* (11.4 %; Figure 6). When the OTU
359 profiles of all four snow studies (this study, Bowers *et al.,* 2009, Hell *et al.,* 2013, Møller *et al.,* 2013) were considered as
360 a single sample, strong similarities were found to assemblages characterized in air ($R = 0.317$, $P = 0.001$; Supplemental
361 Table 3), moderate similarities were found to soil sampled assemblages ($R = 0.689$, $P = 0.001$; Supplemental Table 3) and
362 marine sampled assemblages were found to be dissimilar ($R = 0.919$, $P = 0.001$; Supplemental Table 3). However, on

363 considering each snow study separately, only snow assemblages sampled by Bowers *et al*. (2009) were found to be
364 highly similar to those from air, although, due to having few observations, this analysis was not found to be statistically
365 significant (R = 0.334, P = 0.141; Supplemental Table 3). Of the similarities found between GrIS snow and atmospheric
366 aerosols, OTUs that were most closely related to *Pseudomonadales* and *Burkholderiales* taxa accounted for the greatest
367 percentage of similarities to NW profiles (35.1 % and 20.5 % respectively), and *Burkholderiales* contributed towards
368 85.8 % of the similarities found to SW profiles (Figure 6).
369
370 Both NW and SW GrIS snow assemblages were calculated to lack overlap with assemblages generated from marine
371 environments (R = 0.92, P = 0.001 and R = 0.98, P = 0.001 respectively; Supplemental Table 3). Of the similarities that
372 were identified *Pseudomonadales* (30.0 %), *Actinomycetales* (13.9 %), *Sphingomonadales* (11.1 %) and *Burkholderiales*
373 (9.6 %), and *Burkholderiales* (45.5 %), *Sphingobacteriales* (23.8 %) and *Pseudomonadales* (6.3 %) related OTUs were
374 found to contributed the most towards overlap with NW and SW snow profiles respectively (Figure 6).
375
376 *Discussion*
377 Permanent snow surfaces and annual snowmelt events influence local ecosystem functioning and have widespread
378 ecological, hydrological and climatological impacts (Jones, 1999, Serreze *et al.*, 2006, Vavrus, 2007). The GrIS is the
379 Earth's second largest body of ice, covering 1.7 million $km^2$, receiving ~ 600 $km^3$ of snow each year, and supporting both
380 permanent and seasonal snow covers (Fettweis, 2007, Van Den Broeke *et al.*, 2008). However, despite the magnitude
381 and global importance of the GrIS landmass, this study is the first to report on the structure, diversity and likely origins
382 of GrIS surface snow microbiota. These data provide valuable information on this climate sensitive, underexplored niche,
383 and are of relevance when considering the interconnectivity of the GrIS supraglacial environment to downstream
384 englacial, subglacial, periglacial and marine ecosystems.
385
386 Surface snow samples from NW and SW regions of the GrIS were found to contain diverse bacterial and eukaryotic
387 assemblages, with species of archaea also present (Figures 2, 3, 4). Bacterial taxonomic richness estimates from rarefied
388 NW and SW GrIS snow assemblages (i.e. 357 – 1756 OTUs per sample; Supplemental Table 1) were similar to those

389 calculated from supraglacial cryoconite niches (Cameron *et al.*, 2012), and arable and grassland soils (Hughes *et al.*, 2001,
390 Torsvik *et al.*, 2002). All but one of these bacterial assemblages were dominated by *Proteobacteria*, with NW
391 assemblages being largely composed of *Alpha-*, *Beta-* and *Gammaproteobacteria,* and SW assemblages being strongly
392 represented by the *Betaproteobacteria* (Figure 2), similar to the communities of other supraglacial snow, ice and
393 cryoconite environments (e.g. Amato *et al.*, 2007, Larose *et al.*, 2010, Harding *et al.*, 2011, Cameron *et al.*, 2012,
394 Edwards *et al.*, 2013, Hell *et al.*, 2013). The eukaryotic alpha diversities of rarefied NW and SW GrIS snow transect
395 samples were high (i.e. 836 to 1252 OTUs per sample), especially when compared to a previously constructed small
396 clone library of High-Arctic snow eukaryotes by Bachy *et al.* (38 clones; 2011), where only four eukaryotic OTUs were
397 identified. However, the large eukaryotic alpha diversity estimates calculated within this current study may have been
398 influenced by the *de novo* OTU picking methodology that was applied, which allows for OTU clusters to be generated
399 based on similarities to each other, rather than on a framework of reference sequences. In contrast, alpha diversity
400 indices of heavily rarefied archaeal assemblages were low (i.e. 17 to 38 OTUs per sample when rarefied to 70 sequences
401 per sample), with phylotypes originating from *Thaumarchaeota* and *Euryarchaeota* phyla; this is similar to the archaeal
402 community structure of supraglacial cryoconite holes (Cameron *et al.*, 2012, Edwards *et al.*, 2013). Archaeal alpha
403 diversity estimates calculated on OTU profiles prior to rarefaction were up to an order of magnitude higher (SW.1-3; 365
404 OTUs estimated per sample; Supplemental Table 2), and were found to be similar to diversity estimates of archaeal
405 communities found within the surface waters of the North Sea (Wemheuer *et al.,* 2012). Cyanobacteria were neither a
406 major component of the GrIS surface snow biota analyzed within this current study, nor of Svalbard (Amato *et al.*, 2007,
407 Hell *et al.*, 2013), Canadian Arctic (Harding *et al*., 2011) and Tibetan (Liu *et al.*, 2009) snow communities reported in
408 other studies. An exception to this was noted within the SW.2 GrIS snow bacterial profile, where *Phormidium pristleyi*
409 related OTUs dominated (49 % relative abundance; Figure 2); perhaps as a result of local environmental conditions or
410 biotic sources. Larose *et al.* (2010) similarly found anomalies in the presence of cyanobacteria within Svalbard snow
411 samples, suggesting that there is a necessity to further develop biogeographical datasets, to determine robust ecological
412 patterns between local, regional and inter-landmass snow environments.
413

414 Early summer NW and SW GrIS surface snow bacterial compositions were found to cluster by location, with the
415 exception of the SW.2 sample, however, trends in diversity along the transects were not identified. In contrast, studies
416 of GrIS surface ice bacteria, sampled along a transect by Telling *et al.* (2012), found that the relative abundance of
417 nitrogen fixation genes initially increased with distance from the ice margin, implying that community structure and
418 function fluctuate spatially. Additionally, studies of Himalayan and Alaskan snow revealed relationships between altitude
419 and algal abundance and diversity (Yoshimura *et al.*, 1997, Nozomu, 2013). NW and SW GrIS locations had differing
420 cation and anion compositions (Supplemental Table 1), and at both locations concentrations of $Na^+$ and $Cl^-$ were found
421 to increase with proximity to marine waters. Similarly, maritime influences on snow chemistries have previous been
422 shown to vary with altitude and distance from marine sources within the Terra Nova Bay region of Antarctica (Udisti *et*
423 *al*., 1999). The unique nature of the August sampled NW GrIS snow biotic profile (NW.4; 11 % similarity to other
424 samples) within this current study hints towards surface snow assemblages either undergoing compositional changes in
425 response to environmental conditions, as outlined in snow studies by Segawa *et al.* (2005), Hell *et al.* (2013) and Larose
426 *et al.* (2013c), or experiencing temporal variability based on the deposition of biotic assemblages alongside snow, similar
427 to the temporal studies of airborne microbial communities by Fierer *et al.* (2008). Bacterial assemblages amplified from
428 SW GrIS snow and adjacent supraglacial lakes lacked similarity. Additionally, neither bacterial nor eukaryotic sequence
429 profiles from SW GrIS snow, sampled in May 2012, had resemblances to the longer-established supraglacial cryoconite
430 communities, sampled within the same GrIS region by Cameron *et al.* (2012) in late August 2008, despite potentially
431 being seeded by similar sources. While this lack of identity to communities from surrounding biomes once again
432 suggests that snow microbiota are heavily influenced by the biotic assemblages present at the time of snow deposition,
433 further temporal, spatial and geochemical investigations are undoubtedly essential to ascertain the factors that drive
434 the structure and diversity of GrIS surface snow microbiota.
435
436 When the bacterial assemblages in snow were compared to communities from marine, soil, supraglacial, freshwater and
437 air samples, strong to moderate ANOSIM test resemblances were calculated to profiles of soil bacteria (NW GrIS snow; R
438 = 0.48, P = 0.001, SW GrIS snow; R = 0.72, P = 0.001, all four snow studies compiled; R = 0.69, P = 0.001; Supplemental
439 Table 3), suggesting that the biotic composition of the sampled surface snow, was largely influenced by the translocation

440 of terrestrially originating microorganisms. Similarities between snow and air sampled assemblages were additionally
441 found (Supplemental Table 3), which was largely driven by the presence of *Pseudomonadales* and *Burkholderiales* orders
442 (Figure 6). OTUs which were found to be closely related to taxa previously described as having ice nucleating activities
443 (reviewed in Christner *et* al., 2008 and reported in Joly *et al*., 2013) contributed 5.9 ± 10.0 % of the total number of snow
444 sequences analyzed, and of these, 5 OTUs that were related to *Pseudomonas fluorescens* contributed 5.7 ± 9.9 %. While
445 the full extent of bacterial taxonomies and strains that possess ice nucleating properties remains unknown (Christner *et*
446 *al*., 2008), the significant proportion of bacteria with the potential to perform ice nucleation investigated within this
447 current study further suggest the connection between microbial air transport and snow deposition. Lesser resemblances
448 were found between snow and marine sampled assemblages (NW GrIS snow; R = 0.92, P = 0.001, SW GrIS snow; R = 0.98,
449 P = 0.001, all four snow studies compiled; R = 0.92, P = 0.001; Supplemental Table 3). Despite this, chemical analyses of
450 snow samples revealed strong maritime signatures (Supplemental Figure 1), and furthermore, correlations were
451 calculated between OTU profiles and $Na^+$ and $Cl^-$ chemical compositions (ρ = 0.852, ρ = 0.718 respectively). Interestingly,
452 studies of Canadian Arctic bacterial snow clone libraries by Harding *et al.* (2011), have previously identified the presence
453 of *Glaciecola*, *Colwellia*, *Loktanella* and *Polaribacter* genera, which are typically found in cold oceanic waters.
454
455 While the biogeographical analysis used within this current study is a useful tool to speculate on the origins and seeding
456 mechanisms of snow sampled environments, it is important to note the confinements that are imposed due to the
457 limited availability of comparable published sequence data. In addition, methodological discrepancies in sampling and
458 nucleic acid extractions, amplifications and sequencing between studies reduces their comparability. Study bias effects
459 were most apparent among the Bowers *et al*. (2009) snow and air sampled sequences, which were found to cluster
460 independently from other studies, regardless of environmental type (Figure 5), and which swayed the calculated
461 ANOSIM relationship between air and compiled snow assemblages (Supplemental Table 3). Despite this, the clustering
462 of assemblages, generated through other independent investigations, by environmental type, is nonetheless convincing
463 of the effectiveness of this methodology (Figure 5).
464

465 While investigations into the biogeochemical activities of snow were not studied, the high diversity of bacterial and
466 eukaryotic assemblages identified lends itself towards the potential for a wide range of metabolic activities to be
467 performed. Snow microbial communities from other geographic locations have been found to be biogeochemically
468 active, with noted roles in carbon and nitrogen cycling (Felip *et al.*, 1995, Larsen *et al.*, 2007, Telling *et al.*, 2012, Larose
469 *et al.*, 2013a). Considering the extent of the GrIS surface snow biome, if active, snow communities contribute
470 noteworthy carbon and nitrogen inputs to global budgets. Despite the low abundance of cells associated with snow
471 sampled from the NW region of Greenland ($3.8 \times 10^2$ cells $ml^{-1}$), when considered alongside estimates of GrIS surface
472 runoff in 2010 (~ 400 $km^3$ $yr^{-1}$; Bamber *et al.*, 2012), the total cellular abundance associated with GrIS supraglacial runoff
473 entering downstream environments is likely to be in the region of $1.5 \times 10^{20}$ cells $yr^{-1}$; with a carbon equivalent of 4.5 ±
474 1.85 Mg C $yr^{-1}$, and a nitrogen equivalent of 0.9 ± 0.2 Mg N $yr^{-1}$ [based on mean cellular carbon (30.2 ± 12.3 fg C $cell^{-1}$)
475 and nitrogen (5.8 ± 1.5 fg C $cell^{-1}$) contents of surface coastal bacterial assemblages; Fukuda *et al*., 1998]. Similar
476 calculations, done using the higher cellular abundance of snow sampled from the SW GrIS region ($2.6 \times 10^4$ cells $ml^{-1}$),
477 estimated the annual cellular content of GrIS surface meltwater fluxes to be ~ $1.0 \times 10^{22}$ cells $yr^{-1}$, with a carbon
478 equivalent of $3.14 \times 10^2 \pm 1.28 \times 10^2$ Mg C $yr^{-1}$, and a nitrogen equivalent of 60.3 ± 15.6 Mg N $yr^{-1}$. These assemblies of
479 cells will likely have little impact on the surrounding biologically and nutritionally rich open ocean environments, where
480 biomass equivalents are calculated to be found in 0.8 – 5.6 $km^3$ of Arctic Ocean waters (based on marine abundance
481 measurements by Bowman *et al.*, 2012). However, microbiota originating from GrIS surface snow may, nonetheless,
482 provide valuable nutritional and genetic resources to biotic niches within more proximal downstream englacial,
483 subglacial, periglacial, estuary and coastline environments.
484
485 In summary, the sampled GrIS surface snow environment was found to contain distinct and diverse biotic assemblages,
486 containing bacteria, eukaryotes and archaea; with strong representation from bacterial *Alpha*-, *Beta*- and
487 *Gammaproteobacteria,* and eukaryotic *Alveolata*, *Fungi*, *Stramenopiles* and *Chloroplastida*. Snow biota resembled soil
488 sampled assemblages, suggesting that these environments are predominantly seeded by wind transported terrestrial
489 sources. The fate of these microorganisms could result in their embedment into the GrIS during snow accumulation, or
490 alternatively they could be relocated away from the GrIS surface through melt and wind processes. As a warming

491 climate continues to increase GrIS surface melt rates annually (Hanna *et al.*, 2008, Bamber *et al.*, 2012), understanding
492 the ecological composition and functionality of the snow environment, and deciphering the impacts of biotic processes
493 on downstream environments, is necessary for establishing its biogeochemical role in polar ecology.
494

495 *Acknowledgments*

496 This work was funded by a National Science Foundation grant (NSF - PLR 1023462, 1023233, 1023351), awarded to K.
497 Junge, B. Christner, B. Hagedorn and R. Sletten; B. Christner was also partially supported by funds from NSF 1241161.
498 Fieldwork in Greenland was supported by Polar Field Services (Colorado, USA), and Lu Liu, Zoe Harrold, Laura Quante
499 and Samuel Foredyce provided field assistance. Shelly Carpenter assisted with abundance analysis and logistical support.
500

501 *Conflict of Interest*

502 The authors declare no conflict of interest.
503

504 *References*

505 Aebly, F.A., Fritz, S.C. (2009) Palaeohydrology of Kangerlussuaq (Søndre Strømfjord), West Greenland during the last ~
506 8000 years. The Holocene 19: 91-104.

507 Aller, J.Y., Kuznetsova, M.R., Jahns, C.J., Kemp, P.F. (2005) The sea surface microlayer as a source of viral and bacterial
508 enrichment in marine aerosols. J Aerosol Sci 36: 801-812.

509 Amaral-Zettler, L.A., McCliment, E.A., Ducklow, H.W., Huse, S.M. (2009) A method for studying protistan diversity using
510 massively parallel equencing of v9 hypervariable regions of small-subunit ribosomal RNA genes. PLoS ONE 4: e6372.

511 Amato, P., Hennebelle, R., Magand, O., Sancelme, M., Delort, A-M., Barbante, C., *et al.* (2007) Bacterial characterization
512 of the snow cover at Spitzberg, Svalbard. FEMS Microbiol Ecol 59: 255-264.

513 Bachy, C., López-García, P., Vereshchaka, A., Moreira, D. (2011) Diversity and vertical distribution of microbial
514 eukaryotes in the snow, sea ice and seawater near the North Pole at the end of the polar night. Front Microbiol 2: 106

515 Bamber, J., van den Broeke, M., Ettema, J., Lenaerts, J., Rignot, E. (2012) Recent large increases in freshwater fluxes from
516 Greenland into the North Atlantic. Geophys Res Lett 39: L19501.

517 Bowers, R.M., Lauber, C.L., Wiedinmyer, C., Hamady, M., Hallar, A.G., Fall, R., *et al.* (2009) Characterization of airborne
518 microbial communities at a high-elevation site and their potential to act as atmospheric ice nuclei. Appl Environ
519 Microbiol 75: 5121-5130.

520 Bowman, J.S., Rasmussen, S., Blom, N., Deming, J.W., Rysgaard, S., Sicheritz-Ponten, T. (2012) Microbial community
521 structure of Arctic multiyear sea ice and surface seawater by 454 sequencing of the 16S RNA gene. ISME J 6: 11-20.

522 Bunge, J. (2011) Estimating the number of species with CatchAll. Biocomputing 121-130.

523 Burrows, S.M., Butler, T., Jöckel, P., Tost, H., Kerkweg, A., Pöschl, U., *et al*. (2009) Bacteria in the global atmosphere –
524 part 2: Modeling of emissions and transport between different ecosystems. Atmos. Chem. Phys. 9: 9281-9297.

525 Cameron, K.A., Hodson, A.J., Osborn, A.M. (2012) Structure and diversity of bacterial, eukaryotic and archaeal
526 communities in glacial cryoconite holes from the Arctic and the Antarctic. FEMS Microbiol Ecol 82: 254-67.

527 Caporaso, J.G., Bittinger, K., Bushman, F.D., DeSantis, T.Z., Andersen, G.L., Knight, R. (2010a) Pynast: A flexible tool for
528 aligning sequences to a template alignment. Bioinformatics 26: 266-267.

529 Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., *et al.* (2010b) Qiime allows
530 analysis of high-throughput community sequencing data. Nature Methods 7: 335-336.

531 Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., *et al.* (2011) Global patterns
532 of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci U S A 108: 4516-4522.

533 Carpenter, E.J., Lin, S., Capone, D.G. (2000) Bacterial activity in South Pole snow. Appl Environ Microbiol 66: 4514-4517.

534 Christner, B.C., Cai, R., Morris, C.E., McCarter, K.S., Foreman, C.M., Skidmore, M.L., *et al.* (2008) Geographic, seasonal,
535 and precipitation chemistry influence on the abundance and activity of biological ice nucleators in rain and snow. Proc
536 Natl Acad Sci U S A 105: 18854-18859.

537 Collins, R.E., Carpenter, S.D. and Deming, J.W. (2008) Spatial heterogeneity and temporal dynamics of particles, bacteria,
538 and pEPS in Arctic winter sea ice. J Marine Syst 74: 902-917.

539 Comiso, J.C. (2006) Arctic warming signals from satellite observations. Weather 61: 70-76.

540 DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K*., et al.* (2006) Greengenes, a chimera-checked
541 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 72: 5069-5072.

542 Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26: 2460-2461.

543 Edwards, A., Pachebat, J.A., Swain, M., Hegarty, M., Hodson, A.J., Irvine-Fynn, T.D*., et al.* (2013) A metagenomic
544 snapshot of taxonomic and functional diversity in an alpine glacier cryoconite ecosystem. Environ Res Lett 8: 035003.

545 Felip, M., Sattler, B., Psenner, R., Catalan, J. (1995) Highly active microbial communities in the ice and snow cover of high
546 mountain lakes. Appl Environ Microbiol 61: 2394-401.

547 Fettweis, X. (2007) Reconstruction of the 1979–2006 Greenland Ice Sheet surface mass balance using the regional
548 climate model MAR. The Cryosphere Discuss 1: 123-168.

549 Fierer, N., Liu, Z., Rodríguez-Hernández, M., Knight, R., Henn, M., Hernandez, M.T. (2008) Short-term temporal variability
550 in airborne bacterial and fungal populations. Appl Environ Microbiol 74: 200-207.

551 Finlay, B.J., Fenchel, T. (2004) Cosmopolitan metapopulations of free-living microbial eukaryotes. Protist 155: 237-244.

552 Fukuda, R., Ogawa, H., Nagata, T., Koike, I. (1998) Direct determination of carbon and nitrogen contents of natural
553 bacterial assemblages in marine environments. Appl Environ Microbiol 64: 3352-3358.

554 Haas, B.J., Gevers, D., Earl, A.M., Feldgarden, M., Ward, D.V., Giannoukos, G*., et al.* (2011) Chimeric 16S rRNA sequence
555 formation and detection in Sanger and 454-pyrosequenced PCR amplicons. Genome Res 21: 494-504.

556 Hanna, E., Huybrechts, P., Steffen, K., Cappelen, J., Huff, R., Shuman, C*., et al.* (2008) Increased runoff from melt from
557 the Greenland Ice Sheet: A response to global warming. J Climate 21: 331-341.

558 Harding, T., Jungblut, A.D., Lovejoy, C., Vincent, W.F. (2011) Microbes in High Arctic snow and implications for the cold
559 biosphere. Appl Environ Microbiol 77: 3234-3243.

560 Hartmann, M., Howes, C.G., Veldre, V., Schneider, S., Vaishampayan, P.A., Yannarell, A.C., Quince, C., Johansson, P.,
561 Björkroth, K.J. and Abarenkov, K. (2011) V-REVCOMP: Automated high ‐ throughput detection of reverse

562 complementary 16S rRNA gene sequences in large environmental and taxonomic datasets. FEMS Microbiol Letters 319:
563 140-145.

564 Hell, K., Edwards, A., Zarsky, J., Podmirseg, S.M., Girdwood, S., Pachebat, J.A., *et al.* (2013) The dynamic bacterial
565 communities of a melting High Arctic glacier snowpack. ISME J 7: 1814-1826.

566 Hodson, A.J., Mumford, P.N., Kohler, J., Wynn, P.M. (2005) The High Arctic glacial ecosystem: New insights from nutrient
567 budgets. Biogeochem 72: 233-256.

568 Hodson, A.J., Anesio, A.M., Tranter, M., Fountain, A., Osborn, M., Priscu, J., *et al.* (2008) Glacial ecosystems. Ecol
569 Monographs 78: 41-67.

570 Huber, J.A., Mark Welch, D.B., Morrison, H.G., Huse, S.M., Neal, P.R., Butterfield, D.A., *et al.* (2007) Microbial population
571 structures in the deep marine biosphere. Science 318: 97-100.

572 Hughes, J.B., Hellmann, J.J., Ricketts, T.H., Bohannan, B.J.M. (2001) Counting the uncountable: Statistical approaches to
573 estimating microbial diversity. Appl Environ Microbiol 67: 4399-4406.

574 Irvine-Fynn, T.D.L., Edwards, A., Newton, S., Langford, H., Rassner, S.M., Telling, J., *et al.* (2012) Microbial cell budgets of
575 an Arctic glacier surface quantified using flow cytometry. Environ Microbiol 14: 2998-3012.

576 Joly, M., Attard, E., Sancelme, M., Deguillaume, L., Guilbaud, C., Morris, C.E., *et al.* (2013) Ice nucleation activity of
577 bacteria isolated from cloud water. Atmospheric Environ 70: 392-400.

578 Jones, H.G. (1999) The ecology of snow-covered systems: A brief overview of nutrient cycling and life in the cold.
579 Hydrological Process 13: 2135-2147.

580 Larose, C., Berger, S., Ferrari, C., Navarro, E., Dommergue, A., Schneider, D., *et al.* (2010) Microbial sequences retrieved
581 from environmental samples from seasonal Arctic snow and meltwater from Svalbard, Norway. Extremophiles 14: 205-
582 212.

583 Larose, C., Aurélien, D., Timothy, M.V. (2013a) Microbial nitrogen cycling in Arctic snowpacks. Environ Res Lett 8: 035004.

584 Larose, C., Dommergue, A., Vogel, T. (2013b) The dynamic Arctic snow pack: An unexplored environment for microbial
585 diversity and activity. Biol 2: 317-330.

586 Larose, C., Prestat, E., Cecillon, S., Berger, S., Malandain, C., Lyon, D., Ferrari, C., Schneider, D., Dommergue, A. and Vogel,
587 T.M. (2013c) Interactions between snow chemistry, mercury inputs and microbial population dynamics in an arctic
588 snowpack. PloS one 8 e79972.

589 Larsen, K.S., Grogan, P., Jonasson, S., Michelsen, A. (2007) Respiration and microbial dynamics in two subarctic
590 ecosystems during winter and spring thaw: Effects of increased snow depth. Arct Antarct Alp Res 39: 268-276.

591 Liu, Y., Yao, T., Jiao, N., Kang, S., Xu, B., Zeng, Y.*, et al.* (2009) Bacterial diversity in the snow over Tibetan plateau glaciers.
592 Extremophiles 13: 411-423.

593 Lohmann, U., Feichter, J. (2005) Global indirect aerosol effects: A review. Atmospheric Chem Phys 5: 715-737.

594 Lopatina, A., Krylenkov, V., Severinov, K. (2013) Activity and bacterial diversity of snow around russian antarctic stations.
595 Res Microbiol 164: 949-958.

596 Möhler, O., DeMott, P.J., Vali, G., Levin, Z. (2007) Microbiol and atmospheric processes: The role of biological particles in
597 cloud physics. Biogeosci 4: 1059-1071.

598 Møller, A.K., Søborg, D.A., Al-Soud, W.A., Sørensen, S.J., Kroer, N. (2013) Bacterial community structure in High-Arctic
599 snow and freshwater as revealed by pyrosequencing of 16S rRNA genes and cultivation. Polar Res 32: 17390

600 Nghiem, S.V., Hall, D.K., Mote, T.L., Tedesco, M., Albert, M.R., Keegan, K.*, et al.* (2012) The extreme melt across the
601 Greenland Ice Sheet in 2012. Geophys Res Lett 39: L20502

602 Nozomu, T. (2013) Seasonal and altitudinal variations in snow algal communities on an Alaskan glacier (Gulkana Glacier
603 in the Alaska range). Environ Res Lett 8: 035002.

604 Pratt, K.A., DeMott, P.J., French, J.R., Wang, Z., Westphal, D.L., Heymsfield, A.J., *et al.* (2009) In situ detection of
605 biological particles in cloud ice-crystals. Nature Geosci 2: 398-401.

606 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P.*, et al.* (2013) The SILVA ribosomal RNA gene database
607 project: Improved data processing and web-based tools. Nucleic Acids Res 41: D590-D596.

608 Rousseau, D-D., Schevin, P., Duzer, D., Cambon, G., Ferrier, J., Jolly, D.*, et al.* (2005) Pollen transport to Southern
609 Greenland: New evidences of a late spring long distance transport. Biogeosci Discuss 2: 829-847.

610 Sattler, B., Puxbaum, H., Psenner, R. (2001) Bacterial growth in supercooled cloud droplets. Geophys Res Lett 28: 239-
611 242.

612 Schütte, U.M., Abdo, Z., Bent, S.J., Williams, C.J., Schneider, G.M., Solheim, B., *et al.* (2009) Bacterial succession in a
613 glacier foreland of the High Arctic. ISME J 3: 1258-68.

614 Segawa, T., Miyamoto, K., Ushida, K., Agata, K., Okada, N., Kohshima, S. (2005) Seasonal change in bacterial flora and
615 biomass in mountain snow from the Tateyama Mountains, Japan, analyzed by 16S rRNA gene sequencing and real-time
616 PCR. Appl Environ Microbiol 71: 123-130.

617 Serreze, M.C., Barrett, A.P., Slater, A.G., Woodgate, R.A., Aagaard, K., Lammers, R.B., *et al.* (2006) The large-scale
618 freshwater cycle of the Arctic. J Geophys Res: Oceans 111: C11010.

619 Smith, D.J., Timonen, H.J., Jaffe, D.A., Griffin, D.W., Birmele, M.N., Perry, K.D., *et al.* (2013) Intercontinental dispersal of
620 bacteria and archaea by transpacific winds. Appl Environ Microbiol 79: 1134-1139.

621 Tedesco, M. (2007) A new record in 2007 for melting in Greenland. EOS, Trans American Geophys Union 88: 383.

622 Telling, J., Anesio, A.M., Tranter, M., Irvine-Fynn, T., Hodson, A., Butler, C., *et al.* (2011) Nitrogen fixation on Arctic
623 glaciers, Svalbard. J Geophys Res: J Geophys Res 116: G03039.

624 Telling, J., Stibal, M., Anesio, A.M., Tranter, M., Nias, I., Cook, J., *et al.* (2012) Microbial nitrogen cycling on the Greenland
625 Ice Sheet. Biogeosci 9: 2431-2442.

626 Thomas, W.H., Duval, B. (1995) Sierra Nevada, California, U.S.A., snow algae: Snow albedo changes, algal-bacterial
627 interrelationships, and ultraviolet radiation effects. Arct Alp Res 27: 389-399.

628 Torsvik, V., Øvreås, L., Thingstad, T.F. (2002) Prokaryotic diversity--magnitude, dynamics, and controlling factors. Science
629 296: 1064-1066.

630 Udisti, R., Becagli, S., Castellano, E., Traversi, R., Vermigli, S., Piccardi, G. (1999) Sea-spray and marine biogenic seasonal
631 contribution to snow composition at Terra Nova Bay, Antarctica. Annals Glaciol 29: 77-83.

632 van den Broeke, M., Smeets, P., Ettema, J., Munneke, P.K. (2008) Surface radiation balance in the ablation zone of the
633 West Greenland Ice Sheet. J Geophys Res: Atmospheres 113: D13105.

634 Vavrus, S. (2007) The role of terrestrial snow cover in the climate system. Climate Dynamics 29: 73-88.

635 V. Wintzingerode, F., Göbel, U.B. and Stackebrandt, E. (1997) Determination of microbial diversity in environmental

636 samples: Pitfalls of PCR-based rRNA analysis. FEMS Microbiol. Rev. 21: 213-229.

637 Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R. (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences

638 into the new bacterial taxonomy. Appl Environ Microbiol 73: 5261-5267.

639 Weidick, A. (1995) Greenland. In Satellite image atlas of glaciers of the world. Williams, R.S., Jr, and Ferrigno, J.G. (eds).

640 US Geol. Surv. Prof. Pap. 1386-C, C1–C93.

641 Wemheuer, B., Wemheuer, F., Daniel, R. (2012) RNA-based assessment of diversity and composition of active archaeal

642 communities in the German Bight. Archaea 2012: 695826

643 Wynn, P.M., Hodson, A.J., Heaton, T.H.E., Chenery, S.R. (2007) Nitrate production beneath a High Arctic glacier, Svalbard.

644 Chem Geology 244: 88-102.

645 Yoshimura, Y., Kohshima, S., Ohtani, S. (1997) A community of snow algae on a Himalayan glacier: Change of algal

646 biomass and community structure with altitude. Arct Alp Res 29: 126-137.

647 Zheng, Z., Schwartz, S., Wagner, L., Miller, W. (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol 7:

648 203-14.

649

650 *Table and Figure Legends*

651 Table 1 - Location and biochemical details of samples.

652 Table 2 - Details of studies included in biogeographical analysis. n.p. - not published.

653 Figure 1 - Map of sampling sites. Crosses indicate biotic assemblages that were sampled within this current study. Details

654 of samples can be found within Table 1. Circles indicate the sampling sites of sequencing studies performed by other

655 authors, which were utilized within this current analysis. Details of referenced studies can be found within Table 2.

656 Dotted line represents the Arctic Circle.

657 Figure 2 - Relative abundance of class level bacterial taxonomies based on PCR amplifications of 16S rRNA gene
658 sequences.

659 Figure 3 - Relative abundance of genus level archaeal taxonomies based on PCR amplifications of 16S rRNA gene
660 sequences, using V6-major and V6-minor primers sets with pooled extractions from NW and SW snow transect samples.

661 Figure 4 - Relative abundance of kingdom [phylum / subkingdom] eukaryotic taxonomies based on PCR amplifications of
662 18S rRNA gene sequences, using euk1 and euk2 primers sets with pooled extractions from NW and SW snow transect
663 samples.

664 Figure 5 - Multidimensional scaling plot of square-root transformed Bray-Curtis indices of OTU profiles from multiple
665 compiled studies of differing environmental origins. Circles represent frozen freshwater samples. Triangles represent all
666 other environmental sample types. Letters represent the author of each study, as detailed in Table 2. C - NW and C - SW
667 indicates samples from the current study. Stress factor; 0.16.

668 Figure 6 - SIMPER calculated percentage contribution of order level taxonomies to similarities between A) NW and B) SW
669 bacterial GrIS snow transect assemblages and bacterial communities of other environmental origins. Taxa without >
670 0.5 % contribution in at least one profile are not shown.

Table 1 - Location and biochemical details of samples

| Sample | Sample type | Region | Location | Sample date | pH | EC µS $cm^{-1}$ | $Na^+$ µg $L^{-1}$ | $Cl^-$ µg $L^{-1}$ | $Mg^{2+}$ µg $L^{-1}$ | $K^+$ µg $L^{-1}$ | $Ca^{2+}$ µg $L^{-1}$ | $F^-$ µg $L^{-1}$ | $SO_4^{2-}$ µg $L^{-1}$ | $NO_3^-$ µg $L^{-1}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NW.1 | Snow transect | NW Greenland - Thule | 76.44 N, 67.85 W | 9 June 2011 | 5.6 | 2.4 | 16.5 | - | - | - | 14.0 | 2.5 | 195.9 | 72.9 |
| NW.2 | Snow transect | NW Greenland - Thule | 76.48 N, 68.10 W | 9 June 2011 | 5.2 | 3.3 | 157.9 | 283.3 | 16.9 | 7.3 | 3.6 | 2.1 | 184.3 | 116.2 |
| NW.3 | Snow transect | NW Greenland - Thule | 76.51 N, 68.18 W | 9 June 2011 | 5.4 | 3.7 | 260.2 | 500.3 | 26.4 | 10.3 | 19.5 | 2.1 | 150.5 | 67.3 |
| NW.4 | August snow | NW Greenland - Thule | 76.46 N, 67.93 W | 31 August 2011 | 5.4 | 2.0 | 39.9 | 156.4 | 8.4 | 4.3 | 12.4 | 2.1 | 94.2 | 41.5 |
| NW.IM | Ice margin snow | NW Greenland - Thule | 76.53 N, 68.19 W | 13 June 2011 | - | - | 468.4 | 1114.5 | 131.9 | 31.0 | 270.2 | 298.6 | 991.9 | 519.1 |
| SW.1 | Snow transect | SW Greenland - Kangerlussuaq | 67.01 N, 48.62 W | 18 May 2012 | 5.3 | 2.4 | 22.0 | 138.2 | 4.9 | 6.2 | 22.4 | 2.9 | 109.9 | 126.9 |
| SW.2 | Snow transect | SW Greenland - Kangerlussuaq | 67.12 N, 49.37 W | 18 May 2012 | 5.1 | 2.3 | 26.4 | 151.4 | 5.1 | 10.7 | 23.3 | 3.3 | 139.8 | 175.9 |
| SW.3 | Snow transect | SW Greenland - Kangerlussuaq | 67.23 N, 49.89 W | 18 May 2012 | 5.6 | 1.7 | - | 16.4 | - | - | - | 2.3 | 74.1 | 74.1 |
| SW.2.lake | Supraglacial lake | SW Greenland - Kangerlussuaq | 67.12 N, 49.37 W | 18 May 2012 | - | - | - | - | - | - | - | - | - | - |
| SW.3.lake | Supraglacial lake | SW Greenland - Kangerlussuaq | 67.23 N, 49.89 W | 18 May 2012 | 5.2 | 7.3 | 264.1 | 488.8 | 37.6 | 42.9 | 72.1 | - | 378.4 | 300.9 |

- indicates data not available

1 Table 2 - Details of studies included in biogeographical analysis. n.p. - not published.

2

| Environment | Region | Location | Study accession number | Reference | Key |
|---|---|---|---|---|---|
| **Snow** | Colorado, USA | 40.5N 108.7W | Qiime database - Study 314 ** | Bowers *et al.*, 2009 | Bw |
| | Svalbard | 78.1N 15.4E | PRJEB1743 | Hell *et al.*, 2013 | H |
| | NE Greenland | 81.6N 16.7W | SRP003408 | Møller et al, 2013 | M |
| | **NW and SW Greenland** | **See Table 1** | **PRJEB4904** | **This study** | **C - NW, C - SW** |
| **Slush** | Svalbard | 78.1N 15.4E | PRJEB1743 | Hell *et al.*, 2013 | H |
| **Supraglacial Ice** | Svalbard | 78.1N 15.4E | PRJEB1743 | Hell *et al.*, 2013 | H |
| **Supraglacial Lake** | **SW Greenland** | **See Table 1** | **PRJEB4904** | **This study** | **C - SW** |
| **Marine Water** | High Arctic | 88.7N 158.9W, 88.5N 129.3W, 88.7N 58.5W | SRP006990 | Bowman *et al.*, 2012 | Bm |
| | NW Passage | 78.7N 104.9W | Earth Microbiome Project - Study 723* | n.p. (Plymouth Marine Laboratory) | P |
| | Canadian Basin | n.a. | SRP018324 | n.p. (Zhou) | Z |
| **Sea Ice** | High Arctic | 89.5N 129.3W, 88.7N 69.8W | SRP006990 | Bowman *et al.*, 2012 | Bm |
| | NW Passage | 78.7N 104.9W | Earth Microbiome Project - Study 723* | n.p. (Plymouth Marine Laboratory) | P |
| **Soil** | NW Greenland, Norway, Canadian Arctic | 76N 68W, 70N 19E, 63N 68W, 73N 78W, 82N 62W, 79N 90W | SRP017487 | Bell *et al.*, 2013 | Bl |
| | E Greenland | 74.5N 20.5W | Earth Microbiome Project - Study 1034* | n.p. (Gittel) | G |
| | Svalbard | 78.9N 11.8E | SRP002015 | Schütte *et al.*, 2009 | S |
| **Periglacial lake** | NE Greenland | 81.6N 16.6W | SRP003408 | Møller et al, 2013 | M |
| **Air** | Colorado, USA | 40.5N 108.7W | Qiime database - Study 314 ** | Bowers *et al.*, 2009 | Bw |
| | Antarctica | 78.1S 163.8E | PRJEB1657 | n.p. (International Centre for Terrestrial Antarctic Research) | I |

* http://www.microbio.me/emp/

** http://www.microbio.me/qiime/fusebox.psp

3

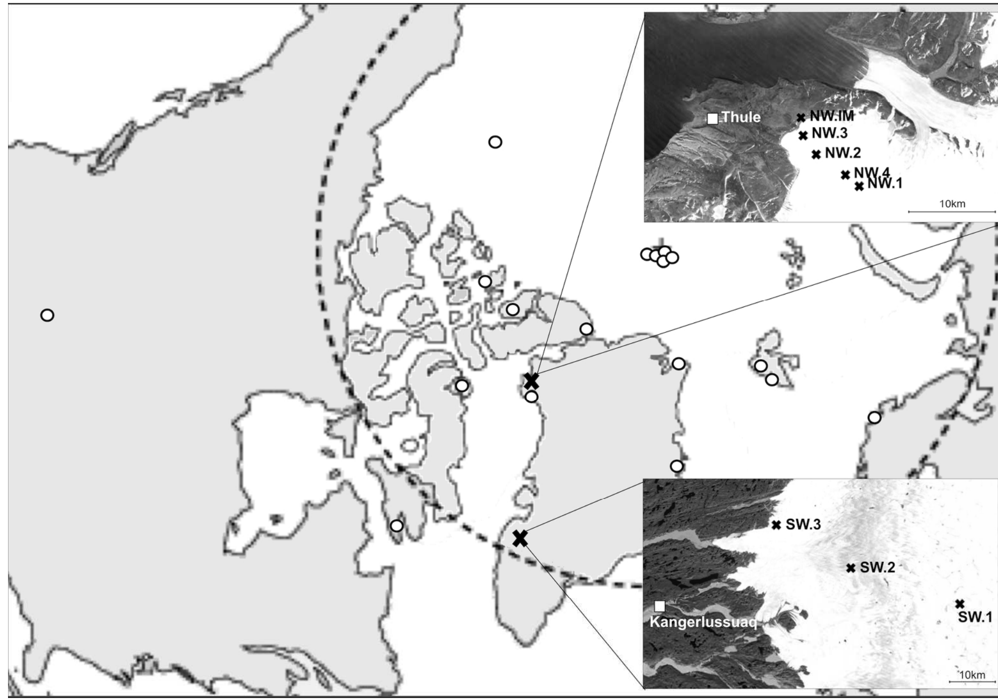



Figure 1 - Map of sampling sites. Crosses indicate biotic assemblages that were sampled within this current study. Details of samples can be found within Table 1. Circles indicate the sampling sites of sequencing studies performed by other authors, which were utilized within this current analysis. Details of referenced studies can be found within Table 2. Dotted line represents the Arctic Circle.
508x353mm (100 x 100 DPI)

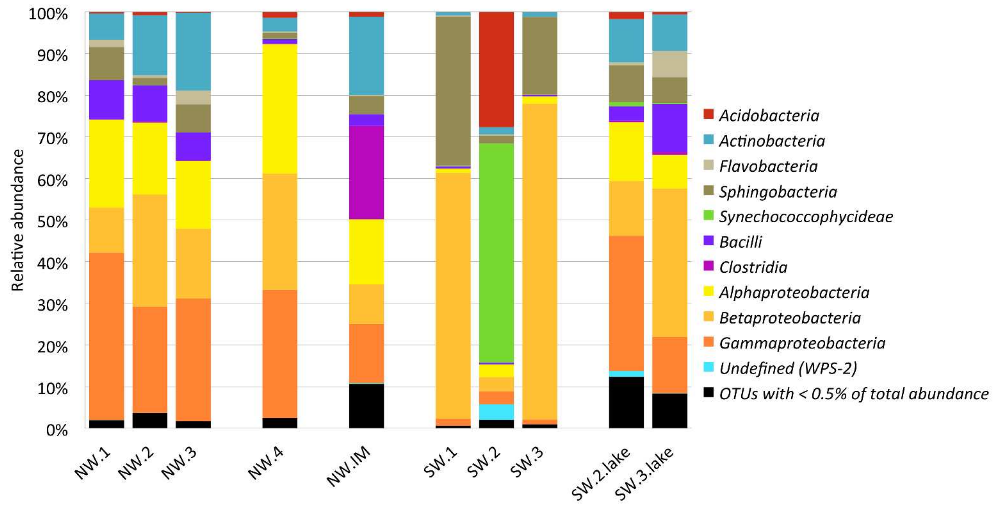



Figure 2 - Relative abundance of class level bacterial taxonomies based on PCR amplifications of 16S rRNA gene sequences.

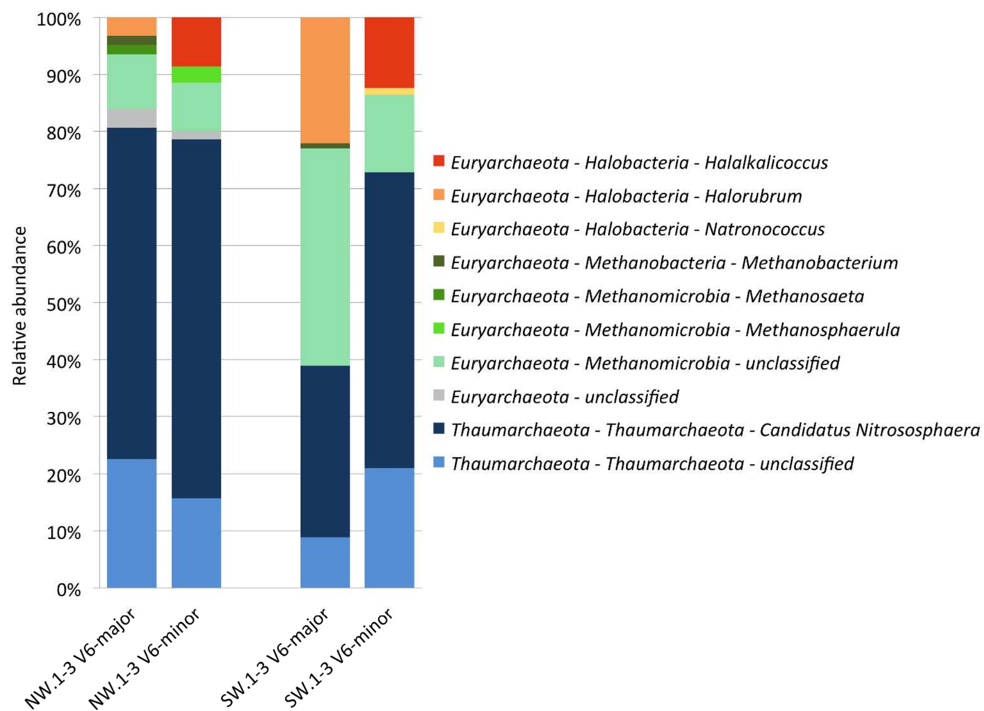



Figure 3 - Relative abundance of genus level archaeal taxonomies based on PCR amplifications of 16S rRNA gene sequences, using V6-major and V6-minor primers sets with pooled extractions from NW and SW snow transect samples.

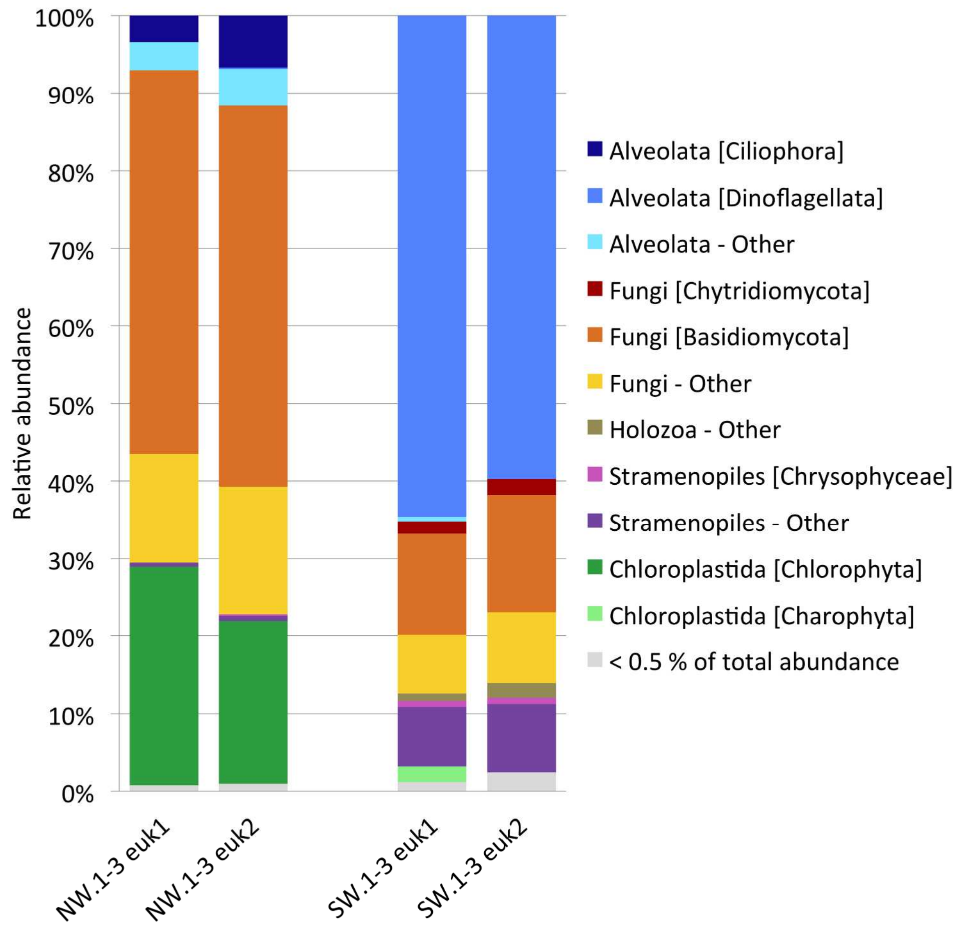



Figure 4 - Relative abundance of kingdom [phylum / subkingdom] eukaryotic taxonomies based on PCR amplifications of 18S rRNA gene sequences, using euk1 and euk2 primers sets with pooled extractions from NW and SW snow transect samples.

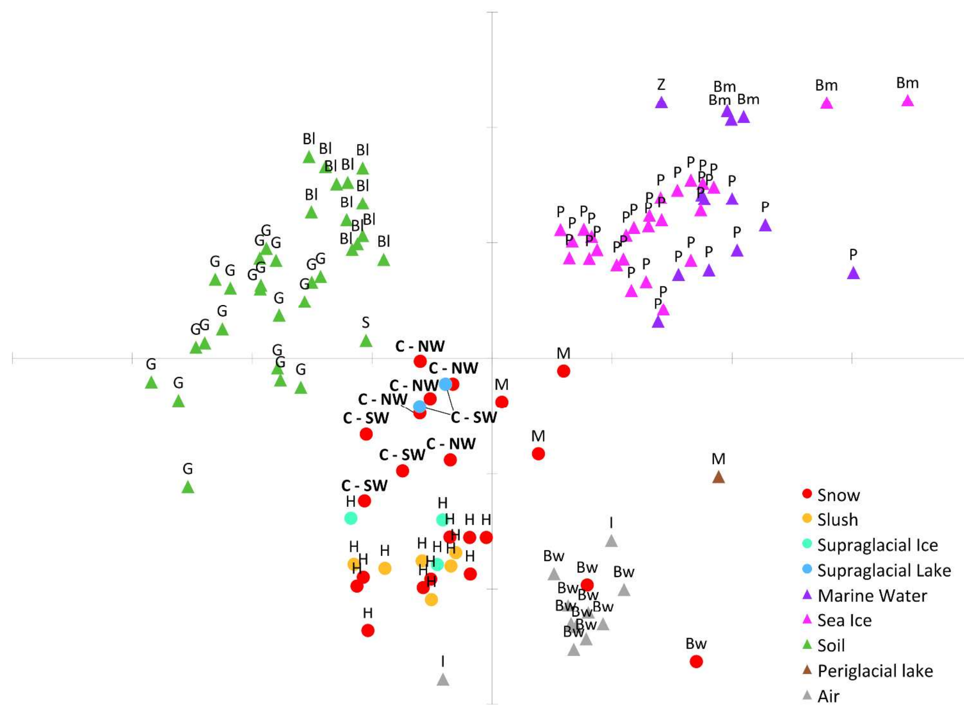



Figure 5 - Multidimensional scaling plot of square-root transformed Bray-Curtis indices of OTU profiles from multiple compiled studies of differing environmental origins. Circles represent frozen freshwater samples. Triangles represent all other environmental sample types. Letters represent the author of each study, as detailed in Table 2. C - NW and C - SW indicates samples from the current study. Stress factor; 0.16.

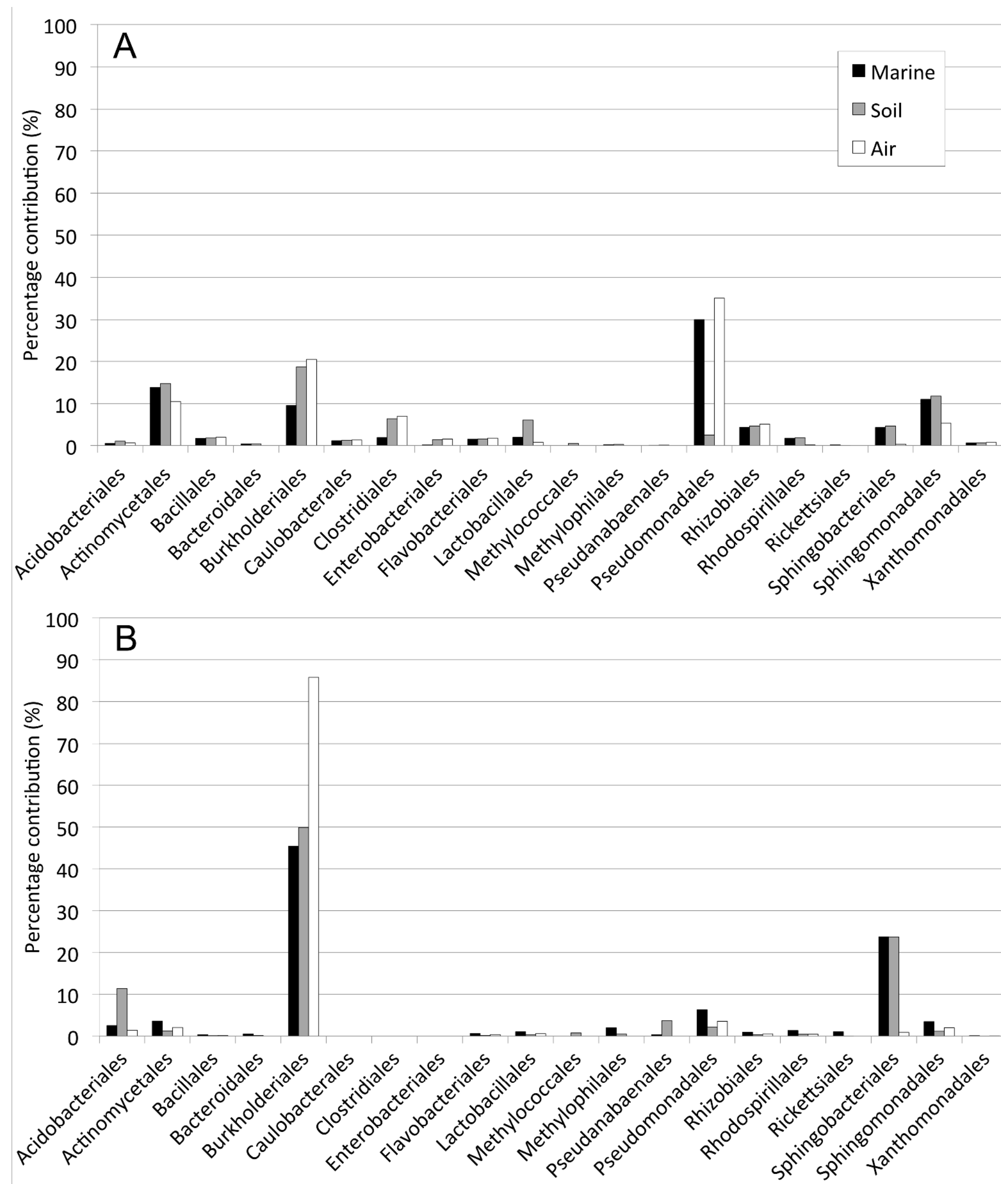



Figure 6 - SIMPER calculated percentage contribution of order level taxonomies to similarities between A) NW and B) SW bacterial GrIS snow transect assemblages and bacterial communities of other environmental origins. Taxa without > 0.5 % contribution in at least one profile are not shown.

1 *Supplemental Table 1*

2 Number of bacterial sequences per sample pre and post quality filtering, and CatchAll diversity estimates pre and post

3 profile rarefaction to 2404 sequences per sample

4

| Amplicon | Sequences per sample pre QF | Sequences per sample post QF | Percentage sequence loss through QF (%) | Diversity estimates pre rarefaction | Diversity estimates post rarefaction |
|---|---|---|---|---|---|
| NW.1 | 53620 | 35161 | 34.4 | 1669.9 | 1756.1 |
| NW.2 | 8410 | 5636 | 33.0 | 1209.9 | 1232.2 |
| NW.3 | 72559 | 56803 | 21.7 | 1457.6 | 464.6 |
| NW.4 | 61404 | 22268 | 63.7 | 1137.1 | 555.7 |
| NW.IM | 20410 | 8153 | 60.1 | 1349.9 | 1175.6 |
| SW.1 | 29468 | 25587 | 13.2 | 1235.8 | 531.9 |
| SW.2 | 14466 | 3077 | 78.7 | 532.2 | 357.4 |
| SW.3 | 19355 | 12901 | 33.3 | 1018.3 | 382.4 |
| SW.2.lake | 32271 | 11361 | 64.8 | 1661.1 | 852.4 |
| SW.3.lake | 77656 | 2404 | 96.9 | 836.9 | 843.7 |

5

1 *Supplemental Table 2*

2 Number of archaeal sequences per sample post quality filtering, and CatchAll diversity estimates pre and post profile

3 rarefaction to 70 sequences per sample

4

| Amplicon | Sequences per sample | Diversity estimates pre rarefaction | Diversity estimates post rarefaction |
|---|---|---|---|
| NW.1-3 V6-major | 70 | 38.3 | 38.3 |
| NW.1-3 V6-minor | 327 | 170.6 | 37.5 |
| SW.1-3 V6-major | 7638 | 229 | 16.5 |
| SW.1-3 V6-minor | 16752 | 365 | 35.6 |

5

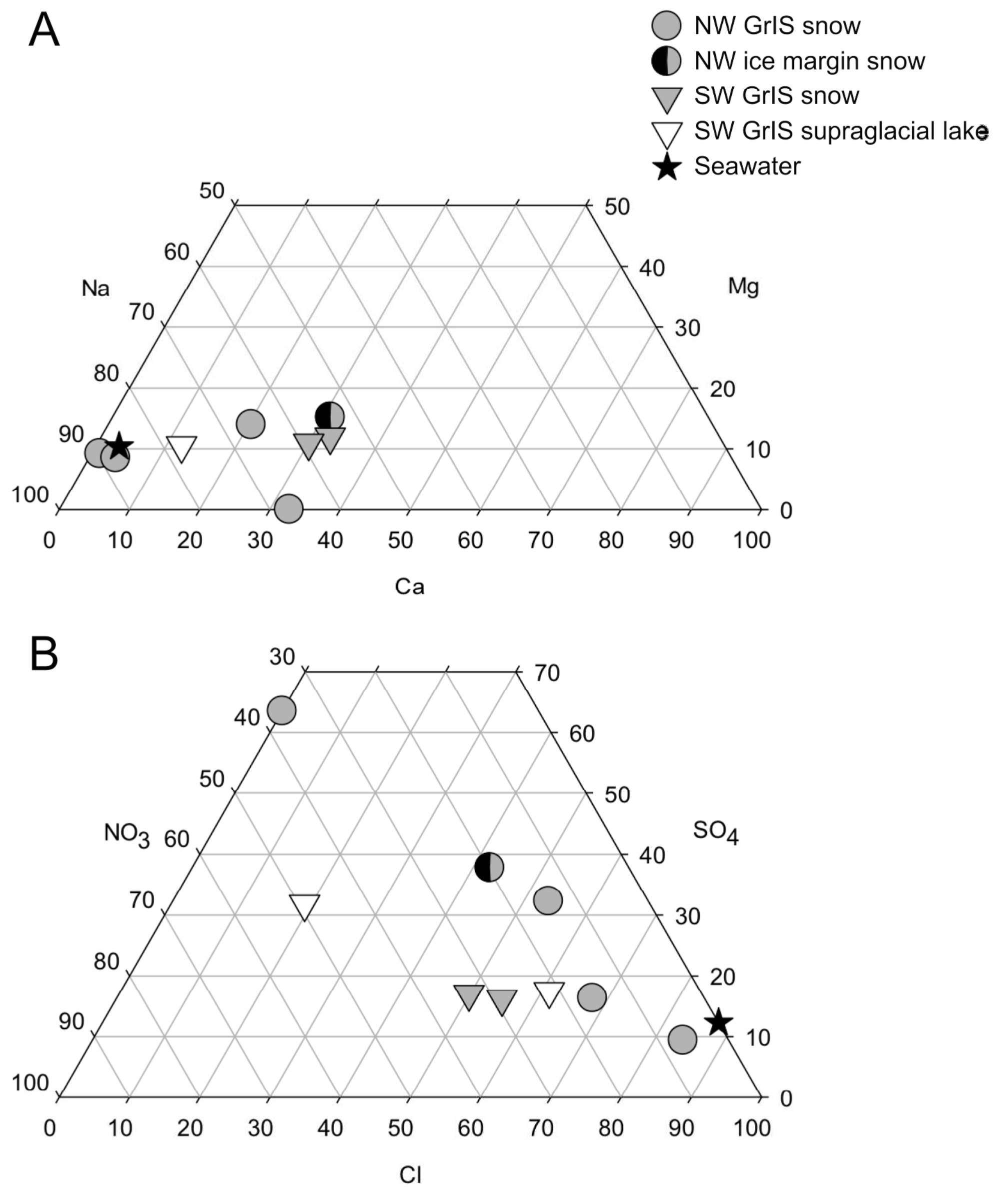
1 *Supplemental Table 3*

2 Pairwise ANOSIM tests of square root transformed Bray-Curtis similarities between snow bacterial OTU assemblages and
3 bacterial assemblages sampled from marine (marine water and sea ice), soil and air environments. Compiled snow
4 analyses are highlighted in gray and consisted of OTU assemblages from all five snow locations (4 snow studies). Results
5 with significance levels of ≤ 0.003 are shown in gray text.

6

| | | **Snow** | | | | | |
|---|---|---|---|---|---|---|---|
| | **Snow sampling location** | **Compiled snow** | **NW GrIS** | **SW GrIS** | **Svalbard** | **NE Greenland** | **Colorado** |
| | Reference | | This Study | This study | Hell *et al.* 2013 | Møller *et al*. 2013 | Bowers *et al.* 2009 |
| **Marine** | R | 0.919 | 0.919 | 0.975 | 0.989 | 0.931 | 0.999 |
| | Significance level | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| **Soil** | R | 0.689 | 0.475 | 0.720 | 0.856 | 0.656 | 0.932 |
| | Significance level | 0.001 | 0.001 | 0.001 | 0.001 | 0.002 | 0.002 |
| **Air** | R | 0.317 | 0.860 | 0.957 | 0.782 | 0.714 | 0.334 |
| | Significance level | 0.001 | 0.002 | 0.003 | 0.001 | 0.008 | 0.141 |

7




Supplemental Figure 1:
Percentage ratio composition of A) anions and B) cations sampled from NW and SW GrIS snow and supraglacial lakes. Symbols plotted as zero indicate data not available. Seawater sample is referenced from Stumm and Morgan (1996).

1 *Supplementary Table and Figure Legends*

2 *Supplemental Table 1*

3 Number of bacterial sequences per sample pre and post quality filtering, and CatchAll diversity estimates pre and post

4 profile rarefaction to 2404 sequences per sample

5

6 *Supplemental Table 2*

7 Number of archaeal sequences per sample post quality filtering, and CatchAll diversity estimates pre and post profile

8 rarefaction to 70 sequences per sample

9

10 *Supplemental Table 3*

11 Pairwise ANOSIM tests of square root transformed Bray-Curtis similarities between snow bacterial OTU assemblages and

12 bacterial assemblages sampled from marine (marine water and sea ice), soil and air environments. Compiled snow

13 analyses are highlighted in gray and consisted of OTU assemblages from all five snow locations (4 snow studies). Results

14 with significance levels of ≤ 0.003 are shown in gray text.

15

16 *Supplemental Figure 1*

17 Percentage ratio composition of A) anions and B) cations sampled from NW and SW GrIS snow and supraglacial lakes.

18 Symbols plotted as zero indicate data not available. Seawater sample is referenced from Stumm and Morgan (1996).

19