Tunturi virus isolates and metagenome-assembled viral genomes provide insights into the virome of *Acidobacteriota* in Arctic tundra soils

Tatiana Demina[1,2*], Heli Marttila[1], Igor S. Pessi[1,2,3], Minna K. Männistö[4], Bas E. Dutilh[5,6], Simon Roux[7], Jenni Hultman[1,4]

[1]Department of Microbiology, Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland;

[2]Helsinki Institute of Sustainability Science (HELSUS), University of Helsinki, Helsinki, Finland;

[3]Finnish Environment Institute (Syke), Helsinki, Finland;

[4]Natural Resources Institute Finland (Luke), Finland;

[5]Institute of Biodiversity, Faculty of Biological Sciences, Cluster of Excellence Balance of the Microverse, Friedrich Schiller University Jena, Jena, Germany;

[6]Theoretical Biology and Bioinformatics, Science for Life, Utrecht University, Utrecht, the Netherlands;

[7]DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

*corresponding author

Abstract

Acidobacteriota are phylogenetically and physiologically diverse soil bacteria that play important roles in key ecological processes. Still, surprisingly little is known about their viruses. Here, we report five virus isolates, Tunturi 1-5, that were obtained from Arctic tundra soils, Kilpisjärvi, Finland (69°N), using *Tunturibacter* spp. strains as hosts. The new virus isolates have tailed particles with podo- (Tunturi 1, 2, 3), sipho- (Tunturi 4), or myovirus-like (Tunturi 5) morphologies. The dsDNA genomes of the viral isolates are 63–98 kbp long, except Tunturi 5, which is a jumbo phage with a 309-kbp genome. Tunturi 1 and Tunturi 2 share 88% overall nucleotide identity, while the other three are not related to one another. Over half of the open reading frames in Tunturi genomes have no homologs in the NCBI nr database. Viral diversity in Kilpisjärvi soils was further assessed using a metagenomic approach. A total of 1938 viral operational taxonomic units (vOTUs) were extracted from bulk soil metagenomes, of which 46 represented >90% complete virus genomes. Almost all vOTUs (98%) were assigned to the class *Caudoviricetes*. For 338 vOTUs, putative hosts were predicted, including *Acidobacteriota* and other common soil phyla involved in carbon and nitrogen cycling. We also observed 16 putative *Terriglobia*-associated proviruses in Kilpisjärvi soils, being related to proviruses originating from soils and other distant biomes. Approximately genus- or higher-level similarities could be observed between Tunturi viruses, Kilpisjärvi vOTUs, and other soil vOTUs, suggesting some shared viral diversity across soils. On a local scale, viral communities, including acidobacterial vOTUs, were habitat-specific, being driven by the same environmental factors as their host communities: soil moisture, SOM, C, N content and C:N ratio. This study represents a comprehensive analysis of *Acidobacteriota*-associated viruses residing in Arctic tundra soils, providing isolates as laboratory models for future studies and adding insights into the viral diversity and virus-host interactions in these climate-critical soils.

Introduction

Huge amounts of carbon and nitrogen are accumulated in Arctic permafrost-affected tundra soils and are expected to be released with increasing temperatures due to climate change (Turetsky et al. 2019). Despite seemingly harsh environmental conditions, Arctic soils host diverse and active microbial communities that decompose soil organic matter and contribute to greenhouse gas emissions (Malard and Pearce 2018; Woodcroft et al. 2018; Kwon et al. 2019; Taş et al. 2018; Pessi et al. 2022). Arctic soil microbial communities have been shown to experience compositional and functional changes with permafrost thawing (Hultman et al. 2015; Deng et al. 2015; Jansson and Taş 2014; Scheel et al. 2022; Ernakovich et al. 2022; Coolen et al. 2011). However, the Arctic covers a large geographic area with diverse environments, and more studies are needed to understand and predict the responses of microorganisms to climate change, including the increased amounts of soil carbon and nitrogen being released from Arctic permafrost. In particular, the contribution of viruses to ecosystem-wide cycling processes in soils has been little explored compared to aquatic environments (Jansson 2023; Suttle 2007; Pratama and van Elsas 2018).

Soil viruses are abundant and diverse (Jansson 2023; Roux and Emerson 2022) and have various impacts on their hosts (Braga et al. 2020; Albright et al. 2022). By lysing microbial cells, viruses indirectly affect biogeochemical cycling in soil (Y. Wang et al. 2022). With lysogenic conversion, i.e., the expression of genes encoded by a lysogen, temperate viruses may support fitness of their hosts, e.g., by protecting them from other viral infections (Bondy-Denomy et al. 2016) or increasing their ability to form biofilms (Schuch and Fischetti 2009). Moreover, a variety of auxiliary metabolic genes (AMGs) have been predicted in soil viruses (Jansson and Wu 2023), and some of these AMGs involved in carbon cycling have been experimentally shown to be active (Emerson et al. 2018; R. Wu et al. 2022). Finally, experimental warming of tundra soils from the Alaskan permafrost region resulted in increased abundance of viral AMGs (glycoside hydrolases) in warmed soils (Ji et al. 2023).

Acidobacteriota are widespread and ubiquitous in various environments, and are especially prominent in acidic soils (Kalam et al. 2020; Jones et al. 2009; Männistö et al. 2013; Giguere et al. 2021; A. Kielak et al. 2009). Despite their abundance, a relatively limited number of *Acidobacteriota* species have been isolated and characterized, as their cultivation is often challenging (Davis, Sangwan, and Janssen 2011; Overmann, Abt, and Sikorski 2017). While information on their ecological functions remains fragmentary, soil *Acidobacteriota* seem to have significant roles in carbon (Rawat et al. 2012; A. M. Kielak et al. 2016; Pankratov et al. 2011; Coluccia and Besaury 2023; Belova, Ravin, et al. 2018), nitrogen (Pessi et al. 2022) and sulfur (Hausmann et al. 2018) cycling. Moreover, *Acidobacteriota* include many putative keystone taxa, i.e., strong drivers of microbiome structure and functioning, in soils and as part of plant-associated microbiota (Banerjee, Schlaeppi, and van der Heijden 2018).

Acidobacteriota are phylogenetically diverse and comprise 15 taxonomic classes (Dedysh and Yilmaz 2018). The members of the class *Terriglobia* are typically dominant and active in Arctic tundra and boreal forest soils (Männistö et al. 2013; Männistö, Tiirola, and Häggblom 2007; Pessi et al. 2022; Viitamäki et al. 2022) and are represented by several cultured isolates (Männistö et al. 2011; 2012; Belova, Suzina, et al. 2018; Belova, Ravin, et al. 2018). Recently, the genus *Tunturibacter* has been described in the class *Terriglobia* (Messyasz et al. 2024). The type species, *Tunturibacter lichenicola*, was originally known as *Edaphobacter*

lichenicola (Belova, Suzina, et al. 2018). *Tunturibacter* representatives are Gram-negative aerobic rods, produce extracellular polysaccharide-like substances and are able to hydrolyze various polysaccharides (Messyasz et al. 2024).

Surprisingly little is known about viruses that infect *Acidobacteriota*. With the use of genomic data, proviral sequences have been predicted in acidobacterial genomes (Eichorst et al. 2018), and viral populations extracted from metagenomic datasets have been putatively linked to acidobacterial hosts (Emerson et al. 2018; Trubl et al. 2018; 2021; Ji et al. 2023). Here, we report, to the best of our knowledge, the first isolates of five viruses that infect *Acidobacteriota*, obtained from tundra soils in Kilpisjärvi, Finnish Lapland. The host strains originate from the same area (Kilpisjärvi) and belong to the genus *Tunturibacter* (Messyasz et al. 2024). In addition to virus isolation, we explored virus-host interactions in Kilpisjärvi meadow and fen soils by extracting viral operational taxonomic units (vOTUs) from bulk metagenomes and linking them to their putative microbial hosts.

Materials and Methods

Soil samples

The sampling sites are located in the oroarctic mountain tundra area in Kilpisjärvi, northwestern Finland (69.04°N, 20.79°E) (Viitamäki et al. 2022; Pessi et al. 2022). The main vegetation cover in the sites were fens and meadows (Supplementary Table S1). Metagenomes from soil cores that were collected in July 2017 and July 2018 (Pessi et al. 2022) were analyzed for viral sequences (see below, Identifying viral sequences in metagenomes). For virus isolation, fresh samples from the same area were collected in April 2021. Snow depth ranged from 72 to 99 cm and the mean air temperature was -2.9°C in Kilpisjärvi in April 2021 (https://en.ilmatieteenlaitos.fi/download-observations). Snow and frozen plant material were removed and samples were chiseled from the top 5 cm of soil surface, spooned into ziplock bags and stored at 4°C. All sampling tools were sterilized with 70% ethanol prior usage.

Bacterial strains and growth conditions

The 18 acidobacterial strains used as potential viral hosts in this study (Table S2) were previously isolated from Kilpisjärvi, Finland (Männistö et al. 2012; Messyasz et al. 2024). The tested strains belonged to four different genera of the class *Terriglobia*: *Tunturibacter*, *Granulicella*, *Acidicapsa*, and *Terriglobus*. The strains were grown in DSMZ medium 1284, containing 0.5 g/L glucose, 0.1 g/L yeast extract (Neogen, Lansing, USA), 0.1 g/L casamino acids (MP Biomedicals,Solon, USA), 0.04 g/L $MgSO_4$ x 7 H_2O, and 0.02 g/L $CaCl_2$ x 2 H_2O (https://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium1284.pdf), pH 5.5. For plates and top-layer agar, 15 and 4 g of agar were added per 1 L, respectively. All culturing was done aerobically at room temperature (RT).

Virus isolation

Soil samples were resuspended in the DSMZ medium 1284 broth (approximately 1:3 ratio [wet weight]) and incubated with aeration at RT for 30 min. The supernatants (Table Eppendorf centrifuge, 2500 g, RT, 30 min) were filtered (0.22 µL LLG Syringe Filters

Spheros filters) and applied to freshly grown strains in the plaque assay: 100-150 µL of the supernatant were mixed with 300 µL of the host culture and 3 mL of the soft agar (46°C), and spread as a top layer on agar plates. The plates were incubated at RT and monitored for plaque formation regularly. The observed single plaques were picked up, resuspended in the DSMZ medium 1284 broth and subjected to the plaque assay, which was repeated three consecutive times to ensure the purity of virus isolates.

Preparation of agar stocks

The top layers of semi-confluent plates were collected and mixed with the DSMZ medium 1284 broth (3 mL per plate), incubated aerobically at RT for 1 h, and the agar as well as cell debris were removed by centrifugation (F15-6x100y, 10000 g, 4°C, 30 min). The supernatant was collected, filtered (0.22 µL LLG Syringe Filters Spheros filters), and stored at 4°C. Stocks were titrated by the plaque assay method as described above.

Virus host range testing

To determine viral host ranges, stocks were first subjected to the spot test. Plates having 300 µL of the bacterial liquid culture and 3 mL of the soft agar as a top layer were prepared, and 7-µL drops of undiluted and 100-fold diluted virus stocks were applied onto them. The drops of broth containing no virus sample were used as a negative control. The plates were incubated at RT and monitored for the growth inhibition. When inhibition was observed, the virus-host pair was additionally tested by the plaque assay with a range of dilutions to verify the spot test results.

Virus purification

Viruses were precipitated from agar stocks by mixing with polyethylene glycol 8000 (PEG 8000, Thermo Scientific, final concentration 10% [w/v]) and NaCl (final concentration 0.5 M) and incubated with stirring at 4°C for 1 h. The pellets (F15-6x100y, 10000 g, 4°C, 30 min) were resuspended by adding the SM buffer (50 mM MES, pH 5.5; 100 mM NaCl; 8 mM $MgSO_4$) in the amount of ~1.5 % (v/v) of the original stock volume. If the resuspended pellets were highly viscous, the resuspension step was repeated with more SM buffer and in some cases, DNase (Stemcell Technologies) was also applied to reduce the viscosity (final concentration of 300 µg/mL). The resuspended samples were further pelleted (F15-6x100y, 10000 g, 4°C, 10 min) and either the supernatant or both the supernatant and the pellet (separately) were used for purification by rate-zonal ultracentrifugation in 10-30 % (w/v) sucrose gradients in the SM buffer. The light-scattering zones observed after ultracentrifugation (Sorvall AH629 112142.4 g or TH641 103557.6 g, 10°C, 20-60 min) were collected and pelleted (Sorvall T1270, 113488.6 g, 4°C, 3 h). In case no clear light-scattering bands could be observed, the gradients were fractionated, and the fractions with highest virus titers were used for pelleting. Pellets were resuspended in 50-100 µL of the SM-buffer, titrated by plaque assay and stored at 4°C.

Electron microscopy

The samples for transmission electron microscopy (TEM) were prepared by applying a drop of a PEG-precipitated or purified virus sample on the Mesh 200 cu grid for 1 min and rinsing

it twice with ultrapure water. The samples were stained by applying a drop of Vitroease (Thermo Scientific) or 3% (w/v) uranyl acetate (pH 4.5) for 1 min, which was repeated twice. The images were taken with the JEOL 1400 electron microscope operating at 80 kV at the Electron Microscopy Unit, Institute of Biotechnology, University of Helsinki. The size of viral particles was measured with the ImageJ program (Schneider, Rasband, and Eliceiri 2012). The head size was measured as the distance between opposite vertices of icosahedral particles, except for the virus Tunturi 4, which had a prolonged icosahedral head.

Genome sequencing and annotation of virus isolates

For the DNA extraction from virus stocks, the protocol by Santos (Santos 1991) was used with modifications as described in (Laanto et al. 2015). The extracted DNA was purified using the GeneJET Genomic DNA purification Kit (Thermo Scientific). The purified DNA was sequenced using the Nextera XP kit and Illumina MiSeq (paired end, 325 bp + 285 bp) at the DNA Sequencing and Genomics Laboratory, Institute of Biotechnology, University of Helsinki. The quality of raw Illumina reads was assessed with FastQC v. 0.11.8 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Cutadapt v. 2.7 was used for removing adaptors and trimming reads (Martin 2011). The virus isolate genomes were assembled using Spades v. 3.15.0 (Bankevich et al. 2012).

Geneious Prime v. 2021.2.2 (https://www.geneious.com) was used for the analyses of the viral genomes. Open reading frames (ORFs) were predicted using Glimmer3 v. 1.5 (Delcher et al. 2007), and their functions were predicted using BLASTx (Altschul et al. 1990) searches against the NCBI non-redundant (nr) protein sequences database (E-value threshold 0.001, identity threshold 20%, searches dated Sep-Oct 2022) and HHpred searches (probability threshold of 95%, searches dated Oct 2023) (Gabler et al. 2020). Viral genome sequences produced circular maps, and in each virus isolate, ORFs were numbered starting from the ORF putatively encoding the terminase large subunit. tRNA genes were predicted using tRNAscan-SE v. 2.0 using bacterial search mode (Lowe and Chan 2016). The programs fastANI v. 1.33 (Jain et al. 2018) and pyani v. 0.2.12 (Pritchard et al. 2015) were used for calculating average nucleotide identities (ANI) between the virus genomes. Overall nucleotide identities were calculated using Emboss Stretcher (Madeira et al. 2024). Intergenomic similarities were calculated with VIRIDIC (Moraru, Varsani, and Kropinski 2020). Pairwise similarities between the genomes were visualized using Easyfig v. 2.2.2 with BLASTn E-value threshold 0.001 (Sullivan, Petty, and Beatson 2011).

The virus isolates genomes were searched against the IMG/VR v. 4 (Camargo, Nayfach, et al. 2023) using BLASTn v. 2.13.0 with the E-value threshold 1e-5. Similarities between genomes were visualized with Circoletto (BLASTn E-value threshold 1e-5) (Darzentas 2010). To detect sequences related to virus isolates in Kilpisjärvi metagenomes, amino acid sequences from the five isolates were clustered with MMseqs2 v. 13.45111 (Steinegger and Söding 2017) to generate a non-redundant protein set (≥50% identity, ≥90% coverage). Quality-filtered metagenomic reads were then mapped to the set of non-redundant proteins with Diamond v. 2.0.15 (Buchfink, Xie, and Huson 2015) using the E-value threshold 1e-5.

Identifying viral sequences in metagenomes

Previous metagenomic data from Kilpisjärvi fens and meadows soils (Pessi et al. 2022) were analyzed for the presence of viral sequences. Raw reads were quality-checked and trimmed

as described in (Pessi et al. 2022). Each of the 22 samples (Table S1) was assembled separately using metaSpades v. 3.14.1 (k-mers 55, 99 and 127) (Bankevich et al. 2012). QUAST v. 5.0.2 (Mikheenko et al. 2018) was used for the quality assessment of the assemblies. Quality-filtered metagenomic reads were mapped to the assemblies using Bowtie v. 2.4.1 (Langmead and Salzberg 2012). For the identification of viral contigs, Virsorter v. 2.2 (Guo et al. 2021), PPR-Meta (Fang et al. 2019) and DeepVirFinder v. 1.0. (Ren et al. 2020) were used within the What-the-Phage pipeline (Marquet et al. 2022). These tools were recently benchmarked as having high sensitivity and precision (L.-Y. Wu et al. 2024). The contigs identified as viral by the three tools with scores/p-values sum >0.75 in were selected. In addition, geNomad v. 1.4.0 (Camargo, Roux, et al. 2023) was used for extracting viral sequences from the metagenomic assemblies. The predictions from the four tools were combined, their taxonomy was predicted using geNomad v. 1.4.0 (Camargo, Roux, et al. 2023), and the resulting contigs were checked for quality and completeness with CheckV v. 0.8.1 (Nayfach et al. 2021). The contigs that were ≥ 5 kbp long or predicted as ≥ 50% complete, had at least one viral gene, and no more than 1.5 host to viral genes ratio were selected for the final set of viral contigs. The set was dereplicated using CheckV -anicalc and -aniclust functions and all contigs within 95% average nucleotide identity and 85% alignment fraction were assigned to the same vOTU cluster, following the suggested standard thresholds (Roux et al. 2019).

Linking vOTUs to their putative hosts

Putative hosts for vOTUs were predicted with iPHoP v. 1.3.3 (Roux et al. 2023) using the default database (iPHoP_db_Aug23_rw) and a custom database including 796 metagenome-assembled genomes (MAGs) previously obtained from Kilpisjärvi metagenomes (Pessi et al. 2022). For proviruses, both proviral sequences and their corresponding larger contigs with remaining host regions were used as an input. The Kilpisijärvi MAGs were classified using GTDB-tk v. 2.3.2 (Chaumeil et al. 2022) with the GTDB release 214 (https://data.gtdb.ecogenomic.org/). Proviral vOTUs were predicted both by geNomad v. 1.4.0 (Camargo, Roux, et al. 2023) and CheckV v. 0.8.1 (Nayfach et al. 2021) and those assigned to *Acidobacteriota* hosts were further explored by comparing to previously reported acidobacterial proviruses (Eichorst et al. 2018) and UViGs that were retrieved from the IMG/VR v. 4 (Camargo, Nayfach, et al. 2023) with *Acidobacteriota* as a host. In addition, acidobacterial proviral vOTUs were searched against the NCBI nr database (discontinious MegaBLAST, searches dated Feb-Mar 2024, E-value threshold 0.001, query coverage threshold 10%) (Altschul et al. 1990), and for bacterial genomes returned as hits satisfying the thresholds, proviral regions were predicted with geNomad v. 1.4.0 (Camargo, Roux, et al. 2023) and further used for comparisons. The sequence similarities were analyzed using Circoletto (BLASTn E-value threshold 1e-5) (Darzentas 2010) and genome-to-genome comparisons were visualized with Easyfig v. 2.2.2 (BLASTn E-value threshold 0.001) (Sullivan, Petty, and Beatson 2011). Proteins encoded on high-quality vOTUs that were assigned to *Acidobacteriota* hosts were annotated using BLASTx (Altschul et al. 1990) searches against the NCBI nr protein sequences database (E-value threshold 0.001, searches dated Jul 2023 or Mar 2024) and HHpred searches (probability threshold 95%, searches dated Jul 2023 or Mar 2024) (Gabler et al. 2020).

Predicting AMGs

For AMGs predictions in the isolated viral genomes and vOTUs, DRAM-v v. 0.1.2 (Shaffer et al. 2020) was used as an app at the Kbase platform (Arkin et al. 2018) with the viral sequences preprocessed by the Virsorter v. 1.0.5 (Roux et al. 2015) app. AMGs of interest were searched against IMG/VR Virus Protein DB v. 4 (BLASTp, E-value threshold of 1e-5, coverage threshold of 60%, identity threshold of 35%) (Camargo, Nayfach, et al. 2023). For further analyses, selected protein sequences were aligned using Muscle v. 5.1 (Edgar 2022) and the trees constructed with Jukes-Cantor genetic distance model, the neighbour-joining tree build method, bootstrap resampling method, 500 replicates, and 50% support threshold using Geneious Tree Builder (https://www.geneious.com). Trees were visualized using ggtree v. 3.12.0 in R (Yu et al. 2017) (https://cran.r-project.org).

Whole-genome comparisons using VConTACT

Kilpisjärvi virus isolates genomes and vOTUs ≥10 kbp were used in the whole genome gene-sharing network analysis by VConTACT v. 2.0 (Bin Jang et al. 2019) together with previously reported vOTUs from peat permafrost microbial communities from Stordalen Mire, Sweden (Emerson et al. 2018), as well as acidobacterial proviruses identified in this study. The NCBI ProkaryoticViralRefSeq211-Merged database was used to resolve taxonomic clustering. The network was visualized with Cytoscape v. 3.9.1. (Shannon et al. 2003).

vOTUs communities

Quality-filtered metagenomic reads were mapped to acidobacterial vOTUs with Bowtie v. 2.4.4 (Langmead and Salzberg 2012) and the mapping output was sorted and indexed with SAMtools v. 1.16.1 (Danecek et al. 2021). CoverM v. 0.6.1 (https://github.com/wwood/CoverM) was then used to count the number of reads mapping to each vOTU, considering only matches with ≥95% identity and ≥75% coverage. CoverM v. 0.6.1 was also used to compute the fraction of each vOTU that was covered by at least one read (horizontal coverage). Abundance values were normalized to reads mapped per kilobase of contig per million reads (RPKM), and the abundance of vOTUs with <50% horizontal coverage was set to zero.

Statistical analyses were done with the package vegan v. 2.6.4 in R (https://github.com/vegandevs/vegan, https://cran.r-project.org). Differences in vOTU abundances between meadows and fens samples were visualized using principal coordinates analysis (PCoA), and confirmed with permutational ANOVA (PERMANOVA) with 9999 permutations. For both, vOTU abundances were square-rooted and distances were computed using the Bray-Curtis metric. The contribution of soil physicochemical variables was also verified with PERMANOVA. The relationships between the total abundance of *Acidobacteriota*-associated vOTUs and soil physicochemical variables and abundance of *Acidobacteriota* MAGs (Pessi et al. 2022) were verified with linear regression.

Results

Virus isolation

Five new virus isolates, which we called Tunturi 1-5, were obtained from the Kilpisjärvi soil samples on *Tunturibacter psychrotolerans* and *T. empetritectus* strains originating from the same area (Table 1). Clear plaques of 1-5 mm in diameter were observed after 5-8 days of incubation. Stocks with titers reaching 10^8-10^9 plaque forming units per mL (PFUs/mL) could be obtained for the isolates. Although the initial spot tests for the virus-host range with 18 acidobacterial strains previously isolated from Kilpisjärvi soils (Table S2) showed several inhibition zones, only one additional virus-host pair could be verified by plaque assay. Virus Tunturi 4 could infect *Granulicella* sp. J1AC2, albeit with plating efficiency lower than the one with its original isolation host (10^7 PFU/mL and 10^8 PFU/mL, respectively).

Table 1. Viruses isolated in this study.

Virus name	Host strain	Sample site	Morpho-type*, size	Virus genome				
				Length, bp	GC%	ORFs	tRNA genes	GenBank acc. number
Tunturi 1	*Tunturibacter psychrotolerans* X5P2	12217 (fen)	Podo-, head 73.2 nm, tail 23.5 nm	63,169	55.3	109	0	PP887698
Tunturi 2	*Tunturibacter psychrotolerans* X5P2	12217 (fen)	Podo-, head 64.4 nm, tail 24.3 nm	63,277	55.1	108	0	PP885685
Tunturi 3	*Tunturibacter psychrotolerans* X5P2	181 (meadow)	Podo-, head 76.7 nm, tail 21.1 nm	97,608	51.3	152	3	PP885686
Tunturi 4	*Tunturibacter empetritectus* M8UP27	12222 (meadow)	Sipho-, head 94.1x54.2 nm, tail 204.1 nm	88,042	55.4	115	1	PP885687
Tunturi 5	*Tunturibacter psychrotolerans* X4BP1	12222 (meadow)	Myo-, head 112.7 nm, tail 167.0 nm	308,711	58.4	348	43	PP885688

*Podo-, podovirus (short non-contractile tail); sipho-, siphovirus (long non-contractile tail); myo-, myovirus (long contractile tail). For viruses Tunturi 1-3 and 5, the head size was measured as the distance between opposite capsid vertices. The head of Tunturi 4 is a 94.1-nm long and 54.2-nm wide prolate icosahedron.

The electron micrographs of purified virus isolates showed that all five viruses displayed tailed particles with icosahedral heads, varying in size (Figure 1, Table 1). Viruses Tunturi 1-3 displayed icosahedral heads (~64-77 nm) and short (~21-24 nm) non-contractile tails of the podovirus morphotype. Tunturi 4 demonstrated a ~94-nm long and a ~54-nm wide elongated (prolate) icosahedral head and a ~204-nm long flexible tail, featuring the siphovirus morphotype. Tunturi 5 had the largest head (~113 nm) and a ~167-nm long contractile tail typical for myoviruses, and both extended and contracted tail conformations were observed.

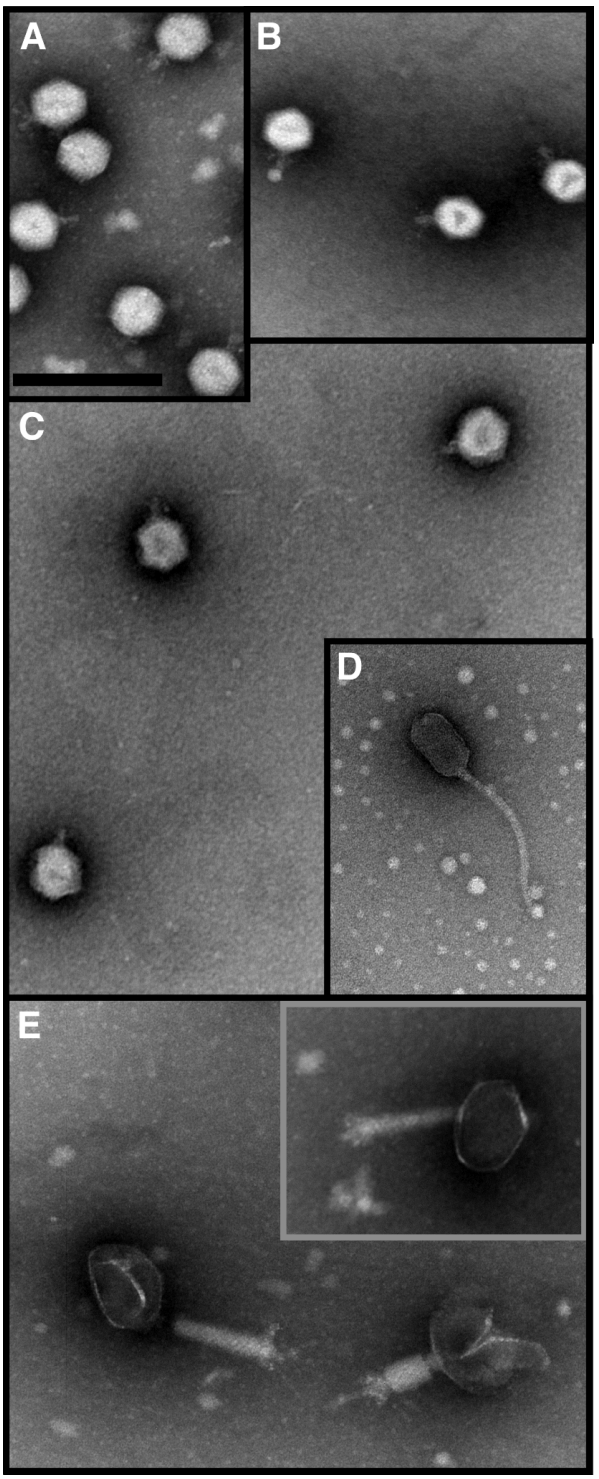


Figure 1. Electron micrographs of the five Tunturi virus isolates: (A) Tunturi 1, (B) Tunturi 2, (C) Tunturi 3, (D) Tunturi 4, (E) Tunturi 5. Scale bar in (A), 200 nm, for all sections.

Genomic characterization of virus isolates

The genome length ranged from ~63 to ~98 kbp for the isolates Tunturi 1-4, while the Tunturi 5 genome was ~309 kbp long (Table 1). The GC content varied from 51.3 to 58.4%, and the virus genomes were predicted to contain from 108 to 348 ORFs (Table 1, Tables S3-S7), tightly packed in the genomes (1.1-1.7 ORFs/kbp). The genomes of Tunturi 1 and

Tunturi 2 were clearly related, having an ANI of 97.8% and overall nucleotide identity of 87.8% (Figure 2A). The other three genomes were not similar to one another. Based on the analyses of intergenomic similarities by VIRIDIC, all five isolates represent different species, but Tunturi 1 and Tunturi 2 clustered to the same genus.

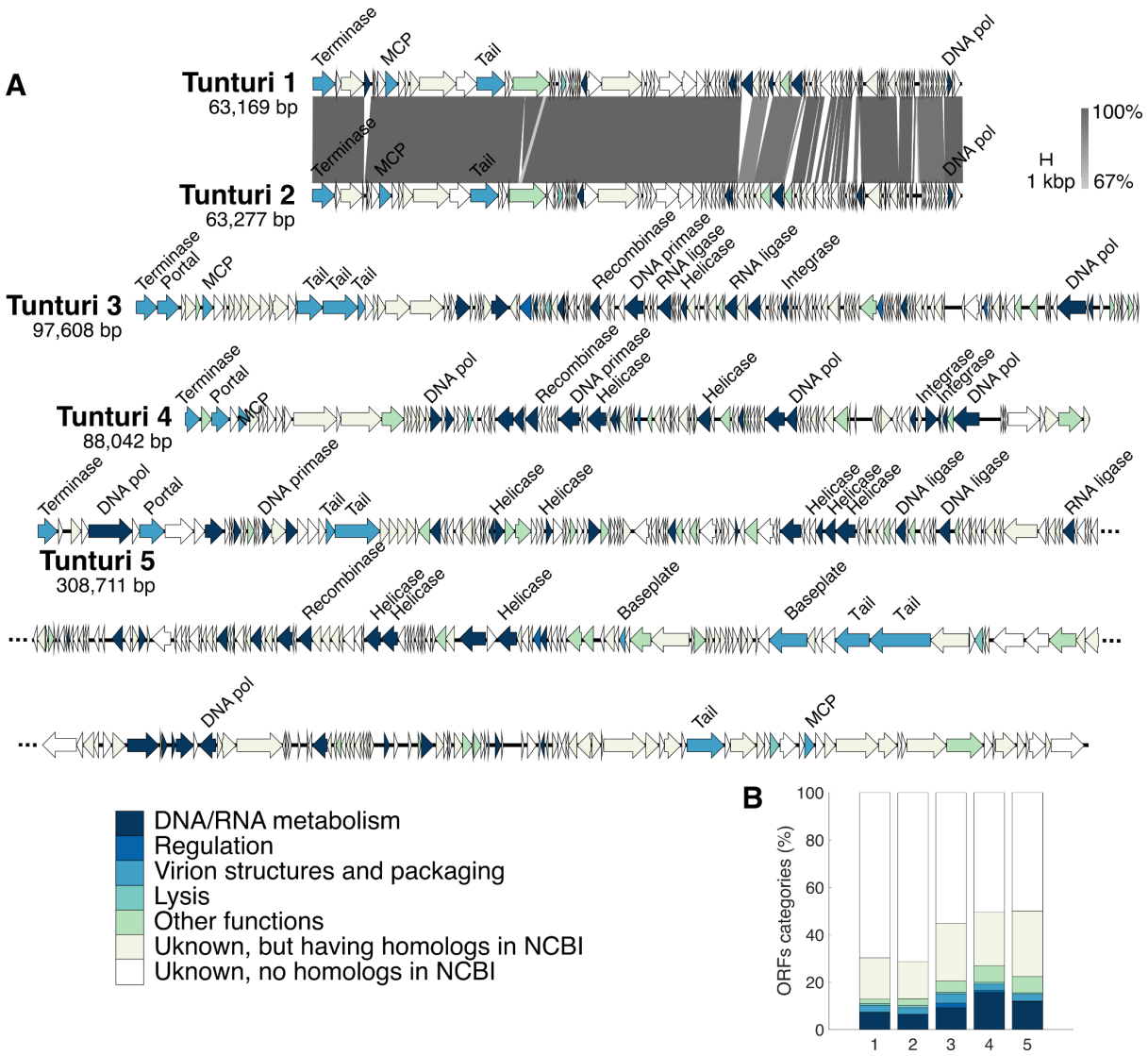


Figure 2. (A) Genomes of Tunturi 1-5. ORFs are shown as arrows and are colored according to the functional categories: see the key inset. Similarities between Tunturi 1 and Tunturi 2 genomes (BLASTn) are visualized by the shades of gray. (B) Distribution of ORFs according to the functional categories, same color code as in A. 1, Tunturi 1; 2, Tunturi 2; 3, Tunturi 3; 4, Tunturi 4; 5, Tunturi 5.

The majority of ORFs (73-87%) had no predicted functions, being either similar to other proteins of unknown functions available from the NCBI nr database or having no significant similarity to any other sequences in the NCBI nr database (Figure 2B). The functions that could be predicted were related to DNA/RNA metabolism, virion structural elements, lysis or other functions typically found in phage genomes (Tables S3-S7). Lysis-related genes were predicted in all five isolates, while recombinase and/or integrase(s) could be annotated in Tunturi 3-5 genomes. In Tunturi 5, gp47 (annotated as dUTP diphosphatase), gp54 (phosphate starvation-inducible protein PhoH) and gp215 (nicotinamide

phosphoribosyltransferase) were considered as AMGs by DRAM-v. No AMGs were detected by DRAM-v in the other four Tunturi viruses genomes. Tunturi 1 and 2 genomes had no predicted tRNA genes. Tunturi 3 encoded three tRNAs (Asn, Phe and unknown) and Tunturi 4 encoded one tRNA (unknown). In contrast, Tunturi 5 genome contained 43 tRNA genes having the anticodon sequences of 20 different amino acids: Leu (4 tRNA genes), Cys (1), Tyr (1), Ser (3), Asn (1), Gln (2), Gly (2), Thr (3), Pro (3), His (1), Ala (3), Phe (1), Arg (5), Trp (1), Asp (1), Met (1), Ile (2), Lys (2), Glu (2), Val (1) and three unknown ones.

When Tunturi genomes were searched against the IMG/VR database, many hits to soil metagenomes could be retrieved. Longer stretches of similarities (typically ≤ 70% nt identity) were observed, for example, against Arctic soil microbial communities from a glacier forefield, Greenland (Tunturi 1 and 2) and peat permafrost microbial communities from Stordalen Mire, Sweden (Tunturi 3 and 5) (Figure S2, Table S8). However, the hits were not limited to the Arctic, and included tropical soils from Puerto Rico (Tunturi 1 and 2) and soils from Indiana, Colorado and Washington in the USA (Tunturi 3 and 4). When comparing Tunturi viruses against Kilpisjärvi metagenomes, a small number of reads (up to 0.025%) could be mapped to Tunturi protein coding sequences (CDSs, up to 50% CDSs per viral genome) (Supplementary Figure S1). Tunturi genomes were further compared to Kilpisjärvi vOTUs in the whole-genome analyses using VConTACT v 2.0 (see below).

Tunturi genomes were further tested with iPHoP to assess the performance of the tool on viruses with a known host. For Tunturi 1 and 2 hosts were predicted from the genus *Edaphobacter* (class *Terriglobia*). These matches were meaningful, taking into account that the isolation hosts of these viruses, i.e., *T. psychrotolerans* strains, were indeed formerly associated with the genus *Edaphohacter.* The new nomenclature has been proposed only recently and thus has not yet been reflected in the Genome Taxonomy Database (used for MAGs classification when including them into the custom iPHoP database) and the default iPHoP database. No hosts were predicted for Tunturi 3-5 using iPHoP. The number of predictions obtained using the Tunturi genomes was consistent with benchmarks of iPHoP on soil viruses (Roux et al. 2023) .

Metagenome-derived vOTUs

The 22 assembled metagenomes from fen and meadow Kilpisjärvi soil samples produced 491,604-9,096,281 contigs per sample (Table S9). The reads were mapped back to the metagenomic assemblies with the 31-79 % overall alignment rate (Table S9). From these metagenomes, 1938 vOTUs were extracted, from which 184 were of medium or high quality, including 46 vOTUs predicted as >90% complete virus genomes (Table S10). The vOTUs ranged from 1.5-208 kb in length (median 8.5 kb, note that each vOTU consisted of a single contig). Based on the geNomad taxonomic assignments, the majority of vOTUs (1899 = 98%) were classified as tailed phages belonging to the class *Caudoviricetes*, from which four vOTUs could be further classified to the family *Herelleviridae* and one to *Straboviridae*. Among other predicted classes, *Tectiliviricetes* (3), *Malgrandaviricetes* (2), *Polintoviricetes* (1), *Megaviricetes* (1), *Faserviricetes* (1), and *Herviviricetes* (1) were found. In addition,1 vOTU was assigned to the kingdom *Bamfordvirae* with no further levels of classification and 29 vOTUs stayed unclassified.

Virus-host linkages

In total, 338 vOTUs could be linked to putative hosts (Figure 3, Table S10), from which 102 were predicted when adding Kilpisjärvi MAGs to the default iPHoP database, highlighting potential connections to local microbial hosts. The largest predicted group of hosts was the phylum *Acidobacteriota* (83), followed by *Actinomycecota* (67) and *Pseudomonadota* (58). Among acidobacterial hosts, the families *SbA1* (42) and *Bryobacteraceae* (18) were represented the most. In addition to bacteria, 16 archaeal hosts were also predicted: *Halobacteriota* (10), *Thermoproteota* (4), *Asgardarchaeota* (1), and *Aenigmatarchaeota* (1).

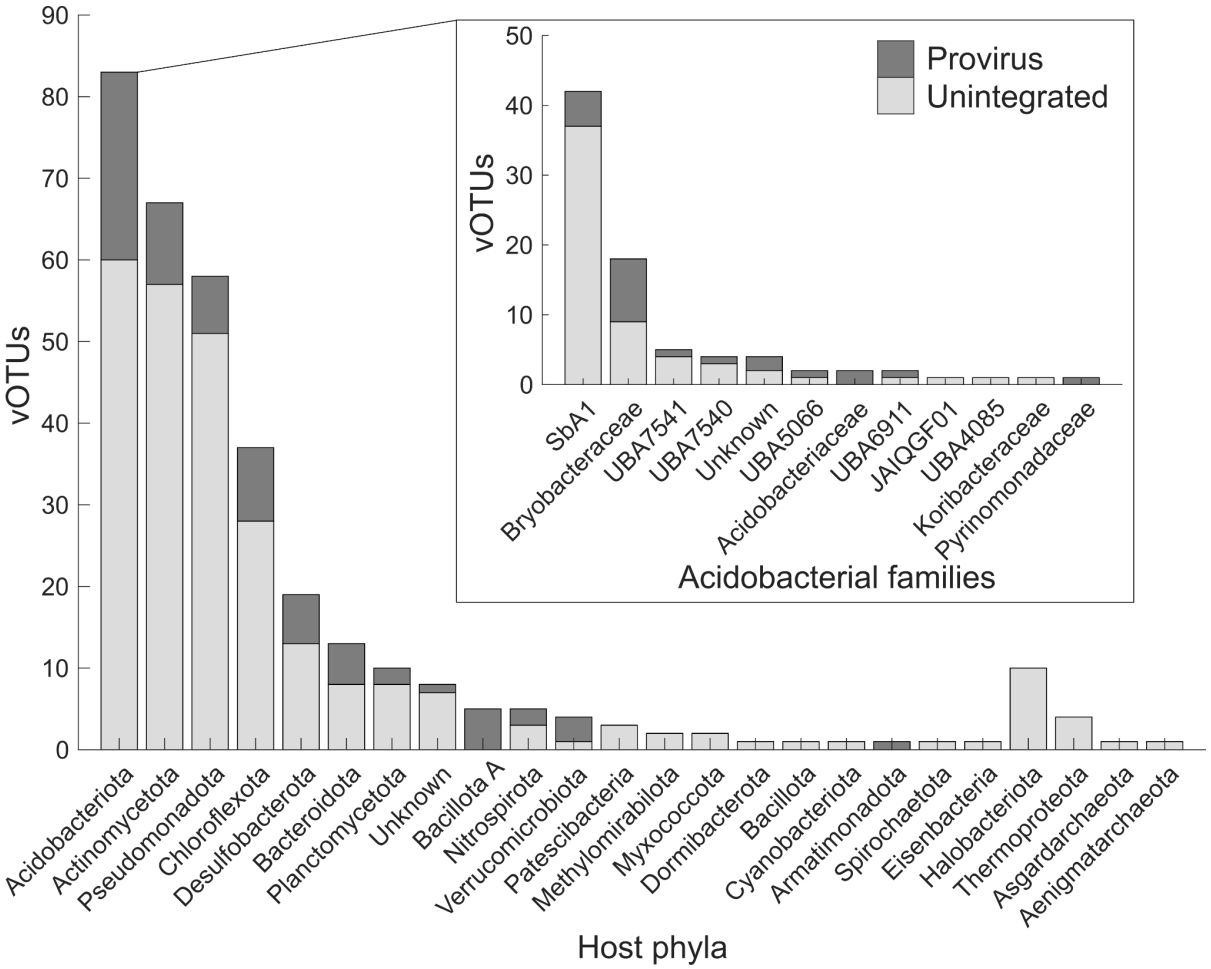


Figure 3. Host predictions for vOTUs obtained in this study (phylum level). The inset shows the distribution of 83 predicted *Acidobacteriota* hosts at the family level. Proviruses stacks are highlighted with dark gray.

Among vOTUs that were predicted to infect *Acidobacteriota*, three were of high quality: 96-100% complete, 44-45 kbp, all classified as *Caudoviricetes*, and putatively assigned to the host family *SbA1* (Figure S3). Functions could be predicted for some of the ORFs in these three vOTUs, but the vast majority (68-88 %) remained unknown (Figure S3, Tables S11-S13).

Predicted proviruses

In total, 114 vOTUs were identified as proviruses, and for 74 of them, bacterial hosts could be predicted (Table S10). No archaeal hosts were predicted for proviral vOTUs. The three largest groups among predicted proviral hosts were *Acidobacteriota* (23 vOTUs assigned), *Actinomycetota* (10), and *Pseudomonadota* (7) (Figure 3, Table S10). In acidobacterial proviruses predicted in this study, two groups with shared sequence similarity and gene order could be identified: (i) a new group (Figure 4) and (ii) the one related to the previously reported acidobacterial prophages (Eichorst et al. 2018) (Figure S4). All Kilpisjärvi vOTUs belonging to these two groups were classified as *Caudoviricetes*.

In addition to Kilpisjärvi *Acidobacteriota*-assigned proviral vOTUs, group (i) (Figure 4) included one proviral vOTU (o12214_NODE_850) from the Kilpisjärvi dataset, which was not assigned with any host, but clustered together with other acidobacterial proviral vOTUs from this group in VConTACT analysis (see below). The group also included Stordalen vOTUs (Emerson et al. 2018) from the same cluster, and with iPHoP, all these Stordallen vOTUs could be putatively assigned to acidobacterial hosts, family *Bryobacteraceae.* Two UViGs from the IMG/VR database also belonged to the group: IMGVR_UViG_639633060_000001 (*Candidatus* Solibacter usitatus Ellin6076) and IMGVR_UViG_2522125054_000002 (*Bryobacterales bacterium* KBS 96). In addition, a few proviral regions predicted in acidobacterial genomes retrieved from NCBI shared similarities with these proviruses. These NCBI references included *Paludibaculum fermentans* P105 (NZ_CP063849) isolated from a littoral wetland of a boreal lake on Valaam Island (Karelia, Russia) (Kulichevskaya et al. 2014) and three MAGs annotated as *Bryobacteraceae* bacterium (OY764903), *Bryobacterales* bacterium (OY764153), and *Paludibaculum* sp. (OY762268) originating from freshwater ciliate *Cyclidium porcatum*, UK (genome assembly GCA_963668605.1), marine sponge *Geodia parva*, Norway (GCA_963667885.1), and freshwater ciliate *Heterometopus palaeformis*, Croatia (GCF_963665245.1) metagenomes, respectively. Interestingly, all hosts that are known or predicted for viral genomes in this group belong to the class *Terriglobia*, order *Bryobacterales* (except o12220_NODE_5196 provirus, which was assigned with the host only at the class level, *Terriglobia*).

The group (ii) (Figure S4) included Kilpisjärvi *Acidobacteriota*-assigned proviral vOTUs and a few (6-18% complete) acidobacterial vOTUs from the same dataset that were not recognised as proviruses by geNomad or CheckV. These vOTUs, however, formed one cluster in the VConTACT analysis. In addition, the group also included Stordalen vOTUs (Emerson et al. 2018) from the same cluster. Almost all these (16 out of 17) Stordalen vOTUs could be predicted with acidobacterial hosts, including families *SbA*, *Koribacteraceae*, and *Acidobacteriaceae*. Similar proviral regions were also predicted in *Granulicella* sp. WH15 (NZ_CP042596) isolated from decaying wood in association with the white-rot fungus *Hypholoma fasciculare* (Netherlands) (Valásková et al. 2009) and a MAG annotated as *Granulicella* sp.(OY843766) from a lichen *Cladonia squamosa* metagenome, UK (genome assembly GCA_947623385.2, (Vancaester E 2024)). All hosts that could be predicted for viral genomes in this group belonged to the class *Terriglobia*, order *Terriglobales*.

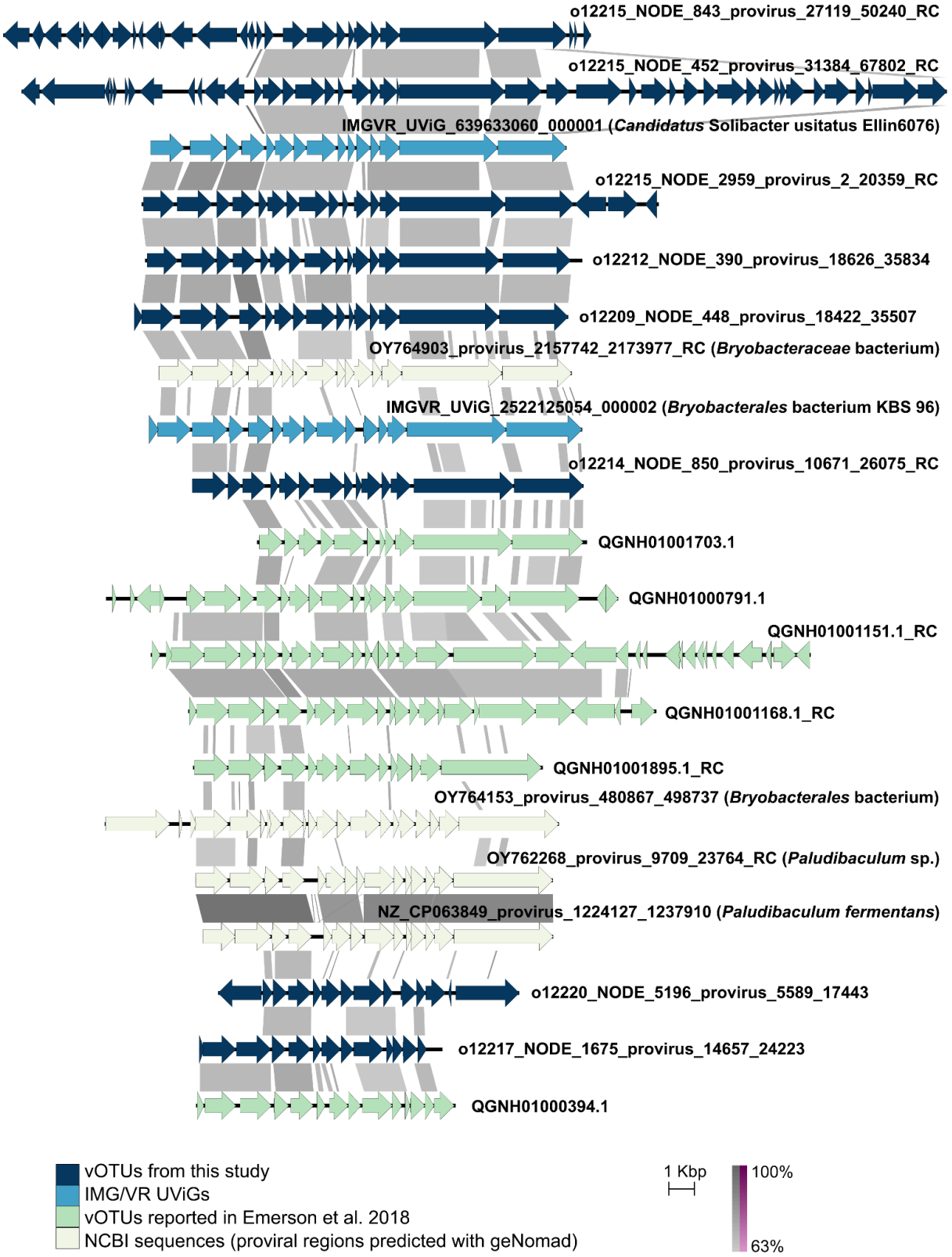


Figure 4. Putative *Acidobacteriota*-associated proviruses identified in this study and related IMG/VR UViGs, vOTUs reported in Emerson et al. 2018, and NCBI references. ORFs are shown as arrows and similarities between genomes (BLASTn) are in shades of gray (direct) or purple (invert). RC, reverse complement.

AMGs predicted in vOTUs

Using DRAM-v, 51 AMGs could be predicted in 36 Kilpisjärvi vOTUs (Table S14). Most of the detected hits were one per vOTU, but eight vOTUs had more than one AMG predicted. For example, the vOTU m12208_NODE_384, which was putatively 100% complete, but with no host prediction, had eight putative AMGs. All of the vOTUs, for which AMGs were predicted, were assigned to the class *Caudoviricetes* and nine vOTUs had putative hosts, including two archaeal ones from to the order *Methanosarcinales*. Overall, the predicted AMGs categories included transporters, carbohydrate utilization, organic nitrogen transformation, and miscellaneous functions.

Genes encoding glycosyl transferases 2 (GT2s) were predicted in five vOTUs, two of which had putative hosts: one from the archaeal genus *Methanosarcina* (phylum *Halobacteriota*) and the other one from the bacterial genus *Gallionella* (phylum *Pseudomonadota*). The search of these two putative GT2s sequences against the IMG/VR Virus Protein DB v. 4 retrieved many hits to glycosyltransferases predicted to be involved in cell wall biosynthesis in UViGs from various sampling locations (Figure S5, Tables S15, S16). In case of the putative GT2 in *Methanosarcina*-assigned vOTU (o12207_NODE_200), retrieved hits had only ~48.5% identities the highest (with the 60% coverage threshold) and represented diverse samples types, including saline water, activated sludge, *Cronobacter* sp. EKM102R isolate, city subway wood, freshwater, marine water, groundwater, hypersaline lake sediments, and soils. Soil samples originated from Risofladan (Finland) and Marcell Experimental Forest (USA). In contrast, the putative GT2 in *Gallionella*-assigned vOTU (o12215_NODE_2432) was similar to GT2s or hypothetical proteins in UViGs mostly from soil, including various soil types and geographically distant locations (Figure S5B, Table S16). The best hit was to a glycosyltransferase involved in cell wall biosynthesis in a UViG from permafrost microbial communities from Nunavut, Canada (100% coverage, 89% identity). In addition to soils, GT2 in o12215_NODE_2432 retrieved hits from freshwater, groundwater, and marine water.

Whole genome comparisons

In the whole genome gene-sharing network analysis by VConTACT with NCBI ProkaryoticViralRefSeq211-Merged database (Figure 5, Table S17), Tunturi 1 and Tunturi 2 clustered together. Tunturi 5 was a singleton. Tunturi 3 clustered with one vOTU from our dataset, o12215_NODE_6138 (*Caudoviricetes*, no host), and three vOTUs from the Stordalen dataset (QGNH01000767.1, QGNH01001143.1, QGNH01001870.1) (Emerson et al. 2018) (Figure S6). Tunturi 4 also clustered with one Stordalen vOTU (QGNH01000831.1) (Emerson et al. 2018). Using iPHoP, no hosts could be predicted for these four Stordalen vOTUs. All putative proviruses shown in Figures 4 and S4 formed two respective clusters, VC_16_0 and VC_349_0 (except three vOTUs that were not included into the analysis due to being <10 kbp long) (Figure 5, Table S17).

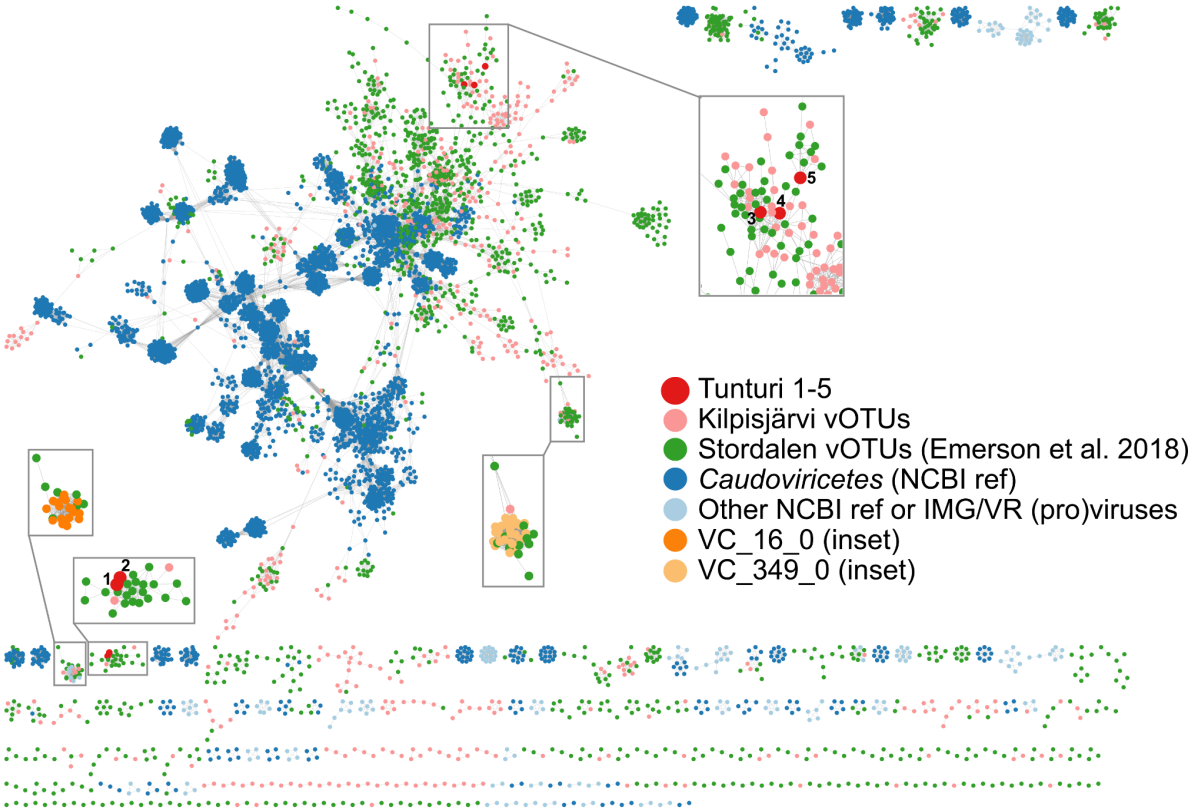


Figure 5. VConTACT v. 2.0 gene-sharing network showing virus sequences as dots and connections between genomes as lines. The sequences are color-coded (see the color key). The insets show Tunturi viruses, highlighted with numbers 1-5 for Tunturi 1-5, respectively; and virus clusters VC_16_0 and VC_349_0 comprising putative proviruses shown in Figures 4 and S4, respectively.

Ecology of viral communities

According to PCoA/db−RDA, Kilpisjärvi acidobacterial MAGs and vOTUs, including *Acidobacteriota*-assigned ones, formed different communities in fens and meadows sites (Figure S7). Environmental variables that contributed to these differences most were soil moisture, SOM, C, N content and C:N ratio, but not soil layer (organic/mineral) or pH. A few acidobacterial vOTUs were abundant in both types of ecosystems, but many were specific for meadows or fens (Figure S8). Abundances of acidobacterial vOTUs across samples linearly correlated with C, N, and SOM content values (Figure S9), as well as abundances of *Acidobacteriota* MAGs (Figure S10).

Discussion

In this study, we presented five viral isolates and 1938 metagenome-assembled vOTUs from Arctic tundra soils. To the best of our knowledge, the viruses isolated on *Tunturibacter* strains are the first *Acidobacteriota*-infecting virus isolates reported. Following the suggested criteria for genome-based phage taxonomy (Turner, Kropinski, and Adriaenssens 2021), we propose that the five isolates described in this study represent five different species, belonging to the class *Caudoviricetes*. Based on the set of strains tested in this study, the

Tunturi viruses have a narrow host range, which may be limited to their original isolation hosts, *Tunturibacter* strains, or may include other genera, as in case of Tunturi 4 being able to infect *Granulicella* sp. J1AC2. These virus-host pairs could be used as laboratory models for future studies, including developing genetic tools for the research on *Acidobacteriota*, which play important roles in key ecological processes in soil and other ecosystems. Resolving molecular details of the interactions between *Acidobacteriota* and their viruses could help understand factors affecting their dynamics and ecosystem functions in particular in climate-critical Arctic tundra soils.

From the five isolates, Tunturi 5 had the largest head and the largest genome, ~309 kbp, which qualifies Tunturi 5 as a jumbo phage (>200 kbp, (Yuan and Gao 2017)). Jumbo phages have been isolated from various environments, but more frequently from water environments, rather than soils (Yuan and Gao 2017). Metagenomics-based studies also reveal the presence of jumbo phages across ecosystems (Al-Shayeb et al. 2020). It has been observed that larger phage genomes tend to have more tRNA genes (Al-Shayeb et al. 2020; Bailly-Bechet, Vergassola, and Rocha 2007). Indeed, Tunturi 5 had 43 tRNA genes and also an ORF coding for a tRNA nucleotidyltransferase/poly(A) polymerase (gp 276), an enzyme involved in tRNA modification and maturation. Larger sets of tRNA genes in larger phage genomes seem to represent codons that are highly used by phages, while being rare in host genomes, and thus, may contribute to higher efficiency in phage protein translation (Bailly-Bechet, Vergassola, and Rocha 2007). In addition, the large genome of Tunturi 5 contains three potential AMGs, including the PhoH-like protein, which has been previously observed in phages residing in marine, paddy and wetland ecosystems (Goldsmith et al. 2011; X. Wang et al. 2016; Li et al. 2019) and might be involved in phosphate metabolism regulation under low-phosphate conditions (Hsieh and Wanner 2010).

The five Tunturi isolates displayed tailed particles, with all three different tail types, and the vast majority of Kilpisjärvi vOTUs were classified as tailed viruses belonging to the class *Caudoviricetes*. Microscopy-based studies have shown that different soil types may be dominated by different virus morphotypes (Williamson, Radosevich, and Wommack 2005; Swanson et al. 2009). In omics studies, *Caudoviricetes* typically represent a large fraction of those dsDNA viral populations that can be classified (Emerson et al. 2018; Camargo, Nayfach, et al. 2023; Coclet et al. 2023; ter Horst et al. 2021; Ji et al. 2023).

Most numerous host phyla for Kilpisjärvi vOTUs, i.e., *Acidobacteriota*, *Actinomycecota*, and *Pseudomonadota*, are abundant and active in Arctic soils (Viitamäki et al. 2022; Pessi et al. 2022; Woodcroft et al. 2018; Tveit et al. 2013; Johnston et al. 2016), as well as in other ecosystems (Banerjee, Schlaeppi, and van der Heijden 2018). *Acidobacteriota* were the largest group of hosts predicted for viral populations in Kilpisjärvi, similarly to other (sub)arctic soils: Stordalen Mire soils (Trubl et al. 2018; Emerson et al. 2018) and the active layer of Alaskan tundra soils at CiPEHR warming experiment (Ji et al. 2023). Although relatively few archaeal hosts could be predicted for Kilpisjärvi vOTUs, most of them were methanogens belonging to the classes *Methanosarcinia* and *Methanomicrobia*, and methanogenic archaea are abundant in Kilpisjärvi fens (Pessi et al. 2022). Thus, virus-host links resolved here include microbes involved in key cycling processes, such as *Acidobacteriota* and methanogens. Moreover, putative carbon-active AMGs (GT2s involved in cell wall biosynthesis) predicted in Kilpisjärvi vOTUs suggest an impact on the metabolism of their hosts, including microaerophilic Fe(II)-oxidizing bacteria belonging to the family *Gallionellaceae*, which drive carbon sequestration in wetlands (Dong, Chen, et al. 2024;

Dong, Wang, et al. 2024) and anaerobic archaea of the genus *Methanosarcina*, which are prominent methanogens in Arctic soil habitats (Wagner and Liebner 2010). Related viral GT2s seem to be common in terrestrial environments, but not limited to those, and their functional (experimental) characterization is needed to verify their functions and assess the impact on the host metabolism.

The Tunturi viruses demonstrated lytic infection cycles, but their genomes contained ORFs for putative integrases and recombinases. Near-complete Kilpisjärvi vOTUs linked to *Acidobacteriota* also contained recombinase/integrase-encoding ORFs and ~28% of Kilpisjärvi Acidobacteria-linked vOTUs were recognised as proviruses. Lysogeny is common in soils (Williamson et al. 2007; 2008), and the majority of prophages predicted in acidobacterial genomes have been found for strains originating from soil (Eichorst et al. 2018). A large set of *Terriglobia*-associated putative proviruses described here included Kilpisjärvi and Stordalen vOTUs, as well as proviral sequences in acidobacterial strains/MAGs originating from distant geographical locations, being not limited to soil, but originating also from, e.g., ciliate-, sponge-, and lichen-associated biomes, which suggests that these proviruses are fairly widespread in *Acidobacteriota* across various environments.

The majority of ORFs in the Tunturi genomes and near-complete *Acidobacteriota*-linked vOTUs remained of unknown functions, highlighting how little is known about genetic diversity in *Acidobacteriota*-infecting phages and, more generally, in soil viruses. Despite only a very small fraction of Kilpisjärvi metagenomic reads could be mapped to Tunturi CDSs, multiple genus- or higher-level links between Tunturi viruses, Kilpisjärvi vOTUs and Stordalen vOTUs could be found with the network-based whole genome gene-sharing profiles by vConTACT. Using IMG/VR, viral sequences related to the Tunturi viruses could be also detected in metagenomes from other Arctic and temperate soils, suggesting some shared viral diversity and functions across soils. Global species-level sequence conservation across soil habitats has been observed when viromes from boreal peatland in northern Minnesota were compared with the PIGEON database having viral sequences from diverse ecosystems (ter Horst et al. 2021). It is, however, unclear whether the observed patterns truly represent biological diversity or are biased because of the available (and yet limited mostly to peats) deeply sequenced virome data (ter Horst et al. 2021). Similarly, shared viral clusters have been reported when comparing viromes from Alaskan tundra permafrost and Stordalen Mire (Ji et al. 2023).

At the same time, at a local scale, soil viruses show habitat-specificity (Emerson et al. 2018; ter Horst et al. 2021). Consistent with this observation, viral communities, including *Acidobacteriota*-associated vOTUs, differed in Kilpisjärvi meadows and fens. The abundance of *Acidobacteriota*-associated vOTUs correlated with the abundance of *Acidobacteriota* MAGs, and same environmental factors (soil moisture, SOM, C, N content and C:N ratio) contributed to the differences in acidobacterial and their associated viral communities across the two types of sites. It has been shown that, for example, soil moisture (Emerson et al. 2018; R. Wu, Davison, Gao, et al. 2021; R. Wu, Davison, Nelson, et al. 2021; R. Wu et al. 2023; Williamson, Radosevich, and Wommack 2005) and soil depth (Ji et al. 2023; Liang et al. 2020; 2019) can be strong factors in modulating viral abundances and life styles. However, it is difficult to assess the causal relationships between environmental variables and dynamics in viral and host populations, as environmental variables can be interconnected, e.g., changing soil moisture can cause shifts in soil pH (Zárate-Valdez, Zasoski, and Läuchli 2006), and changes in viral communities can be linked to the

processes going on with their hosts rather than directly environmental impacts. Our current understanding of soil ecology still lacks a clear view of how multiple biotic and abiotic factors collectively drive viral communities in soil (Jansson and Wu 2023; Roux and Emerson 2022). This study adds insights into the diversity of viral communities and factors shaping them in climate-critical Arctic regions.

Data availability

Tunturi genome sequences are available from the NCBI with the following accession numbers: PP887698, PP885685-PP885688 for Tunturi 1-5, respectively. Kilpisjärvi vOTUs sequences described here can be downloaded from https://figshare.com/articles/dataset/Kilpisj_rvi_vOTUs/25976386?file=46841947.

Acknowledgements

We acknowledge Riina Ihonen for technical assistance. We thank Prof. Max Häggblom for valuable discussions about *Tunturibacter* classification. We acknowledge Electron Microscopy Unit (EMBI), DNA Genomics and Sequencing core facility, and Biocomplex unit (member of Instruct-ERIC Centre Finland, FINStruct, and Biocenter Finland), Helsinki Institute of Life Science (HiLIFE), University of Helsinki. We acknowledge CSC – IT Center for Science, Finland, for computational resources as well as for technical support. We also thank the support from the Research Council of Finland (TD: grant 330977, JH: 335354, 314114, 308128) and Kone Foundation (TD). The work conducted by the U.S. Department of Energy Joint Genome Institute (https://ror.org/04xm1d337), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231. BED was supported by the Marie Skłodowska-Curie Actions Innovative Training Networks grant agreement no. 955974 (VIROINF), the European Research Council (ERC) Consolidator grant 865694: DiversiPHI, the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2051 – Project-ID 390713860, and the Alexander von Humboldt Foundation in the context of an Alexander von Humboldt-Professorship founded by German Federal Ministry of Education and Research.

References

Albright, Michaeline B. N., La Verne Gallegos-Graves, Kelli L. Feeser, Kyana Montoya, Joanne B. Emerson, Migun Shakya, and John Dunbar. 2022. "Experimental Evidence for the Impact of Soil Viruses on Carbon Cycling during Surface Plant Litter Decomposition." *ISME Communications* 2 (1): 1–8. https://doi.org/10.1038/s43705-022-00109-4.

Al-Shayeb, Basem, Rohan Sachdeva, Lin-Xing Chen, Fred Ward, Patrick Munk, Audra Devoto, Cindy J. Castelle, et al. 2020. "Clades of Huge Phages from across Earth's Ecosystems." *Nature* 578 (7795): 425–31. https://doi.org/10.1038/s41586-020-2007-4.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. "Basic Local Alignment Search Tool." *Journal of Molecular Biology* 215 (3): 403–10. https://doi.org/10.1016/S0022-2836(05)80360-2.

Arkin, Adam P., Robert W. Cottingham, Christopher S. Henry, Nomi L. Harris, Rick L. Stevens, Sergei Maslov, Paramvir Dehal, et al. 2018. "KBase: The United States Department of Energy Systems Biology Knowledgebase." *Nature Biotechnology* 36 (7): 566–69. https://doi.org/10.1038/nbt.4163.

Bailly-Bechet, Marc, Massimo Vergassola, and Eduardo Rocha. 2007. "Causes for the Intriguing Presence of tRNAs in Phages." *Genome Research* 17 (10): 1486–95. https://doi.org/10.1101/gr.6649807.

Banerjee, Samiran, Klaus Schlaeppi, and Marcel G. A. van der Heijden. 2018. "Keystone Taxa as Drivers of Microbiome Structure and Functioning." *Nature Reviews Microbiology* 16 (9): 567–76. https://doi.org/10.1038/s41579-018-0024-1.

Bankevich, Anton, Sergey Nurk, Dmitry Antipov, Alexey A. Gurevich, Mikhail Dvorkin, Alexander S. Kulikov, Valery M. Lesin, et al. 2012. "SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing." *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology* 19 (5): 455–77.

https://doi.org/10.1089/cmb.2012.0021.
Belova, Svetlana E., Nikolai V. Ravin, Timofey A. Pankratov, Andrey L. Rakitin, Anastasia A. Ivanova, Alexey V. Beletsky, Andrey V. Mardanov, Jaap S. Sinninghe Damsté, and Svetlana N. Dedysh. 2018. "Hydrolytic Capabilities as a Key to Environmental Success: Chitinolytic and Cellulolytic Acidobacteria From Acidic Sub-Arctic Soils and Boreal Peatlands." *Frontiers in Microbiology* 9:2775. https://doi.org/10.3389/fmicb.2018.02775.
Belova, Svetlana E., Natalia E. Suzina, W. Irene C. Rijpstra, Jaap S. Sinninghe Damsté, and Svetlana N. Dedysh. 2018. "Edaphobacter Lichenicola Sp. Nov., a Member of the Family Acidobacteriaceae from Lichen-Dominated Forested Tundra." *International Journal of Systematic and Evolutionary Microbiology* 68 (4): 1265–70. https://doi.org/10.1099/ijsem.0.002663.
Bin Jang, Ho, Benjamin Bolduc, Olivier Zablocki, Jens H. Kuhn, Simon Roux, Evelien M. Adriaenssens, J. Rodney Brister, et al. 2019. "Taxonomic Assignment of Uncultivated Prokaryotic Virus Genomes Is Enabled by Gene-Sharing Networks." *Nature Biotechnology* 37 (6): 632–39. https://doi.org/10.1038/s41587-019-0100-8.
Bondy-Denomy, Joseph, Jason Qian, Edze R. Westra, Angus Buckling, David S. Guttman, Alan R. Davidson, and Karen L. Maxwell. 2016. "Prophages Mediate Defense against Phage Infection through Diverse Mechanisms." *The ISME Journal* 10 (12): 2854–66. https://doi.org/10.1038/ismej.2016.79.
Braga, Lucas P. P., Aymé Spor, Witold Kot, Marie-Christine Breuil, Lars H. Hansen, João C. Setubal, and Laurent Philippot. 2020. "Impact of Phages on Soil Bacterial Communities and Nitrogen Availability under Different Assembly Scenarios." *Microbiome* 8 (1): 52. https://doi.org/10.1186/s40168-020-00822-z.
Buchfink, Benjamin, Chao Xie, and Daniel H. Huson. 2015. "Fast and Sensitive Protein Alignment Using DIAMOND." *Nature Methods* 12 (1): 59–60. https://doi.org/10.1038/nmeth.3176.
Camargo, Antonio Pedro, Stephen Nayfach, I.-Min A. Chen, Krishnaveni Palaniappan, Anna Ratner, Ken Chu, Stephan J. Ritter, et al. 2023. "IMG/VR v4: An Expanded Database of Uncultivated Virus Genomes within a Framework of Extensive Functional, Taxonomic, and Ecological Metadata." *Nucleic Acids Research* 51 (D1): D733–43. https://doi.org/10.1093/nar/gkac1037.
Camargo, Antonio Pedro, Simon Roux, Frederik Schulz, Michal Babinski, Yan Xu, Bin Hu, Patrick S. G. Chain, Stephen Nayfach, and Nikos C. Kyrpides. 2023. "Identification of Mobile Genetic Elements with geNomad." *Nature Biotechnology*, September. https://doi.org/10.1038/s41587-023-01953-y.
Chaumeil, Pierre-Alain, Aaron J Mussig, Philip Hugenholtz, and Donovan H Parks. 2022. "GTDB-Tk v2: Memory Friendly Classification with the Genome Taxonomy Database." *Bioinformatics* 38 (23): 5315–16. https://doi.org/10.1093/bioinformatics/btac672.
Coclet, Clement, Patrick O. Sorensen, Ulas Karaoz, Shi Wang, Eoin L. Brodie, Emiley A. Eloe-Fadrosh, and Simon Roux. 2023. "Virus Diversity and Activity Is Driven by Snowmelt and Host Dynamics in a High-Altitude Watershed Soil Ecosystem." *Microbiome* 11 (1): 237. https://doi.org/10.1186/s40168-023-01666-z.
Coluccia, Marion, and Ludovic Besaury. 2023. "Acidobacteria Members Harbour an Abundant and Diverse Carbohydrate-Active Enzymes (Cazyme) and Secreted Proteasome Repertoire, Key Factors for Potential Efficient Biomass Degradation." *Molecular Genetics and Genomics* 298 (5): 1135–54. https://doi.org/10.1007/s00438-023-02045-x.
Coolen, Marco J. L., Jeroen van de Giessen, Elizabeth Y. Zhu, and Cornelia Wuchter. 2011. "Bioavailability of Soil Organic Matter and Microbial Community Dynamics upon Permafrost Thaw." *Environmental Microbiology* 13 (8): 2299–2314. https://doi.org/10.1111/j.1462-2920.2011.02489.x.
Danecek, Petr, James K. Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O. Pollard, Andrew Whitwham, et al. 2021. "Twelve Years of SAMtools and BCFtools." *GigaScience* 10 (2): giab008. https://doi.org/10.1093/gigascience/giab008.
Darzentas, Nikos. 2010. "Circoletto: Visualizing Sequence Similarity with Circos." *Bioinformatics (Oxford, England)* 26 (20): 2620–21. https://doi.org/10.1093/bioinformatics/btq484.
Davis, Kathryn E. R., Parveen Sangwan, and Peter H. Janssen. 2011. "Acidobacteria, Rubrobacteridae and Chloroflexi Are Abundant among Very Slow-Growing and Mini-Colony-Forming Soil Bacteria." *Environmental Microbiology* 13 (3): 798–805. https://doi.org/10.1111/j.1462-2920.2010.02384.x.
Dedysh, Svetlana N., and Pelin Yilmaz. 2018. "Refining the Taxonomic Structure of the Phylum Acidobacteria." *International Journal of Systematic and Evolutionary Microbiology* 68 (12): 3796–3806. https://doi.org/10.1099/ijsem.0.003062.
Delcher, Arthur L., Kirsten A. Bratke, Edwin C. Powers, and Steven L. Salzberg. 2007. "Identifying Bacterial Genes and Endosymbiont DNA with Glimmer." *Bioinformatics (Oxford, England)* 23 (6): 673–79. https://doi.org/10.1093/bioinformatics/btm009.
Deng, Jie, Yunfu Gu, Jin Zhang, Kai Xue, Yujia Qin, Mengting Yuan, Huaqun Yin, et al. 2015. "Shifts of Tundra Bacterial and Archaeal Communities along a Permafrost Thaw Gradient in Alaska." *Molecular Ecology* 24 (1): 222–34. https://doi.org/10.1111/mec.13015.
Dong, Leheng, Manjia Chen, Chengshuai Liu, Yahui Lv, Xugang Wang, Qinkai Lei, Yujuan Fang, and Hui Tong. 2024. "Microbe Interactions Drive the Formation of Floating Iron Films in Circumneutral Wetlands." *The Science of the Total Environment* 906 (January):167711. https://doi.org/10.1016/j.scitotenv.2023.167711.
Dong, Leheng, Xugang Wang, Hui Tong, Yahui Lv, Manjia Chen, Jiahui Li, and Chengshuai Liu. 2024. "Distribution and Correlation of Iron Oxidizers and Carbon-Fixing Microbial Communities in Natural Wetlands." *The Science of the Total Environment* 912 (February):168719. https://doi.org/10.1016/j.scitotenv.2023.168719.
Edgar, Robert C. 2022. "High-Accuracy Alignment Ensembles Enable Unbiased Assessments of Sequence Homology and Phylogeny." bioRxiv. https://doi.org/10.1101/2021.06.20.449169.
Eichorst, Stephanie A., Daniela Trojan, Simon Roux, Craig Herbold, Thomas Rattei, and Dagmar Woebken. 2018. "Genomic Insights into the Acidobacteria Reveal Strategies for Their Success in Terrestrial Environments." *Environmental Microbiology* 20 (3): 1041–63. https://doi.org/10.1111/1462-2920.14043.
Emerson, Joanne B., Simon Roux, Jennifer R. Brum, Benjamin Bolduc, Ben J. Woodcroft, Ho Bin Jang, Caitlin M. Singleton, et al. 2018. "Host-Linked Soil Viral Ecology along a Permafrost Thaw Gradient." *Nature Microbiology* 3 (8): 870–80. https://doi.org/10.1038/s41564-018-0190-y.
Ernakovich, Jessica G., Robyn A. Barbato, Virginia I. Rich, Christina Schädel, Rebecca E. Hewitt, Stacey J. Doherty, Emily D. Whalen, et al. 2022. "Microbiome Assembly in Thawing Permafrost and Its Feedbacks to Climate." *Global Change Biology* 28 (17): 5007–26. https://doi.org/10.1111/gcb.16231.
Fang, Zhencheng, Jie Tan, Shufang Wu, Mo Li, Congmin Xu, Zhongjie Xie, and Huaiqiu Zhu. 2019. "PPR-Meta: A Tool for Identifying Phages and Plasmids from Metagenomic Fragments Using Deep Learning." *GigaScience* 8 (6): giz066. https://doi.org/10.1093/gigascience/giz066.

Gabler, Felix, Seung-Zin Nam, Sebastian Till, Milot Mirdita, Martin Steinegger, Johannes Söding, Andrei N. Lupas, and Vikram Alva. 2020. “Protein Sequence Analysis Using the MPI Bioinformatics Toolkit.” *Current Protocols in Bioinformatics* 72 (1): e108. https://doi.org/10.1002/cpbi.108.

Giguere, Andrew T., Stephanie A. Eichorst, Dimitri V. Meier, Craig W. Herbold, Andreas Richter, Chris Greening, and Dagmar Woebken. 2021. “Acidobacteria Are Active and Abundant Members of Diverse Atmospheric H2-Oxidizing Communities Detected in Temperate Soils.” *The ISME Journal* 15 (2): 363–76. https://doi.org/10.1038/s41396-020-00750-8.

Goldsmith, Dawn B., Giuseppe Crosti, Bhakti Dwivedi, Lauren D. McDaniel, Arvind Varsani, Curtis A. Suttle, Markus G. Weinbauer, Ruth-Anne Sandaa, and Mya Breitbart. 2011. “Development of phoH as a Novel Signature Gene for Assessing Marine Phage Diversity▿.” *Applied and Environmental Microbiology* 77 (21): 7730–39. https://doi.org/10.1128/AEM.05531-11.

Guo, Jiarong, Ben Bolduc, Ahmed A. Zayed, Arvind Varsani, Guillermo Dominguez-Huerta, Tom O. Delmont, Akbar Adjie Pratama, et al. 2021. “VirSorter2: A Multi-Classifier, Expert-Guided Approach to Detect Diverse DNA and RNA Viruses.” *Microbiome* 9 (1): 37. https://doi.org/10.1186/s40168-020-00990-y.

Hausmann, Bela, Claus Pelikan, Craig W. Herbold, Stephan Köstlbacher, Mads Albertsen, Stephanie A. Eichorst, Tijana Glavina Del Rio, et al. 2018. “Peatland Acidobacteria with a Dissimilatory Sulfur Metabolism.” *The ISME Journal* 12 (7): 1729–42. https://doi.org/10.1038/s41396-018-0077-1.

Horst, Anneliek M. ter, Christian Santos-Medellín, Jackson W. Sorensen, Laura A. Zinke, Rachel M. Wilson, Eric R. Johnston, Gareth Trubl, et al. 2021. “Minnesota Peat Viromes Reveal Terrestrial and Aquatic Niche Partitioning for Local and Global Viral Populations.” *Microbiome* 9 (1): 233. https://doi.org/10.1186/s40168-021-01156-0.

Hsieh, Yi-Ju, and Barry L. Wanner. 2010. “Global Regulation by the Seven-Component Pi Signaling System.” *Current Opinion in Microbiology* 13 (2): 198–203. https://doi.org/10.1016/j.mib.2010.01.014.

Hultman, Jenni, Mark P. Waldrop, Rachel Mackelprang, Maude M. David, Jack McFarland, Steven J. Blazewicz, Jennifer Harden, et al. 2015. “Multi-Omics of Permafrost, Active Layer and Thermokarst Bog Soil Microbiomes.” *Nature* 521 (7551): 208–12. https://doi.org/10.1038/nature14238.

Jain, Chirag, Luis M. Rodriguez-R, Adam M. Phillippy, Konstantinos T. Konstantinidis, and Srinivas Aluru. 2018. “High Throughput ANI Analysis of 90K Prokaryotic Genomes Reveals Clear Species Boundaries.” *Nature Communications* 9 (1): 5114. https://doi.org/10.1038/s41467-018-07641-9.

Jansson, Janet K. 2023. “Soil Viruses: Understudied Agents of Soil Ecology.” *Environmental Microbiology* 25 (1): 143–46. https://doi.org/10.1111/1462-2920.16258.

Jansson, Janet K., and Neslihan Taş. 2014. “The Microbial Ecology of Permafrost.” *Nature Reviews Microbiology* 12 (6): 414–25. https://doi.org/10.1038/nrmicro3262.

Jansson, Janet K., and Ruonan Wu. 2023. “Soil Viral Diversity, Ecology and Climate Change.” *Nature Reviews Microbiology* 21 (5): 296–311. https://doi.org/10.1038/s41579-022-00811-z.

Ji, Mengzhi, Xiangyu Fan, Carolyn R. Cornell, Ya Zhang, Mengting Maggie Yuan, Zhen Tian, Kaili Sun, Rongfeng Gao, Yang Liu, and Jizhong Zhou. 2023. “Tundra Soil Viruses Mediate Responses of Microbial Communities to Climate Warming.” *mBio* 14 (2): e0300922. https://doi.org/10.1128/mbio.03009-22.

Johnston, Eric R., Luis M. Rodriguez-R, Chengwei Luo, Mengting M. Yuan, Liyou Wu, Zhili He, Edward A. G. Schuur, et al. 2016. “Metagenomics Reveals Pervasive Bacterial Populations and Reduced Community Diversity across the Alaska Tundra Ecosystem.” *Frontiers in Microbiology* 7:579. https://doi.org/10.3389/fmicb.2016.00579.

Jones, Ryan T., Michael S. Robeson, Christian L. Lauber, Micah Hamady, Rob Knight, and Noah Fierer. 2009. “A Comprehensive Survey of Soil Acidobacterial Diversity Using Pyrosequencing and Clone Library Analyses.” *The ISME Journal* 3 (4): 442–53. https://doi.org/10.1038/ismej.2008.127.

Kalam, Sadaf, Anirban Basu, Iqbal Ahmad, R. Z. Sayyed, Hesham Ali El-Enshasy, Daniel Joe Dailin, and Ni Luh Suriani. 2020. “Recent Understanding of Soil Acidobacteria and Their Ecological Significance: A Critical Review.” *Frontiers in Microbiology* 11:580024. https://doi.org/10.3389/fmicb.2020.580024.

Kielak, Anna M., Cristine C. Barreto, George A. Kowalchuk, Johannes A. van Veen, and Eiko E. Kuramae. 2016. “The Ecology of Acidobacteria: Moving beyond Genes and Genomes.” *Frontiers in Microbiology* 7:744. https://doi.org/10.3389/fmicb.2016.00744.

Kielak, Anna, Agata S. Pijl, Johannes A. van Veen, and George A. Kowalchuk. 2009. “Phylogenetic Diversity of Acidobacteria in a Former Agricultural Soil.” *The ISME Journal* 3 (3): 378–82. https://doi.org/10.1038/ismej.2008.113.

Kulichevskaya, Irina S., Natalia E. Suzina, W. Irene C. Rijpstra, Jaap S. Sinninghe Damsté, and Svetlana N. Dedysh. 2014. “Paludibaculum Fermentans Gen. Nov., Sp. Nov., a Facultative Anaerobe Capable of Dissimilatory Iron Reduction from Subdivision 3 of the Acidobacteria.” *International Journal of Systematic and Evolutionary Microbiology* 64 (Pt 8): 2857–64. https://doi.org/10.1099/ijs.0.066175-0.

Kwon, Min Jung, Ji Young Jung, Binu M. Tripathi, Mathias Göckede, Yoo Kyung Lee, and Mincheol Kim. 2019. “Dynamics of Microbial Communities and CO2 and CH4 Fluxes in the Tundra Ecosystems of the Changing Arctic.” *Journal of Microbiology* 57 (5): 325–36. https://doi.org/10.1007/s12275-019-8661-2.

Laanto, Elina, Jaana K. H. Bamford, Janne J. Ravantti, and Lotta-Riina Sundberg. 2015. “The Use of Phage FCL-2 as an Alternative to Chemotherapy against Columnaris Disease in Aquaculture.” *Frontiers in Microbiology* 6:829. https://doi.org/10.3389/fmicb.2015.00829.

Langmead, Ben, and Steven L. Salzberg. 2012. “Fast Gapped-Read Alignment with Bowtie 2.” *Nature Methods* 9 (4): 357–59. https://doi.org/10.1038/nmeth.1923.

Li, Xiang, Yan Sun, Junjie Liu, Qin Yao, and Guanghua Wang. 2019. “Survey of the Bacteriophage phoH Gene in Wetland Sediments in Northeast China.” *Scientific Reports* 9 (1): 911. https://doi.org/10.1038/s41598-018-37508-4.

Liang, Xiaolong, Regan E. Wagner, Jie Zhuang, Jennifer M. DeBruyn, Steven W. Wilhelm, Fang Liu, Lu Yang, Margaret E. Staton, Andrew C. Sherfy, and Mark Radosevich. 2019. “Viral Abundance and Diversity Vary with Depth in a Southeastern United States Agricultural Ultisol.” *Soil Biology and Biochemistry* 137 (October):107546. https://doi.org/10.1016/j.soilbio.2019.107546.

Liang, Xiaolong, Yingyue Zhang, K. Eric Wommack, Steven W. Wilhelm, Jennifer M. DeBruyn, Andrew C. Sherfy, Jie Zhuang, and Mark Radosevich. 2020. “Lysogenic Reproductive Strategies of Viral Communities Vary with Soil Depth and Are Correlated with Bacterial Diversity.” *Soil Biology and Biochemistry* 144 (May):107767. https://doi.org/10.1016/j.soilbio.2020.107767.

Lowe, Todd M., and Patricia P. Chan. 2016. “tRNAscan-SE On-Line: Integrating Search and Context for Analysis of Transfer RNA Genes.” *Nucleic Acids Research* 44 (W1): W54–57. https://doi.org/10.1093/nar/gkw413.

Madeira, Fábio, Nandana Madhusoodanan, Joonheung Lee, Alberto Eusebi, Ania Niewielska, Adrian R N Tivey, Rodrigo Lopez, and Sarah Butcher. 2024. “The EMBL-EBI Job Dispatcher Sequence Analysis Tools Framework in 2024.” *Nucleic Acids Research*, April, gkae241. https://doi.org/10.1093/nar/gkae241.

Malard, Lucie A., and David A. Pearce. 2018. “Microbial Diversity and Biogeography in Arctic Soils.” *Environmental Microbiology Reports* 10 (6): 611–25. https://doi.org/10.1111/1758-2229.12680.

Männistö, Minna K., Emilia Kurhela, Marja Tiirola, and Max M. Häggblom. 2013. “Acidobacteria Dominate the Active Bacterial Communities of Arctic Tundra with Widely Divergent Winter-Time Snow Accumulation and Soil Temperatures.” *FEMS Microbiology Ecology* 84 (1): 47–59. https://doi.org/10.1111/1574-6941.12035.

Männistö, Minna K., Suman Rawat, Valentin Starovoytov, and Max M. Häggblom. 2011. “Terriglobus Saanensis Sp. Nov., an Acidobacterium Isolated from Tundra Soil.” *International Journal of Systematic and Evolutionary Microbiology* 61 (Pt 8): 1823–28. https://doi.org/10.1099/ijs.0.026005-0.

Männistö, Minna K., Suman Rawat, Valentin Starovoytov, and Max M. Häggblom. 2012. “Granulicella Arctica Sp. Nov., Granulicella Mallensis Sp. Nov., Granulicella Tundricola Sp. Nov. and Granulicella Sapmiensis Sp. Nov., Novel Acidobacteria from Tundra Soil.” *International Journal of Systematic and Evolutionary Microbiology* 62 (Pt_9): 2097–2106. https://doi.org/10.1099/ijs.0.031864-0.

Männistö, Minna K., Marja Tiirola, and Max M. Häggblom. 2007. “Bacterial Communities in Arctic Fjelds of Finnish Lapland Are Stable but Highly pH-Dependent.” *FEMS Microbiology Ecology* 59 (2): 452–65. https://doi.org/10.1111/j.1574-6941.2006.00232.x.

Marquet, Mike, Martin Hölzer, Mathias W. Pletz, Adrian Viehweger, Oliwia Makarewicz, Ralf Ehricht, and Christian Brandt. 2022. “What the Phage: A Scalable Workflow for the Identification and Analysis of Phage Sequences.” *GigaScience* 11 (November):giac110. https://doi.org/10.1093/gigascience/giac110.

Martin, Marcel. 2011. “Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads.” *EMBnet.Journal* 17 (1): 10–12. https://doi.org/10.14806/ej.17.1.200.

Messyasz, Adriana, Minna K. Männistö, Lee J. Kerkhof, and Max M. Häggblom. 2024. “Genome Analysis and Description of Tunturibacter Gen. Nov. Expands the Diversity of Terriglobia in Tundra Soils.” *Environmental Microbiology* 26 (5): e16640. https://doi.org/10.1111/1462-2920.16640.

Mikheenko, Alla, Andrey Prjibelski, Vladislav Saveliev, Dmitry Antipov, and Alexey Gurevich. 2018. “Versatile Genome Assembly Evaluation with QUAST-LG.” *Bioinformatics* 34 (13): i142–50. https://doi.org/10.1093/bioinformatics/bty266.

Moraru, Cristina, Arvind Varsani, and Andrew M. Kropinski. 2020. “VIRIDIC-A Novel Tool to Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses.” *Viruses* 12 (11): 1268. https://doi.org/10.3390/v12111268.

Nayfach, Stephen, Antonio Pedro Camargo, Frederik Schulz, Emiley Eloe-Fadrosh, Simon Roux, and Nikos C. Kyrpides. 2021. “CheckV Assesses the Quality and Completeness of Metagenome-Assembled Viral Genomes.” *Nature Biotechnology* 39 (5): 578–85. https://doi.org/10.1038/s41587-020-00774-7.

Overmann, Jörg, Birte Abt, and Johannes Sikorski. 2017. “Present and Future of Culturing Bacteria.” *Annual Review of Microbiology* 71 (1): 711–30. https://doi.org/10.1146/annurev-micro-090816-093449.

Pankratov, Timofey A., Anastasia O. Ivanova, Svetlana N. Dedysh, and Werner Liesack. 2011. “Bacterial Populations and Environmental Factors Controlling Cellulose Degradation in an Acidic Sphagnum Peat.” *Environmental Microbiology* 13 (7): 1800–1814. https://doi.org/10.1111/j.1462-2920.2011.02491.x.

Pessi, Igor S., Sirja Viitamäki, Anna-Maria Virkkala, Eeva Eronen-Rasimus, Tom O. Delmont, Maija E. Marushchak, Miska Luoto, and Jenni Hultman. 2022. “In-Depth Characterization of Denitrifier Communities across Different Soil Ecosystems in the Tundra.” *Environmental Microbiome* 17 (1): 30. https://doi.org/10.1186/s40793-022-00424-2.

Pratama, Akbar Adjie, and Jan Dirk van Elsas. 2018. “The ‘Neglected’ Soil Virome - Potential Role and Impact.” *Trends in Microbiology* 26 (8): 649–62. https://doi.org/10.1016/j.tim.2017.12.004.

Pritchard, Leighton, Rachel H. Glover, Sonia Humphris, John G. Elphinstone, and Ian K. Toth. 2015. “Genomics and Taxonomy in Diagnostics for Food Security: Soft-Rotting Enterobacterial Plant Pathogens.” *Analytical Methods* 8 (1): 12–24. https://doi.org/10.1039/C5AY02550H.

Rawat, Suman R., Minna K. Männistö, Yana Bromberg, and Max M. Häggblom. 2012. “Comparative Genomic and Physiological Analysis Provides Insights into the Role of Acidobacteria in Organic Carbon Utilization in Arctic Tundra Soils.” *FEMS Microbiology Ecology* 82 (2): 341–55. https://doi.org/10.1111/j.1574-6941.2012.01381.x.

Ren, Jie, Kai Song, Chao Deng, Nathan A. Ahlgren, Jed A. Fuhrman, Yi Li, Xiaohui Xie, Ryan Poplin, and Fengzhu Sun. 2020. “Identifying Viruses from Metagenomic Data Using Deep Learning.” *Quantitative Biology (Beijing, China)* 8 (1): 64–77. https://doi.org/10.1007/s40484-019-0187-4.

Roux, Simon, Evelien M Adriaenssens, Bas E Dutilh, Eugene V Koonin, Andrew M Kropinski, Mart Krupovic, Jens H Kuhn, et al. 2019. “Minimum Information about an Uncultivated Virus Genome (MIUViG).” *Nature Biotechnology* 37 (1): 29–37. https://doi.org/10.1038/nbt.4306.

Roux, Simon, Antonio Pedro Camargo, Felipe H. Coutinho, Shareef M. Dabdoub, Bas E. Dutilh, Stephen Nayfach, and Andrew Tritt. 2023. “iPHoP: An Integrated Machine Learning Framework to Maximize Host Prediction for Metagenome-Derived Viruses of Archaea and Bacteria.” *PLoS Biology* 21 (4): e3002083. https://doi.org/10.1371/journal.pbio.3002083.

Roux, Simon, and Joanne B. Emerson. 2022. “Diversity in the Soil Virosphere: To Infinity and Beyond?” *Trends in Microbiology* 30 (11): 1025–35. https://doi.org/10.1016/j.tim.2022.05.003.

Roux, Simon, Francois Enault, Bonnie L. Hurwitz, and Matthew B. Sullivan. 2015. “VirSorter: Mining Viral Signal from Microbial Genomic Data.” *PeerJ* 3:e985. https://doi.org/10.7717/peerj.985.

Santos, M. A. 1991. “An Improved Method for the Small Scale Preparation of Bacteriophage DNA Based on Phage Precipitation by Zinc Chloride.” *Nucleic Acids Research* 19 (19): 5442. https://doi.org/10.1093/nar/19.19.5442.

Scheel, Maria, Athanasios Zervas, Carsten S. Jacobsen, and Torben R. Christensen. 2022. “Microbial Community Changes in 26,500-Year-Old Thawing Permafrost.” *Frontiers in Microbiology* 13:787146. https://doi.org/10.3389/fmicb.2022.787146.

Schneider, Caroline A., Wayne S. Rasband, and Kevin W. Eliceiri. 2012. “NIH Image to ImageJ: 25 Years of Image Analysis.” *Nature Methods* 9 (7): 671–75. https://doi.org/10.1038/nmeth.2089.

Schuch, Raymond, and Vincent A. Fischetti. 2009. “The Secret Life of the Anthrax Agent Bacillus Anthracis: Bacteriophage-Mediated Ecological Adaptations.” *PLoS ONE* 4 (8): e6532. https://doi.org/10.1371/journal.pone.0006532.

Shaffer, Michael, Mikayla A. Borton, Bridget B. McGivern, Ahmed A. Zayed, Sabina Leanti La Rosa, Lindsey M. Solden, Pengfei Liu, et al. 2020. “DRAM for Distilling Microbial Metabolism to Automate the Curation of Microbiome Function.” *Nucleic*

Acids Research 48 (16): 8883–8900. https://doi.org/10.1093/nar/gkaa621.
Shannon, Paul, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. 2003. "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks." *Genome Research* 13 (11): 2498–2504. https://doi.org/10.1101/gr.1239303.
Steinegger, Martin, and Johannes Söding. 2017. "MMseqs2 Enables Sensitive Protein Sequence Searching for the Analysis of Massive Data Sets." *Nature Biotechnology* 35 (11): 1026–28. https://doi.org/10.1038/nbt.3988.
Sullivan, Mitchell J., Nicola K. Petty, and Scott A. Beatson. 2011. "Easyfig: A Genome Comparison Visualizer." *Bioinformatics (Oxford, England)* 27 (7): 1009–10. https://doi.org/10.1093/bioinformatics/btr039.
Suttle, Curtis A. 2007. "Marine Viruses — Major Players in the Global Ecosystem." *Nature Reviews Microbiology* 5 (10): 801–12. https://doi.org/10.1038/nrmicro1750.
Swanson, M.m., G. Fraser, T.j. Daniell, L. Torrance, P.j. Gregory, and M. Taliansky. 2009. "Viruses in Soils: Morphological Diversity and Abundance in the Rhizosphere." *Annals of Applied Biology* 155 (1): 51–60. https://doi.org/10.1111/j.1744-7348.2009.00319.x.
Taş, Neslihan, Emmanuel Prestat, Shi Wang, Yuxin Wu, Craig Ulrich, Timothy Kneafsey, Susannah G. Tringe, Margaret S. Torn, Susan S. Hubbard, and Janet K. Jansson. 2018. "Landscape Topography Structures the Soil Microbiome in Arctic Polygonal Tundra." *Nature Communications* 9 (1): 777. https://doi.org/10.1038/s41467-018-03089-z.
Trubl, Gareth, Ho Bin Jang, Simon Roux, Joanne B. Emerson, Natalie Solonenko, Dean R. Vik, Lindsey Solden, et al. 2018. "Soil Viruses Are Underexplored Players in Ecosystem Carbon Processing." *mSystems* 3 (5): e00076-18. https://doi.org/10.1128/mSystems.00076-18.
Trubl, Gareth, Jeffrey A. Kimbrel, Jose Liquet-Gonzalez, Erin E. Nuccio, Peter K. Weber, Jennifer Pett-Ridge, Janet K. Jansson, Mark P. Waldrop, and Steven J. Blazewicz. 2021. "Active Virus-Host Interactions at Sub-Freezing Temperatures in Arctic Peat Soil." *Microbiome* 9 (October):208. https://doi.org/10.1186/s40168-021-01154-2.
Turetsky, Merritt R., Benjamin W. Abbott, Miriam C. Jones, Katey Walter Anthony, David Olefeldt, Edward A. G. Schuur, Charles Koven, et al. 2019. "Permafrost Collapse Is Accelerating Carbon Release." *Nature* 569 (7754): 32–34. https://doi.org/10.1038/d41586-019-01313-4.
Turner, Dann, Andrew M. Kropinski, and Evelien M. Adriaenssens. 2021. "A Roadmap for Genome-Based Phage Taxonomy." *Viruses* 13 (3): 506. https://doi.org/10.3390/v13030506.
Tveit, Alexander, Rainer Schwacke, Mette M Svenning, and Tim Urich. 2013. "Organic Carbon Transformations in High-Arctic Peat Soils: Key Functions and Microorganisms." *The ISME Journal* 7 (2): 299–311. https://doi.org/10.1038/ismej.2012.99.
Valásková, Vendula, Wietse de Boer, Paulien J. A. Klein Gunnewiek, Martin Pospísek, and Petr Baldrian. 2009. "Phylogenetic Composition and Properties of Bacteria Coexisting with the Fungus Hypholoma Fasciculare in Decaying Wood." *The ISME Journal* 3 (10): 1218–21. https://doi.org/10.1038/ismej.2009.64.
Vancaester E, Blaxter M. 2024. "MarkerScan: Separation and Assembly of Cobionts Sequenced alongside Target Species in Biodiversity Genomics Projects [Version 1; Peer Review: 1 Approved, 3 Approved with Reservations]" 9 (33). https://doi.org/10.12688/wellcomeopenres.20730.1.
Viitamäki, Sirja, Igor S. Pessi, Anna-Maria Virkkala, Pekka Niittynen, Julia Kemppinen, Eeva Eronen-Rasimus, Miska Luoto, and Jenni Hultman. 2022. "The Activity and Functions of Soil Microbial Communities in the Finnish Sub-Arctic Vary across Vegetation Types." *FEMS Microbiology Ecology* 98 (8): fiac079. https://doi.org/10.1093/femsec/fiac079.
Wagner, D., and S. Liebner. 2010. "Methanogenesis in Arctic Permafrost Habitats." In *Handbook of Hydrocarbon and Lipid Microbiology*, edited by Kenneth N. Timmis, 655–63. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-77587-4_52.
Wang, Xinzhen, Junjie Liu, Zhenhua Yu, Jian Jin, Xiaobing Liu, and Guanghua Wang. 2016. "Novel Groups and Unique Distribution of Phage phoH Genes in Paddy Waters in Northeast China." *Scientific Reports* 6 (1): 38428. https://doi.org/10.1038/srep38428.
Wang, Yajiao, Yu Liu, Yuxing Wu, Nan Wu, Wenwen Liu, and Xifeng Wang. 2022. "Heterogeneity of Soil Bacterial and Bacteriophage Communities in Three Rice Agroecosystems and Potential Impacts of Bacteriophage on Nutrient Cycling." *Environmental Microbiome* 17 (1): 17. https://doi.org/10.1186/s40793-022-00410-8.
Williamson, Kurt E., Mark Radosevich, David W. Smith, and K. Eric Wommack. 2007. "Incidence of Lysogeny within Temperate and Extreme Soil Environments." *Environmental Microbiology* 9 (10): 2563–74. https://doi.org/10.1111/j.1462-2920.2007.01374.x.
Williamson, Kurt E., Mark Radosevich, and K. Eric Wommack. 2005. "Abundance and Diversity of Viruses in Six Delaware Soils." *Applied and Environmental Microbiology* 71 (6): 3119–25. https://doi.org/10.1128/AEM.71.6.3119-3125.2005.
Williamson, Kurt E., Jennifer B. Schnitker, Mark Radosevich, David W. Smith, and K. Eric Wommack. 2008. "Cultivation-Based Assessment of Lysogeny among Soil Bacteria." *Microbial Ecology* 56 (3): 437–47. https://doi.org/10.1007/s00248-008-9362-2.
Woodcroft, Ben J., Caitlin M. Singleton, Joel A. Boyd, Paul N. Evans, Joanne B. Emerson, Ahmed A. F. Zayed, Robert D. Hoelzle, et al. 2018. "Genome-Centric View of Carbon Processing in Thawing Permafrost." *Nature* 560 (7716): 49–54. https://doi.org/10.1038/s41586-018-0338-1.
Wu, Ling-Yi, Yasas Wijesekara, Gonçalo J. Piedade, Nikolaos Pappas, Corina P. D. Brussaard, and Bas E. Dutilh. 2024. "Benchmarking Bioinformatic Virus Identification Tools Using Real-World Metagenomic Data across Biomes." *Genome Biology* 25 (1): 97. https://doi.org/10.1186/s13059-024-03236-4.
Wu, Ruonan, Michelle R. Davison, Yuqian Gao, Carrie D. Nicora, Jason E. Mcdermott, Kristin E. Burnum-Johnson, Kirsten S. Hofmockel, and Janet K. Jansson. 2021. "Moisture Modulates Soil Reservoirs of Active DNA and RNA Viruses." *Communications Biology* 4 (1): 992. https://doi.org/10.1038/s42003-021-02514-2.
Wu, Ruonan, Michelle R. Davison, William C. Nelson, Emily B. Graham, Sarah J. Fansler, Yuliya Farris, Sheryl L. Bell, et al. 2021. "DNA Viral Diversity, Abundance, and Functional Potential Vary across Grassland Soils with a Range of Historical Moisture Regimes." *mBio* 12 (6): e0259521. https://doi.org/10.1128/mBio.02595-21.
Wu, Ruonan, Michelle R. Davison, William C. Nelson, Montana L. Smith, Mary S. Lipton, Janet K. Jansson, Ryan S. McClure, Jason E. McDermott, and Kirsten S. Hofmockel. 2023. "Hi-C Metagenome Sequencing Reveals Soil Phage-Host Interactions." *Nature Communications* 14 (1): 7666. https://doi.org/10.1038/s41467-023-42967-z.
Wu, Ruonan, Clyde A. Smith, Garry W. Buchko, Ian K. Blaby, David Paez-Espino, Nikos C. Kyrpides, Yasuo Yoshikuni, et al. 2022. "Structural Characterization of a Soil Viral Auxiliary Metabolic Gene Product - a Functional Chitosanase." *Nature Communications* 13 (1): 5485. https://doi.org/10.1038/s41467-022-32993-8.
Yu, Guangchuang, David K. Smith, Huachen Zhu, Yi Guan, and Tommy Tsan-Yuk Lam. 2017. "Ggtree: An r Package for

Visualization and Annotation of Phylogenetic Trees with Their Covariates and Other Associated Data." *Methods in Ecology and Evolution* 8 (1): 28–36. https://doi.org/10.1111/2041-210X.12628.
Yuan, Yihui, and Meiying Gao. 2017. "Jumbo Bacteriophages: An Overview." *Frontiers in Microbiology* 8 (March):403. https://doi.org/10.3389/fmicb.2017.00403.
Zárate-Valdez, José L., Robert J. Zasoski, and André E. Läuchli. 2006. "SHORT-TERM EFFECTS OF MOISTURE CONTENT ON SOIL SOLUTION pH AND SOIL EH." *Soil Science* 171 (5): 423. https://doi.org/10.1097/01.ss.0000222887.13383.08.

Supplementary Figures

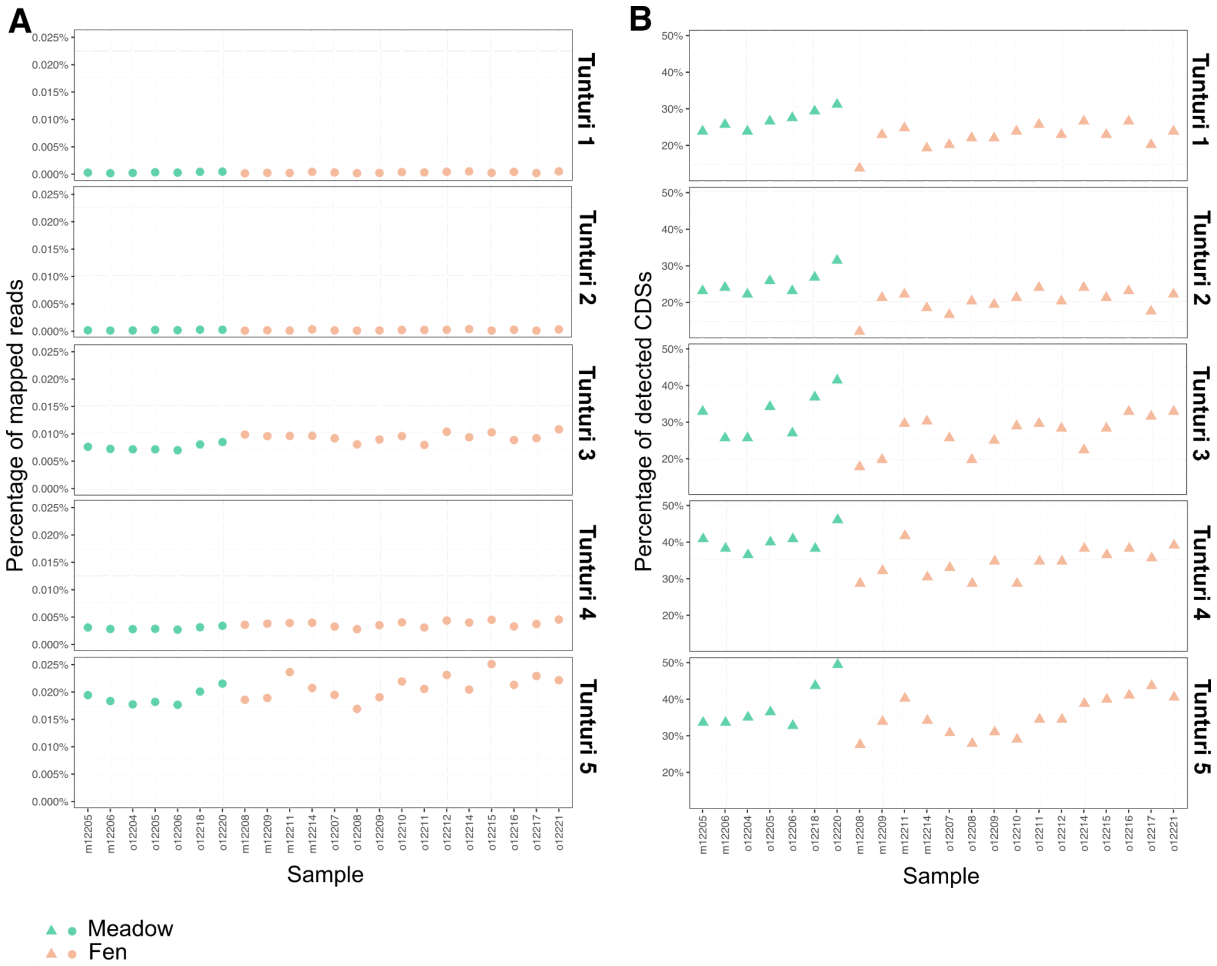


Figure S1. Distribution of the five isolates across the 22 Kilpisjärvi meadow and fen metagenomes shown as (A) the percentage of mapped reads and (B) the percentage of detected CDSs.

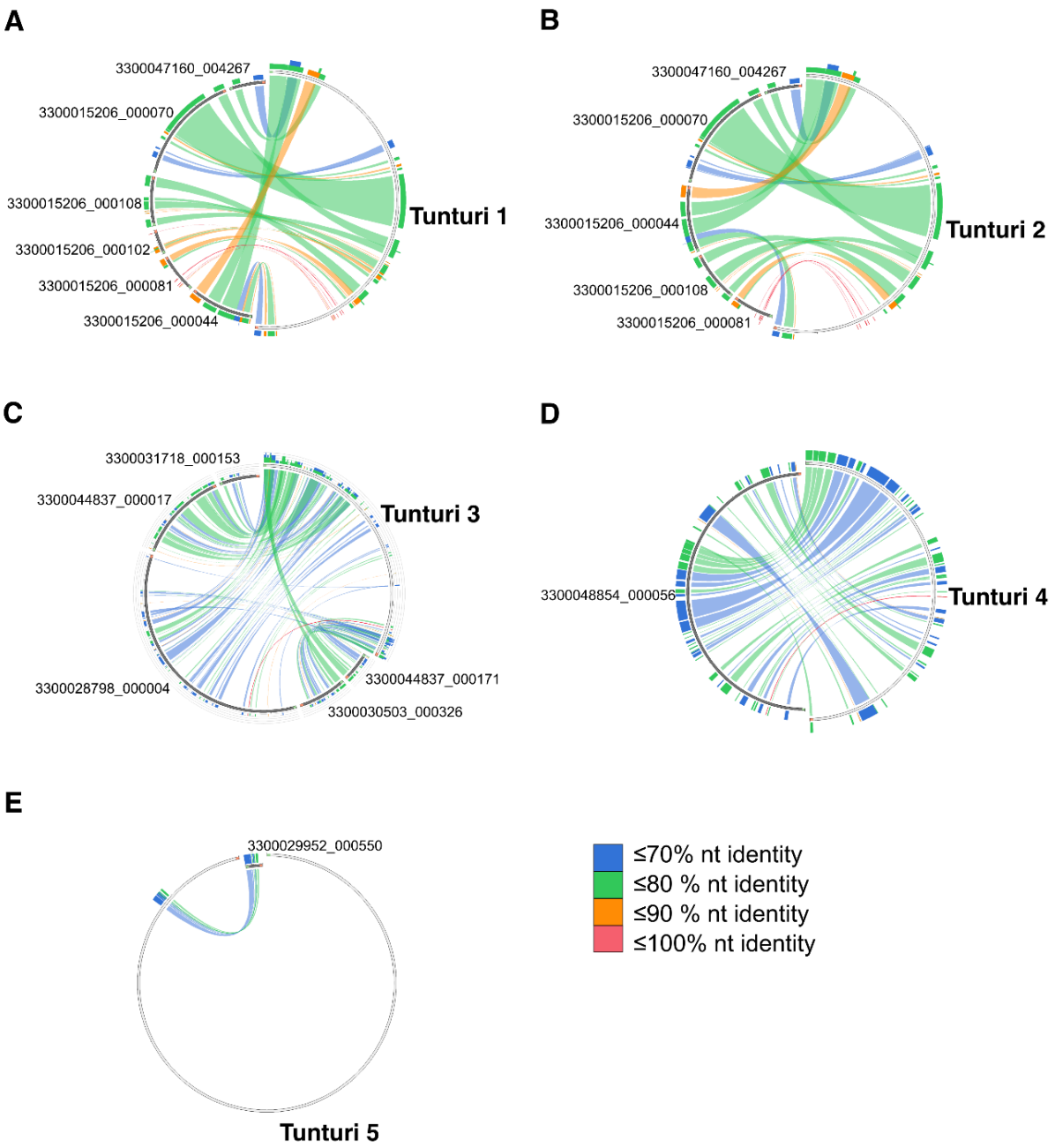


Figure S2. Similarities between the genomes of Tunturi 1-5 viruses and UViGs retrieved from the IMG/VR database. UViGs IDs are labeled skipping the “IMGVR_UViG_” prefix, for more information about UViGs, see Table S16. Ribbons are colored by % identity (see the key). Minimum and maximum % identities: (A) 64.71 and 100.00 , (B) 64.96 and 100.00, (C) 65.11 and 95.83, (D) 64.79 and 90.32, (E) 64.69 and 79.45. Orientation is clockwise for all sequences.

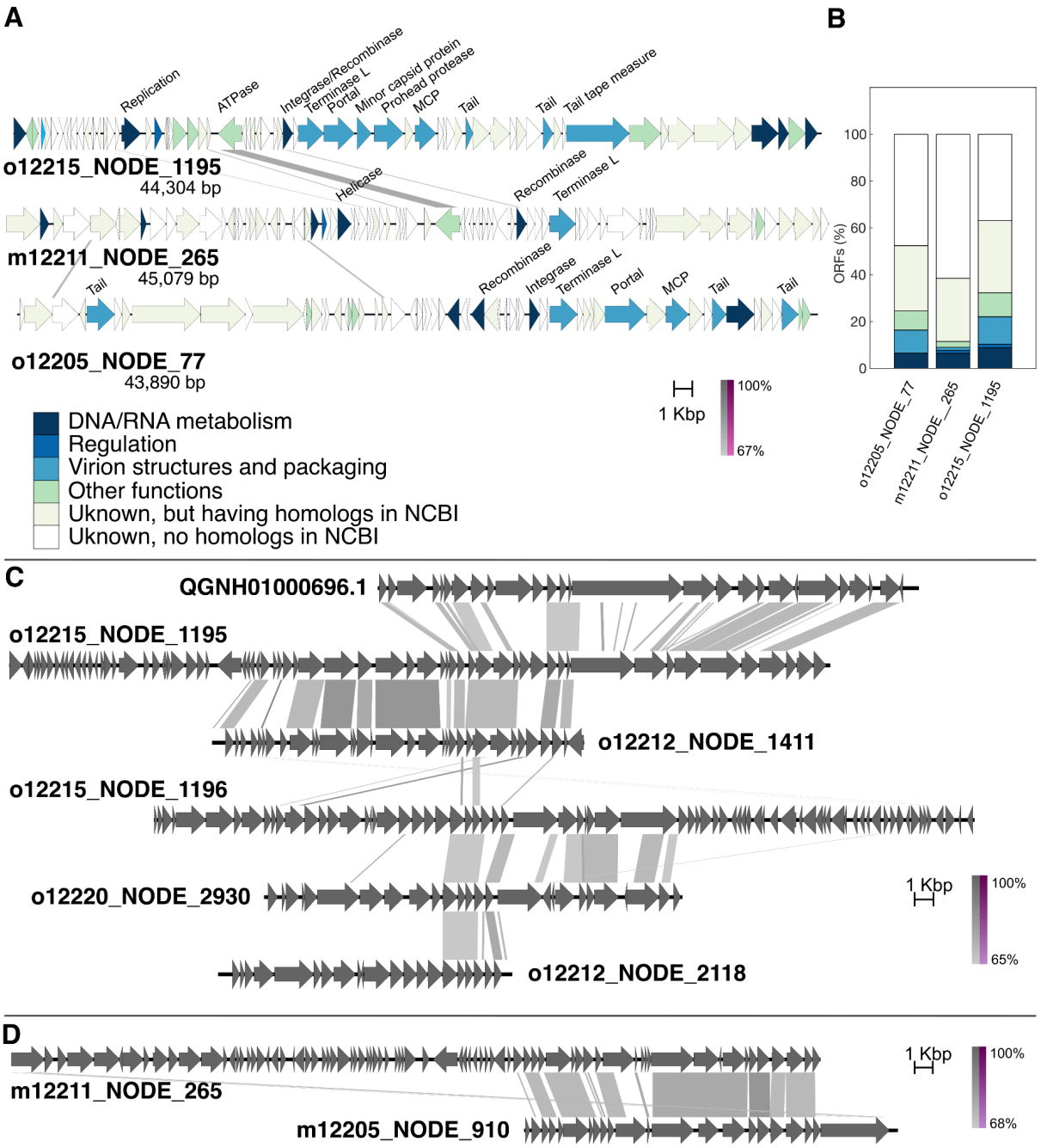
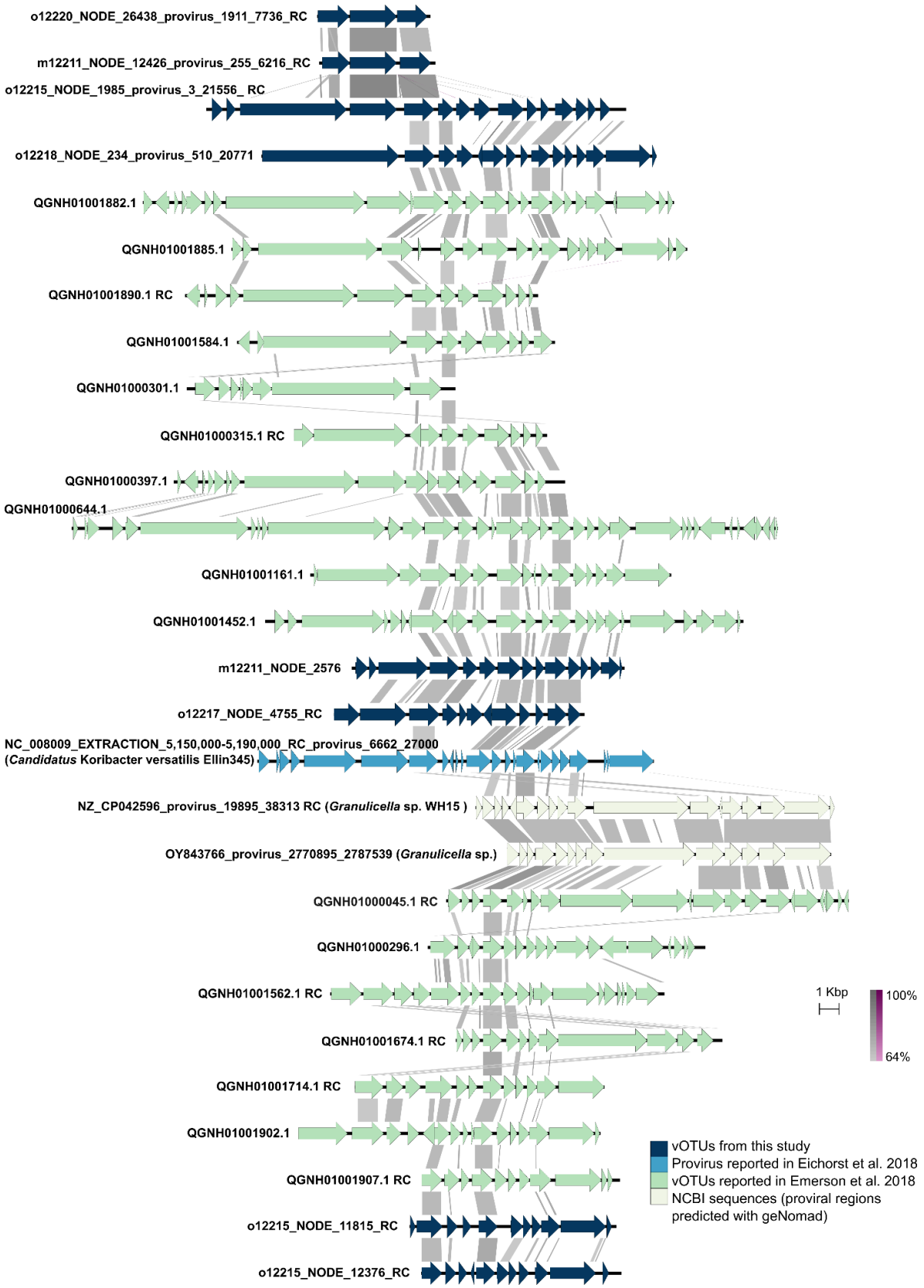


Figure S3. High-quality (96-100% complete) vOTUs assigned to *Acidobacteriota*. (A) Genomes with ORFs shown as arrows and colored according to the functional categories (see the color key). (B) Distribution of ORFs according to the functional categories, same color key as in (A). Similarities between the *Acidobacteriota*-assigned complete vOTUs (C) o12215_NODE_1195 and (D) m12211_NODE_265 to their respective viral cluster members determined in the VConTACT analyses. Similarities between the genomes (BLASTn) are shown with the shades of gray (direct) or purple (invert).



FigureS4. Putative acidobacterial proviruses identified in this study and related vOTUs reported in Emerson et al. 2018, *Candidatus* Koribacter versatilis provirus reported in Eichorst et al. 2018, and NCBI references. ORFs are shown as arrows and similarities between genomes (BLASTn) are in shades of gray (direct) or purple (invert). RC, reverse complement.

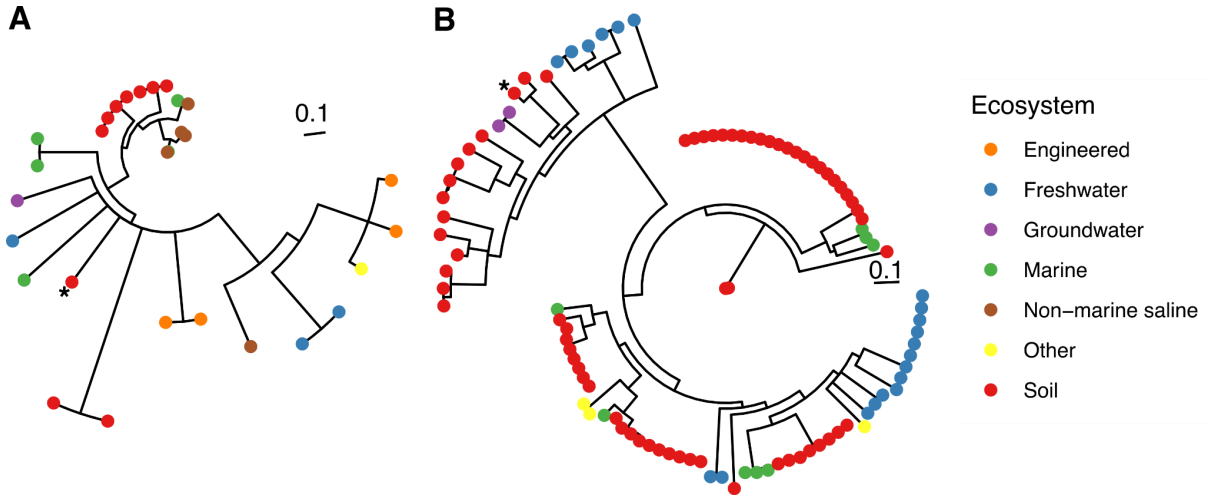


Figure S5. Phylogenetic trees of putative GT2s in (A) *Methanosarcina*-assigned o12207_NODE_200 and (B) *Gallionella*-assigned o12215_NODE_2432 and their respective BLASTp hits retrieved from the IMG/VR Virus Protein DB v. 4. For more information about UViGs that the hits represent, see Table S15. Color code refers to the sites that IMG/VR UViGs originate from. GT2s identified in this study are marked with an asterisk. Scale bar, substitutions per site.

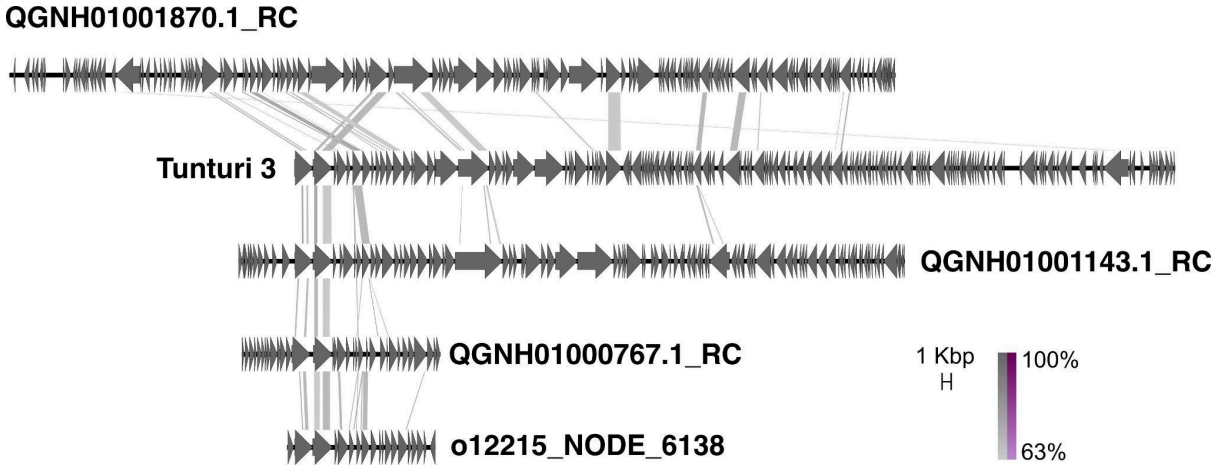


Figure S6. VConTACT viral cluster comprising the isolate Tunturi 3, o12215_NODE_6138 (vOTU obtained in this study) and three vOTUs reported in Emerson et al 2018. Similarities between the genomes (BLASTn) are shown with the shades of gray (direct) or purple (invert).

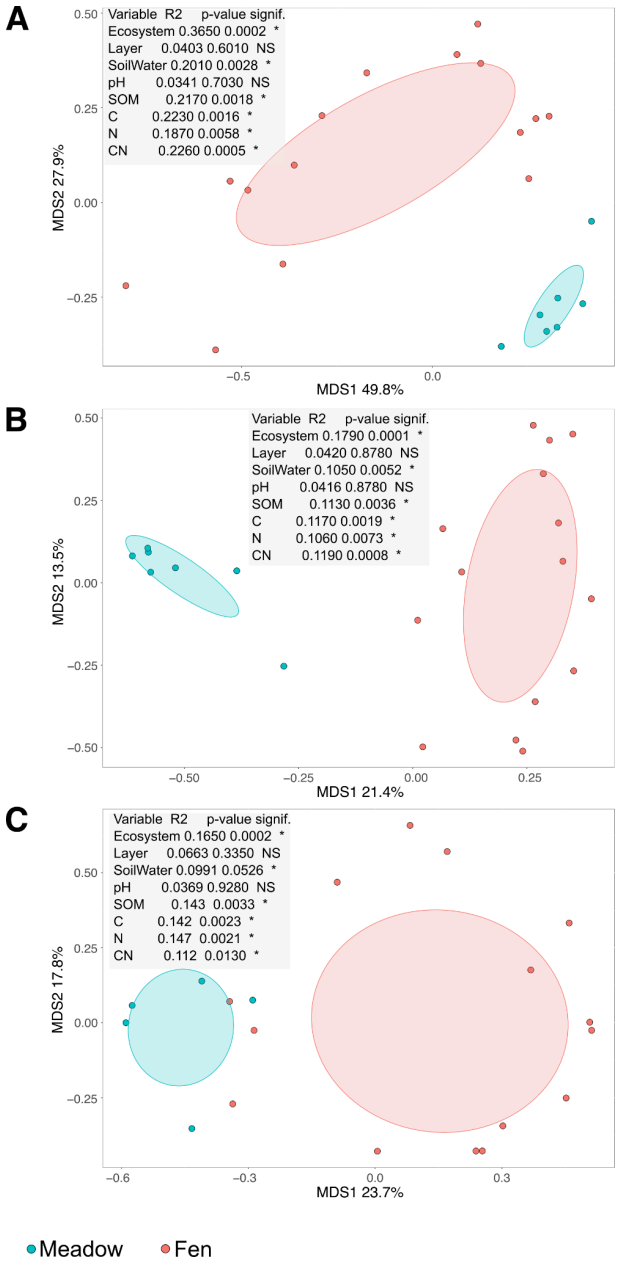


Figure S7. PCoA of (A) Kilpisjärvi *Acidobacteriota* MAGs (n=172) (Pessi et al. 2022), (B) all vOTUs (n=1938), and (C) *Acidobacteriota*-associated vOTUs (n=83) in 22 Kilpisjärvi meadows and fens metagenomes. In (B) and (C) the minimum 50% horizontal coverage was applied. Ellipses show one standard deviation around the group's centroid (95% confidence interval). The R2 and p-values were obtained separately for each variable.

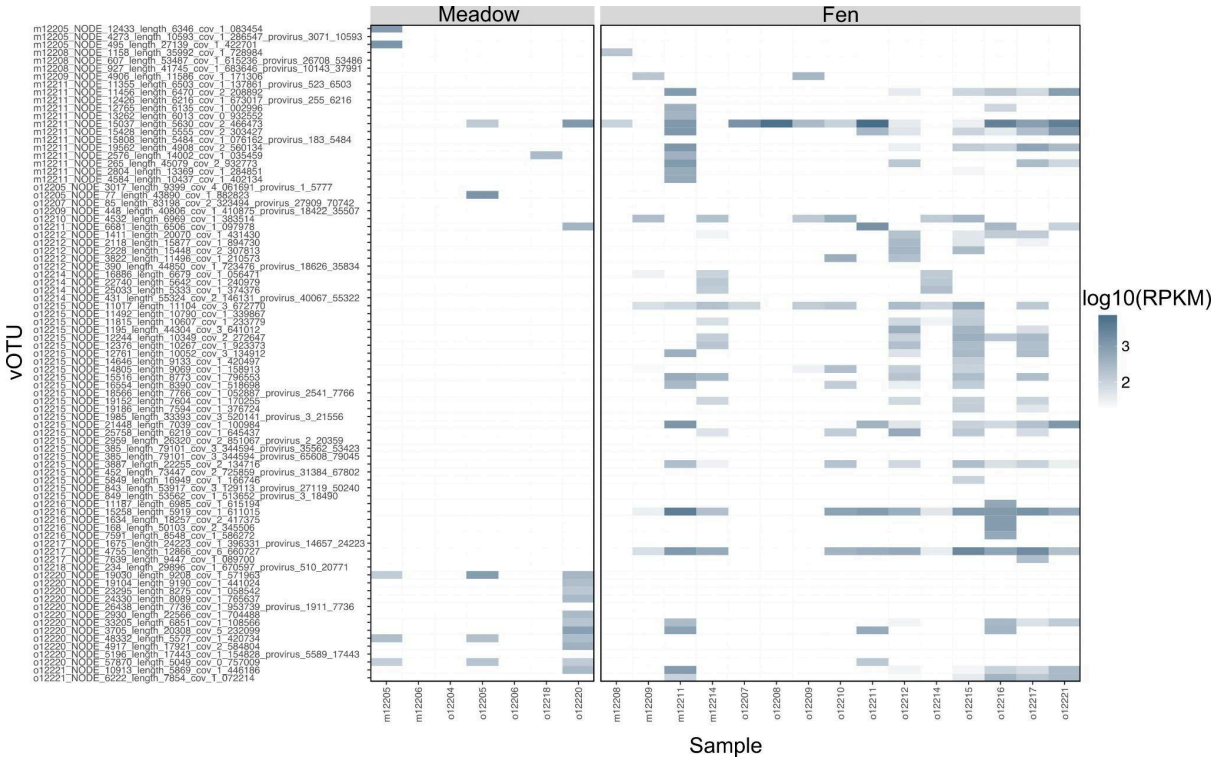


Figure S8. Distribution of *Acidobacteriota*-associated vOTUs across 22 Kilpisjärvi fen and meadows metagenomes, 50% horizontal coverage applied.

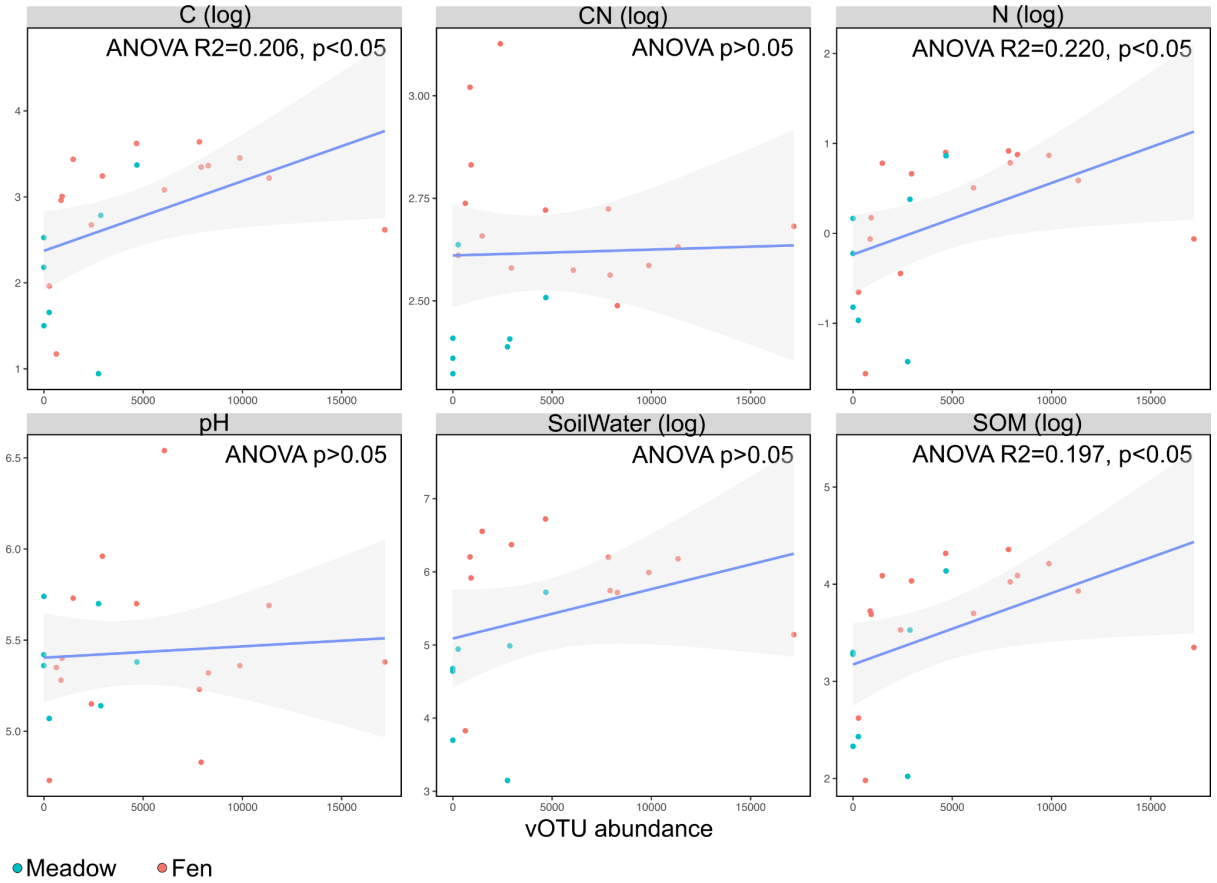


Figure S9. Correlations between abundances of *Acidobacteriota*-associated vOTUs and environmental variables across 22 Kilpisjärvi meadow and fen sites. The lines show the fitted linear models and the 95% confidence interval (shaded area).

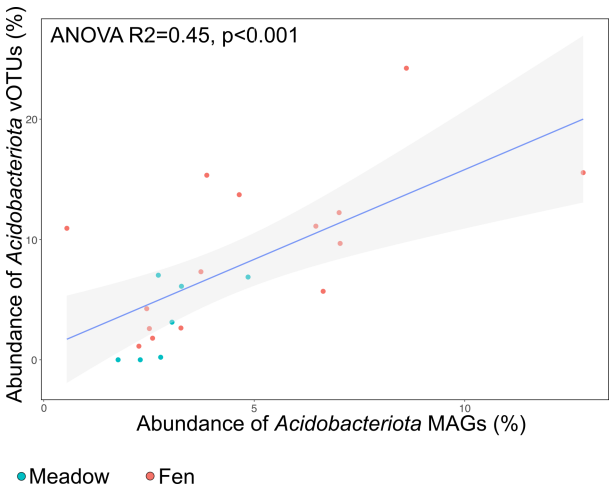


Figure S10. Correlation between the abundances of *Acidobacteriota* MAGs and *Acidobacteriota*-associated vOTUs across 22 Kilspisjärvi metagenomes. The abundance of *Acidobacteriota* MAGs was calculated as percentages relative to the total number of reads (Pessi et al. 2022). The abundance of *Acidobacteriota*-associated vOTUs was calculated as percentages relative to the number of reads mapped to all (1938) vOTUs.The line shows the fitted linear models and the 95% confidence interval (shaded area).