

Impact of storage and extraction methods on peat soil microbiomes

Dylan Cronin^{1,2,*}, Yueh-Fen Li^{2,*}, Paul Evans³, Gene W. Tyson⁴, Ben J. Woodcroft⁴, Virginia I. Rich^{1,2,5} and IsoGenie 2016 and 2019 Field Teams

¹ Center of Microbiome Science, Ohio State University, Columbus, Ohio, United States

² Department of Microbiology, Ohio State University, Columbus, Ohio, United States

³ School of Chemistry and Molecular Biosciences, The Australian Centre for Ecogenomics, University of Queensland, St Lucia, Queensland, Australia

⁴ Centre for Microbiome Research, School of Biomedical Sciences, Translational Research Institute, Queensland University of Technology, Woolloongabba, Queensland, Australia

⁵ Byrd Polar and Climate Research Center, Ohio State University, Columbus, Ohio, United States

* These authors contributed equally to this work.

ABSTRACT

Recovered microbial community structure is known to be influenced by sample storage conditions and nucleic acid extraction methods, and the impact varies by sample type. Peat soils store a large portion of soil carbon and their microbiomes mediate climate feedbacks. Here, we tested three storage conditions and five extraction protocols on peat soils from three physicochemically distinct habitats in Stordalen Mire, Sweden, revealing significant methodological impacts on microbial (here, meaning bacteria and archaea) community structure. Initial preservation method impacted alpha but not beta diversity, with in-field storage in LifeGuard buffer yielding roughly two-thirds the richness of in-field flash-freezing or transport from the field on ice (all samples were stored at -80°C after return from the field). Nucleic acid extraction method impacted both alpha and beta diversity; one method (*the PowerSoil Total RNA Isolation kit with DNA Elution Accessory kit*) diverged from the others (*PowerMax Soil DNA Isolation kit-High Humic Acid Protocol*, and three variations of a modified *PowerMax Soil DNA/RNA isolation kit*), capturing more diverse microbial taxa, with divergent community structures. Although habitat and sample depth still consistently dominated community variation, method-based biases in microbiome recovery for these climatologically-relevant soils are significant, and underscore the importance of methodological consistency for accurate inter-study comparisons, long-term monitoring, and consistent ecological interpretations.

Submitted 5 April 2024

Accepted 2 December 2024

Published 23 December 2024

Corresponding author

Virginia I. Rich, rich.270@osu.edu

Academic editor

Liang Wang

Additional Information and
Declarations can be found on
page 20

DOI 10.7717/peerj.18745

© Copyright

2024 Cronin et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Ecosystem Science, Microbiology, Soil Science, Climate Change Biology

Keywords Microbiome, Soil, Storage, Extraction, Methods, Peatland, Stordalen

INTRODUCTION

Microbiome data are shaped by the methods employed for sample storage and processing. Differential biases among methods can overshadow genuine ecologically-driven differences in community structure (e.g., *Kennedy et al., 2014*; *Hermans, Buckley & Lear, 2018*; *Elie et al., 2023*; *Galla et al., 2024*). For example, a meta-analysis of human

microbiota 16S rRNA gene sequencing studies revealed significant impacts from variation in extraction protocol, 16S rRNA gene target region, and sequencing platform (Lozupone *et al.*, 2013). Such findings underscore the non-negligible biases introduced at different stages of data acquisition, necessitating careful consideration when conducting research or comparing results across studies. In this study, our specific focus is on sample storage protocols at the time of sample collection and on nucleic acid extraction.

Habitat physicochemistry and cell densities greatly impact methodological choices and outcomes. Soils can be particularly challenging due to the presence of co-extracted inhibitors such as humic acids (the dominant component of humic substances, a variable organic component of soil and decaying organic matter). While humic acids and humic substances lack consistent definitions, and may be created by soil chemical extractions (Lehmann & Kleber, 2015), they can bind nucleic acids and block some enzymes' binding sites, inhibiting PCR and other enzymatic reactions (Tebbe & Vahjen, 1993; Albers *et al.*, 2013; Sidstedt, Rådström & Hedman, 2020). Humic acids can also interact with preservation or extraction chemicals, such as the ammonium sulfate in RNAlater, decreasing nucleic acid yield (Rissanen *et al.*, 2010; Wnuk *et al.*, 2020). Peat soils are particularly high in humic acids, and have sufficient cell densities (*e.g.*, 10^8 – 10^9 cells/gram, Woodcroft *et al.*, 2018) that specialized low-biomass methods are not required (*i.e.*, they provide a high-humic test case for 'standard' soil storage and extraction methods).

Common storage methods for soil samples range from simple collection into a chilled cooler and then freezer storage, to the addition of preservation buffers such as LifeGuard (LG) or RNAlater followed by freezer or room temperature storage, to flash-freezing in liquid nitrogen (LN), then freezer storage. While LN is frequently favored for DNA and RNA preservation and is also the preferred method of the Earth Microbiome Project (EMP) ([dx.doi.org/10.17504/protocols.io.pfqdjmww](https://doi.org/10.17504/protocols.io.pfqdjmww)), a large-scale initiative to create standards and characterize microbial life in different environments, transport of LN to field sites is not always feasible and can pose safety challenges. Storage buffers such as RNAlater offer convenience in the field and sample stability during transport but can alter microbial community structure and lower DNA/RNA yield (Rissanen *et al.*, 2010; McCarthy *et al.*, 2015; Smenderovac *et al.*, 2024). The most cost-effective approach is placing samples on ice or frozen gel packs in the field for transport to a freezer, however this does not arrest transcription so is suitable for DNA analysis only.

There is no universal methodology for nucleic acid extraction from soil samples due to soil's wide range of physicochemical properties (*e.g.*, pH, texture, organic content, inorganic matter composition, alkalinity, *etc.*) and potential presence of reaction-inhibitors such as humic acids. The general steps of extraction of nucleic acids are cell lysis, separation of nucleic acids from proteins and other impurities, and precipitation and purification of nucleic acids. Traditional extractions employ organic solvents, are time-consuming and low throughput, and often yield insufficient or impure extracts for downstream analyses. Additionally, user differences can introduce significant variation in results. Commercial nucleic acid extraction products (*e.g.*, the Qiagen DNeasy PowerSoil DNA Isolation kit, originally the Mobio PowerSoil DNA Isolation kit) exist that are

tailored to different sample types and designed to minimize inhibitor co-extraction, reduce protocol duration, increase throughput, and improve standardization. Generally, kits can differ by homogenization and lysis method (physical and/or chemical), contaminant minimization approach (e.g., proprietary technologies for removal of native co-extracted chemicals that act as downstream enzymatic inhibitors), and nucleic acid retrieval method (e.g., silica membrane or ion-exchange columns). A major branch point in extraction kit choice is the option for DNA and RNA co-extraction, which facilitates parallel examination of metagenomes and metatranscriptomes by avoiding differential bias of separate DNA and RNA protocols.

To address storage and extraction impact on microbiome profiles in the consequential habitat of permafrost-associated peat, we evaluated three storage conditions and five DNA/RNA co-extraction protocols. In this study, our focus is on permafrost-associated peatlands, as the accurate characterization of their microbiota is relevant to improved predictions of climate feedbacks in these rapidly changing systems (e.g., [McCalley et al., 2014](#)). We tested samples from Stordalen Mire, a thawing permafrost ecosystem which has been the focus of multi-omics studies for more than a decade ([Mondav et al., 2014](#); [Woodcroft et al., 2018](#); [Singleton et al., 2018](#); [Martinez et al., 2019](#)). Understanding the impact of storage and extraction is particularly crucial for long-term studies because kits may be discontinued, making comparisons over time particularly difficult. Therefore, the ability to leverage alternative solutions becomes a significant consideration. Due to the longevity of these studies, and their role as a major source for microbiome data for such habitats, the methods compared here include the approaches used regularly at the site. Soils were sampled from three thaw-stage habitats with distinct peat soils: an aerobic palsa, a partially inundated bog, and a fully inundated fen.

MATERIALS AND METHODS

Overview

Replicate tubes of peat from two depths in palsa, bog, and fen were preserved and extracted in multiple ways, the extracted DNA and RNA yield and quality were measured, and basic microbiome profiles were assessed *via* 16S rRNA gene amplicon sequencing ([Fig. 1](#)). Palsa, bog, and fen reflect stages of permafrost thaw, with a marked shift in habitat features: dry raised palsas slump into ombrotrophic bogs with perched water tables and then fully-thawed and inundated minerotrophic fens ([Malmer et al., 2005](#)); vegetation shifts concomitantly from ericaceous shrubs to Sphagnum moss-dominated and then sedge-dominated ([Hough et al., 2022](#)); thaw progression decreases peat C/N ratios and increases organic matter humification indices, with changing dissolved organic matter profiles ([Hodgkins et al., 2014](#); [Cory et al., 2022](#); [Wilson et al., 2022](#)); microbiomes show increasing portions of anaerobic lineages, changing diversity, and shifts in carbon-degradation and methane cycling potential ([Mondav et al., 2017](#); [Woodcroft et al., 2018](#); [Singleton et al., 2018](#); [Martinez et al., 2019](#); [Ellenbogen et al., 2024](#); [McGivern et al., 2024](#); [Sun et al., 2024](#)) and activities ([Wilson et al., 2021](#); [Fofana et al., 2022](#)); and carbon fluxes shift, with thaw-associated increases in carbon dioxide uptake and in methane emissions ([Malmer et al., 2005](#); [McCalley et al., 2014](#); [Varner et al., 2021](#)).

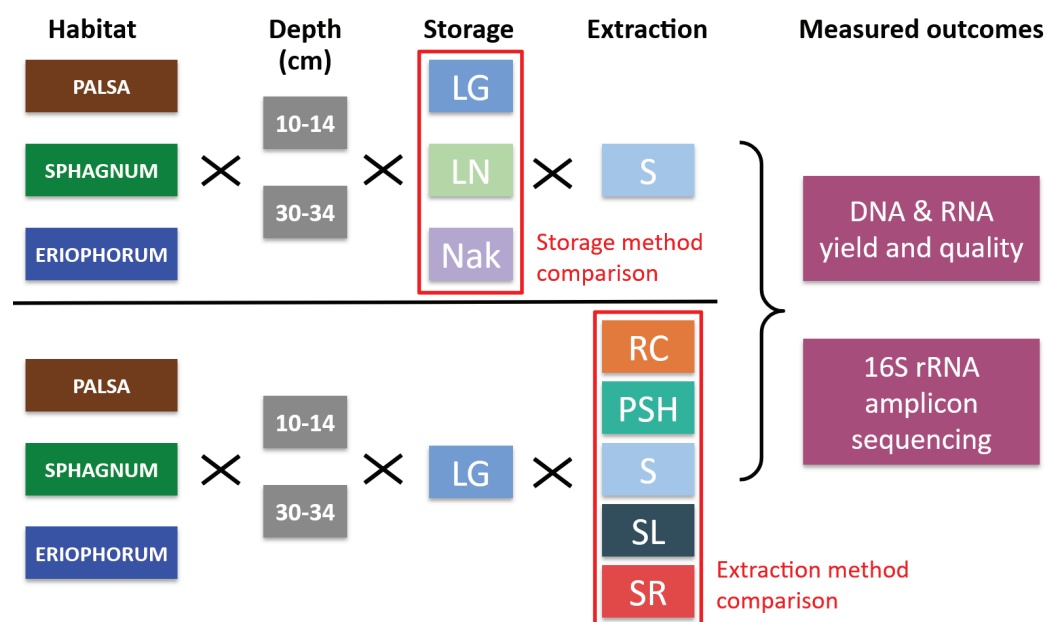


Figure 1 Overview of Methods. Soil from all three habitats, from two depths, were stored under three conditions and tested with a single extraction protocol ('S'). Material stored in LifeGuard was extracted by five methods. Triplicate extractions were performed for each method combination. Extraction success and differential bias was characterized by DNA and RNA yield and quality, and by 16S rRNA amplicon sequencing. Sample Storage Experiment: 3 sites \times 2 depths \times 3 storage methods \times 3 replicates \times 1 extraction method. Nucleic Acids Extraction Experiment: 3 sites \times 2 depths \times 1 storage method \times 3 replicates \times 5 extraction methods. Abbreviations: Storage: LG, LifeGuard buffer; LN, liquid nitrogen; NAK, no buffer; Extraction: S, IsoGenie standard protocol; SL, IsoGenie Standard protocol with LifeGuard removed before extraction; SR, IsoGenie standard protocol with alternative reagents; PSH, PowerMax Soil DNA Isolation kit-High Humic Acid Protocol; RC, PowerSoil Total RNA Isolation kit with DNA elution accessory kit.

Full-size [DOI: 10.7717/peerj.18745/fig-1](https://doi.org/10.7717/peerj.18745/fig-1)

Sample collection

Sampling occurred at Stordalen Mire in Abisko, Sweden, at "AJ's site" from the three major permafrost thawing stages, palsa, bog, and fen (latitude and longitude: palsa = "AJ-Palsa" N 68 21.2708, E 19 02.8048, bog = "AJ-Sphagnum" N 68 21.2663, E 19 02.8016, fen = "AJ-Eriophorum" N 68 21.2660, E 19 02.7959), in July 2016. The field site at Stordalen Mire is under the jurisdiction of the Abisko Scientific Research Station, which is managed by the Swedish Polar Research Secretariat. Sampling there requires application, approval and documentation, and our team filed the standard application to sample peat and porewater. Application approval is issued by Magnus Augner, Deputy Director-General, and Manager of Abisko Scientific Research Station, Swedish Polar Research Secretariat.

Information on sample collection and porewater pH is in [Table S1](#). The porewater pH did not differ substantially between bog and fen (and the dry palsas lacked porewater), although typically at this site bogs are more acidic ([Hodgkins et al., 2014](#)). One push-core per habitat was collected and the peat core was subsampled at 10–14 cm ("shallow") and 30–34 cm ("deep"). To ensure homogeneity, the subsampled peat was initially placed in ziplock bags and gently kneaded. We did not add community standards because our goal was to evaluate relative methodological performances for the resident, complex soil

microbiomes; the richness and diversity of lineages recovered, and to a lesser extent DNA yield and quality, were the salient performance metrics. Spike-ins perform across a limited range of cell types, and generally have limited time to interact with soil physicochemistry, and so their extraction results reflect the interaction of storage conditions and extraction methods for those spike-in cells rather than that of the native microbial cells. With well-homogenized starting material in our experiments, the DNA and community metrics provide the clearest view of relative method performance.

Sample storage experiment

Replicate aliquots of ~4 ml (equivalent to ~2 g) of peat were placed into 15 ml Falcon tubes and subjected to one of three storage conditions:

1. *LifeGuard buffer* (LG; product, Qiagen)—~12 ml of LG was added to the tube, followed by vigorous hand shaking to achieve a thorough mixture of LG and peat. (18 replicate tubes per depth, for downstream comparisons).
2. *Liquid Nitrogen* (LN)—the tube was immediately flash frozen in LN in the field. (six replicate tubes per depth).
3. *No buffer* (NAK)—nothing was added to the tube (nine replicate tubes per depth).

All samples were placed on ice packs frozen from -80°C freezers in coolers and transported to the nearby field station (Abisko Scientific Research Station, Abisko, Sweden within at most 6 h), where they were transferred to a -80°C freezer until shipping. The samples were shipped at -55°C , and then returned to -80°C freezer storage until extraction. The storage time in -80°C was 6–7 months. To compare storage conditions, three replicate vials from each were thawed on ice, and were extracted with the “*Standard* (S)” extraction protocol, a modified PowerMax Soil DNA/RNA isolation kit (Qiagen Cat# 12966-10). Modifications to the manufacturer’s protocol were described in the “Nucleic acids extraction experiment” section. This protocol has been used from 2010 through present for multi-omics studies by several sequential US-based projects (IsoGenie and EMERGE). In 2010 this protocol demonstrated superior yield and quality for DNA and RNA across the three thaw stage peats, compared to several other methods (data not shown).

Nucleic acids extraction experiment

Next, to compare extraction methods, replicate vials from the LG storage condition were thawed on ice (extractions were not performed on all storage methods because the long-term microbiome research at the site has all used LG, so the performance of extraction methods in that background storage condition was most relevant to the ongoing site research). Extractions were performed in triplicate by each of five extraction methods, listed below. All methods were variations of the PowerSoil and PowerMax kits. These kits’ chemistry includes a patented inhibitor removal technology (IRT) designed to reduce humic acid impacts on soil extractions, which was a requirement in these humic-laden peat soils.

1. *PowerSoil Total RNA Isolation kit and DNA Elution Accessory kit* (“RC”, Qiagen Cat# 12866-25 and 12867-25)—The manufacturer’s protocol was followed for this extraction method. This kit employs an affinity column for nucleotide recovery, while the remaining methods use size-exclusion columns (RC—an abbreviation for “RNA Column”; this kit uses an anion-exchange column for nucleic acid recovery, instead of the silica membrane column).

2. *PowerMax Soil DNA Isolation kit-High Humic Acid Protocol* (PSH, Qiagen Cat# 12988-10HH)—This protocol was a modification of the PowerMax Soil DNA Isolation kit (Qiagen Cat# 12988-10) intended to co-extract DNA and RNA and to further minimize humic acids, and provided by previous MoBio technical support, Drs. Suzanne Kennedy, Michelle Carlson and Mr. Yoshiaki Kono. Briefly, to each sample, 10 ml of the bead beating buffer, 1 ml of buffer C1, 0.5 ml of buffer C2, and 5 ml of phenol:chloroform:isoamyl alcohol were added, shaken to mix well and vortexed at maximum speed for 10 min for homogenization and initial cell lysis. Tubes were centrifuged at $4,700\times g$ for 6 min, then the supernatant was transferred to a clean 50 ml tube. One-third of the supernatant volume of buffer C3 was added, mixed by inverting 4–6 times, then the tubes were placed on ice for 5 min. The centrifugation and supernatant transfer was then repeated. To this volume, an equal volume of the buffer C4 and of 100% ethanol were added and mixed well. The mixture was transferred to the spin column in batches, and centrifuged at $4,700\times g$ for 2 min, and the flow through was discarded, until all volume had been passed through the spin column. The column was then washed with freshly-made wash buffer (7.2 ml of buffer C4 mixed with 8.8 ml of 100% ethanol), then washed with buffer C5, then washed by 100% ethanol. All flow-through was discarded, and the column was spun dry at $4,700\times g$ for 8 min to remove any residual ethanol. The column was transferred to a clean 50 ml tube and placed in the hood with the lid open for 10 min to further dry the column. Finally, the nucleic acids were eluted by adding 5 ml DEPC-treated water and spinning at $4,700\times g$ for 2 min.

3. *Standard (S)*—This was a modification of the PowerMax Soil DNA/RNA isolation kit (Qiagen Cat# 12966-10). For the S and both the SR and SL methods described below, an average of 4.5 g peat (the weight of the peat without any additions; SD 2.4 g), or 8.9 g of peat with LG buffer (the weight of the peat with LG buffer; SD 1.4 g), were added to 0.1 mm ceramic bead tubes. All three methods used these modifications to the manufacturer’s protocol: (i) omitting the addition of beta-mercaptoethanol; (ii) tubes with LG received proportionally increased reagent amounts to maintain the concentration and strength of solutions; (iii) an additional ethanol wash of the nucleic acids-bound column was performed to enhance impurity removal.

4. *Standard with alternative reagents (SR)*—The original kit used in the S extraction described above was discontinued, and to preserve comparability of the interannual data series, the teams reconstructed it with identical (according to Qiagen technical support) but differently named versions of the reagents. Specifically, RD1 was replaced with PM1, RD2 with C3, RD3 with PM4, RD4 with C5, and RD5 with DEPC-treated water. Thus, SR was identical to the “S” method, except that the reagents used in the original kit were replaced with identical alternative reagents provided by Mobio/Qiagen.

5. *Standard with LG spin down (SL)*—Identical to the “S” method, except that peat materials stored in LG were centrifuged at 4,700×g for 5 min to separate peat from LG, which was removed before adding the peat materials to the extraction process.

All extractions were performed within a 35-day window of time, except the second extracts of RC used for the “Repetition of Subsets of the Experiment” (see below), which were performed 6 months later. The up-to 6-month difference in −80 °C storage should not impact sequencing results, associated community, and methods interpretations (Laufer *et al.*, 2010; Carroll *et al.*, 2012; Kia *et al.*, 2016; Tap *et al.*, 2019). Randomization of extractions was not performed.

DNA and RNA purification

Extracted nucleic acids were concentrated by ethanol precipitation and resuspended in 100 µl of TE buffer. The integrity of the nucleic acids was checked by agarose gel electrophoresis (1%). For all methods except RC, which sequentially elutes DNA and RNA from an affinity column, DNA and RNA were then purified as follows. Total extracted nucleic acids were aliquoted into two 2 ml tubes, one of which was treated with RNase, the other with DNase, followed by phenol:chloroform purification to remove enzymes and impurities. The DNA and RNA were ethanol-precipitated, the pellets were resuspended in 50 µl and 25 µl of TE buffer respectively. The final purified DNA and RNA were quantified using the Qubit 3.0 system, and the quality of the extracted RNA was evaluated using TapeStation analysis at the Genome Shared Resources (GSR) facility at the Ohio State University. The nucleic acids were stored at −80 °C for downstream sequencing analysis.

16S rRNA gene amplicon sequencing

The microbiome of all samples were characterized in order to test differential biases in community recovery from the storage and extraction methods, and to ensure extracted DNA was of sufficient quality for amplification and sequencing. 16S rRNA gene amplicon sequencing was performed at the Australian Centre of Ecogenomics (ACE) at the University of Queensland, Australia.

At ACE, the V6-V8 region was targeted using the universal primer pair Univ_SSU_926F-1392wR: 926F (5'-AAACTYAAAKGAATTGRCGG-3') and 1392wR (5'-ACGGGCGGTGWGTRC-3') primers (Engelbrektson *et al.*, 2010) modified to contain Illumina specific adapter sequence (the Nextera transposase adapters for tagmentation) (926F:5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAACTYAAAKGAATTGRCGG-3' and 1392wR:5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACGGGCGGTGWGTRC-3'). Preparation of the 16S library followed the workflow outlined by Illumina (#15044223 Rev.B), with the substitution of Q5 Hot Start High-Fidelity 2X Master Mix (New England Biolabs) in standard PCR conditions. PCR amplicons were purified using Agencourt AMPure XP beads (Beckman Coulter), and indexed with unique 8 bp barcodes using the Illumina Nextera XT 384 sample Index Kit A-D (Illumina FC-131-1002) in standard PCR conditions with Q5 Hot Start High-Fidelity 2X Master Mix. Indexed amplicons were pooled in equimolar concentrations and

sequenced on MiSeq Sequencing System (Illumina) using paired-end sequencing with V3 300 bp chemistry according to manufacturer's protocol.

Repetition of subsets of the experiment

For the storage variants of the bog habitat only, in both depths, the LN-preserved communities were outliers. To rule out error or technical biases in amplicon region or sequencing center, the storage comparison experiment was repeated for the bog habitat only, using a different region of the rRNA gene and a different sequencing center. This required new material, so in 2019, the same bog site was revisited, and an additional push-core was collected and again subsampled at 10–14 cm (“shallow”) and 30–34 cm (“deep”), and processed as described above. After extraction, these 18 extracts were sequenced at Argonne National Laboratory, using a different primer set. The V4 region of the 16S rRNA gene was targeted with region-specific primers that include sequencer adapter sequences used in the Illumina flowcell (Caporaso *et al.*, 2011, 2012). The forward amplification primer also contained a twelve base barcode sequence that supports pooling of up to 2,167 different samples in each lane (Caporaso *et al.*, 2011, 2012; Walters *et al.*, 2016). The specific primers used were 515F (Parada)–806R (Apprill), forward-barcoded: FWD: GTGYCAGCMGCCGCGGTAA, REV: GGACTACNVGGGTWTCTAAT (Apprill *et al.*, 2015; Parada, Needham & Fuhrman, 2016). Each 25 µl PCR reaction contained 9.5 µl of MO BIO PCR Water (Certified DNA-Free), 12.5 µl of QuantaBio's AccuStart II PCR ToughMix (2x concentration, 1x final), 1 µl Forward Primer (5 µM concentration, 200 pM final), 1 µl Golay barcode tagged Reverse Primer (5 µM concentration, 200 pM final), and 1 µl of template DNA. The conditions for PCR were initial denaturation at 94 °C for 3 min, followed by 35 cycles of 94 °C for 45 s, 50 °C for 60 s, and 72 °C for 90 s, then a final extension of 10 min at 72 °C to ensure complete amplification. Amplicons were quantified *via* PicoGreen (Invitrogen, Whitefield, Bangalore) on a plate reader (Infinite 200 PRO, Tecan), then pooled at equimolar amounts into a single tube. This pool was purified *via* AMPure XP Beads (Beckman Coulter), quantified *via* fluorometer (Qubit, Invitrogen, Carlsbad, California), and diluted to 2 nM, denatured, and then diluted to a final concentration of 6.75 pM with a 10% PhiX spike for sequencing on the Illumina MiSeq. Amplicons were sequenced by Illumina paired-end sequencing at the Environmental Sample Preparation and Sequencing Facility (ESPSF) at Argonne National Laboratory, on a 151 bp × 12 bp × 151 bp MiSeq run using customized procedures (Caporaso *et al.*, 2012).

This experiment with fresh samples and sequencing of a different 16S rRNA gene region at a different sequencing center did not recapitulate the strong outlier nature of the bog LN samples seen in the original data (Fig. S1). The bog-LN data were therefore excluded from the overall storage analysis. Accessions for all data can be found in Table S2.

In a separate experimental comparison to test the reproducibility of the divergence of the RC extraction results, we sent DNA aliquots from the original 2016 RC and non-RC samples to Argonne to be resequenced. The three RC samples produced no successful sequencing libraries (in a run that otherwise worked), which we attribute to degradation in the freezer over the 7 years since extraction. The initial RC data were therefore retained in the overall analysis.

Data processing

All amplicon sequences were processed through the QIIME2 (v2020.2) bioinformatics platform (Bolyen *et al.*, 2019) for operational taxonomic unit (OTU), amplicon sequence variant (ASV), and taxonomic assignment. Briefly, all samples were demultiplexed prior to import into the QIIME2 platform; idemp (<https://github.com/yhwu/idemp>) was used for the Argonne-sequenced samples. Demultiplexed reads were then imported into QIIME2. Forward reads were used for all subsequent processing and analysis. Amplicon sequence variants (ASVs) were then formed through the DADA2 plugin (Callahan *et al.*, 2016), where trimming was performed for both primer removal and to maintain high quality reads. For ACE sequencing reads, reads were truncated at 250 bp, and 13 bp were trimmed from the left, as parameters for the DADA2 plugin. Average sequence quality of the raw reads was calculated through FastQC (Andrews, 2010) and aggregated through MultiQC (Ewels *et al.*, 2016) (Table S3). Sequences were also clustered at 97% identity OTUs through QIIME2 for separate comparisons to the results based on ASVs. For Argonne sequencing reads, reads were truncated at 150 bp, and using the same left trimming parameterization as above. ASVs were then assigned a taxonomy through the QIIME naive Bayes classifier, leveraging the SILVA (v138.1) 99% 16S-only seven-level majority taxonomy database for training with the appropriate forward and reverse primer pairs. Minimum and maximum length for read extraction for ACE sequenced samples was 200 and 600 bp, respectively. For Argonne sequences, the minimum and maximum length to extract reads for the feature classifier from the QIIME2 plugin were 100 and 400 bp, respectively. All ASVs taxonomically annotated as Mitochondria, Chloroplast, or with a Kingdom-level assignment of “Unassigned” were removed from this analysis. Finally, PICRUSt2 (Douglas *et al.*, 2020) was run on the representative sequences from QIIME2 using the “picrust2_pipeline.py” with the stratified flag enabled.

Statistical analysis

The OTU tables (Table S4), taxonomic classification (Table S5), and PICRUSt2 results were imported into R using the phyloseq package (McMurdie & Holmes, 2014). First, for beta diversity, the phyloseq object was transformed into relative abundances using the microbiome package (Lahti & Shetty, 2017), and ordinations were generated from Bray-Curtis dissimilarity. In our primary analysis, we used the adonis2 test to assess the effect of storage and extraction methods on community distances, with a model formula of $\text{Dist} \sim \text{Habitat} * \text{Depth} * \text{Storage/Extraction}$ using the vegan package (Oksanen, 2022). This assessed differences overall in the data across all habitats and depths. Subsequently, we applied the adonis2 test separately for each habitat and depth to isolate and compare the specific differences in extraction and storage methods within these subcategories. Given the multiple comparisons involved in this procedure, we adjusted the *p*-values using a Benjamini-Hochberg correction.

Second, for alpha diversity, the phyloseq object was randomly rarefied without replacement to an even sequencing depth to match the minimum number of post-QC sample amplicon counts across the dataset. For both storage and extraction comparisons, a Kruskal-Wallis test was performed to determine alpha diversity differences between all

groups, followed by a Wilcoxon rank sum test for pairwise comparisons, adjusted using Benjamini-Hochberg p -value correction.

In the differential abundance analysis, taxa present in less than 10% of samples were initially filtered out to reduce data sparsity. Subsequently, the refined phyloseq dataset, with 'Extraction' designated as the variable of interest, was transformed into a DESeq2 (Love, Huber & Anders, 2014) object. To account for variations in sequencing depth, size factors were estimated using geometric means calculated across all samples for each taxa. The Wald test, employing a parametric fit, was then applied to discern taxa with significant abundance discrepancies across extraction methods. We then applied a 0.01 alpha threshold as significantly different. This filtered list was visually represented in a heatmap, leveraging the pheatmap package in R (Kolde, 2019).

Given the research focus on methane emissions and methanogens at this site, we also examined treatment impact on the methanogen portion of the community, identified based on genus-level taxonomic annotation (Fig. S2A; Table S6). These methanogens were then subset down to the phylum of Euryarchaeota in the shallow bog as well (Figs. S2B, S2C).

For phylogeny, sequence alignment was conducted using the SILVA database version 138.1 for small subunit ribosomal RNA (Quast et al., 2012). The SINA tool (v1.7.2) was employed for the alignment process using the SILVA database as the reference (Pruesse, Peplies & Glöckner, 2012). After alignment, sequences were trimmed using the trimal (v1.4. rev15) tool (Capella-Gutiérrez, Silla-Martínez & Gabaldón, 2009). The specified trimming parameters retained only those columns in the alignment where at least 90% of the sequences had a gap and those with a consensus of at least 60%. The phylogenetic tree was imported into R using the ape package (Paradis & Schliep, 2019) and subsequently midpoint rooted before any further calculations. Weighted and unweighted Unifrac distances were calculated using the rbiom package (Smith, 2024), and Faith's Phylogenetic Diversity was calculated through the picante package (Kembel et al., 2010). Tests for differences in between the communities overall were assessed, as described above, using the adonis2 test.

To assess the impact of different extraction methods on microbial community structure, we calculated pairwise distances between samples using two distance metrics: Bray-Curtis and Weighted UniFrac. Pairwise comparisons were filtered to ensure each comparison had samples from the same habitat and depth, but with differing extraction methods. Significance between extraction methods was determined through pairwise Wilcoxon rank sum tests. To account for the multiplicity of tests and control the false discovery rate, p -values were adjusted using the Benjamini-Hochberg correction. The adjusted p -values were then visualized using compact letter displays, with significant differences between extraction methods represented by unique letters.

To explore the cause underlying the all-habitat divergence of RC-extracted communities and the potential for contamination from other samples during library preparation, we examined the shared and unique lineages recovered by each extraction method, and visualized them via the ComplexUpset package (Krassowski, 2020).

RESULTS

Yield

All methods successfully yielded adequate DNA and RNA according to agarose gel electrophoresis, Qubit measurements, and TapeStation analysis (Fig. S3, and data not shown). Adequacy is defined here as a sufficient quantity to meet the contemporary DNA and RNA input requirements for metagenomic and metatranscriptomic sequencing (100 and 300 ng, respectively). It was not possible to report yield-per-gram peat for samples stored in LG since extractions occurred on a peat-buffer slurry, with a subset of slurry used for extraction. There was a difference in quality for RC extracts, which were visibly brown, and had an average absorbance A260:280 of 1.65, compared to 1.92 for the other methods, indicating lower nucleic acid purity in the RC samples.

Storage

There was no statistically significant difference among storage methods at the level of whole-community dissimilarity based on ASV abundances (Fig. 2A). Variation attributable to habitat, depth, and storage methods was approximately 22% ($p = 0.001$), 14% ($p = 0.001$), and 3% ($p = 0.084$), respectively. When we isolated habitat and depth to specifically compare community structure differences due to storage conditions, a significant difference ($p = 0.048$) emerged only in the shallow palsa after correcting for multiple comparisons (Table S7), attributable to the LG treatment (Fig. 2). In addition, samples stored in LG had significantly lower alpha diversity indices of richness and Shannon's diversity, with an average of ~42% fewer observed ASVs than samples in LN (and no significant differences between NAK and LN storage) (Figs. 2B–2C). In comparing the differential abundance of lineages between LG, LN, and NAK, eight OTUs were consistently depleted in LG samples (Fig. S4). However, when leveraging PICRUSt2 relative abundances of pathways, the storage condition was significant ($p < 0.05$) with variation attributable to habitat, depth, and storage methods being approximately 58% (0.001), 21% ($p = 0.001$), and less than 1% ($p = 0.01$), respectively (Fig. S5).

Extraction

Extraction protocols significantly differed in the community alpha and beta diversities (Fig. 3). Beta diversity variations (determined from ASV abundances) attributed to habitat, depth, and extraction protocols were 14% ($p = 0.001$), 10% ($p = 0.001$), and 8% ($p = 0.001$), respectively. Likewise, the variation attributable to habitat, depth, and extraction protocol for PICRUSt2 relative abundances were 51% ($p = 0.001$), 19% ($p = 0.001$), and 4% ($p = 0.001$) (Fig. S5), respectively. When ASV abundances were subset by both habitat and depth, extraction method was a universally significant ($p < 0.05$) determinant of community structure differences (Table S8). Next, since the RC-derived data were outliers, the analyses were repeated with their removal; habitat and depth then accounted for more of the variation (19% ($p = 0.001$), and 13% ($p = 0.001$), respectively), while extraction was no longer significant (3% of variation ($p = 0.05$)). When subset by depth, extraction's significance diminished further to $p = \sim 0.1$ (Fig. S6). When subset by both habitat and depth, extraction became significant ($p < 0.05$) for half the habitat and depth combinations

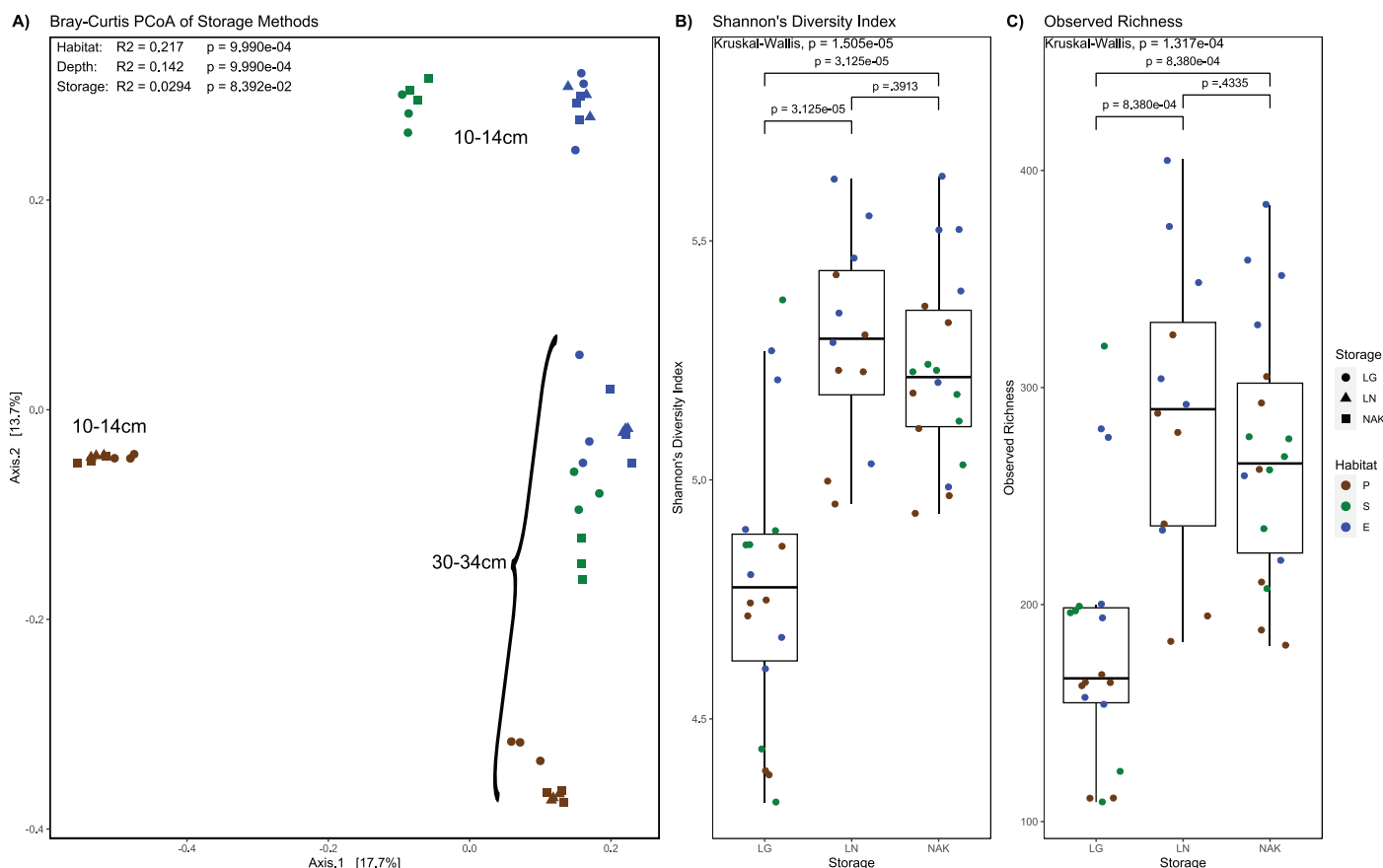


Figure 2 Impacts of storage conditions on microbial community structure and diversity. (A) Principal coordinate analysis (PCoA) based on Bray-Curtis dissimilarities of microbiome profiles (via 16S rRNA amplicon sequencing) for the three storage methods. The R^2 values and associated p -values from an adonis2 test (PERMANOVA), indicating the proportion of variance explained by habitat, depth, and storage, are provided. Storage methods are indicated by shape: circles = LifeGuard (LG), triangles = liquid nitrogen (LN), and squares = on ice without any preservation buffer (NAK). The three habitats are indicated by color, brown = palsa, green = bog, and blue = fen. (B–C) Two alpha diversity metrics, Shannon's diversity (B) and Richness as Observed ASVs (C) were compared across the three storage conditions. Statistical significance for alpha diversity metrics was evaluated by a Kruskal-Wallis non-parametric test between groups; pairwise comparison p -values are denoted above the associated lines, and the significance of overall storage condition impact on microbiome composition is indicated at the top of each panel.

Full-size [DOI: 10.7717/peerj.18745/fig-2](https://doi.org/10.7717/peerj.18745/fig-2)

(both palsa depths, and the shallow fen), and approached significance ($p \sim 0.06$) for shallow bog and deep fen, with comparable R^2 values (>0.3) (Table S9). Quality assessment of the reads for LG-stored samples showed average read quality scores ranging from 34.30 to 37.21 (Table S3).

Significant alpha diversity differences were observed among extraction methods. Notably, the RC method exhibited greater Shannon's diversity and richness for both ASVs (Figs. 3C, 3D) and 97%-identity OTUs (Figs. 3G, 3H). Furthermore, differences among extraction methods were significant ($p < 0.05$) for ASVs (Figs. 3A, 3B) and OTUs (Figs. 3E, 3F), though with lower R^2 values for OTUs ($\sim 75\%$ lower for shallow depths (Figs. 3A, 3E) and $\sim 35\%$ lower for deep depths (Figs. 3B, 3F)).

We then assessed shared and unique lineages among extraction methods. Among the shared lineages, 102 (1.3% of total) and 87 (1.1% of total) ASVs in the shallow and deep

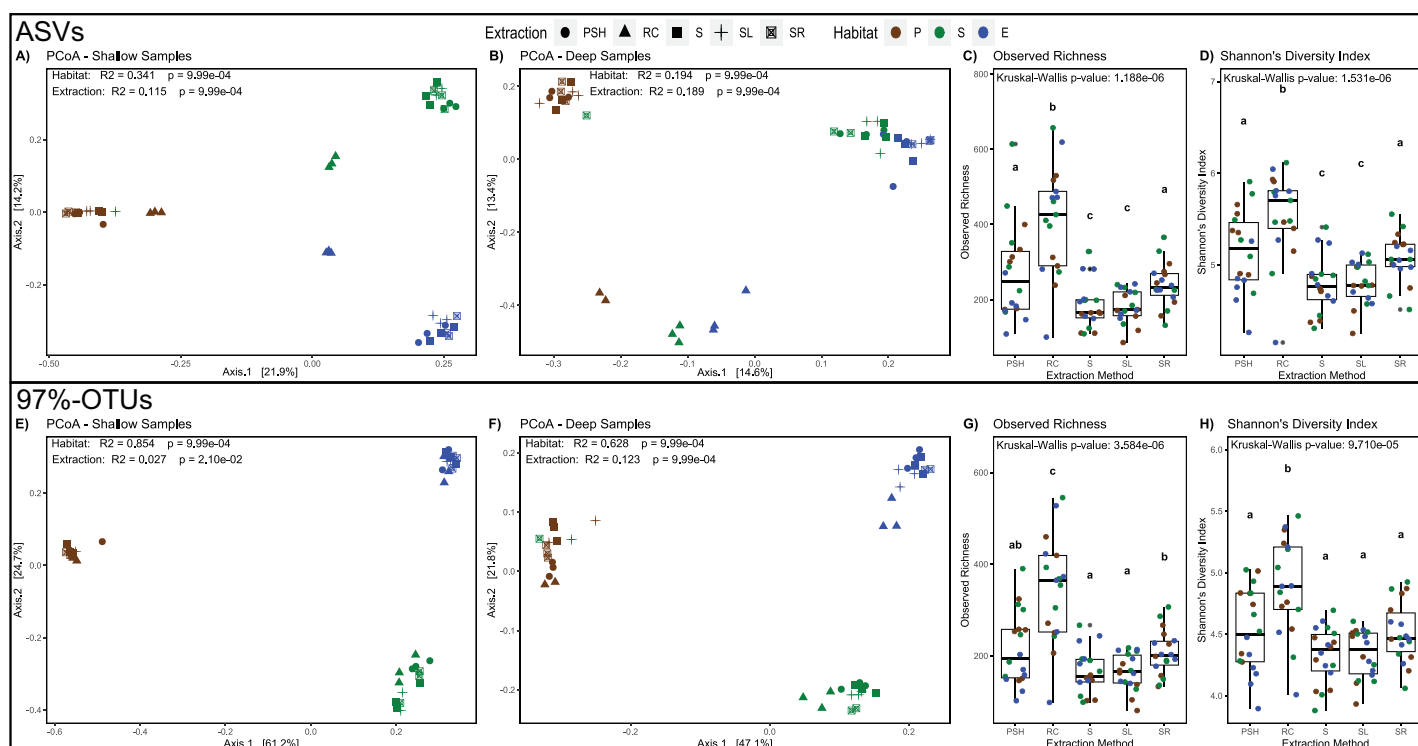


Figure 3 Impacts of extraction protocols on microbial community structure and diversity. (A, B, E, F) Principal coordinate analysis (PCoA) based on Bray-Curtis dissimilarities of microbiome profiles (*via* 16S rRNA amplicon sequencing) for five different extraction methods on samples stored in LifeGuard buffer. The R^2 values and associated p -values from an adonis2 test (PERMANOVA), indicating the proportion of variance explained by habitat and extraction are provided. (A–D) Abundance data calculated from ASVs. (E–H) Abundance data based on 97%-identity OTUs. Shallow samples (10–14 cm), Deep samples (30–34 cm). Extraction methods: S, IsoGenie standard protocol; SL, IsoGenie Standard protocol with LifeGuard removed before extraction; SR, IsoGenie standard protocol with alternative reagents; PSH, PowerMax Soil DNA Isolation kit–High Humic Acid Protocol; RC, PowerSoil Total RNA Isolation kit with DNA elution accessory kit. Two alpha diversity metrics, Shannon's diversity (D, H) and observed richness (C, G) were compared across the five extraction protocols. Statistical significance for alpha diversity metrics was evaluated by a Wilcoxon rank sum tests between groups, adjusted *via* the Benjamini-Hochberg correction; significant p -values ($p < 0.05$) are denoted above the associated groups, where a significant difference between two extraction methods is indicated by a different lettering. Groups with the same letter display insignificant differences ($p > 0.05$). Common legend is applicable to both plots. [Full-size DOI: 10.7717/peerj.18745/fig-3](https://doi.org/10.7717/peerj.18745/fig-3)

depths, respectively, displayed differential abundance across methods, spanning 20 phyla (Figs. 4A, 4C). RC samples exhibited the highest relative abundance of differentially abundant ASVs, with these ASVs representing an average of ~15% of the community and reaching up to ~23% in the shallow bog. In contrast, the other methods showed a lower average contribution of ~7% (Figs. 4A, 4C). These differentially abundant ASVs formed clusters primarily driven by habitat, with the RC extraction method creating distinct clusters in both the shallow and deep depths (Figs. 4B, 4D). A subset of ASVs differentially abundant in RC extraction, including Euryarchaeota in the shallow bog, showed consistent structure across all extraction methods except RC (Fig. 4, Fig. S2B). At 97%-identity OTUs, nearly all 30 ASVs for Euryarchaeota in the shallow bog collapsed into a single OTU (Fig. S2C). These OTUs displayed relatively stable total abundances across extraction methods (Fig. S2C). Among the unique lineages, the RC method had ~3–6.5 times more method-unique lineages than any other method, however all the method-unique lineages comprised <5% of the total abundance-weighted communities (Fig. S7). Excluding RC, the

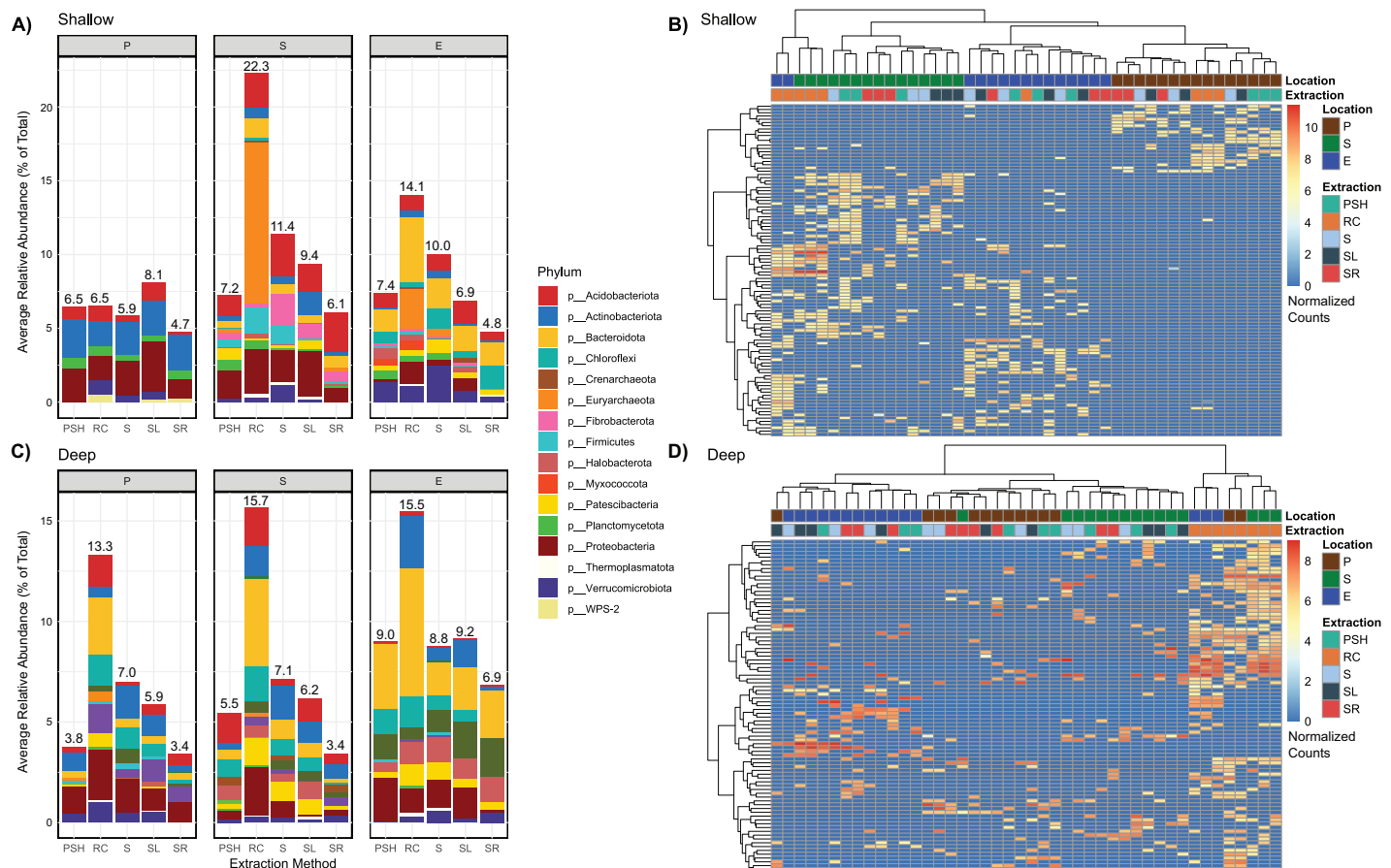


Figure 4 Recovery of differentially abundant lineages across extraction methods. (A, C) Total differentially-recovered lineages for each of the five extraction methods, for shallow (A) and deep (C) samples. Stacked barcharts show the average relative abundance of differentially-recovered lineages (*i.e.*, either over- or under-represented lineages), colored by phylum of bacteria or archaea. (B, D) Heatmap of the significantly differentially abundant ASVs (rows), with sample annotation for location/habitat and extraction method (columns), for shallow (B) and deep (D). Rows and columns are hierarchically clustered using the Ward.D2 method, and the data were normalized using log-transformed counts from DESeq2. Note slightly different scale for B and D. Extraction methods: S, IsoGenie standard protocol; SL, IsoGenie Standard protocol with LifeGuard removed before extraction; SR, IsoGenie standard protocol with alternative reagents; PSH, PowerMax Soil DNA Isolation kit–High Humic Acid Protocol; RC, PowerSoil Total RNA Isolation kit with DNA elution accessory kit. [Full-size DOI: 10.7717/peerj.18745/fig-4](https://doi.org/10.7717/peerj.18745/fig-4)

other four extraction methods exhibited ~4% distinct ASVs on average (Fig. S7). The 1,552 RC-unique lineages were equally distributed among the three habitats (435, 462, and 452 were unique to the palsa, bog, and fen, respectively); only seven RC-unique ASVs were shared between all three habitats (Fig. S8).

We tested whether there was a phylogenetic structure underlying the community differences caused by the extraction method. Faith's PD is a bulk measure that sums all branch lengths and thereby detects differences at coarser taxonomic resolution; a Kruskal-Wallis test showed a significant impact of extraction method (Fig. S9) on community Faith's PD. Unifrac distance is a distance metric which includes total branch length but accounts for shared structure between samples; the adonis2 test indicated a significant impact of extraction on both abundance-weighted and unweighted community Unifrac distances, with roughly twice the R^2 for unweighted than weighted (Table S10).

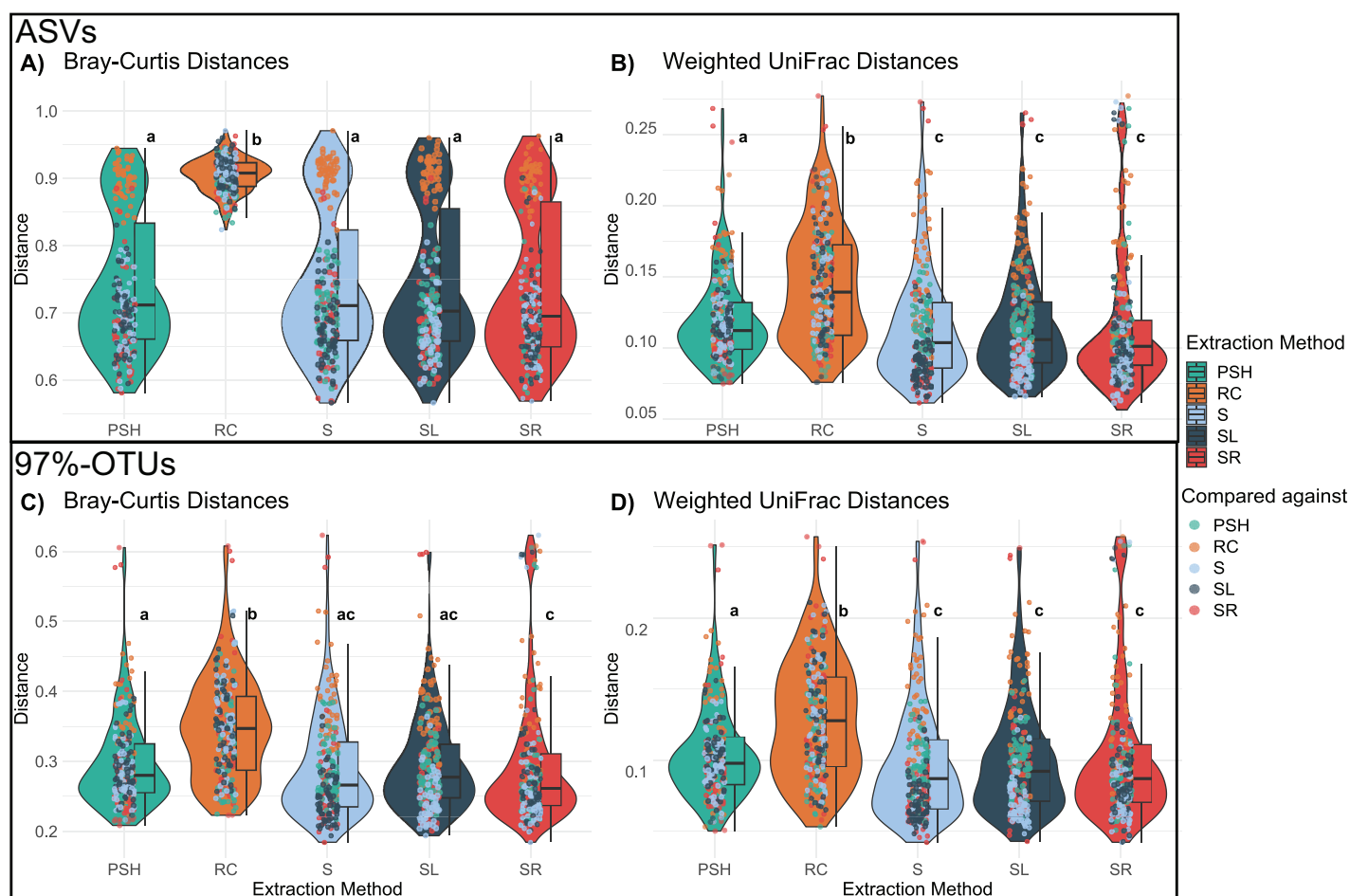


Figure 5 Comparative analysis of microbial community distances based on different extraction methods. Boxplots and violin plots illustrating the pairwise Bray-Curtis distances (A, C) and Weighted UniFrac distance (B, D) between microbial communities for various extraction methods for both ASVs (top row; A, B) and 97%-OTUs (bottom row; C, D). Each extraction method is compared against others within the same habitat and depth strata ($n = 213$ for all groups except RC, which $n = 204$). Significant pairwise comparisons between extraction methods are denoted with letters, indicating adjusted p -values post multiple testing correction (Wilcoxon rank-sum tests with Benjamini-Hochberg correction). The jittered points represent comparisons against a specific extraction method, colored according to the compared method. Note different y-axis scales. Extraction methods: S, IsoGenie standard protocol; SL, IsoGenie Standard protocol with LifeGuard removed before extraction; SR, IsoGenie standard protocol with alternative reagents; PSH, PowerMax Soil DNA Isolation kit–High Humic Acid Protocol; RC, PowerSoil Total RNA Isolation kit with DNA elution accessory kit.

Full-size [DOI: 10.7717/peerj.18745/fig-5](https://doi.org/10.7717/peerj.18745/fig-5)

We compared the Bray-Curtis and weighted Unifrac distances between extracts from the same material (Figs. 5A, 5B) at the resolution of ASVs, and although RC's Bray-Curtis distance to all other extracts was significantly higher than the other methods', the magnitude of this difference was much lower by weighted Unifrac. When calculating the same metrics for 97%-identity OTUs (Figs. 5C, 5D), the median Bray-Curtis distance of the RC samples to the others fell by roughly two-thirds (from ~ 0.92 for ASVs to ~ 0.35 for OTUs). Notably, for ASVs and 97%-identity OTUs, only the weighted Unifrac comparison for PSH was also significantly different from the other methods (Figs. 5B, 5D).

DISCUSSION

The quality and quantity of extracted DNA and RNA are crucial determinants of the success of many downstream molecular analyses. In this study, we ascertained that irrespective of the chosen storage condition or extraction method, both DNA and RNA yields met contemporary sequencing requirements (Fig. S3; while long-read sequencing platforms require 1–2 ug of starting DNA for amplification-free library preparation, there are also low biomass-input protocols that can use as little as 200 ng template DNA, a yield threshold met for these soils by nearly all combinations of storage and extraction methods tested). Given the presence of LG in the input mixtures, a statistical comparison of yield per gram of peat was precluded. In several other systems, studies suggest that different storage (Rissanen *et al.*, 2010; McCarthy *et al.*, 2015; Elie *et al.*, 2023) and extraction (Simister, Schmitt & Taylor, 2011; Corcoll *et al.*, 2017; Galla *et al.*, 2024) methods can significantly impact DNA and RNA quality and yield. Nonetheless, these methods often still produce yields that meet modern standards, as we observed. The broader consequences on microbial community representation were less uniform.

While storage preservatives have been reported to have some impacts on microbial community structure and DNA/RNA yield (Rissanen *et al.*, 2010; McCarthy *et al.*, 2015), for our soils, the three tested storage conditions had limited impact (accounting for only 3% of the variation), and were overshadowed by differences in habitat and depth (Fig. 2). This result was also consistent when leveraging the metabolic inferences, where storage accounted for less than 1% of the variation (Fig. S5). However, LG did reduce alpha diversity (Fig. 2) with consistently depleted lineages (Fig. S4), which likely accounted for the significant beta diversity difference among storage treatments in the shallow palsa and near-significant differences in the deep palsa and deep fen (Table S7). This is consistent with Tatangelo *et al.* (2014) observation of fewer tRFLP bands (implying lower diversity) from extracts of soil preserved in LG compared to other storage conditions or those processed fresh without buffer. Recently, Smenderovac *et al.* (2024) also observed the LG-preserved soils yielded lower beta diversity of the bacterial community compared to the freshly extracted and immediate frozen samples. The most parsimonious explanation for these collective results is that LG does not work well on a subset of lineages, such that those cells degrade during storage and are absent from characterization. LG's non-negligible impact on soil microbial community recovery from soils should be considered during experimental design and analyses, especially when biodiversity is consequential to the study. If only DNA is required, and reliable frozen transport of soils to the processing location is assured, then our results support using either LN or no buffer (and freezing within 6 h), as these methods yield no significant difference in alpha or beta diversity. If RNA is required, and/or reliable frozen transportation to processing location is uncertain, then our results suggest a careful weighing of the tradeoffs of the protective qualities of LG with their impacts on community recovery.

In examining the influence of extraction protocols (focused on samples stored in LG, based on its long-term use at this site), our analyses underscore the potentially large influence of the selected method on microbial community characterizations. The RC

extraction protocol, which was the only one using an affinity column, was significantly different than the other methods in both alpha and beta diversity metrics (Fig. 3), with higher richness and Shannon's diversity. This signal persisted in ASVs, 97%-identity OTUs, and metabolic inferences (Fig. 3, Fig. S5). The remaining extraction protocols used identical silica membrane size exclusion columns and exhibited some significant differences in alpha and beta diversity (Fig. 3, Tables S8, S9), which were minor compared to habitat and depth differences overall. The recovery of differentially abundant ASVs across extraction methods was not limited to low-abundance lineages, and spanned an array of phyla (Fig. 4). Despite the dominant impact of habitat and depth in samples' beta diversities (Fig. 3), the ASVs that were differentially abundant by method collectively accounted for a substantial portion of each sample (Fig. 4); with the highest average contribution to relative abundance being in the RC communities. Incorporating phylogenetic metrics (Table S10, Fig. 5) decreased the divergence of the RC communities from the others, suggesting that the RC extraction did not recover so many widely unrelated lineages, but in fact produced a similar phylogenetic recovery of the community. Overall, while the relative abundance of specific taxa may fluctuate strongly depending on the extraction approach, the phylogenetic structure in the community is more stable, albeit with a greater diversity of lineages (Fig. S9). The impact of extraction was also lesser with the RC results removed, where there remained some significant variation among the other four extraction methods, although the impact was not nearly as strong nor as consistent (Figs. 3A, 3B; Fig. S6, Table S9). The variation among extraction methods, excluding RC (~4% distinct ASVs for each method, Fig. S7), was less than the biological variation observed among habitats and depths (Figs. 3A, 3B; Fig. S6). These significant but diminished impacts, relative to biological variation, reflect some previously reported results, including studies that utilized different extraction kits and those conducted in other environments (e.g., Mackenzie, Waite & Taylor, 2015; Soliman et al., 2017). Additionally, recent work by Galla et al. (2024) highlighted significant variations in microbial communities based on extraction protocols, particularly in bulk soil and rhizosphere samples, with lesser, although often significant, effects observed in other ecosystems (such as mammalian feces). While differences in microbial community structure and alpha diversity across extraction methods and kits are well-documented (e.g., Mahmoudi, Slater & Fulthorpe, 2011; Mackenzie, Waite & Taylor, 2015; Soliman et al., 2017; Hermans, Buckley & Lear, 2018; Pearman et al., 2020), the magnitude and consistency of these effects often vary across systems and protocols.

The pronounced difference in the community profiles of the extraction protocols could be attributed to several factors, including differential cleanliness of extracts based on the methods' physicochemistry, or differential lysis efficiency for certain microbial taxa. One might hypothesize that the RC method, with its affinity column, might yield cleaner extracts, but this was not observed - they were brown and had poor A260:280 ratios. Previous studies have shown that Powersoil kits often yield higher-quality DNA (Mahmoudi, Slater & Fulthorpe, 2011). However, the correlation between A260:280 ratios and microbial richness has been inconsistent across methods and appears to depend on sample origin (Galla et al., 2024). In addition, some RC DNA extracts-which are

sequentially eluted from the affinity column rather than DNase-treated as in the other methods - did exhibit minor RNA contamination. Theoretically, if rRNAs were amplified during amplicon library preparation, this could lead to lower-abundance but highly active lineages being present in the resulting dataset; their phylogenetic coherence with the rest of the community might be consistent with this. However, Taq DNA polymerase has only minimal ability to use RNA as a template, and the buffer used with it should allow no amplification from RNA (Qiagen, 2024).

Enhanced lysis efficiency by the RC protocol causing its higher richness is possible; the five methods employ similar lysis chemistry (all are part of the same PowerSoil/PowerMax kit series; the lysis buffers (called SR1,2,3 for RC, C1,2,3 for PSH, SR, and RD1,2 a.k.a. PM1,C3 for S and SL) contain at a minimum SDS and Qiagen's proprietary Inhibitor Removal Technology (IRT)) and all use bead-beating with 10 minutes of vortexing at maximum speed. However, RC uses sharp-edged silica carbide particles, while the others used 0.1mm glass beads. The pronounced beta diversity difference of the RC-derived community profiles *via* ASVs was significantly reduced when using 97%-identity clustered OTUs instead (Figs. 3, 5). This suggests that there are major differences in the structure of RC-extracted samples at roughly the strain-level, but much more similar species-level structure; although, the differences in beta diversity were still significant (Figs 3, 5). Differences in bead-beating lysis efficiency have been previously reported to alter genus- and species-level community recovery (Zhang *et al.*, 2021). RC-extracted samples consistently displayed significantly higher alpha diversity, regardless of the method of clustering (ASVs or OTUs), indicating that lower-abundance lineages (impacting alpha more than beta diversity) that diverged beyond the strain-level were being differentially recovered by RC (Fig. 3); indeed RC-unique lineages comprised <5% of the mean relative abundance (Fig. S7). Furthermore, when examining Euryarchaeota ASVs in the shallow bog, the structure from all extraction methods, except RC, is consistent (Fig. S2B). However, the majority of the Euryarchaeota ASVs clustered into a single 97%-identity OTU (Fig. S2C), suggesting that the RC extraction method selectively recovers a distinct subset of near strain-level variants. Meanwhile, the total abundance of these lineages (at the OTU level) remains fairly consistent. This is somewhat consistent with previous reports using different Powersoil Isolation kits than those used here, which recovered similar relative abundances of Archaea compared to other kits in soils, when clustered at 97%-identity (Soliman *et al.*, 2017), but differential recovery of Archaeal lineages across extraction methods when looking more granularly at the level of ASVs (Galla *et al.*, 2024). In addition to the bead differences in RC lysis, a possible explanation of these results could be poor quality reads. When examined, all quality scores maintained above acceptable levels for the region analyzed (Table S3). Moreover, the probability of obtaining consistent errors in sequencing across multiple samples, given this quality, is extremely low. Thus, differences in sequence quality were likely not the cause of the community-wide differences in community structure.

A final consideration is a batch effect, as the RC was sequenced at the same facility on a different date from the other samples. Our predominant concern was possible contamination of those samples during library preparation or sequencing. Although the

decreased separation of RC samples when phylogeny was accounted for suggested this was unlikely, we checked the ASVs that were unique to the RC communities from the palisade, bog and fen samples to see if they overlapped significantly, which would strongly suggest contamination; they did not (Figs. S7, S8). A remaining batch effect is run-to-run variability on the MiSeq instrument; while some run-to-run variation is known (e.g. Wen *et al.*, 2017), aberrant runs (*i.e.*, which produce a markedly different community composition) have been reported (Song *et al.*, 2018; Yeh *et al.*, 2018) but appear rare (e.g., 1 in 50, Yeh *et al.*, 2018). It is possible that the RC results represent such an aberrant sequencing run.

A potential limitation of the extraction methods comparison was its testing only on LG-stored samples, given that LG reduced microbial community richness and consistently underrepresented some lineages (Fig. 2, Fig S4). LG storage was chosen for the extraction experiment as it has been used for fourteen years of multi-omics sampling campaigns at the site, so extraction performance in that background-and with those site's soils-was the priority for this study. There is no evidence in the literature that residual LG would interact with either the RC's different beads (silica carbide particles *vs.* the other methods' glass beads) or its anion-exchange column (*vs.* the other methods' size exclusion columns), such that the aberrant results with RC would be an artifact restricted to LG-stored samples. However, given LG's reduced richness, and RC's enhanced richness, it is possible that the impact of RC could be reduced with other storage methods that preserved a greater portion of the community (*i.e.*, in the LG-preserved samples, the RC-unique species were low abundance; these might be higher in non-LG-stored samples, and thus not uniquely extracted by RC's harsher bead-beating).

Recommendations based on our results: The optimal method for preservation and extraction of field peat soil samples depends on field and transport considerations and on research questions; it may also differ among types of peat soils, though that was not observed here for three physicochemically distinct permafrost-associated soils. For sample preservation, in the field as in the lab, LN is ideal as it inactivates cellular functions near-instantly, allowing for both field-representative RNA and DNA sequencing. If only DNA sequencing is required, storing samples on ice with same-day transfer to a -80°C freezer is a good choice, as it produced similar alpha and beta diversity metrics across all three different habitats. Storage at -20°C might also be adequate but was not tested here. If RNA sequencing is required but LN is unavailable and/or transport conditions are uncertain, sample preservation in LG represents a clear trade-off between biologic inactivation in the field with protection of sample integrity during transport, and a decreased recovery of lineages (~42% fewer in our samples). For extraction, if the goal is ecological comparisons among samples, we recommend PSH. While the true wild communities were unknown such that each method's accuracy at extracting them could not be determined, the S, SL, SR, and PSH extractions produced similar microbial community structures, with PSH and SR yielding slightly higher diversities, while PSH currently costs roughly half as much. If, however, the goal is solely to maximize recovered biodiversity, the RC extraction protocol may be optimal. However, it is not a good choice for ecological comparisons as it produced the highest variation compared to other methods in alpha and beta diversities as well as in

the biogeochemically consequential Euryarchaeota, and some samples may require additional cleanup (and associated loss) before sequencing. The magnitude of difference of the RC results from the other methods also highlights the consequentiality here of column choice in nucleic acid recovery (as RC uses an affinity column and the others use size-exclusion columns).

CONCLUSIONS

Choices made in sample storage and extraction can influence the interpreted ecology of a sample. In no case were the impacts of storage or extraction greater than that of habitat and depth, but in some cases the proportion of variance explained by extraction approached known environmental drivers. The impact of extraction in particular could impact data interpretation (*e.g.*, tracking of particular lineages of interest, evaluating finer-scale ecological variation, metabolic inferences).

ACKNOWLEDGEMENTS

We gratefully acknowledge Drs. Suzanne Kennedy and Michelle Carlson, and Mr. Yoshiaki Kono, who were previously part of the MO BIO Laboratories technical support team, for sharing the high humic acids protocol. We also thank three anonymous reviewers and reviewer Aditya Bandla for providing feedback on the submitted version of this manuscript. Additionally, we acknowledge Ruth K. Varner, Moira Hough, Apryl Perry, Nicole Irwin-Raab, Joachim Jansen, and Kathryn A. Bennett as members of the IsoGenie 2016 and 2019 Field Teams.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This research is a contribution of the EMERGE Biology Integration Institute, funded by the National Science Foundation, Biology Integration Institutes Program, grant # 2022070. This study was also supported by the Genomic Science Program of the United States Department of Energy Office of Biological and Environmental Research, grant #s DE-SC0010580 and DE-SC0016440. Quality control of RNA used in this publication was performed at The Ohio State University Comprehensive Cancer Center Core Facilities, which is supported by Center and by the National Institutes of Health grant # P30 CA016058. The Abisko Scientific Research Station (where samples were collected) is managed by the Swedish Polar Research Secretariat with support from the Swedish Research Council's grant # 4.3-2021-00164. Ben J Woodcroft was supported by an Australian Research Council Future Fellowship, grant # FT210100521. There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

EMERGE Biology Integration Institute Biology Integration Institutes: # 2022070.

Genomic Science Program of the United States Department of Energy Office of Biological and Environmental Research: #s DE-SC0010580 and DE-SC0016440.

The Ohio State University Comprehensive Cancer Center Core Facilities.

National Institutes of Health: # P30 CA016058.

The Abisko Scientific Research Station.

Swedish Polar Research Secretariat with support from the Swedish Research Council's: # 4.3-2021-00164.

Australian Research Council Future Fellowship: # FT210100521.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Dylan Cronin conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Yueh-Fen Li conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Paul Evans conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Gene W. Tyson conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Ben J. Woodcroft conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Virginia I. Rich conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

Field Study Permissions

The following information was supplied relating to field study approvals (*i.e.*, approving body and any reference numbers):

The field site at Stordalen Mire is under the jurisdiction of the Abisko Scientific Research Station, which is managed by the Swedish Polar Research Secretariat. Sampling there requires application, approval and documentation, and our team filed the standard application to sample peat and porewater. Application approval is issued by Magnus Augner, Deputy Director-General, and Manager of Abisko Scientific Research Station, Swedish Polar Research Secretariat, Magnus.Augner@polar.se, Phone: +46-70-6398004

Data Availability

The following information was supplied regarding data availability:

The raw sequence reads are available at NCBI SRA: [PRJNA1085467](https://www.ncbi.nlm.nih.gov/sra/PRJNA1085467).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.18745#supplemental-information>.

REFERENCES

- Albers CN, Jensen A, Bælum J, Jacobsen CS. 2013. Inhibition of DNA polymerases used in Q-PCR by structurally different soil-derived humic substances. *Geomicrobiology Journal* 30(8):675–681 DOI 10.1080/01490451.2012.758193.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available at <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Apprill A, McNally S, Parsons R, Weber L. 2015. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology* 75(2):129–137 DOI 10.3354/ame01753.
- Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolk T, Kreps J, Langille MGI, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Priesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, Van Der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, Von Hippel M, Walters W, Wan Y, Wang M, Warren J. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME2. *Nature Biotechnology* 37(8):852–857 DOI 10.1038/s41587-019-0209-9.
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods* 13(7):581–583 DOI 10.1038/nmeth.3869.
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25(15):1972–1973 DOI 10.1093/bioinformatics/btp348.
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal* 6(8):1621–1624 DOI 10.1038/ismej.2012.8.
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences of the United States of America* 108(supplement_1):4516–4522 DOI 10.1073/pnas.1000080107.
- Carroll IM, Ringel-Kulka T, Siddle JP, Klaenhammer TR, Ringel Y. 2012. Characterization of the fecal microbiota using high-throughput sequencing reveals a stable microbial community during storage. *PLOS ONE* 7(10):e46953 DOI 10.1371/journal.pone.0046953.

- Corcoll N, Österlund T, Sinclair L, Eiler A, Kristiansson E, Backhaus T, Eriksson KM. 2017. Comparison of four DNA extraction methods for comprehensive assessment of 16S rRNA bacterial diversity in marine biofilms using high-throughput sequencing. *FEMS Microbiology Letters* 364(14):e0132783 DOI 10.1093/femsle/fnx139.
- Cory AB, Chanton JP, Spencer RGM, Rich McC VI, IsoGenie Project Coordinators CK, Wilson RM. 2022. Quantifying the inhibitory impact of soluble phenolics on carbon mineralization from sphagnum-rich peatlands. *PLOS ONE* 17(2):e0252743 DOI 10.1371/journal.pone.0252743.
- Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI. 2020. PICRUSt2 for prediction of metagenome functions. *Nature Biotechnology* 38(6):685–688 DOI 10.1038/s41587-020-0548-6.
- Elie C, Perret M, Hage H, Sentausa E, Hesketh A, Louis K, Fritah-Lafont A, Leissner P, Vachon C, Rostaing H, Reynier F, Gervasi G, Saliou A. 2023. Comparison of DNA extraction methods for 16S rRNA gene sequencing in the analysis of the human gut microbiome. *Scientific Reports* 13(1):10279 DOI 10.1038/s41598-023-33959-6.
- Ellenbogen JB, Borton MA, McGivern BB, Cronin DR, Hoyt DW, Freire-Zapata V, McCalley CK, Varner RK, Crill PM, Wehr RA, Chanton JP, Woodcroft BJ, Tfaily MM, Tyson GW, Rich VI, Wrighton KC. 2024. Methylophony in the Mire: direct and indirect routes for methane production in thawing permafrost. *mSystems* 9:e00698-23 DOI 10.1128/msystems.00698-23.
- Engelbrektson A, Kunin V, Wrighton KC, Zvenigorodsky N, Chen F, Ochman H, Hugenholtz P. 2010. Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *The ISME Journal* 4(5):642–647 DOI 10.1038/ismej.2009.153.
- Ewels P, Magnusson M, Lundin S, Källér M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32(19):3047–3048 DOI 10.1093/bioinformatics/btw354.
- Fofana A, Anderson D, McCalley CK, Hodgkins S, Wilson RM, Cronin D, Raab N, Torabi M, Varner RK, Crill P, Saleska SR, Chanton JP, Tfaily MM, Rich VI. 2022. Mapping substrate use across a permafrost thaw gradient. *Soil Biology and Biochemistry* 175(11):108809 DOI 10.1016/j.soilbio.2022.108809.
- Galla G, Praeg N, Rzehak T, Sprecher E, Colla F, Seeber J, Illmer P, Hauße HC. 2024. Comparison of DNA extraction methods on different sample matrices within the same terrestrial ecosystem. *Scientific Reports* 14(1):8715 DOI 10.1038/s41598-024-59086-4.
- Hermans SM, Buckley HL, Lear G. 2018. Optimal extraction methods for the simultaneous analysis of DNA from diverse organisms and sample types. *Molecular Ecology Resources* 18(3):557–569 DOI 10.1111/1755-0998.12762.
- Hodgkins SB, Tfaily MM, McCalley CK, Logan TA, Crill PM, Saleska SR, Rich VI, Chanton JP. 2014. Changes in peat chemistry associated with permafrost thaw increase greenhouse gas production. *Proceedings of the National Academy of Sciences of the United States of America* 111(16):5819–5824 DOI 10.1073/pnas.1314641111.
- Hough M, McCabe S, Vining SR, Pickering Pedersen E, Wilson RM, Lawrence R, Chang K-Y, Bohrer G, Coordinators TI, Riley WJ, Crill PM, Varner RK, Blazewicz SJ, Dorrepaal E, Tfaily MM, Saleska SR, Rich VI. 2022. Coupling plant litter quantity to a novel metric for litter quality explains C storage changes in a thawing permafrost peatland. *Global Change Biology* 28(3):950–968 DOI 10.1111/gcb.15970.

- Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26(11):1463–1464 DOI 10.1093/bioinformatics/btq166.
- Kennedy NA, Walker AW, Berry SH, Duncan SH, Farquarson FM, Louis P, Thomson JM, Satsangi J, Flint HJ, Parkhill J, Lees CW, Hold GL. 2014. The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing. *PLOS ONE* 9(2):e88982 DOI 10.1371/journal.pone.0088982.
- Kia E, Mackenzie BW, Middleton D, Lau A, Waite DW, Lewis G, Chan Y-K, Silvestre M, Cooper GJS, Poppitt SD, Taylor MW. 2016. Integrity of the human faecal microbiota following long-term sample storage. *PLOS ONE* 11(10):e0163666 DOI 10.1371/journal.pone.0163666.
- Kolde R. 2019. pheatmap: Pretty Heatmaps. Available at <https://github.com/raivokolde/pheatmap>.
- Krassowski M. 2020. krassowski/complex-upset. Zenodo DOI 10.5281/zenodo.3700590.
- Lahti L, Shetty S. 2017. Tools for microbiome analysis in R. Available at <http://microbiome.github.com/microbiome>.
- Lauber CL, Zhou N, Gordon JI, Knight R, Fierer N. 2010. Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples: influence of short-term storage conditions on microbiota. *FEMS Microbiology Letters* 307(1):80–86 DOI 10.1111/j.1574-6968.2010.01965.x.
- Lehmann J, Kleber M. 2015. The contentious nature of soil organic matter. *Nature* 528:60–68 DOI 10.1038/nature16069.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15(12):550 DOI 10.1186/s13059-014-0550-8.
- Lozupone CA, Stombaugh J, Gonzalez A, Ackermann G, Wendel D, Vázquez-Baeza Y, Jansson JK, Gordon JI, Knight R. 2013. Meta-analyses of studies of the human microbiota. *Genome Research* 23(10):1704–1714 DOI 10.1101/gr.151803.112.
- Mackenzie BW, Waite DW, Taylor MW. 2015. Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences. *Frontiers in Microbiology* 6:130 DOI 10.3389/fmicb.2015.00130.
- Mahmoudi N, Slater GF, Fulthorpe RR. 2011. Comparison of commercial DNA extraction kits for isolation and purification of bacterial and eukaryotic DNA from PAH-contaminated soils. *Canadian Journal of Microbiology* 57:623–628 DOI 10.1139/w11-049.
- Malmer N, Johansson Tö, Olsrud M, Christensen TR. 2005. Vegetation, climatic changes and net carbon sequestration in a North-Scandinavian subarctic mire over 30 years. *Global Change Biology* 11(11):1895–1909 DOI 10.1111/j.1365-2486.2005.01042.x.
- Martinez MA, Woodcroft BJ, Ignacio Espinoza JC, Zayed AA, Singleton CM, Boyd JA, Li Y-F, Purvine S, Maughan H, Hodgkins SB, Anderson D, Sederholm M, Temperton B, Bolduc B, Saleska SR, Tyson GW, Rich SSR VI, Tyson GW, Rich VI. 2019. Discovery and ecogenomic context of a global Caldiserica-related phylum active in thawing permafrost, Candidatus Cryoserica phylum nov., Ca. Cryoserica class nov., Ca. Cryosericales ord. nov., Ca. Cryoseriaceae fam. nov., comprising the four species Cryosericum septentrionale gen. nov. sp. nov., Ca. C. hinesii sp. nov., Ca. C. odellii sp. nov., Ca. C. terrychapinii sp. nov. *Systematic and Applied Microbiology* 42(1):54–66 DOI 10.1016/j.syapm.2018.12.003.
- McCalley CK, Woodcroft BJ, Hodgkins SB, Wehr RA, Kim E-H, Mondav R, Crill PM, Chanton JP, Rich VI, Tyson GW, Saleska SR. 2014. Methane dynamics regulated by microbial community response to permafrost thaw. *Nature* 514(7523):478–481 DOI 10.1038/nature13798.

- McCarthy A, Chiang E, Schmidt ML, Denef VJ. 2015. RNA preservation agents and nucleic acid extraction method bias perceived bacterial community composition. *PLOS ONE* 10(3):e0121659 DOI 10.1371/journal.pone.0121659.
- McGivern BB, Cronin DR, Ellenbogen JB, Borton MA, Knutson EL, Freire-Zapata V, Bouranis JA, Bernhardt L, Hernandez AI, Flynn RM, Woyda R, Cory AB, Wilson RM, Chanton JP, Woodcroft BJ, Ernakovich JG, Tfaily MM, Sullivan MB, Tyson GW, Rich VI, Hagerman AE, Wrighton KC. 2024. Microbial polyphenol metabolism is part of the thawing permafrost carbon cycle. *Nature Microbiology* 9(6):1454–1466 DOI 10.1038/s41564-024-01691-0.
- McMurdie PJ, Holmes S. 2014. Waste not, want not: why rarefying microbiome data is inadmissible. *PLOS Computational Biology* 10(4):e1003531 DOI 10.1371/journal.pcbi.1003531.
- Mondav R, McCalley CK, Hodgkins SB, Frolking S, Saleska SR, Rich VI, Chanton JP, Crill PM. 2017. Microbial network, phylogenetic diversity and community membership in the active layer across a permafrost thaw gradient. *Environmental Microbiology* 19(8):3201–3218 DOI 10.1111/1462-2920.13809.
- Mondav R, Woodcroft BJ, Kim E-H, McCalley CK, Hodgkins SB, Crill PM, Chanton J, Hurst GB, VerBerkmoes NC, Saleska SR, Hugenholtz P, Rich VI, Tyson GW. 2014. Discovery of a novel methanogen prevalent in thawing permafrost. *Nature Communications* 5(1):3212 DOI 10.1038/ncomms4212.
- Oksanen J. 2022. Vegan: community ecology package. R package. Available at <https://github.com/vegandevs/vegan>.
- Parada AE, Needham DM, Fuhrman JA. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology* 18(5):1403–1414 DOI 10.1111/1462-2920.13023.
- Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35(3):526–528 DOI 10.1093/bioinformatics/bty633.
- Pearman JK, Keeley NB, Wood SA, Laroche O, Zaiko A, Thomson-Laing G, Biessy L, Atalah J, pochon X. 2020. Comparing sediment DNA extraction methods for assessing organic enrichment associated with marine aquaculture. *PeerJ* 8:e10231 DOI 10.7717/peerj.10231.
- Pruesse E, Peplies J, Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28(14):1823–1829 DOI 10.1093/bioinformatics/bts252.
- Qiagen. 2024. Can Taq DNA Polymerase use RNA as a template, and generate false positives in no-RT controls?. Available at <https://www.qiagen.com/us/resources/faq?id=65b9b97f-1961-492f-8510-2de502a15159&lang=en>.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41(D1):D590–D596 DOI 10.1093/nar/gks1219.
- Rissanen AJ, Kurhela E, Aho T, Oittinen T, Tiirola M. 2010. Storage of environmental samples for guaranteeing nucleic acid yields for molecular microbiological studies. *Applied Microbiology and Biotechnology* 88(4):977–984 DOI 10.1007/s00253-010-2838-2.
- Sidstedt M, Rådström P, Hedman J. 2020. PCR inhibition in qPCR, dPCR and MPS—mechanisms and solutions. *Analytical and Bioanalytical Chemistry* 412(9):2009–2023 DOI 10.1007/s00216-020-02490-2.
- Simister RL, Schmitt S, Taylor MW. 2011. Evaluating methods for the preservation and extraction of DNA and RNA for analysis of microbial communities in marine sponges. *Journal of Experimental Marine Biology and Ecology* 397(1):38–43 DOI 10.1016/j.jembe.2010.11.004.

- Singleton CM, McCalley CK, Woodcroft BJ, Boyd JA, Evans PN, Hodgkins SB, Chanton JP, Frolking S, Crill PM, Saleska SR, Rich VI, Tyson GW. 2018. Methanotrophy across a natural permafrost thaw environment. *The ISME Journal* 12(10):2544–2558 DOI 10.1038/s41396-018-0065-5.
- Smenderovac E, Emilson C, Rheault K, Brazeau É, Morency M-J, Gagné P, Venier L, Martineau C. 2024. Drying as an effective method to store soil samples for DNA-based microbial community analyses: a comparative study. *Scientific Reports* 14(1):1725 DOI 10.1038/s41598-023-50541-2.
- Smith D. 2024. rbiom: Read/Write, Analyze, and Visualize ‘BIOM’ Data. Available at <https://cmmr.github.io/rbiom/>.
- Soliman T, Yang S-Y, Yamazaki T, Jenke-Kodama H. 2017. Profiling soil microbial communities with next-generation sequencing: the influence of DNA kit selection and technician technical expertise. *PeerJ* 5:e4178 DOI 10.7717/peerj.4178.
- Song Z, Schlatter D, Gohl DM, Kinkel LL. 2018. Run-to-run sequencing variation can introduce taxon-specific bias in the evaluation of fungal microbiomes. *Phytobiomes Journal* 2(3):165–170 DOI 10.1094/PBIOMES-09-17-0041-R.
- Sun CL, Pratama AA, Gazitúa MC, Cronin D, McGivern BB, Wainaina JM, Vik DR, Zayed AA, Bolduc B, Wrighton KC, Rich VI, Sullivan MB. 2024. IsoGenie project field teams 2010-2017, the IsoGenie and EMERGE Projects Coordinators, Virus ecology and 7-year temporal dynamics across a permafrost thaw gradient. *Environmental Microbiology* 26(8):e16665 DOI 10.1111/1462-2920.16665.
- Tap J, Cools-Portier S, Pavan S, Druesne A, Öhman L, Törnblom H, Simren M, Derrien M. 2019. Effects of the long-term storage of human fecal microbiota samples collected in RNAlater. *Scientific Reports* 9(1):601 DOI 10.1038/s41598-018-36953-5.
- Tatangelo V, Franzetti A, Gandolfi I, Bestetti G, Ambrosini R. 2014. Effect of preservation method on the assessment of bacterial community structure in soil and water samples. *FEMS Microbiology Letters* 356(1):32–38 DOI 10.1111/1574-6968.12475.
- Tebbe CC, Vahjen W. 1993. Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant DNA from bacteria and a yeast. *Applied and Environmental Microbiology* 59(8):2657–2665 DOI 10.1128/aem.59.8.2657-2665.1993.
- Varner RK, Crill PM, Frolking S, McCalley CK, Burke SA, Chanton JP, Holmes ME, Saleska S, Coordinators IP, Palace MW. 2021. Permafrost thaw driven changes in hydrology and vegetation cover increase trace gas emissions and climate forcing in Stordalen Mire from 1970–2014. *Philosophical Transactions of the Royal Society A* 380(2215):20210022 DOI 10.1098/rsta.2021.0022.
- Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, Gilbert JA, Jansson JK, Caporaso JG, Fuhrman JA, Apprill A, Knight R. 2016. Improved bacterial 16S rRNA Gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems* 1(1):e0000915 DOI 10.1128/mSystems.00009-15.
- Wen C, Wu L, Qin Y, Van Nostrand JD, Ning D, Sun B, Xue K, Liu F, Deng Y, Liang Y, Zhou J. 2017. Evaluation of the reproducibility of amplicon sequencing with Illumina MiSeq platform. *PLOS ONE* 12(4):e0176716 DOI 10.1371/journal.pone.0176716.
- Wilson RM, Hough MA, Verbeke BA, Hodgkins SB, Coordinators IG, Chanton JP, Saleska SD, Rich VI, Tfaily MM. 2022. Plant organic matter inputs exert a strong control on soil organic matter decomposition in a thawing permafrost peatland. *Science of the Total Environment* 820(2):152757 DOI 10.1016/j.scitotenv.2021.152757.

- Wilson RM, Zayed AA, Crossen KB, Woodcroft B, Tfaily MM, Emerson J, Raab N, Hodgkins SB, Verbeke B, Tyson G, Crill P, Saleska S, Chanton JP, Rich IGPC VI, Team IGPF. 2021. Functional capabilities of microbial communities to carry out large scale geochemical processes are maintained during ex situ anaerobic incubation. *PLOS ONE* 16(2):e0245857 DOI 10.1371/journal.pone.0245857.
- Wnuk E, Waśko A, Walkiewicz A, Bartmiński P, Bejger R, Mielnik L, Bieganski A. 2020. The effects of humic substances on DNA isolation from soils. *PeerJ* 8(8):e9378 DOI 10.7717/peerj.9378.
- Woodcroft BJ, Singleton CM, Boyd JA, Evans PN, Emerson JB, Zayed AAF, Hoelzle RD, Lamberton TO, McCalley CK, Hodgkins SB, Wilson RM, Purvine SO, Nicora CD, Li C, Frolking S, Chanton JP, Crill PM, Saleska SR, Rich VI, Tyson GW. 2018. Genome-centric view of carbon processing in thawing permafrost. *Nature* 560(7716):49–54 DOI 10.1038/s41586-018-0338-1.
- Yeh Y-C, Needham DM, Sieradzki ET, Fuhrman JA. 2018. Taxon disappearance from microbiome analysis reinforces the value of mock communities as a standard in every sequencing run. *mSystems* 3:e00023-18 DOI 10.1128/mSystems.00023-18.
- Zhang B, Brock M, Arana C, Dende C, Van Oers NS, Hooper LV, Raj P. 2021. Impact of Bead-beating intensity on the genus- and species-level characterization of the gut microbiome using amplicon and complete 16S rRNA gene sequencing. *Frontiers in Cellular and Infection Microbiology* 11:678522 DOI 10.3389/fcimb.2021.678522.