

Glacial microbiota are hydrologically connected and temporally variable

Karen A. Cameron ^{1,2,3,4*} Oliver Müller ⁵
Marek Stibal ^{2,3,6} Arwyn Edwards ¹ and
Carsten Suhr Jacobsen ^{2,3,7}

¹*Institute of Biological, Rural and Environmental Sciences (IBERS), Aberystwyth University, Aberystwyth, SY23 3DD, UK.*

²*Center for Permafrost (CENPERM), University of Copenhagen, Copenhagen, 1350, Denmark.*

³*Department of Geochemistry, Geological Survey of Denmark and Greenland (GEUS), Copenhagen, 1350, Denmark.*

⁴*School of Geographical and Earth Sciences, University of Glasgow, Glasgow, G12 8QQ, UK.*

⁵*Department of Biological Sciences, University of Bergen, Bergen, 5006, Norway.*

⁶*Department of Ecology, Faculty of Science, Charles University, Prague, 128 44, Czech Republic.*

⁷*Department of Environmental Science, Aarhus University, Roskilde, 4000, Denmark.*

Summary

Glaciers are melting rapidly. The concurrent export of microbial assemblages alongside glacial meltwater is expected to impact the ecology of adjoining ecosystems. Currently, the source of exported assemblages is poorly understood, yet this information may be critical for understanding how current and future glacial melt seasons may influence downstream environments. We report on the connectivity and temporal variability of microbiota sampled from supraglacial, subglacial and periglacial habitats and water bodies within a glacial catchment. Sampled assemblages showed evidence of being biologically connected through hydrological flowpaths, leading to a meltwater system that accumulates prokaryotic biota as it travels downstream. Temporal changes in the connected assemblages were similarly observed. Snow assemblages changed markedly throughout the sample period, likely reflecting changes in the

surrounding environment. Changes in supraglacial meltwater assemblages reflected the transition of the glacial surface from snow-covered to bare-ice. Marked snowmelt across the surrounding periglacial environment resulted in the flushing of soil assemblages into the riverine system. In contrast, surface ice within the ablation zone and subglacial meltwaters remained relatively stable throughout the sample period. Our results are indicative that changes in snow and ice melt across glacial environments will influence the abundance and diversity of microbial assemblages transported downstream.

Introduction

Climate warming is leading to changes in glacial environments around the world. Enhanced glacial melt is expanding ice sheet and large glacier microbial habitats, such as wet snow, exposed surface ice, hydrologically influenced subglacial environments and glacial forefields (e.g. Bradley *et al.*, 2014; Kohler *et al.*, 2017; Stibal *et al.*, 2017; Ryan *et al.*, 2019). Environmental conditions within these habitats are also changing and will continue to change, through increased temperatures, rainfall, cloud cover and water content (e.g. Vavrus *et al.*, 2009; Bintanja and Andry, 2017; Södergren *et al.*, 2017; Overland *et al.*, 2019), which may impact the form and function of *in situ* biological communities. Glacial meltwater exports microbial assemblages (Dubnick *et al.*, 2017; Cameron *et al.*, 2017b; Kohler *et al.*, 2020), sediments (Hudson *et al.*, 2014; Overeem *et al.*, 2017) and nutrients (Hawkings *et al.*, 2015) from glaciers to downstream environments where they have ecological impacts, such as seeding new ecosystems (e.g. Rime *et al.*, 2016) and fertilizing coastal productivity (e.g. Hood *et al.*, 2009). Increased glacial melt is expected to result in the elevated release of microbiota (Cameron *et al.*, 2017b), with potential consequences including altered microbial compositions of fjord surface waters (Gutiérrez *et al.*, 2015), increased heterotrophy in fjords (Paulsen *et al.*, 2017) and changes in methane cycling within buried fluvial sediments (Cameron *et al.*, 2017a).

In many cases, it is challenging to ascertain the sources of glacial meltwater microbiota. Arctic and Alpine

Received 10 January, 2020; revised 15 April, 2020; accepted 2 May, 2020. *For correspondence. E-mail kac32@aber.ac.uk; Tel. +44 (0) 141 330 4782.

supraglacial assemblages have long been considered to be inoculated from wind-borne particles, blown in from surrounding or distant environments (e.g. Nordenskiöld, 1872; Harding *et al.*, 2011; Edwards *et al.*, 2013; Cameron *et al.*, 2015; Franzetti *et al.*, 2017b). Studying this is, however, problematic due to the low biomass of polar air (Šanti-Temkiv *et al.*, 2018), the high spatial and temporal variability of borderless air bodies (Fierer *et al.*, 2008), and the selective glacial conditions that subsequently shape surface ice assemblages (Edwards *et al.*, 2013; Stibal *et al.*, 2015a). For subglacial systems, their comparative inaccessibility, and their vast and complex hydrological networks make studying microbial transport and transfer particularly challenging. As a solution, studies have resorted to identifying phototrophic assemblages that are typical of sunlit supraglacial environments within meltwater released from dark subglacial catchments (Cameron *et al.*, 2017b; Žárský *et al.*, 2018). While the presence of these phototrophs is a convincing indicator of supraglacial microbiota transfer, beyond this, these analyses provide limited insight into the spatial mixing and temporal export of glacial microbiota through melt events. Comparisons between supraglacial and subglacial microbial assemblages have also been performed in the hope of subtracting supraglacial assemblages from sampled subglacial meltwater outflows (Dieser *et al.*, 2014). However, this approach overlooks generalist species that might be present in both environments, and it assumes that the supraglacial assemblages sampled will be the same as those routed to, and then sampled from, the subglacial environment. Correlation between proglacial river water physicochemical measurements and their microbial composition has furthermore been performed in the hope of identifying increasingly subglacial microbiota throughout a melt season (Dubnick *et al.*, 2017). However, in these investigations, the stability of the proglacial microbial composition throughout the melt season highlighted the possibility that low abundance subglacial assemblages are drowned out by other more abundant sources. Beyond the proglacial environment, the complexity of contributing microbial inputs increases further, adding to the challenge of teasing the microbiology of meltwater sources apart (Hauptmann *et al.*, 2016).

In addition to spatial variations in meltwater inputs, contributing meltwater compositions can change with time. At the glacial surface within the Arctic and Alpine environments, this may be due to seasonal and annual changes in microbial composition (Stibal *et al.*, 2015a; Franzetti *et al.*, 2017a; Pittino *et al.*, 2018), especially when going from snow-covered to bare-ice supraglacial conditions (Musilova *et al.*, 2015). In the subglacial environment, an expanding drainage area throughout the melt season may also impact the composition of

microbiota that is released (Dubnick *et al.*, 2017; Kohler *et al.*, 2017). Downstream of the glacier, permafrost, lake and river microbial assemblages can all experience seasonal changes (e.g. Crump *et al.*, 2003; Schostag *et al.*, 2015; Cavaco *et al.*, 2019), adding a further layer of complexity to disentangling meltwater contributions.

In light of these challenges, this study aimed to investigate the interconnectivity and temporal variability of microbial assemblages within an Arctic glacial catchment that encompasses the Lyngmarksbræ ice cap on Disko Island, Greenland. The study was performed by comparing sequence profiles and qPCR calculated prokaryotic abundance estimations of microbial assemblages sampled at high spatial and temporal resolution (Fig. 1; Table S1). Sequence profiles were used as fingerprints to investigate the ecological connectivity of snow, exposed surface ice (termed 'ice' herein), dark aggregates of biotic and abiotic material found on the ice surface (termed 'ice debris' herein), and supraglacial, subglacial and periglacial meltwater throughout the melt season. Microbial assemblages from the air, soil and marine sources were furthermore considered for their potential as contributing sources. Our hypothesis was that microbial habitats are connected hydrologically through meltwater flows that are responsible for transporting temporally variable microbial assemblages downstream. Understanding the ecological impacts that transported glacial microbiota have on downstream ecosystems will provide insight into the ongoing and ever-increasing regional and global consequences of climate warming.

Results

Abundance analysis

Snow prokaryotic abundance tended to increase over the season (May to August 2015; see Table S1 for the dates that each sample type was available for collection); however, high spatial variability over meter-scale distances meant that this increase was not found to be significant (linear regression; $R^2 = 0.080$, $p = 0.069$, $n = 42$; Fig. 2). When snow samples were grouped into two groups by date (before and after the 16-day sampling break that occurred from June 25, day of year (DOY) 176 until July 11, DOY 192), a significant difference in abundance was found (*t*-test; $t = 2.389$, $p = 0.022$, $n = 27$ and 15). On average, snow was found to have the lowest prokaryotic abundance of all the sample types tested (mean over the whole sample season; $2.45 \times 10^3 \pm 9.31 \times 10^3$ copies ml^{-1} , $n = 42$). Ice samples, which were available for collection between July and August, were similarly found to be highly variable in their prokaryotic abundance, with on-day variability often ranging by over an order of

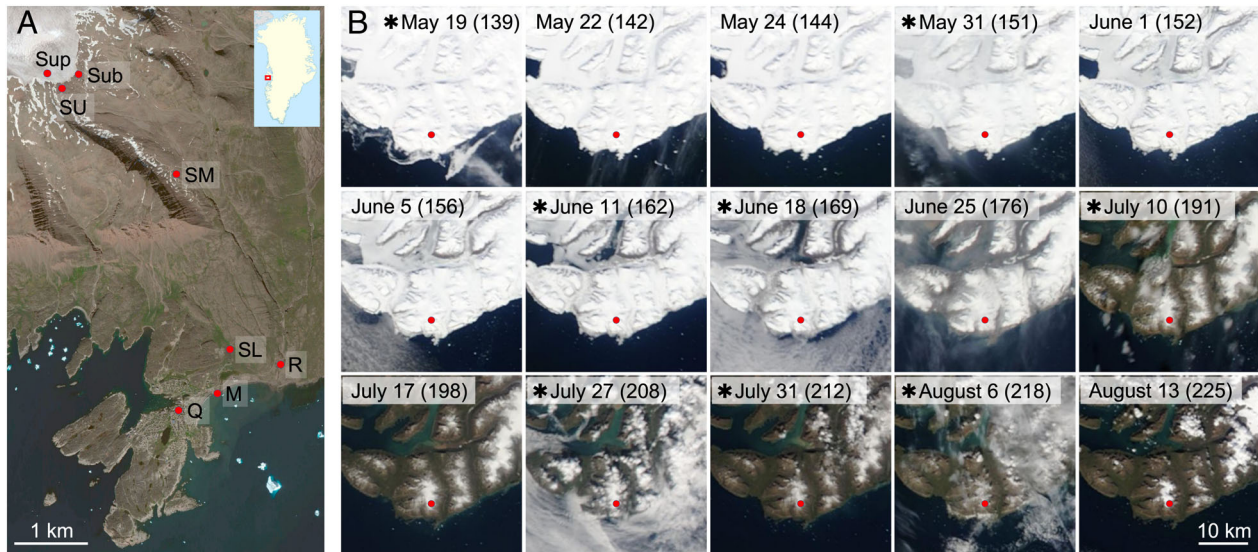


Fig 1. (A) Satellite image (Maxar Technologies, 2019, CO, USA) of Disko Island (inset map) sample site locations on July 31, 2015. Supraglacial (Sup) site encompasses snow, ice, supraglacial meltwater and air sample approximate locations. Subglacial (Sub), Red River (R), marine (M) and soil upper (SU), -middle (SM) and -lower (SL) sites are also marked. The town of Qeqertarsuaq (Q) is shown. (B) Terra and Aqua MODIS satellite images showing snow cover on glacial sampling days. Red circles mark the approximate supraglacial sample locations. *depicts that clear satellite imagery was not available on the day of sampling; therefore, the closest cloud-free day has been shown. Day of year 2015 for each satellite image is shown in brackets. [Color figure can be viewed at wileyonlinelibrary.com]

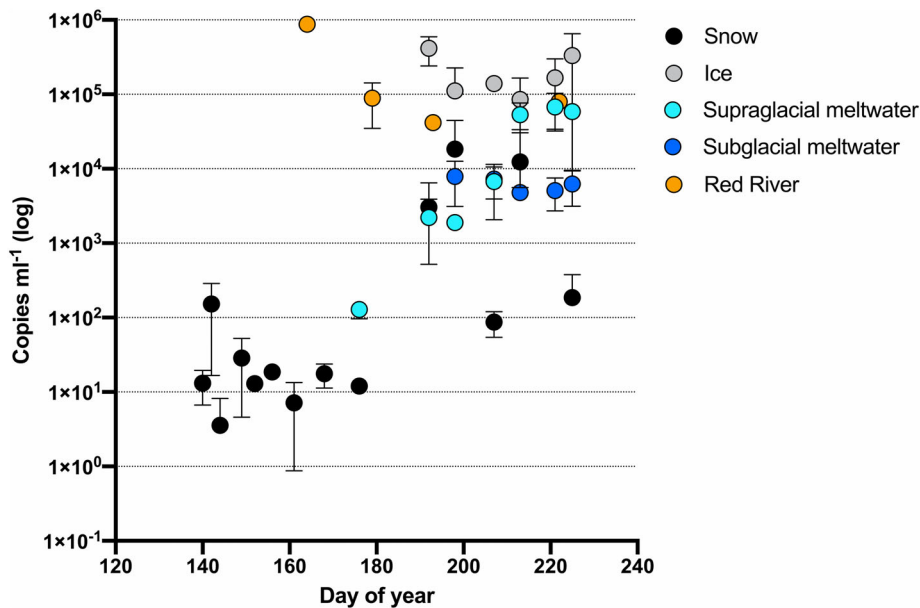


Fig 2. qPCR calculated 16S rRNA gene copy abundance of transect samples across the field season. Error bars show standard deviation. For each data point, $n = 3$, except for supraglacial meltwater on DOY 176, 198 and 207, and ice and snow on DOY 198, where $n = 2$. [Color figure can be viewed at wileyonlinelibrary.com]

magnitude. The average prokaryotic abundance in ice was $2.05 \times 10^5 \pm 1.85 \times 10^5$ copies ml^{-1} ($n = 19$), and no general increase in abundance was calculated by date (linear regression; $R^2 = 0.007$, $p = 0.733$, $n = 19$). Supraglacial meltwater prokaryotic abundance was found to increase over the sampling season (June to August; linear regression; $R^2 = 0.621$, $p = 0.0001$, $n = 18$). This started at $1.28 \times 10^2 \pm 3.20 \times 10^1$ copies ml^{-1} on June

25 and increased by three orders of magnitude to $5.87 \times 10^4 \pm 4.38 \times 10^3$ copies ml^{-1} by August 13 (DOY 225). In contrast, prokaryotic abundance in meltwater coming from the subglacial environment remained constant throughout the sample season (July to August), with a mean of $6.25 \times 10^3 \pm 2.93 \times 10^3$ copies ml^{-1} ($n = 15$), which was almost an order of magnitude lower than the final supraglacial meltwater samples that were taken.

Red River samples started at $8.74 \times 10^5 \pm 1.50 \times 10^5$ copies ml^{-1} on June 13, when discharge levels were low ($<10 \text{ m s}^{-1}$; Fig. S1), however, after this, abundance dropped by an order of magnitude for the rest of the sampling season. Ice debris and soil had several orders of magnitude greater 16S rRNA gene copy abundance than the glacial transect samples ($3.48 \times 10^8 \pm 1.44 \times 10^8$ copies ml^{-1} , $n = 3$, and $1.28 \times 10^9 \pm 1.46 \times 10^9$ copies ml^{-1} , $n = 9$ respectively) and marine samples had an average abundance of $2.31 \times 10^5 \pm 1.50 \times 10^5$ copies ml^{-1} ($n = 15$).

Sequence and operational taxonomic unit generation

After library clean up, and prior to SourceTracker contamination removal and rarefaction, a total of 2 185 687 sequences were available across 133 samples (excluding air samples and control samples). These sequences grouped into 126 465 operational taxonomic units (OTUs) (see Table S2 for the number of sequences per sample). The SourceTracker contamination removal step, which considered Sterivex control samples as potential sources of contamination, removed 456 850 potential contaminants, leaving a total sequence pool of 1 728 837 sequences (20.9% sequence reduction). After rarefaction to 2101 sequences per sample, six samples were discarded due to being below the sequence cut-off point, leaving 127 samples, with sequences grouping into a total of 53 905 OTUs. Of these, 1041 were from the GG dataset and 52 864 were *de novo* OTUs.

Species richness and alpha diversity

Catchall (which computes the best model to estimate species richness; Bunge, 2011), chao1 (a measure of species richness that gives weight to low abundance species) and Faith's Phylogenetic Diversity (which takes into account phylogenetic similarities between OTUs) analyses found that species richness and alpha diversity increased several fold along the hydrological flow path (Fig. 3). Shannon calculated diversity (which takes into account species evenness as well as richness) and the number of observed OTUs showed an increasing trend in diversity down the hydrological transect, with the exception of snow, which was found to be more diverse than ice, ice debris and supraglacial meltwater (Fig. 3). Mean catchall calculated richness in early- (DOY 140, 142, 144), mid- (DOY 149, 152, 156, 161, 168) and late- (DOY 176, 192, 198, 207, 213, 225) season snow samples were significantly different, with an increase in richness observed between early- and mid-season snow samples

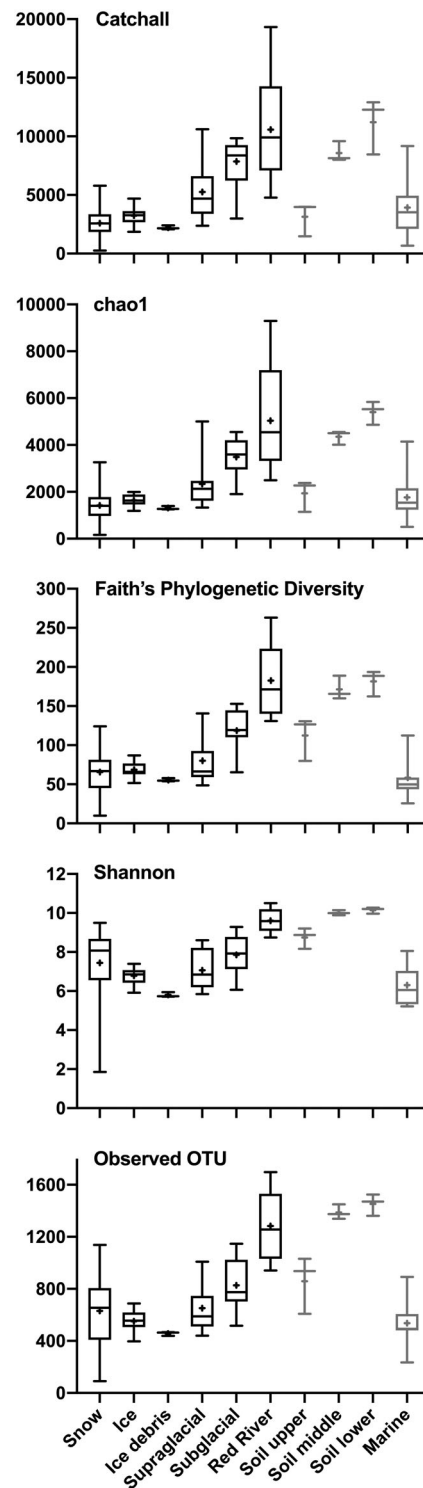


Fig 3. Box and whiskers plots of calculated diversity for each sample type throughout the season. Boxes extend from the 25th to the 75th percentile, whiskers extend from the minimum to the maximum value, box lines plot the median and crosses plot the mean. Grey data points denote sample groups not on the hydrological transect.

(Welch's *t*-test $t = 2.520$, $p = 0.0205$, $n = 9$ and 14), and a drop in richness between mid- and late-season snow samples (Welch's *t*-test $t = 3.047$, $p = 0.0049$, $n = 14$ and 19).

Data ordination

Unimodal ordination methods were used to identify patterns of OTU distribution and abundance from samples taken along the meltwater transect, and to predict the proportion of variability that is influenced by environmental factors. End member samples (from ice debris, soil and marine environments) were omitted from these analyses to avoid skewing the results. Detrended correspondence analysis (DCA) identified that there was a total variation of 20.04% in the data, and of this, only 12.36% could be explained within the first four axes. To identify the significance of potential factors that may be influencing changes in microbial assemblages, a canonical correspondence analysis (CCA) was performed using sample type (snow, ice, supraglacial meltwater, subglacial meltwater and Red River), sampling date (as DOY), 16S rRNA gene copy abundance and whether the samples were solid or liquid as the explanatory variables. Overall, there was found to be a low global total variation of 19.93% (pseudo- $F = 4.0$, $p = 0.002$ on the first axis). Interactive forward selection identified that explanatory variables accounted for 13.3% of this variation, with snow, ice and Red River sample type contributing towards 9.6% of the explained variation (pseudo- $F = 1.5$ – 3.8 , $p = 0.002$) and DOY accounting for 2.3% (pseudo- $F = 2.5$, $p = 0.002$). After early season samples (prior to DOY 192) were excluded, ensuring that sample group collection dates overlapped, a CCA analysis revealed that global variation was reduced to 12.30%, with 82.8% of this variation being attributed to sample type (pseudo- $F = 1.9$ – 4.0 , $p = 0.002$) and 16.6% of being attributed to DOY (pseudo- $F = 2.7$, $p = 0.002$).

Data comparison

Analysis of variance (ANOVA) in sample group means, relative to the variance of within-group means, revealed that samples, grouped by sample type along the melt transect, were significantly different from each other. This was true for Bray–Curtis dissimilarity analysis of square-root transformed data with down-weighted high abundance OTUs ($F(4,95) = 7.222$, $p = 0.001$), for weighted UniFrac dissimilarity analysis, which utilizes abundance-based phylogenetic distance relatedness ($F(4,95) = 15.185$, $p = 0.001$), and for unweighted UniFrac dissimilarity analysis, which utilizes presence–absence based phylogenetic distance relatedness ($F(4,95) = 6.024$, $p = 0.001$ respectively).

Analysis of similarities (ANOSIM) analyses of weighted UniFrac distances were used to calculate the differences between samples grouped by sample type (Table S3). Here, comparisons of assemblages that result in an R -value of 1 are considered to be highly dissimilar, R values of 0 are considered to have a low dissimilarity, and middle-range R values (0.3–0.7) are considered to be moderately dissimilar. The snow was found to be moderately dissimilar to downstream sample type groups along the hydrological transect ($R = 0.488$ – 0.620 , $p = 0.001$). Ice was found to be moderately dissimilar to the supraglacial meltwater sample group ($R = 0.528$, $p = 0.001$); however, dissimilarities to subglacial and Red River sample groups increased down the hydrological transect. Low dissimilarities were found between supraglacial and subglacial meltwater sample groups ($R = 0.176$, $p = 0.016$), whereas moderate dissimilarities were found between supraglacial or subglacial meltwater sample groups and the Red River sample group ($R = 0.566$, $p = 0.001$; $R = 0.646$, $p = 0.001$ respectively). Snow was found to have a moderate dissimilarity to soil ($R = 0.378$, $p = 0.002$); supraglacial meltwater, subglacial meltwater and Red River samples also had moderate dissimilarities to soil ($R = 0.583$ – 0.766 , $p = 0.001$); and ice samples were found to be highly dissimilar to soil ($R = 0.978$, $p = 0.001$). When the soil was split into upper, middle and lower sampled groups, supraglacial meltwater, subglacial meltwater and Red River groups all had the lowest dissimilarity to upper soils ($R = 0.394$, $p = 0.002$; $R = 0.636$, $p = 0.001$; $R = 0.744$, $p = 0.002$ respectively); however, caution should be taken when considering these analyses due to small soil group sizes ($n = 3$). Marine samples were found to be somewhat dissimilar to samples collected along the hydrological transect ($R = 0.713$ – 0.979 , $p = 0.001$). ANOSIM analysis of ice samples to ice debris samples showed that these sample types were moderately dissimilar ($R = 0.579$, $p = 0.001$).

As the 16S rRNA gene copy abundance of snow, supraglacial meltwater and Red River samples were temporally variable, and as CCA found that sample type followed by DOY contributed to the sampled assemblage composition, weighted UniFrac distances were used to determine whether each sample type should be subdivided into groups based on temporal assemblage shifts. Cluster analyses were used to visualize average weighted UniFrac distances between groups of samples (Fig. S2). Based on this, snow was sub-divided into three groups, namely early- (DOY 140, 142, 144), mid- (DOY 149, 152, 156, 161, 168) and late- (DOY 176, 192, 198, 207, 213, 225) season snow. Ice assemblages were found to be relatively similar; therefore, this group was not sub-divided. Supraglacial meltwater samples were subdivided into early- (DOY 176, 192) and late-season

groups (DOY 198, 207, 213), as were subglacial meltwater samples (DOY 198, 207, 213 and DOY 221, 225 respectively). Samples taken during the peak Red River discharge (DOY 179) were grouped separately from those that were taken during low flow. Late-season snow and late-season supraglacial samples could have been further subdivided into groups, however, late-season snow did not split logically by date and reduction of group size risks compromising the robustness of downstream analyses, therefore further subdivisions were not used.

ANOSIM analyses of weighted UniFrac distances were used to calculate differences between sub-divided sample type groups. These analyses produced several notable results (Table S4); however, with the reduced sample group sizes, caution must be placed on the robustness of these analyses. All snow sample groups were found to be moderately dissimilar to soil samples ($R = 0.562\text{--}0.758$, $p = 0.001$), and late snow had moderate dissimilarities to marine samples ($R = 0.679$, $p = 0.001$). Ice samples were found to have moderate dissimilarities to the late snow samples ($R = 0.441$, $p = 0.001$), both of which were sampled in July and August. However, strong dissimilarities were found between ice and the earlier snow sampled assemblages ($R = 0.965$, $p = 0.001$). Similarly, supraglacial and subglacial meltwater samples were calculated to have less dissimilarity to late snow samples than to the earlier snow samples. Early supraglacial meltwater samples were calculated to have low dissimilarities to late snow assemblages ($R = 0.291$, $p = 0.012$) and high dissimilarity to ice ($R = 1$, $p = 0.001$), whereas late supraglacial meltwater samples were moderately to highly dissimilar to both late snow and ice ($R = 0.690$, $p = 0.001$; $R = 0.720$, $p = 0.001$ respectively). Subglacial and supraglacial meltwater samples were moderately dissimilar to each other where sampling dates of the groups crossed ($R = 0.367\text{--}0.583$, $p = 0.001$). In contrast to this, early supraglacial meltwater and late subglacial meltwater samples, which were sampled during different months, were found to be highly dissimilar ($R = 0.992$, $p = 0.001$). Red River low flow samples were generally found to be unique; however, moderate dissimilarities between peak flow samples and late snow ($R = 0.413$, $p = 0.003$) and early supraglacial meltwater ($R = 0.497$, $p = 0.018$) were calculated. Two-way crossed ANOSIM analyses that considered sample type and DOY groupings were largely found to be non-significant, likely due to the small group sizes (data not shown).

OTUs that were shared between sample groups were used as an indicator of the potential transfer of assemblages along the hydrological transect. For this, two analyses were performed on samples that had the potential

to be hydrologically connected on or after DOY 192; one that included OTUs that contributed towards $\geq 0.1\%$ of the total sample group assembly, allowing for the transfer of dominant assemblages to be investigated (Fig. 4a), and one that included OTUs that contributed towards $< 0.1\%$ of the total sample group assembly, allowing for the transfer of the remaining rare biosphere to be considered (Fig. 4b). Each sample group had 87–109 dominant OTUs that contributed towards $\geq 0.1\%$ of the total group sequence profile. A large proportion of dominant supraglacial meltwater OTUs was similarly found within the sampled subglacial meltwater assemblages (53% of the dominant OTU supraglacial meltwater profile). High OTU homogeneity was also observed between the sampled snow and ice assemblages, with 43 OTUs being found within the dominant OTU profiles of both sample groups, equating to 49% of the dominant snow OTU profile and 39% of the dominant ice OTU profile. Supraglacial meltwater shared 28% and 27% of dominant snow and ice OTUs respectively. Lower OTU sharing was seen between the subglacial and the Red River dominant OTU profiles (16% of dominant subglacial OTU profile shared). Potential OTU accumulation down the hydrological transect was in part evidenced through high levels of OTU sharing between adjacent sample group assemblages, and lesser OTU sharing with assemblages sampled from further upstream. For example, the majority of the 51 dominant OTUs that were shared between subglacial meltwater assemblages and all upstream assemblages (i.e. dominant OTUs identified from supraglacial meltwater, snow or ice assemblages) were similarly shared between just dominant subglacial and supraglacial meltwater assemblages (47 OTUs; Fig. 4a), with the remaining four OTUs being found within the less abundant supraglacial meltwater assemblages. Further upstream, 32 of these dominant subglacial meltwater OTUs were shared with dominant snow or ice assemblages (26 OTUs each; Fig. 4a). The same potential accumulative effect was seen with the Red River assemblages, where 14 dominant OTUs were shared between dominant Red River assemblages and all dominant upstream assemblages, and of these, all 14 OTUs were found within the dominant subglacial assemblages (Fig. 4a) and four OTUs were found within the dominant supraglacial meltwater, snow or ice assemblages (two OTUs per assemblage group type, Fig. 4a). Lower percentages of OTUs were shared between sample groups when the rare biosphere was considered (Fig. 4b). This may have been an artefact of PCR bias, favouring amplification of dominant assemblages, or it may have been the result of working with rarefied datasets, which preferentially removes dominant assem-

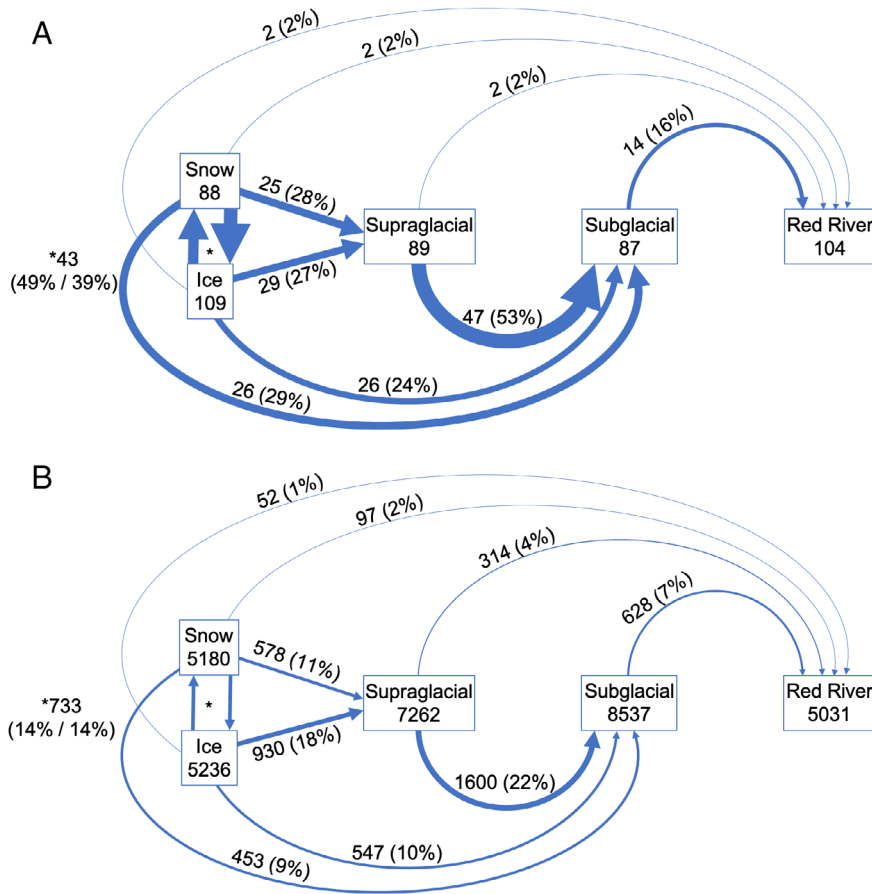


Fig 4. Shared OTUs along the hydrological transect. (A) Details the dominant assemblages with OTUs that contributed towards $\geq 0.1\%$ of the total sample group assembly, (B) details the rare assemblages with OTUs that contributed towards $< 0.1\%$ of the total sample group assembly. Numbers in boxes depict the total number of OTUs within the sample group. Numbers alongside the connectors depict the number of shared OTUs between the groups, and the percentage of shared source OTUs, shown in brackets. Arrow widths depict the percentage of shared source OTUs. [Color figure can be viewed at wileyonlinelibrary.com]

blages. As with the dominant OTUs, the greatest number of rare OTUs that were shared were calculated between supraglacial and subglacial meltwater assemblages (22% of the rare OTU supraglacial meltwater profile was shared), and between ice and supraglacial meltwater assemblages (18% of the rare OTU ice profile was shared).

Source contributions

SourceTracker was used to identify the likely source contributions of each sample type. None of the inputted potential sources were identified as being inoculants for the early-season snow samples (Fig. 5a). The SourceTracker model indicated that mid-season snow was likely inoculated from the soil that predominately resembled the upper sampled soil; however, $70.76 \pm 11.19\%$ of the source was unknown. DOY 176 snow samples were indicated as having marine assemblages as a potential source ($19.00 \pm 14.43\%$) and DOY 192 snow was indicated as having marine influences, albeit to a lesser extent ($1.72 \pm 2.96\%$). Ice accounted for

the greatest proportion of snow sources on or after DOY 192 ($59.03 \pm 25.63\%$), and vice versa for late-season snow as a potential source for ice ($72.20 \pm 7.30\%$; Fig. 5b). However, when ice was excluded as a potential source for late-season snow, or vice versa, the potential contributing sources were largely unknown (Figs S3b, S3e). SourceTracker analyses indicated that the first two supraglacial meltwater sampling dates had notable contributions from late-season snow (DOY 176, 192; $26.51 \pm 13.87\%$; Fig. 5c), but the majority of contributions were calculated as unknown ($61.75 \pm 12.49\%$). Mid-way through the supraglacial meltwater sampling both late-season snow and ice feature as the major contributors of these samples (DOY 198, 207; $25.09 \pm 7.65\%$, $60.03 \pm 8.98\%$ respectively). However, the final three supraglacial meltwater sampling events were solely dominated by ice contributions (DOY 213, 221, 225; $81.76 \pm 3.85\%$). Subglacial meltwater samples were calculated to have relatively stable contributions from supraglacial meltwater and upper soils ($58.62 \pm 15.29\%$, $13.11 \pm 9.64\%$ respectively; Fig. 5d). Red River samples were calculated to have steady but small contributions

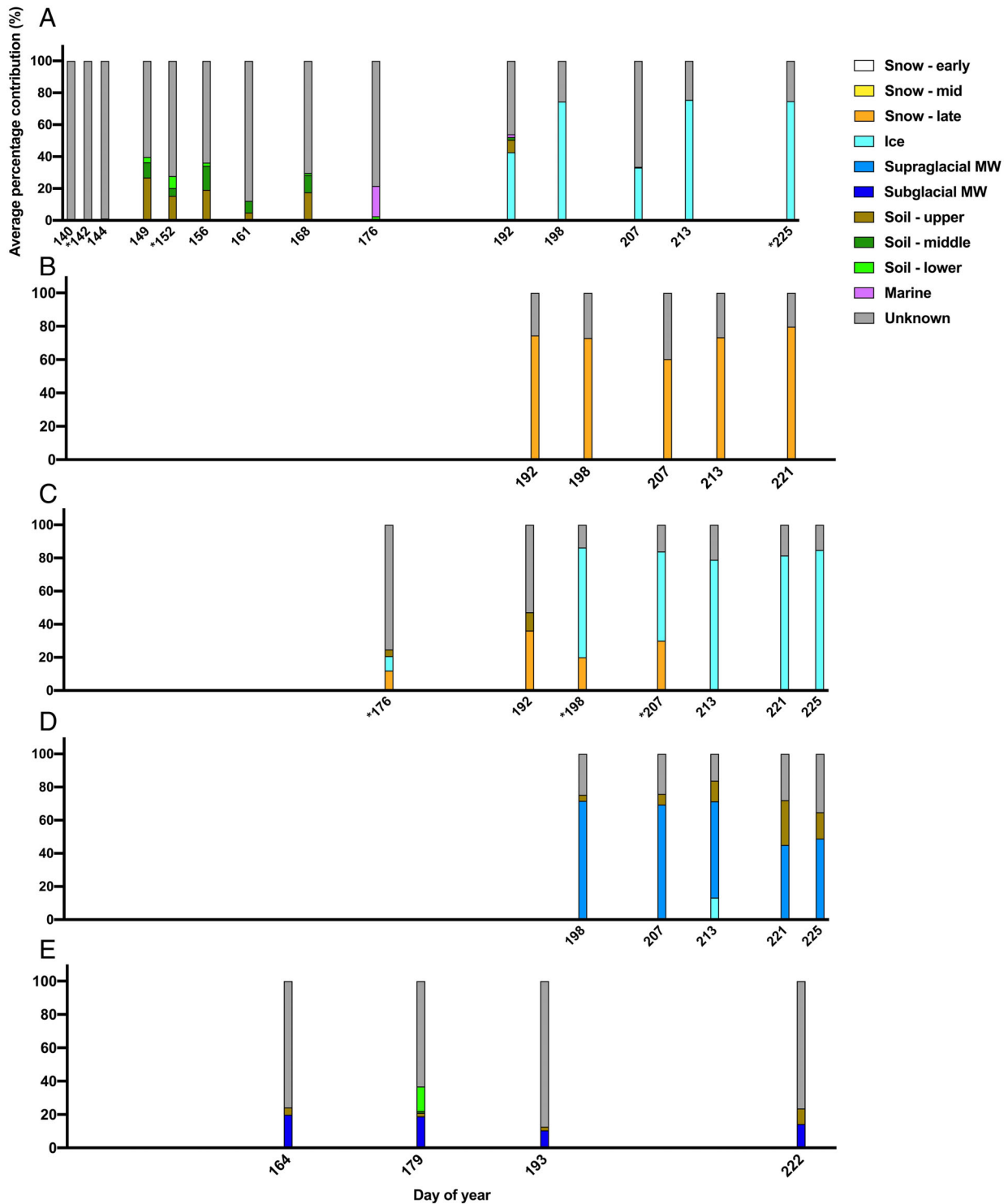


Fig 5. SourceTracker calculated the average percentage of source contributions to (A) snow, (B) ice, (C) supraglacial meltwater, (D) subglacial and (E) Red River sink samples. $n = 3$ for each bar, except those marked * where $n = 2$. [Color figure can be viewed at wileyonlinelibrary.com]

from subglacial meltwaters and upper soils throughout the sample season ($15.87 \pm 4.53\%$, $4.53 \pm 3.89\%$ respectively; Fig. 5e). During peak Red River flow the

lower soil signature was also indicated as a source influence for Red River assemblages (DOY 179; $14.77 \pm 2.90\%$). The remaining sources of Red River

assemblages were calculated as unknown ($75.65 \pm 9.35\%$).

Discussion

The connectivity of glacial microbial assemblages along hydrological flow-paths

In this study, we investigated changes in assemblage abundance and composition between hydrologically connected microbial habitats within a glacial catchment and throughout a melt season. Our results are indicative that microbial assemblages from supraglacial, subglacial and periglacial environments are transported cumulatively along a hydrological transect, contributing to the biomass of meltwater as it travels downstream. Our studies therefore support the concept that glacial environments are biologically, as well as hydrologically, connected. Low ANOSIM calculated dissimilarities between adjoining assemblages and higher ANOSIM calculated dissimilarities between more distant assemblages; high percentages of shared OTUs between adjoining sample types and lower percentages of shared OTUs between more distant sample types; SourceTracker calculated potential adjacent sources; and low variability in replicate assemblage structures support this conclusion. Lower dissimilarities between sample groups when unweighted UniFrac dissimilarity analysis was performed in comparison to weighted UniFrac dissimilarity analysis furthermore suggests that rare assemblages are present throughout this system, whereas dominant assemblages are more responsible for distinctions between the sample groups. Dierer *et al.* (2014) noted resemblances between the OTU profiles of supra- and sub-glacial meltwaters, with 61% of dominant supraglacial meltwater OTUs (with $\geq 1\%$ relative abundance) being found within subglacial outflow waters. While many of these OTUs may be generalist species to both environments, the identification of cyanobacterial phototrophs emerging from the subglacial environment (Dierer *et al.*, 2014; Cameron *et al.*, 2017b; Žárský *et al.*, 2018; this study: 0.11% relative abundance, data not shown) supports the argument that supraglacial meltwater assemblages are exported to the subglacial domain. Investigations by Dubnick *et al.* (2017) additionally provide compelling biogeochemical evidence of subglacial assemblages contributing to glacial meltwaters as they travel through this environment. Microbial studies of a meltwater transect along a glacier and fjord system in Svalbard also hint at this same connected relationship, although robust comparisons were hindered due to limited sample numbers (Thomas *et al.*, 2019).

As glacier-fed rivers are typically fast flowing, and as meltwater travels through environments that are limited in their storage capacity and lacking physical resistance to

flow, we support the concept proposed by Hauptmann *et al.* (2016) of an additive riverine system, where contributing water bodies build on the biomass of the waters that they enter into. Although Hauptmann *et al.* did not study the glacier itself, they similarly found a marked downstream increase in assemblage diversity between samples taken from the same glacial meltwater tributary as was used in this study, and those taken further down the transect at a Red River sampling site. In contrast to these findings, studies of non-glacial Arctic rivers with substantial lake and soil storage systems found diversity to reduce with distance from the source (e.g. Crump *et al.*, 2012; Comte *et al.*, 2018), likely as a result of long residence times that facilitate community growth, development and sorting. Evidence of a cumulative boreal freshwater system has been reported; however, the dominance of soil assemblages was found to mask much of the added diversity that streams, rivers and lakes contribute (Ruiz-González *et al.*, 2015). One glacial habitat where the ecological effect of residence time has previously been shown is the subglacial environment. Here, longer residence times during low flow periods result in stronger subglacial signatures in the exported meltwater assemblages (Dubnick *et al.*, 2017). This finding was likely a consequence of reduced subglacial waterbody dilution by supraglacial meltwaters, rather than local sorting or community development. In this current study, increased 'unknown' calculated sources of subglacial meltwater samples towards the end of the sample period, as well as increased 'upper soil' calculated sources as potential analogues of subglacial overridden soils, hint at increased subglacial contributions as the season progressed. Changes in subglacial meltwater volumes and melt-affected areas additionally influence nutrient export, which may in turn have downstream ecological consequences (Hawkings *et al.*, 2015; Kohler *et al.*, 2017).

Temporal variability of microbial assemblages and contributing sources

By sampling the glacial environment over several months of a melt season, temporal assemblage developments were observed. We noted a marked change in snow assemblages from samples taken in mid-May, late-May to mid-June and late-June to August. We recorded an overall increase in snow 16S rRNA gene copy abundance as the season progressed. Given the changes that were additionally observed in the calculated likely sources of these assemblages, from 'unknown', to including soil sources, to including surface ice sources, we suggest that the increase in snow abundance was the result of a combination of changing aeolian inputs and an evolving, growing community. For the mid-May samples, the sources of the snow assemblages could not be

identified using our end-member assemblages. By late-May to mid-June, soil signatures were consistently suggested as the likely source for up to 37% of the snow sampled assemblages. During this period, the steep mountain sides of the surrounding environment started to become exposed due to snow loss (Fig. 1), making it feasible that local soils from these slopes were contributing to this change. Late-season snow samples additionally had moderate ANOSIM dissimilarities to soils, however, their strong resemblances to ice assemblages, and vice versa, may have largely masked the inclusion of soil as a SourceTracker calculated source to late-season snow and ice environments. The composition of microbial assemblages in the atmosphere over southwest Greenland, and in snow sampled from western Greenland and the Canadian High Arctic, have previously been linked to the dispersal of both local and distant sources (Harding *et al.*, 2011; Cameron *et al.*, 2015; Šanti-Temkiv *et al.*, 2018). Interestingly, on DOY 176, the marine environment is calculated as the likely source of 22% of the sampled assemblages. In the period between this sample point and the previous one, strong winds were experienced from the southwest, northwest and east on several days (Fig. S4). Due to the position of Lyngmarksbræ Glacier, winds from any of these directions have the potential to carry marine aerosols onto the glacier without crossing extensive terrestrial landscapes. It is, however, curious that only one sample date was highlighted for its potential marine sources, when other periods of strong winds from these directions were seen throughout the sample season.

From late-June to August, snow assemblages were predominately calculated to be sourced from the ice environment, and vice versa. This period coincides with the exposure of glacial ice after the loss of the winter snowpack (Fig. 1). At this point, we consider three feasible explanations for this relationship: (i) snow and ice environments support similar assemblages as a result of analogous environmental conditions; (ii) ice assemblages are seeding snow assemblages, or vice versa, leading to the homogenisation of these environments; or (iii) both environments are being seeded by the same allochthonous sources. This study was originally designed to test these hypotheses through comparing snow and ice assemblages to microorganisms sampled from the overlying air. However, due to the low biomass of the air samples collected, we were unable to obtain adequate air microbial signatures using our adopted techniques, therefore this line of investigation could not be pursued. Successful studies of Arctic air masses found them to be predominately composed of local soil sources, but also to contain signatures of distant terrestrial, marine and glacial sources (Šanti-Temkiv *et al.*, 2018). This, combined with the changes in snow diversity and 16S rRNA gene

copy abundance that we observed between the early-, mid- and late-season snow, is suggestive that the low biomass late-winter snowpack becomes seeded by local soil sources once lower elevation environments become exposed. We surmise that the snowpack environment then refines the community structure, favouring members that are most tolerant to the snow conditions and diluting out lower abundance or inactive organisms. Interestingly, ice assemblages were found to be moderately dissimilar to ice debris, as has been found in other studies (Cameron *et al.*, 2016), but highly dissimilar to the soil, which may be suggestive of an endogenous system propagating surface ice assemblages; perhaps owing to the abundant, active (Anesio *et al.*, 2009) and well-distributed assemblages associated with ice debris. Considering the ANOSIM calculated dissimilarities that were found between late-season snow samples and all downstream water samples, it would be prudent to further consider the ecological consequences of exported snowmelt. In particular, the impacts of temporal changes in snow, especially with respect to the influences such as aeolian inputs, environmental warming, light regimes and water availability should be considered. Future adaptations of the aforementioned conditions may not only impact the microbial form and function of snow itself but the release of these assemblages within snow meltwaters may impact downstream environments too.

Supraglacial meltwater assemblages reflected the transition from snowmelt to surface ice melt. As the melt season progressed, 16S rRNA gene copy abundance increased by three orders of magnitude, which is suggestive of a shift in meltwater contributions from the less abundant snow assemblages to the more abundant ice assemblages. It may also be the result of cells from the snow environment being filtered out of early-season meltwaters by the physical structure of snow (Björkman *et al.*, 2014; Gokul *et al.*, 2019), only to re-join the meltwater later, once the snow structures disintegrate. Late season 16S rRNA gene copy abundances in ice and supraglacial meltwaters were comparable to previous studies of July and August supraglacial meltwaters from Svalbard (Irvine-Fynn *et al.*, 2012) when an average copy number per cell of 2.54 ± 0.98 were used (calculated using Greengenes OTU 16S rRNA gene copy numbers per cell for each sampled assemblage, data not shown; DeSantis *et al.*, 2006; Cameron *et al.*, 2017b). However, comparisons between our qPCR analysis approach and the flow cytometry approach adopted by Irvine-Fynn *et al.* should be treated with caution due to variability in technique accuracy and reproducibility, with qPCR being viewed less favourably (Stibal *et al.*, 2015b). ANOSIM analysis revealed low dissimilarities between late season snow and early supraglacial meltwaters. This relationship, and snow's relationship to soil, may explain the

moderate dissimilarities that were calculated between early supraglacial meltwater and upper soil assemblages. Late season supraglacial meltwaters were in contrast found to be less dissimilar to ice, although caution must be taken with these analyses, as sample numbers per group were small. Nevertheless, this ANOSIM result was mirrored by changes in the SourceTracker calculated potential sources of supraglacial meltwater, and both analyses tied into the marked reduction of snowpack on the glacial surface. By August, when SourceTracker no longer calculates snow as a likely potential source for supraglacial meltwater, snow at the sample site was limited to a few, small, isolated snowdrift patches.

Meltwater microbial assemblages exiting subglacial environments were found to be relatively stable over a 4-week period. Such stability has similarly been found in other biotic studies of subglacial meltwater (Dubnick *et al.*, 2017; Cameron *et al.*, 2017b). The results of our SourceTracker analyses show that subglacial meltwater sources likely came from the supraglacial environment and upper soils. We propose that supraglacial microbial assemblages are transported in meltwaters through the subglacial environment, as has been suggested previously (Dieser *et al.*, 2014; Dubnick *et al.*, 2017; Žárský *et al.*, 2018; Kohler *et al.*, 2020). These meltwaters then mix with subglacial assemblages at the glacial bed, which are likely similar in composition to the recently deglaciated upper soil samples. It is interesting that our upper soil assemblages, sampled 100 m from the glacial snout, were highlighted as likely sources for these subglacial waters. This provides insight into the potential similarities between glacial forefield soils and subglacial overridden soils, which may play an important role in glacial forefield soil succession after glacial retreat (Bradley *et al.*, 2014). We found that subglacial meltwaters maintained a constant 16S rRNA gene copy abundance over the month-long sample period, which was an order of magnitude less than the final sampled supraglacial meltwaters. This discrepancy between the water bodies during the final three sample points could be the result of (i) cells getting trapped within the physical structures of the subglacial system, (ii) increased meltwater volumes opening up new subglacial channels, thus reducing subglacial water pressure and the abundance of assemblages entrained, or (iii) sampled supraglacial meltwaters being poor representations of the supraglacial meltwater that gets routed to the subglacial environment. Studies of prokaryotic release from a Greenland Ice Sheet subglacial environment recorded one to two orders of magnitude more cells than we found in this current study (Cameron *et al.*, 2017b). As increased meltwater subglacial residence time impacts the composition and abundance of exiting microbial assemblages (Dubnick *et al.*, 2017), the small size of

Lynngmarksbræ Glacier and the short subglacial journey that meltwaters take may, therefore, be a significant factor in the lower 16S rRNA gene copy abundances that we found.

The flow of glacial microbial assemblages through periglacial landscapes

Moderate dissimilarities were calculated between Red River assemblages and upstream snow, supraglacial and subglacial assemblages. Subglacial assemblages were predicted to contribute to around 16% of Red River's sources. In another study, indicator OTUs (i.e. OTUs with a high abundance in one sample group compared with other sample groups) from samples taken downstream of our subglacial outflow site have previously been found within Red River samples (Hauptmann *et al.*, 2016). Studies that considered the transfer of microbial assemblages from a terrestrial snow pack to surrounding freshwater and marine environments found the snow to be a potential seed bank, with 30% of active snow phylotypes being similarly active in downstream lakes, and the majority of snow OTUs being shared with other downstream assemblages (Comte *et al.*, 2018). While this study by Comte *et al.* looked at the transfer of potentially active microorganisms within a glacier-free environment, and other studies have looked at the transfer of active and non-active microbial assemblages within environments downstream of glaciers (e.g. Hauptmann *et al.*, 2016; Cavaco *et al.*, 2019), we believe that our current study is the first to look at the hydrological transfer of glacial snow and ice assemblages to downstream glacial and periglacial ecosystems.

Lower soil assemblages were highlighted as potential sources to Red River during its peak discharge on DOY 179. In the fortnight running up to this event, substantial snow loss was observed from sea level to the mountain summits (Fig. 1). Soil moisture data, measured 2 km from our Red River site and 97 m above sea level, revealed marked increases in soil moisture content, followed by marked decreases preceding this sample point (Fig. S5). The top 5 cm of soil in particular exhibited near-total saturation to 50% saturation in the week preceding DOY 179. Red River 16S rRNA gene copy abundance was an order of magnitude lower when measured on or after this peak melt event, possibly as a result of a meltwater dilution effect. This sequence of events is indicative that microbial assemblages from surface soils were flushed into Red River during this period of high melt, resulting in a shift in Red River assemblage composition. Other studies of Arctic freshwater microbiota have highlighted the importance of soils (Crump *et al.*, 2007; Crump *et al.*, 2012; Ruiz-González *et al.*, 2015) and lakes (Crump *et al.*, 2007; Hauptmann *et al.*, 2016; Niño-García

et al., 2016) in building and adapting freshwater microbial composition. These soil and lake sources, for which we have few end-member samples to include in our analyses, may account for much of the 'unknown' Red River source contributions in our study. Future interconnectivity studies may, therefore, wish to consider a wider range of end-member samples, such as permafrost and mountain soils, lakes, tundra and moraines, to reduce this 'unknown' element of contributing sources. In larger glacial meltwater river systems, these periglacial inputs may be dwarfed, as was seen in a study of Watson River, which drains a portion of the Greenland Ice Sheet (Cameron *et al.*, 2017b). Watson River microbial assemblages were found to remain relatively stable from the glacier to the fjord throughout a melt season, despite passing through 30 km of moraines, permafrost and adjoining lake landscapes, and encompassing the July and August peak melt period.

Future perspectives

As glacial environments continue to experience unprecedented levels of melt, at higher altitudes, across larger areas, and for extended melt periods, it is imperative to consider the ecological consequences of this change, both locally and downstream (Cauvy-Fraunié and Dangles, 2019; Stibal *et al.*, 2020). Alongside the resulting expansion of melt affected glacial microbial habitats, the distances that melt-water travels in the supraglacial, subglacial and periglacial environments will also increase. As a result of climate warming, we anticipate substantial alterations in the microbial assemblages that are exported to adjacent ecosystems, which may impact the ecological services that they provide. Previous studies have indicated that increased and more expansive glacial melt will increase the overall biomass released (Cameron *et al.*, 2017b). Our results highlight that it is not just the subglacial environment that should be considered in this respect (Dubnick *et al.*, 2017; Cameron *et al.*, 2017b), but that changes in supraglacial environments should be accounted for too, with biota from each melt affected habitat potentially accumulating as meltwater travels downstream. In a warming climate where we see rapidly changing environments, it is therefore important to consider the ecological knock-on effects of even the most minimally populated environments.

Experimental procedures

Sample site and temporal changes

Samples were collected around the Lyngmarksbræ Glacier area on Disko Island in West Greenland between May and August 2015 (Supplemental Table 1). This

glacier was chosen due to its proximity to marine, lake and soil environments, due to the possibility of sampling subglacial meltwaters, and due to its small catchment area. The glacier lies to the south of the island, 5 km north of Qeqertarsuaq. The surrounding geology is predominantly Maligât Formation, consisting of basalt lava flows and their equivalent hyaloclastites, over basement gneiss (Pedersen *et al.*, 2003). Discontinuous permafrost soil overrides this at lower elevations, which is densely vegetated with low arctic species. The samples collected on each sampling day depended on availability, which shifted as the melt season progressed (Fig. 1; Supplemental Table 1). Briefly, sea ice clung to the coastline until mid-May. Full snow cover to sea level persisted until the end of May. From early- to mid-June, snow at sea level became patchy. From mid-June snow at higher elevations started to thin until it became very patchy on the higher mountain tiers by early-July. From the end of July until sampling ended in mid-August, snow was confined to the glacier.

Sample collection

All sample types were taken in triplicate on each day of sampling. A supraglacial meltwater sample was lost on DOY 176, DOY 198 and DOY 207, and on DOY 198, a surface ice sample and a snow sample were lost. After these losses, the following samples remained: 48 air samples and three control air samples, 44 snow samples, 17 surface ice samples, three ice debris samples, 18 supraglacial meltwater samples, 15 subglacial meltwater samples, 12 Red River water samples, 15 marine water samples and nine soil samples. All sampling equipment and collection bottles were sterilized using 10% HCl and 70% ethanol. For snow samples, the top 2–4 cm of snow was collected into clean 40 L polyethene sacs using a shovel. Surface ice was collected into 2 L whirlpak (Nasco, WI, USA) bags by chipping and collecting the top 1–3 cm of ice using a chisel. Water flowing on the supraglacial surface and from the subglacial environment were collected into 2 L whirlpak bags. Red River and marine surface water samples were collected from the riverbank and shoreline into 1 L Nalgene bottles (Nalgene, NY, USA), which were rinsed three times with 0.22 µm filtered distilled water and three times with sample water. Soil samples, consisting of the top 1 cm of soil, were collected into 2 ml sterile screw cap tubes using a spatula. The soil was collected from upper-, middle- and lower-soil sampling sites, which were located 0.1, 2 and 4.5 km from the glacier snout respectively (Fig. 1). Air samples were collected in the vicinity of the supraglacial sampling area (Fig. 1) using a BioSampler (SKC, PA, USA) containing 20 ml of 0.22 µm filtered water, attached to a vacuum pump (KNF, Germany, model N950.50

KNDC-B). BioSamplers were run for 5 h per sample with an airflow of 12.5 L min⁻¹, resulting in ~3750 L of air passing through each device. Further details on the sampling sites can be found in Table S1.

Sample preservation

Frozen samples were transported to the Arctic Field Station at Qeqertarsuaq, where they were kept at room temperature until thawed (<24 h), and then filtered through 0.22 µm Sterivex-GP polyethersulfone filters (Merck, Germany). Sterivex filters were then filled with LifeGuard (MO BIO, San Diego, USA), sealed using sterilized caps and frozen at -20 °C until processing. Liquid samples, including liquid from the BioSamplers, were filtered through 0.22 µm Sterivex-GP polyethersulfone filters in the field. These were then filled with LifeGuard and sealed using sterilized caps. Field filtered liquid samples and soil samples were transported back to the Arctic Field Station on ice, where they were immediately frozen at -20 °C. The average volumes filtered through each Sterivex filter are detailed in Table S1.

DNA extraction and sequencing

DNA was extracted from Sterivex filters using the PowerWater Sterivex DNA extraction kit (MO BIO), following the manufacturer's protocol. An unused Sterivex filter was extracted alongside each batch of extractions as a procedural control ($n = 9$). Ice debris and soil samples had their DNA extracted using a PowerSoil DNA extraction kit (MO BIO), following the manufacturer's protocol. All extractions were performed within 5 months of sampling.

Abundance analysis

Prokaryotic 16S rRNA gene copy abundance was quantified using a qPCR set-up with primer pairs 341F (5'-CCTACGGGAGGCAGCAG-3') and 518R (5'-ATTACCGCGGCTGCTGG-3'; (Muyzer *et al.*, 1993). Laboratory conditions, reaction mixtures and cycle programs were set up as detailed in Cameron *et al.* (2017b) and reactions were performed on a CFX96 Touch real-time PCR detection system (Bio-Rad, CA, USA). All qPCR reactions were performed in duplicate.

Sequencing libraries

Forward and reverse reads were joined with 25 base pair overlaps and primer sequences were removed. Libraries were filtered to have a Phred score of ≥ 25 with a maximum of five consecutive low-quality base calls. Primer sequences were used to split the libraries with 0 errors in

base calling. USEARCH61 (Edgar, 2010) was used to check for chimera; 25.5% of sequences were removed through this process. OTUs were defined as sequences that possessed $\geq 97\%$ identity. OTU picking was done using a combined *de novo* and reference-based picking strategy (uclust_ref; QIIME; Caporaso *et al.*, 2010) using the 13/08 GG data set (DeSantis *et al.*, 2006). Singletons were removed. Taxonomy was assigned by training the Ribosomal Database Project classifier (Wang *et al.*, 2007) to use the GG (13/08) dataset to the assigned species level (97% identity) where possible. Chloroplast and mitochondria sequences were removed from the datasets.

The validity of air samples was tested by considering the proportion of each sequence library that contained potential contaminant OTUs, based on nine control Sterivex sequence libraries. OTUs that contributed towards $\geq 0.01\%$ of control Sterivex samples were considered to be potential contaminant OTUs. These potential contaminants consisted of 313 OTUs including 56 that were closely related to the family Bradyrhizobiaceae, 109 to the family Methylobacteriaceae, 38 to the family Sphingomonadaceae, and 52 to the phylum TM7. The majority of sequence libraries from air samples were found to have high compositions of potential contaminant OTUs (42 out of 47 air samples had $\geq 50\%$ contaminant relative abundance), therefore all air samples were removed from further consideration. To identify contaminants in the remaining PowerWater and PowerSoil extracted samples, SourceTracker (Knights *et al.*, 2011) was used in R to identify the likely sources of each OTU. SourceTracker uses Bayesian modelling to estimate contamination from defined source assemblages into tested sink assemblages, where the mixing proportions are unknown. For this, all nine control samples were used as potential sources of each sample type. Source prediction outputs for each OTU in each sample library were used to amend OTU profiles from samples. Samples were finally rarefied to 2101 reads per sample.

Amplicon data are available at The European Bioinformatics Institute under study accession number PRJEB36007 (<http://www.ebi.ac.uk>).

Data analysis

Alpha diversity metrics, including Catchall, chao1 and Faith's Phylogenetic Diversity was calculated using QIIME. DCA and CCA data ordination analyses were calculated using CANOCO 5 (version 5.01). CCA was performed using interactive forward selection with arcsin fractions on untransformed data and by downweighting rare OTUs. ANOVA, ANOSIM and cluster analyses were performed using PRIMER6 (version 6.1.16). Cluster analyses were performed using weighted UniFrac

dissimilarity distances, using a group average cluster mode and a cut off distance of 0.7 for cluster group selections. The conditions of each ANOVA and ANOSIM test is detailed next to the results. SourceTracker was used to test the likely transfer of assemblages from one environment (source) to another (sink). Feasible sources along the hydrological transect were inputted into the analyses of each sample group. It was unclear whether snow and ice were feasible sources for each other, therefore a range of snow and ice source combinations were calculated alongside soil and marine potential sources (Fig. S3). Supraglacial meltwater, subglacial meltwater and Red River sinks were always tested for having snow, ice, soil and marine potential sources. Subglacial meltwater sinks additionally included supraglacial meltwater sources, and Red River sinks additionally considered supraglacial meltwater and subglacial meltwater sources.

Acknowledgements

We thank Dr. Christian Juncher Jørgensen and Kjeld Akaaraq Emil Mølgaard (Akaaraq) at the Arctic Research Station for logistical support and for sharing their wealth of knowledge on the surrounding area. Data from the Greenland Ecosystem Monitoring Programme were provided by the Department of Bioscience, Aarhus University, Denmark, in collaboration with the Department of Geosciences and Natural Resource Management, Copenhagen University, Denmark. We thank Dr. Aviaja Lyberth Hauptmann for discussions on her previous work.

References

Anesio, A.M., Hodson, A.J., Fritz, A., Psenner, R., and Sattler, B. (2009) High microbial activity on glaciers: importance to the global carbon cycle. *Glob Change Biol* **15**: 955–960.

Bintanja, R., and Andry, O. (2017) Towards a rain-dominated Arctic. *Nat Clim Change* **7**: 263–267.

Björkman, M.P., Zarsky, J.D., Kühnel, R., Hodson, A., Sattler, B., and Psenner, R. (2014) Microbial cell retention in a melting high Arctic snowpack, Svalbard. *Arct Antarct Alp Res* **46**: 471–482.

Bradley, J.A., Singarayer, J.S., and Anesio, A.M. (2014) Microbial community dynamics in the forefield of glaciers. *Proc Roy Soc B Biol Sci* **281**: 20140882.

Bunge, J. (2011) Estimating the number of species with CatchAll. *Biocomputing* **2011**: 121–130.

Cameron, K.A., Hagedorn, B., Dieser, M., Christner, B.C., Choquette, K., Sletten, R., *et al.* (2015) Diversity and potential sources of microbiota associated with snow on western portions of the Greenland ice sheet. *Environ Microbiol* **17**: 594–609.

Cameron, K.A., Stibal, M., Hawkings, J.R., Mikkelsen, A.B., Telling, J., Kohler, T.J., *et al.* (2017b) Meltwater export of prokaryotic cells from the Greenland ice sheet. *Environ Microbiol* **19**: 524–534.

Cameron, K.A., Stibal, M., Olsen, N.S., Mikkelsen, A.B., Elberling, B., and Jacobsen, C.S. (2017a) Potential activity of subglacial microbiota transported to anoxic river delta sediments. *Microb Ecol* **74**: 6–9.

Cameron, K.A., Stibal, M., Zarsky, J.D., Gözdereliler, E., Schostag, M., and Jacobsen, C.S. (2016) Supraglacial bacterial community structures vary across the Greenland ice sheet. *FEMS Microbiol Ecol* **92**: fiv 164.

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

Cauvy-Fraunié, S., and Dangles, O. (2019) A global synthesis of biodiversity responses to glacier retreat. *Nat Ecol Evol* **3**: 1675–1685.

Cavaco, M.A., St. Louis, V.L., Engel, K., St. Pierre, K.A., Schiff, S.L., Stibal, M., and Neufeld, J.D. (2019) Freshwater microbial community diversity in a rapidly changing high Arctic watershed. *FEMS Microbiol Ecol* **95**: fiz161.

Comte, J., Culley, A.I., Lovejoy, C., and Vincent, W.F. (2018) Microbial connectivity and sorting in a high Arctic watershed. *ISME J* **12**: 2988–3000.

Crump, B.C., Adams, H.E., Hobbie, J.E., and Kling, G.W. (2007) Biogeography of bacterioplankton in lakes and streams of an arctic tundra catchment. *Ecology* **88**: 1365–1378.

Crump, B.C., Amaral-Zettler, L.A., and Kling, G.W. (2012) Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *ISME J* **6**: 1629–1639.

Crump, B.C., Kling, G.W., Bahr, M., and Hobbie, J.E. (2003) Bacterioplankton community shifts in an arctic lake correlate with seasonal changes in organic matter source. *Appl Environ Microbiol* **69**: 2253–2268.

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

Dieser, M., Broensen, E.L.J.E., Cameron, K.A., King, G.M., Achberger, A., Choquette, K., *et al.* (2014) Molecular and biogeochemical evidence for methane cycling beneath the western margin of the Greenland ice sheet. *ISME J* **8**: 2305–2316.

Dubnick, A., Kazemi, S., Sharp, M., Wadham, J., Hawkings, J., Beaton, A., and Lanoil, B. (2017) Hydrological controls on glacially exported microbial assemblages. *Eur J Vasc Endovasc Surg* **122**: 1049–1061.

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

Edwards, A., Rassner, S.M., Anesio, A.M., Worgan, H.J., Irvine-Fynn, T.D.L., Williams, H.W., *et al.* (2013) Contrasts between the cryoconite and ice-marginal bacterial communities of Svalbard glaciers. *Polar Res* **32**: 19468.

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

Franzetti, A., Navarra, F., Tagliaferri, I., Gandolfi, I., Bestetti, G., Minora, U., *et al.* (2017a) Temporal variability of bacterial communities in cryoconite on an alpine glacier. *Environ Microbiol Rep* **9**: 71–78.

- Franzetti, A., Navarra, F., Tagliaferri, I., Gandolfi, I., Bestetti, G., Minora, U., et al. (2017b) Potential sources of bacteria colonizing the cryoconite of an alpine glacier. *PLoS One* **12**: e0174786.
- Gokul, J.K., Cameron, K.A., Irvine-Fynn, T.D.L., Cook, J.M., Hubbard, A., Stibal, M., et al. (2019) Illuminating the dynamic rare biosphere of the Greenland ice sheet's dark zone. *FEMS Microbiol Ecol* **95**: fiz177.
- Gutiérrez, M.H., Galand, P.E., Moffat, C., and Pantoja, S. (2015) Melting glacier impacts community structure of bacteria, archaea and fungi in a Chilean Patagonia fjord. *Environ Microbiol* **17**: 3882–3897.
- Harding, T., Jungblut, A.D., Lovejoy, C., and Vincent, W.F. (2011) Microbes in high Arctic snow and implications for the cold biosphere. *Appl Environ Microbiol* **77**: 3234–3243.
- Hauptmann, A., Markussen, T., Stibal, M., Olsen, N., Elberling, B., Bælum, J., et al. (2016) Upstream freshwater and terrestrial sources are differentially reflected in the bacterial community structure along a small Arctic river and its estuary. *Front Microbiol* **7**: 1474.
- Hawkings, J., Wadham, J., Tranter, M., Lawson, E., Sole, A., Cowton, T., et al. (2015) The effect of warming climate on nutrient and solute export from the Greenland ice sheet. *Geochem Perspect Lett* **1**: 94–104.
- Hood, E., Fellman, J., Spencer, R.G.M., Hernes, P.J., Edwards, R., D'Amore, D., and Scott, D. (2009) Glaciers as a source of ancient and labile organic matter to the marine environment. *Nature* **462**: 1044–1047.
- Hudson, B., Overeem, I., McGrath, D., Syvitski, J., Mikkelsen, A., and Hasholt, B. (2014) MODIS observed increase in duration and spatial extent of sediment plumes in Greenland fjords. *Cryosphere* **8**: 1161–1176.
- Irvine-Fynn, T.D.L., Edwards, A., Newton, S., Langford, H., Rassner, S.M., Telling, J., et al. (2012) Microbial cell budgets of an Arctic glacier surface quantified using flow cytometry. *Environ Microbiol* **14**: 2998–3012.
- Knights, D., Kuczynski, J., Charlson, E.S., Zaneveld, J., Mozer, M.C., Collman, R.G., et al. (2011) Bayesian community-wide culture-independent microbial source tracking. *Nat Methods* **8**: 761–763.
- Kohler, T.J., Vinšová, P., Falteisek, L., Žárský, J.D., Yde, J. C., Hatton, J.E., et al. (2020) Patterns in microbial assemblages exported from the meltwater of Arctic and sub-Arctic glaciers. *Front Microbiol* **11**: 669.
- Kohler, T.J., Žárský, J.D., Yde, J.C., Lamarche-Gagnon, G., Hawkings, J.R., Tedstone, A.J., et al. (2017) Carbon dating reveals a seasonal progression in the source of particulate organic carbon exported from the Greenland ice sheet. *Geophys Res Lett* **44**: 6209–6217.
- Musilova, M., Tranter, M., Bennett, S.A., Wadham, J.L., and Anesio, A. (2015) Stable microbial community composition on the Greenland ice sheet. *Front Microbiol* **6**: 193.
- Muyzer, G., De Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *App Environ Microbiol* **59**: 695–700.
- Niño-García, J.P., Ruiz-González, C., and del Giorgio, P.A. (2016) Landscape-scale spatial abundance distributions discriminate core from random components of boreal lake bacterioplankton. *Ecol Lett* **19**: 1506–1515.
- Nordenskiöld, A. (1872) Account of an expedition to Greenland in the year 1870. *Geol Mag* **9**: 355–368.
- Overeem, I., Hudson, B.D., Syvitski, J.P.M., Mikkelsen, A.B., Hasholt, B., van den Broeke, M.R., et al. (2017) Substantial export of suspended sediment to the global oceans from glacial erosion in Greenland. *Nat Geosci* **10**: 859–863.
- Overland, J., Dunlea, E., Box, J.E., Corell, R., Forsius, M., Kattsov, V., et al. (2019) The urgency of Arctic change. *Polar Sci* **21**: 6–13.
- Paulsen, M.L., Nielsen, S.E.B., Müller, O., Møller, E.F., Stedmon, C.A., Juul-Pedersen, T., et al. (2017) Carbon bioavailability in a high Arctic fjord influenced by glacial meltwater, NE Greenland. *Front Mar Sci* **4**: 176.
- Pedersen, A.K., Larsen, L., Pedersen, G.K., Heinesen, M., and Dueholm, K. (2003) Geological section along the south and south-west coast of Disko, central West Greenland. 1: 20000 coloured geological sheet. Technical report: The Geological Survey of Denmark and Greenland.
- Pittino, F., Maglio, M., Gandolfi, I., Azzoni, R.S., Diolaiuti, G., Ambrosini, R., and Franzetti, A. (2018) Bacterial communities of cryoconite holes of a temperate alpine glacier show both seasonal trends and year-to-year variability. *Ann Glaciol* **59**: 1–9.
- Rime, T., Hartmann, M., and Frey, B. (2016) Potential sources of microbial colonizers in an initial soil ecosystem after retreat of an alpine glacier. *ISME J* **10**: 1625–1641.
- Ruiz-González, C., Niño-García, J.P., and del Giorgio, P.A. (2015) Terrestrial origin of bacterial communities in complex boreal freshwater networks. *Ecol Lett* **18**: 1198–1206.
- Ryan, J.C., Smith, L.C., van As, D., Cooley, S.W., Cooper, M.G., Pitcher, L.H., and Hubbard, A. (2019) Greenland ice sheet surface melt amplified by snowline migration and bare ice exposure. *Sci Adv* **5**: eaav3738.
- Šanti-Temkiv, T., Gosewinkel, U., Starnawski, P., Lever, M., and Finster, K. (2018) Aeolian dispersal of bacteria in Southwest Greenland: their sources, abundance, diversity and physiological states. *FEMS Microbiol Ecol* **94**: fiy031.
- Schostag, M., Stibal, M., Jacobsen, C.S., Bælum, J., Taş, N., Elberling, B., et al. (2015) Distinct summer and winter bacterial communities in the active layer of Svalbard permafrost revealed by DNA- and RNA-based analyses. *Front Microbiol* **6**: 399.
- Södergren, A.H., McDonald, A.J., and Bodeker, G.E. (2017) An energy balance model exploration of the impacts of interactions between surface albedo, cloud cover and water vapor on polar amplification. *Clim Dyn* **51**: 1639–1658.
- Stibal, M., Box, J.E., Cameron, K.A., Langen, P.L., Yallop, M.L., Mottram, R.H., et al. (2017) Algae drive enhanced darkening of bare ice on the Greenland ice sheet. *Geophys Res Lett* **44**: 11463–11471.
- Stibal, M., Bradley, J.A., Edwards, A., Hotaling, S., Zawierucha, K., Rosvold, J., et al. (2020) Glacial ecosystems are essential to understanding biodiversity responses to glacier retreat. *Nat Ecol Evol* **4**: 686–687.
- Stibal, M., Gözdereliler, E., Cameron, K.A., Box, J.E., Stevens, I.T., Gokul, J.K., et al. (2015b) Microbial

abundance in surface ice on the Greenland ice sheet. *Front Microbiol* **6**: 225.

- Stibal, M., Schostag, M., Cameron, K.A., Hansen, L.H., Chandler, D.M., Wadham, J.L., and Jacobsen, C.S. (2015a) Different bulk and active bacterial communities in cryoconite from the margin and interior of the Greenland ice sheet. *Environ Microbiol Rep* **7**: 293–300.
- Thomas, F.A., Sinha, R.K., and Krishnan, K.P. (2019) Bacterial community structure of a glacio-marine system in the Arctic (Ny-Ålesund, Svalbard). *Sci Total Environ* **718**: 135264.
- Vavrus, S., Waliser, D., Schweiger, A., and Francis, J. (2009) Simulations of 20th and 21st century Arctic cloud amount in the global climate models assessed in the IPCC AR4. *Clim Dyn* **33**: 1099–1115.
- Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**: 5261–5267.
- Žárský, J.D., Kohler, T.J., Yde, J.C., Falteisek, L., Lamarche-Gagnon, G., Hawkings, J.R., *et al.* (2018) Prokaryotic assemblages in suspended and subglacial sediments within a glacierized catchment on Qeqertarsuaq (Disko Island), West Greenland. *FEMS Microbiol Ecol* **94**: fiy100.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Supplemental Fig. 1 Discharge rates from Red River ($\text{m}^3 \text{s}^{-1}$). Measurements were taken every 10 min from day of year (DOY) 169 to DOY 232. Dashed lines depict days of Red River sampling, the red dashed line depicts sampling on DOY 179, which occurred around the period of peak river discharge.

Supplemental Fig. 2: Cluster analyses of weighted unifracs distances between (a) snow, (b) ice, (c) supraglacial meltwater, (d) subglacial meltwater, and (e) Red River sample groups. Data points are shown as day of year 2015. Symbols depict early- (open circle), mid- (grey circle) and late- (black circle) season designated samples for snow, supraglacial meltwater and subglacial meltwater; ungrouped ice samples (triangle); and normal- (open square) and peak- (black square) Red River flow. Group cut off distances of 0.70–0.75 were used.

Supplemental Fig. 3: SourceTracker analyses of average source contributions to a-b) snow, and c-e) ice sinks using a variety of source input options. All analyses include marine and soil potential sources, however other sources considered in the above analyses included: a) ice, b) none, c) snow and ice debris, d) snow, e) none. a) and d) are presented in Fig. 5.

Supplemental Fig. 4: Wind speed and direction from the Greenland Ecosystem Monitoring Programme's Arctic weather station 3 (AWS-3) at Blæsedalen which stands at 90 m above sea level. Half-hourly measurements are shown from May 15 (day of year; DOY 135) to August 13 (DOY 225) 2015. Red dots are from the period DOY 169 to DOY 176.

Supplemental Fig. 5: Soil moisture content (percentage volume) with depth and time at the Greenland Ecosystem Monitoring Programme's Arctic weather station 3 (AWS-3) at Blæsedalen which stands at 90 m above sea level. Dashed lines depict days of Red River sampling, the red dashed line depicts sampling on DOY 179, which occurred around the period of peak river discharge.

Supplemental Table 1 Samples taken from Lyngmarksbræ glacier, including latitude and longitude of the predominant site location and the volumes filtered where appropriate. DOY: Day of year 2015. MW: meltwater. Colours denote split sampling groups with light blue: early-, dark blue: mid-, light orange: late-, dark orange: peak- melt samples. *Soil sample sites included upper- (N 69.28857°, W 053.57557°), middle- (N 69.27705°, W 053.53466°) and lower- (N 69.25494°, W 053.51662°) locations.

Supplemental Table 2 Number of sequences per sample after removal of chimeras, chloroplasts mitochondria, singletons and potential contaminants.

Supplemental Table 3 Heatmap table of ANOSIM analyses comparing sample group pairs. Dark green depicts groups with the greatest calculated similarities and white depicts sample groups with the least calculated similarities. Black text depicts $p \leq 0.01$ and dark blue text depicts $p 0.01 > 0.05$.

Supplemental Table 4 Heatmap table of ANOSIM analyses comparing pairs of cluster defined sample groups. Dark green depicts groups with the greatest calculated similarities and white depicts sample groups with the least calculated similarities. Black text depicts $p \leq 0.01$, dark blue depicts $p 0.01 > 0.05$ and light blue depicts $p \geq 0.05$.