Supporting Information

A Preliminary Investigation of Microbial Communities on the Athabasca Glacier Within Deposited Organic

*Milena Esser[1], Phillip Ankley[1], Caroline Aubry-Wake[2,3], Yuwei Xie[1], Helen Baulch[3,4], Cameron Hoggarth[3], Markus Hecker[1,4], Henner Hollert[5], John P. Giesy[1,6,7], John W. Pomeroy[2,3], and Markus Brinkmann[1,2,3,4]**

[1] Toxicology Centre, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

[2] Centre for Hydrology, University of Saskatchewan, Saskatoon, Canmore, Alberta, Canada

[3] Global Institute for Water Security, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

[4] School of Environment and Sustainability, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

[5] Institute of Ecology, Evolution and Diversity, Goethe University, Frankfurt, Germany

[6] Department of Veterinary Biomedical Sciences, University of Saskatchewan, Saskatoon, Canada

[7] Department of Environmental Sciences, Baylor University, Waco, Texas, USA

*** Corresponding authors**

E-mail address: markus.brinkmann@usask.ca (M. Brinkmann)

Summary of the number of figures and tables: Figures: 7; Tables: 4

CONTENT

SI-1. Preparation for and imaging using scanning electron microscopy

Samples were thawed at room temperature and then fixed in 2.5% glutaraldehyde in 0.1 M sodium cacodylate (NaCAC. pH 7.2) for 2 h before being rinsed twice with 0.1M NaCAC and stored in NaCAC at 4 °C for 24 h. After 24 h, dehydration of samples was performed using 15 min incubations of 25, 50, 75, 95, 100, and 100% ethanol solutions. Samples were substituted with hexamethyldisilazane (HMDS) twice for 15 min each, air-dried, and then mounted on a SEM stub and sputter-coated with a thin layer of gold (Edwards S150B), with 2.5 min coating at 1kv, 20 ma. Samples were then imaged using a SU8010 electron microscope (Hitachi High-Technologies Canada Inc., Etobicoke, Canada) with backscattered electrons (BSE) imaging mode.

SI-2. DNA extraction, PCR amplification and next-generation sequencing

Between 1 and 2 g were subsampled from each sample and centrifuged at 16,000×g and 2 °C for 30 min. The supernatant of each sample was removed, and pellets were extracted with the E.Z.N.A Soil DNA Kit (OMEGA Bio-Tec, Inc., U.S.A.) following the manufacturer's protocol. DNA concentrations of all extracts were measured using the Qubit double-stranded DNA high-sensitivity assay (Thermo Fisher Scientific, U.S.A.) and purity was assessed using the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc., U.S.A.). One extraction blank was conducted with each batch for quality control (QC). Concentrations of DNA from extraction blanks were less than the limit of detection. All extracts were normalized to a DNA concentration of 10 ng/µl before PCR amplification. The hypervariable V3-V4 region of the bacterial 16S rRNA gene was amplified by the use of the universal prokaryotic 341F/R806 primer set (forward primer 5'-CCTACGGGNBGCASCAG-3'(1), reverse primer 5'-GACTACNVGGGTATCTAATCC-3'(2)). To profile compositions of protist communities, the V4 fragment of the eukaryotic 18S rRNA gene was amplified with the TA-Reuk454TWD1/TAReukREV3 primer set by Stoeck et al. (forward primer 5'-CCAGCASCYGCGGTAATTCC-3'; reverse primer 5'-ACTTTCGTTCTTGATYRA-3').(3) PCR negative controls (nuclease-free water as non-template control) were included in each PCR for QC. PCR cycling conditions were comprised of an initial denaturation step lasting 30 sec at 98°C, followed by 28 cycles of denaturation at 98°C for 10 sec, annealing at 58°C for 30 sec and elongation at 72°C for 30 sec, with a final extension step at 72°C for 5 min. PCR products were visualized on a 1.2% agarose gel to verify whether the barcodes were successfully amplification of target barcodes and whether cross-contamination occurred between wells during plate preparation. After confirmation of successful PCR through gel electrophoresis, the amplified

DNA fragments were purified to eliminate PCR residues such as primers, nucleotides, enzymes, and other impurities. Isolation of DNA fragments based on size and with a second purification was achieved through agarose gel electrophoresis (1.5% agarose gel). After 1 h, the gel was examined under UV light and the desired DNA fragments were cut out with a clean, sterile razor blade. Each gel slice was transferred into a 5 ml microcentrifuge tube and extracted using the MicroElute Gel Extraction Kit (Omega Bio-Tec, Inc., U.S.A.) following the manufacturer's protocol. NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolab, U.K.) and the NEBNext Multiple Oligos for Illumina, Index Primer Set 1 (New England Biolab, U.K.) was used for library preparation of pooled samples. The final libraries were quantified using NEBNext Library Quant Kit for Illumina (New England Biolab, U.K.). All libraries were normalized to a concentration of 85 nM and pooled together. Denaturation and final dilution to a concentration of 10 pM were performed according to Protocol A of the MiSeq System Denature and Dilute Libraries Guide 2019. A PhiX control spike-in of 10% was added as sequencing control and to increase the diversity of the sequencing run. Finally, a 600-cycle paired-end sequencing run was performed on an Illumina® MiSeq sequencer (Illumina, San Diego, CA) using a 2x300 bp paired-end Illumina® chemistry kit.

SI-3. Next-generation amplicon sequencing processing

Quality control of the sequencing output was first conducted using FastQC (4). Poor quality and technical sequences were removed using the pre-processing tool Trimmomatic (version 0.32).(5) The USEARCH algorithm (v11)6 was then applied for paired-end read merging. The merged reads were sorted and distinguished by their unique sample tags using the split_libraries.py script implemented in the open-source microbiome bioinformatics platform

QIIME (version 1.8.0).(7) A quality filtering step was performed using the fastq_filter command implemented in USEARCH (version 11).(8) discarding reads with a maximum expected error (maxEE) higher than 1. Strictly identical sequences were grouped to reduce the size of data (Dereplication), followed by a denoising step with the unoise3 algorithm.(9) Sequences with abundances below 8 were discarded in this step. As a result, a set of predicted biological sequences called zero-radius taxonomic units (ZOTUs) was formed. The taxonomic annotation for each ZOTU was assigned using the SILVA database(10) applying the QIIME2 Scikit-learn plugin,11,12 and mitochondrial and chloroplast sequences were removed. To enable a robust comparison of diversity matrices, rarefaction was applied.(13,14) A threshold of 34,697 sequences per sample was set to obtain even sequencing depth for 16S and 5,298 sequences per sample for 18S.

SI-4. Measurement of total organic carbon (TOC) in glacier samples

Total organic carbon (TOC) was quantified by a commercial lab (Bureau Veritas, Edmonton, Canada) and expressed as either percent or mass concentration, depending on sample characteristics. Solid samples (cryoconite and glacier sediment) were analyzed using the in-house laboratory method labeled CAL SOP-00243. The method is the Organic Carbon in Solids by LECO® TruMac CNS based on the analytical method of LECO® Corporation Form No. 203-821-498. Water samples (glacier ice) were analyzed using the in-house laboratory method labeled AB SOP-00087. This method is the measurement of organic carbon by Technicon - Persulfate UV Oxidation based on Methods Manual for Chemical Analysis of Water and Wastes, Method Code 119.

SI-5. Measurement of nutrient concentrations in glacier samples

For total dissolved phosphorus (TDP) and total dissolved nitrogen (TDN) analysis, samples were thawed and then centrifuged at 3,500 rpm for 10 minutes to separate liquid and particulate fractions, with the liquid fraction being collected and filtered through 0.45 μm nylon syringe filters. Subsamples of the filtered water were then digested and analyzed on a SmartChem 170 autoanalyzer. The particulate fraction was dried and 0.0010-0.1306 g diluted with 5 ml of reagent water for total phosphorus (TP) and 5 or 10 ml of reagent water for total nitrogen (TN) before digesting the sample.

TDP and TP were analyzed using a method based on EPA 365.1.15 All forms of phosphorus were converted to orthophosphate using sulfuric acid and ammonium persulphate digestion. Ammonium molybdate and antimony potassium tartrate reacted with phosphorus to form a complex of antimony-phospho-molybdate, which reacts with ascorbic acid to produce a blue color proportional to the orthophosphorus concentration, which was measured at 880 nm. Results from similarly digested reagent water were subtracted from the digested sample TDP results.

TDN and TN were analyzed using a method based on Standard Methods 4500-Norg D,16 in which all forms of nitrogen are converted to nitrate using a potassium persulfate and sodium hydroxide digestion. Nitrate was then reduced to nitrite by the passage of a filtered sample through an open tubular copperized cadmium redactor. Total dissolved nitrogen was then determined as nitrite by diazotizing with sulfanilamide followed by coupling with N-(napthyl)-ethylenediamine dihydrochloride. The resulting magenta solution was then measured on a spectrophotometer at 550 nm with TDN results from similarly processed digested reagent water subtracted from the digested sample. TP and TN values are reported in Table S4.

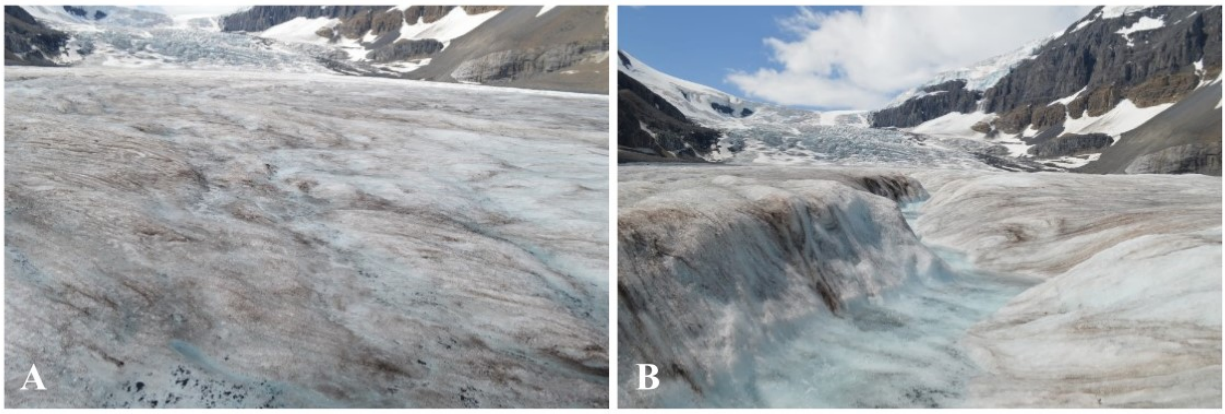

Fig. S1 Athabasca Glacier and target area of sampling. The glacier samples were taken on July 23rd, 2018, at 52.19182, -117.25165.

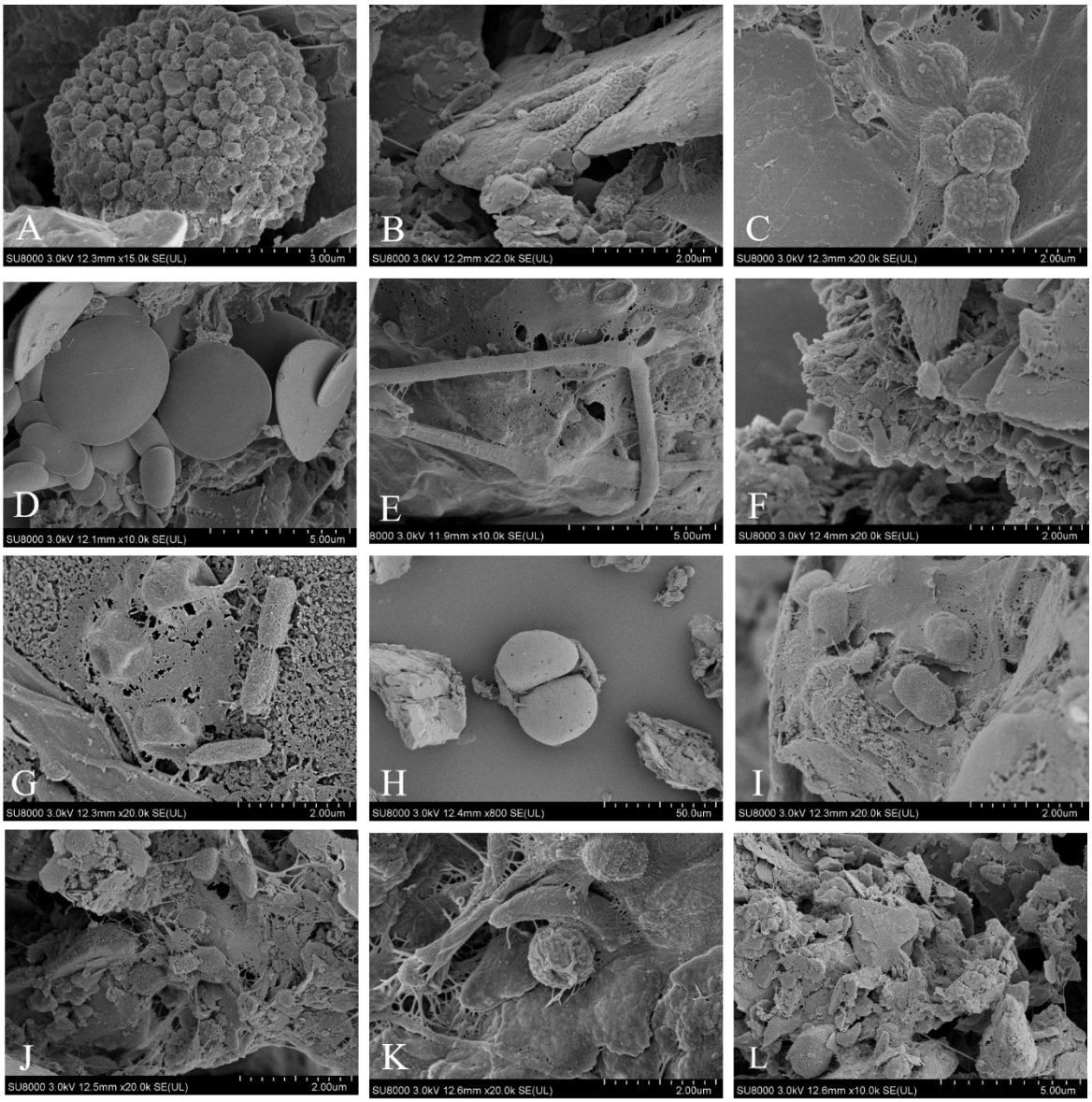

Fig. S2 Additional SEM images of glacier samples from the Athabasca Glacier in Canada of ice (A-C), cryoconite holes (D-K), and sediment (L). A, C: cluster of ultra-microbacteria cells in an EPS matrix. B, G: rod-shaped bacterial cells in EPS. E: filamentous bacterial cells (possibly cyanobacteria). H: pine pollen. I: small, short rod cells. D, J, L: clay/inorganic particles with EPS strands and (nano-sized) microorganisms attached. K: coccoid bacterial cells with polymer. Possible ultra-microbacteria cells attached.

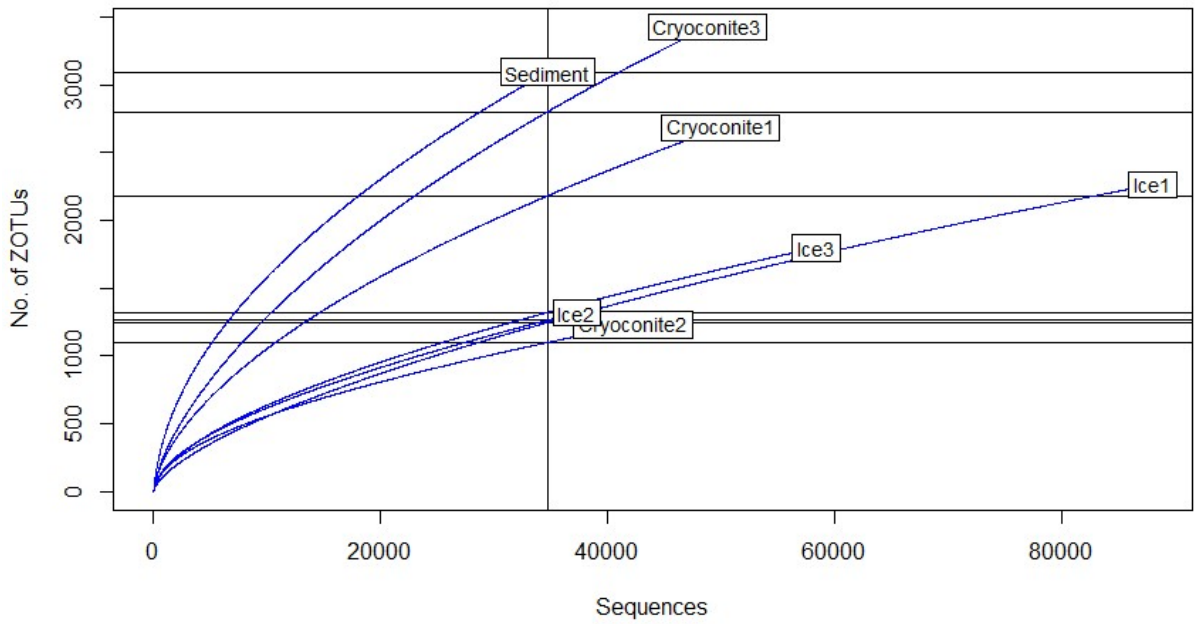

Fig. S3 Rarefaction curves of prokaryotic (16S) ZOTUs. The vertical line marks the number of sequences used for rarefaction (34,697), with one vertical line drawn for each sample for the respective rarefied ZOTU richness. The slope of the rarefaction curve and expected number of ZOTUs (richness) for each sample at the given number of sequences for rarefaction (34,697): *Cryoconite 1*: 0.036 (2,180.39 ZOTUs, standard error 18.54); *Cryoconite 2*: 0.018 (1,098.04 ZOTUs, SE 10.27); *Cryoconite 3*: 0.048 (2,800.57 ZOTUs, SE 20.73); *Ice 1*: 0.023 (1,245.72 ZOTUs, SE 20.86); *Ice 2*: 0.022 (1.262.8 ZOTUs, SE 7.03); *Ice 3*: 0.023 (1,319.8 ZOTUs, SE 16.97); *Sediment*: 0.047 (3094 ZOTUs, SE 0).

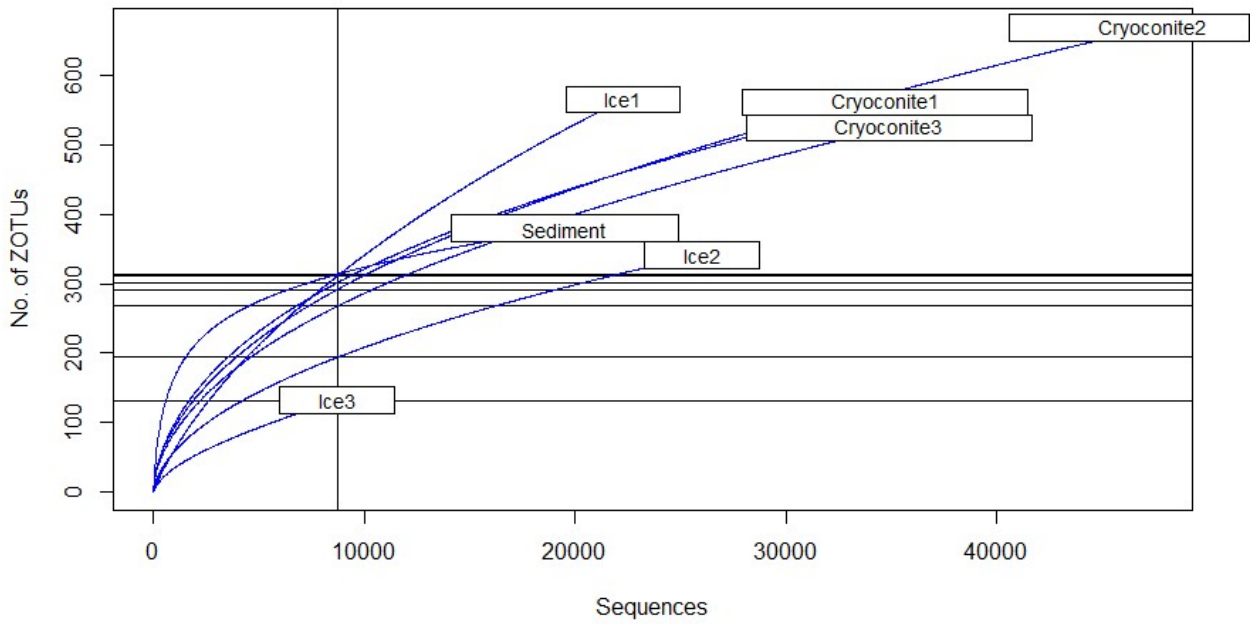

Fig. S4 Rarefaction curve of eukaryotic (18S) ZOTUs. The vertical line marks the number of sequences used for rarefaction (8,704), with one vertical line drawn for each sample for the respective rarefied ZOTU richness. The slope of the rarefaction curve and expected number of ZOTUs (richness) for each sample at the given number of sequences for rarefaction (8,704): *Cryoconite 1*: 0.016 (301.49 ZOTUs, standard error 9.25); *Cryoconite 2*: 0.017 (291.79 ZOTUs, SE 9.84); *Cryoconite 3*: 0.015 (267.7 ZOTUs, SE 9.1); *Ice 1*: 0.024 (311.26 ZOTUs, SE 10.52); *Ice 2*: 0.012 (193.95 ZOTUs, SE 7.64); *Ice 3*: 0.010 (132 ZOTUs, SE 0); *Sediment*: 0.008 (313.79 ZOTUs, SE 5.82).

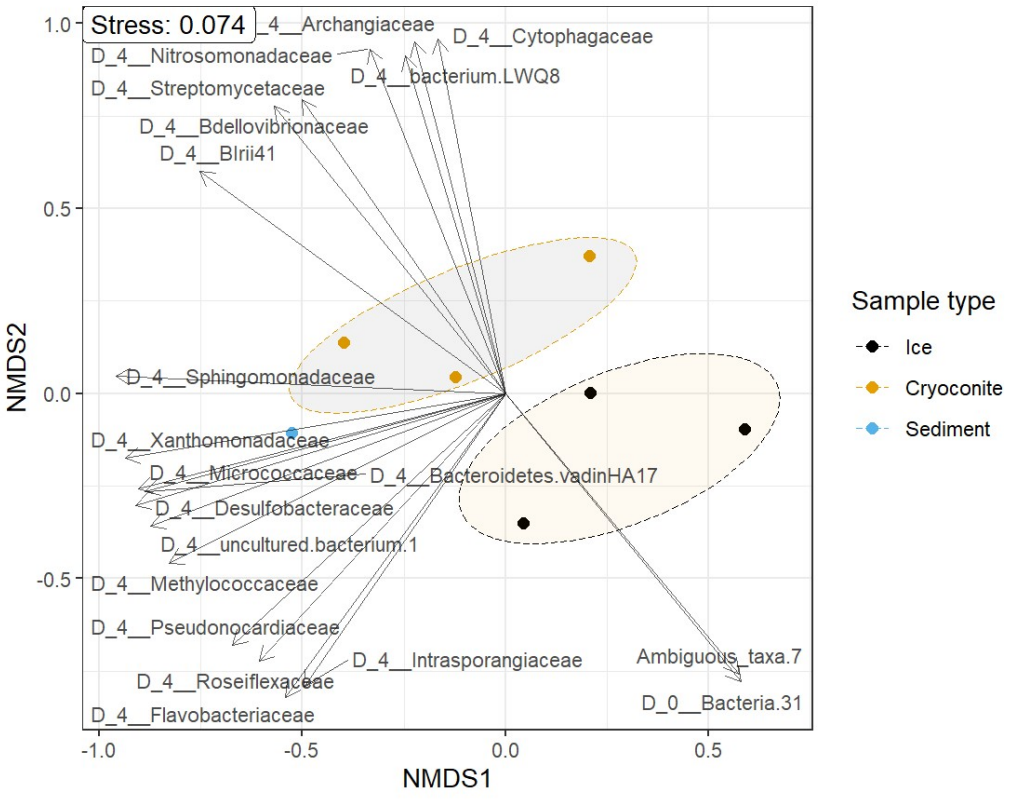

Fig. S5 Non-metric multidimensional scaling (NMDS) plot illustrating separation of samples based on Bray-Curtis dissimilarity in prokaryotic (16S) microbial community structure. Shapes and colors correspond to different sample types, as listed in the legend. Vectors represent significant correlations ($p < 0.01$) of relative abundances of phyla (right) to the distribution of microbial communities plotted. Permutational multivariate analysis of variance using distance matrices (adonis2 function; package vegan; number of permutations: 5,039) indicated that the difference in sample types was significant ($p < 0.02$*).

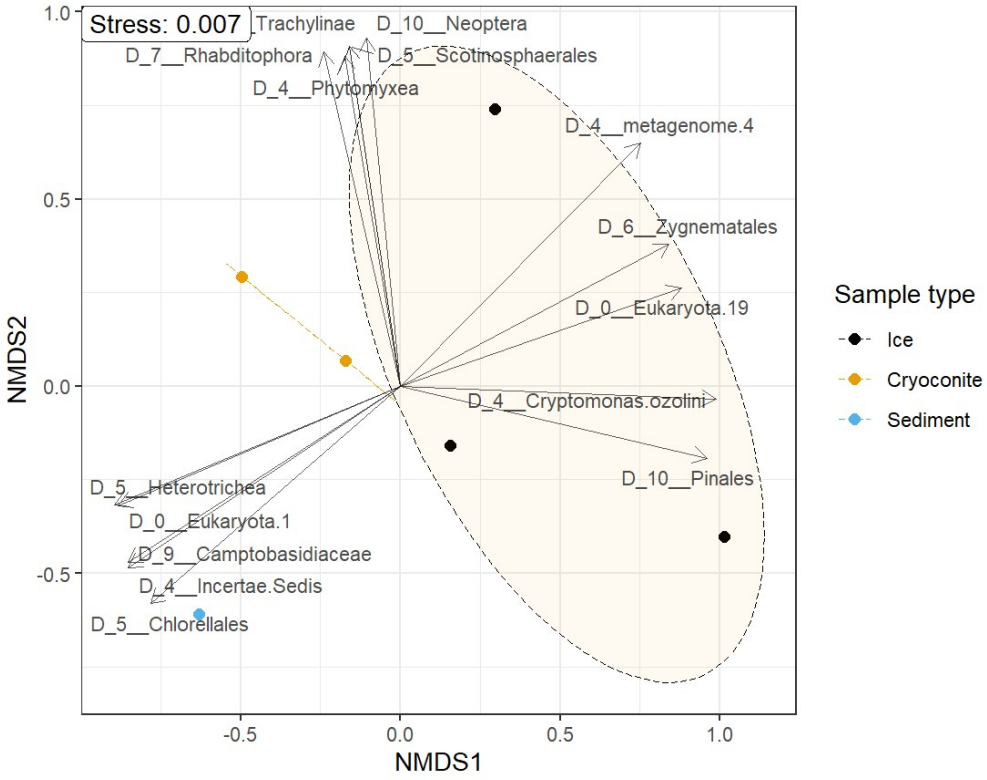

Fig. S5 Non-metric multidimensional scaling (NMDS) plot illustrating separation of samples based on Bray-Curtis dissimilarity in eukaryotic (18S) microbial community structure. Shapes and colors correspond to different sample types, as listed in the legend. Vectors represent significant correlations ($p < 0.01$) of environmental parameters (left)/relative abundances of phyla (right) to the distribution of microbial communities plotted. Permutational multivariate analysis of variance using distance matrices (adonis2 function; package vegan; number of permutations: 5,039) indicated that the difference in sample types was significant ($p < 0.02$*).

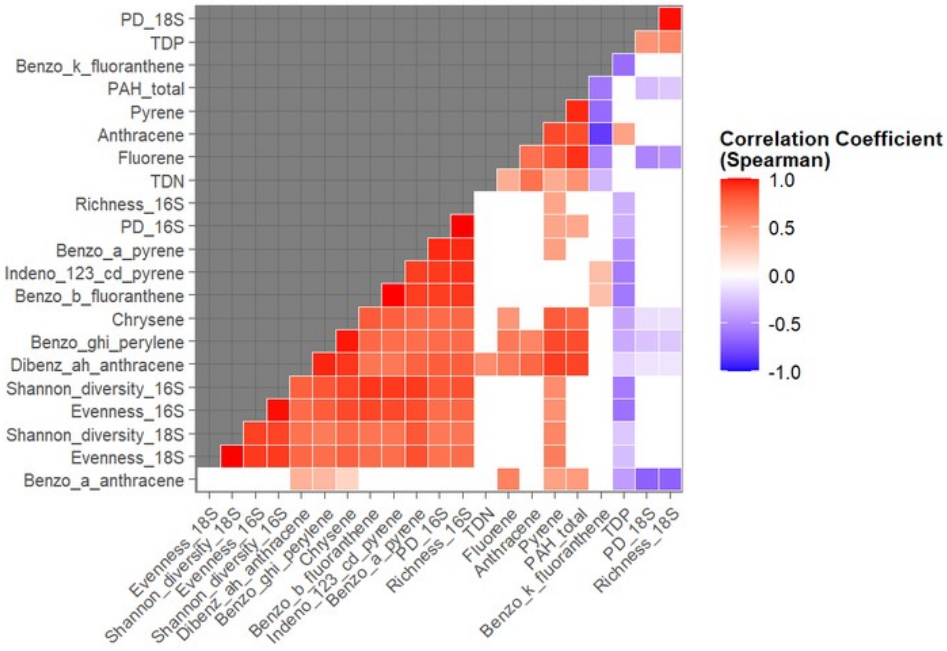

Fig. S7 Correlogram illustrating correlation (Spearman) between diversity and physicochemical parameters. White indicates a p-value > 0.05 and therefore nonsignificant correlation.

Table S1 Alpha diversity indices for glacier samples.

Barcode	Alpha diversity index	Ice (*n*=3)	Cryoconite (*n*=3)	Sediment (*n*=1)
16S (prokaryotic)	Richness	1280.67 ± 42.72	2034.67 ± 869.20	3094
	Faith's PD	78.82 ± 3.92	104.48 ± 26.87	133.81
	Shannon	2.96 ± 0.98	4.56 ± 0.55	6.14
	Evenness	0.41 ± 0.14	0.60 ± 0.04	0.76
18S (eukaryotic)	Richness	214 ± 98.21	281 ± 21.63	320
	Faith's PD	26.08 ± 8.32	33.45 ± 2.43	32.54
	Shannon	0.90 ± 0.44	3.08 ± 0.16	4.15
	Evenness	0.17 ± 0.08	0.55 ± 0.02	0.72

Table S2 Mean relative abundances (%) of microbial taxa found in glacier samples.

Barcode		Ice (*n*=3)	Cryoconite (*n*=3)	Sediment (*n*=1)
16S (prokaryotic)	*Cyanobacteria*	58.3 ± 16.1	29.3 ± 1.55	7.70
	Proteobacteria	17.1 ± 8.32	30.3 ± 4.28	43.1
	Bacteroidetes	11.2 ± 2.13	22.3 ± 4.19	26.0
	Actinobacteria	5.63 ± 1.94	8.16 ± 0.68	9.77
	Archaeplastida	88.7 ± 9.26	25.0 ± 0.86	29.0
18S (eukaryotic)	*SAR*	6.11 ± 6.79	34.3 ± 0.18	31.9
	Opisthokonta	4.14 ± 1.95	27.5 ± 3.77	12.8
	Amoebozoa	0.62 ± 0.66	11.6 ± 4.20	16.5

Table S3 Concentrations (ng PAH/kg) of congeners in glacier samples.

PAH Compound	S_2 (Sediment)	S_3 (Cryoconite)	S_4 (Ice)	S_5 (Cryoconite)	S_6 (Ice)	S_7 (Ice)	SUM (ng/kg)
Acenapthylene	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD
Anthracene	147	50.4	92.7	<LOD	<LOD	<LOD	291
Benzo[*a*]anthracene	5.40	2.03	<LOD	<LOD	<LOD	<LOD	7.43
Benzo[*a*]pyrene	9.09	2.90	<LOD	7.98	1.21	2.55	23.7
Benzo[*b*]fluoranthene	11.2	2.54	0.652	10.9	6.14	2.84	34.3
Benzo[*ghi*]perylene	32.2	2.00	<LOD	6.60	9.43	3.54	53.8
Benzo[*k*]fluoranthene	1.29	2.84	1.55	21.3	22.6	12.8	62.4
Chrysene	64.7	15.1	6.30	24.4	30.5	9.12	150
Dibenz[*a,h*]anthracene	1.20	<LOD	<LOD	0.173	<LOD	<LOD	1.37
Fluorene	345	71.0	189	<LOD	136	208	948
Indeno[123-*cd*]pyrene	8.52	1.41	<LOD	8.68	4.30	1.62	24.5
Phenanthrene	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD
Pyrene	111	37.6	25.0	<LOD	18.1	20.4	212

Table S4 TP and TN measured in glacier samples.

Sample	Solid TP (mg/g)	Solid TN (mg/g)
Ice (n = 3)	1.00 ± 0.03	2.63 ± 0.30
Cryoconite hole (n = 2)	0.89	3.38
Sediment (n = 1)	1.18	3.04

REFERENCES:

1. Muyzer G, de Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol [Internet]. 1993 [cited 2021 Jun 19];59(3):695–700. Available from: https://pubmed.ncbi.nlm.nih.gov/7683183/

2. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci U S A [Internet]. 2011 Mar 15 [cited 2021 Jun 19];108(SUPPL. 1):4516–22. Available from: www.pnas.org/cgi/doi/10.1073/pnas.1000080107

3. Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner HW, et al. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Mol Ecol [Internet]. 2010 Mar [cited 2021 Jun 19];19(SUPPL. 1):21–31. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-294X.2009.04480.x

4. Andrews S. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data [Internet]. 2010 [cited 2021 Jun 19]. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

5. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics [Internet]. 2014 Aug 1 [cited 2021 Jun 19];30(15):2114–20. Available from: https://pubmed.ncbi.nlm.nih.gov/24695404/

6. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics [Internet]. 2010 Aug 12 [cited 2021 Jun 18];26(19):2460–1. Available from: https://pubmed.ncbi.nlm.nih.gov/20709691/

7. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data [Internet]. Vol. 7, Nature Methods. Nat Methods; 2010 [cited 2021 Jun 19]. p. 335–6. Available from: https://pubmed.ncbi.nlm.nih.gov/20383131/

8. Edgar RC, Flyvbjerg H. Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics [Internet]. 2015 Dec 18 [cited 2021 Jun 19];31(21):3476–82. Available from: http://drive5.com/usearch.

9. Edgar R. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. bioRxiv [Internet]. 2016 Oct 15 [cited 2021 Jun 18];081257. Available from: https://doi.org/10.1101/081257

10. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res [Internet]. 2013 Jan 1 [cited 2021 Jun 19];41(D1). Available from: https://pubmed.ncbi.nlm.nih.gov/23193283/

11. Hao J, Ho TK. Machine Learning Made Easy: A Review of Scikit-learn Package in Python Programming Language [Internet]. Vol. 44, Journal of Educational and Behavioral Statistics. SAGE Publications Inc.; 2019 [cited 2021 Jun 19]. p. 348–61. Available from: https://journals.sagepub.com/doi/10.3102/1076998619832248

12. Bokulich N, Dillon M, Bolyen E, Kaehler B, Huttley G, Caporaso J. q2-sample-classifier: machine-learning tools for microbiome classification and regression. J Open Source Softw [Internet]. 2018 Oct 23 [cited 2021 Jun 19];3(30):934. Available from: http://joss.theoj.org/papers/10.21105/joss.00934

13. Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, et al. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 2017 5:1 [Internet]. 2017 Mar 3 [cited 2021 Sep 30];5(1):1–18. Available from: https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-017-0237-y

14. Sanders HL. Marine Benthic Diversity: A Comparative Study. Am Nat [Internet]. 1968 May 15 [cited 2021 Jun 19];102(925):243–82. Available from: https://www.journals.uchicago.edu/doi/abs/10.1086/282541