Truncated denitrifiers dominate the denitrification pathway in tundra soil metagenomes

Igor S. Pessi[1,2], Sirja Viitamäki[1], Anna-Maria Virkkala[3,4], Eeva Eronen-Rasimus[1,5], Tom O. Delmont[6], Maija E. Marushchak[7,8], Miska Luoto[3] and Jenni Hultman[1,2,*]

[1]Department of Microbiology, University of Helsinki, Helsinki, Finland
[2]Helsinki Institute of Sustainability Science (HELSUS), Helsinki, Finland
[3]Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland
[4]Woodwell Climate Research Center, Falmouth, MA, USA
[5]Marine Research Centre, Finnish Environment Institute (SYKE), Helsinki, Finland
[6]Department of Bioinformatics, Genoscope, Paris, France
[7]Department of Biological and Environmental Science, University of Jyväskylä, Finland
[8]Department of Environmental and Biological Sciences, University of Eastern Finland, Kuopio, Finland
*Corresponding author: jenni.hultman@helsinki.fi

Abstract

In contrast to earlier assumptions, there is now mounting evidence for the role of tundra soils as important sources of the greenhouse gas nitrous oxide (N_2O). However, the microorganisms involved in the cycling of N_2O in this system remain largely uncharacterized. In this study, we manually binned and curated 796 metagenome-assembled genomes (MAGs) from tundra soils in northern Finland, covering a wide range of ecosystems from dry heathlands to water-logged fens. We then searched for MAGs harbouring genes involved in denitrification, an important biotic process driving N_2O emissions. Communities of potential denitrifiers differed across soil ecosystems but all were dominated by poorly characterized taxa with truncated denitrification pathways, i.e. lacking one or more denitrification genes. Water-logged fens were dominated by poorly known microorganisms affiliated with the Chloroflexota lineage Ellin6529 and the Acidobacteriota subdivision Gp23. Re-analysis of a previously published metagenomic dataset from soils in northern Sweden suggests that truncated denitrifiers are dominant throughout the tundra biome. With the support of *in situ* measurements of N_2O fluxes, we hypothesize that truncated denitrifiers contribute significantly to the cycling of N_2O in tundra soils.

Introduction

Nitrous oxide (N_2O) is a greenhouse gas (GHG) that has approximately 300 times the global warming potential of carbon dioxide on a 100-year scale [1]. Atmospheric N_2O concentrations have increased by nearly 20% since pre-industrial times, with soils – both natural and anthropogenic – accounting for up to 70% of the global emissions [2]. Despite being nitrogen (N) limited and enduring low temperatures throughout most of the year, tundra soils are increasingly recognized as important sources of N_2O [3–7]. The relative contribution of tundra soils to global GHG emissions is predicted to increase in the future [8, 9], as the warming rate at high latitude environments is more than twice as high than in tropical and temperate regions [10].

Microbial denitrification is an important biotic source of N_2O [11]. Denitrification is a series of enzymatic steps in which soluble nitrate (NO_3^-) or nitrite (NO_2^-) are sequentially reduced to the gaseous products nitric oxide (NO), N_2O, and dinitrogen (N_2). The denitrification trait is common across a wide range of archaea, bacteria, and fungi, most of which are facultative anaerobes that switch to N oxides as electron acceptors in anoxic conditions [12]. Denitrification has been traditionally known as a modular community process performed in synergy by different microbial taxa that execute only a subset of the complete denitrification pathway [12, 13]. With the growing number of microbial genomes sequenced in recent decades, it has become evident that only a fraction of the microorganisms involved in the denitrification pathway encode the enzymatic machinery needed for complete denitrification [14, 15]. For instance, a study that investigated 652 genomes of cultured microorganisms containing denitrification genes showed that only approximately 31% encode the full set of enzymes needed for complete denitrification [14].

Compared to high N_2O-emitting systems such as agricultural and tropical soils, our knowledge of denitrifier communities in tundra soils is limited. As denitrification potentially results in the loss of N to the atmosphere, it enhances the N-limited status of tundra systems thus impacting both microbial and plant productivity [16, 17]. Investigations of denitrifier diversity in the tundra have been largely limited to gene-centric surveys using microarrays, amplicon sequencing, qPCR, and read-based metagenomics, which provide limited information on the taxonomic identity and genomic composition of community members. These studies have shown that tundra denitrifier communities are dominated by members of the Proteobacteria, Actinobacteria, and Bacteroidetes, and that the potential for complete denitrification is usually present at the community level [18–22]. However, it is not known whether the complete

denitrification potential occurs within discrete microbial populations or is widespread throughout populations of truncated denitrifiers.

A better understanding of the ecological, metabolic, and functional traits of denitrifiers is critical for improving current models and mitigating N_2O emissions [23]. This invariably relies on the characterization of the so-called uncultured majority, i.e. microorganisms that have not been cultured to date but which comprise a high proportion of the microbial diversity in complex ecosystems [24, 25]. Genome-resolved metagenomics is a powerful tool to access the genomes of uncultured microorganisms and has provided important insights into carbon cycling processes in tundra soils [26–28]. However, it has not yet been applied to investigate the mechanisms driving N_2O cycling in the tundra. Here, we used genome-resolved metagenomics to obtain a more in-depth picture of the diversity and metabolic capabilities of potential denitrifiers in tundra soils. We analysed 1.4 Tb and 21.5 Gb of metagenomic reads from 69 soil samples taken from different ecosystems in an area of mountain tundra biome in Kilpisjärvi, northern Finland. By linking denitrification genes to specific microbial populations, we show that denitrifier communities in the tundra are dominated by microorganisms with truncated denitrification pathways, i.e. harbouring only a subset of the genes required for complete denitrification. By coupling an in-depth characterization of denitrifier communities with *in situ* measurements of N_2O fluxes, we hypothesize that N_2O cycling in tundra soils is driven primarily by microorganisms with truncated denitrification pathways.

Methods

Study area and sampling for metagenomic analysis

The Saana Nature Reserve (69.04°N, 20.79°E) is located in Kilpisjärvi, northern Finland **(Suppl. Fig. S1a)**. The area is part of the mountain tundra biome and is characterized by a mean annual temperature of –1.9°C and annual precipitation of 487 mm [29]. Our study sites are distributed across Mount Saana and Mount Korkea-Jehkas and the valley in between **(Suppl. Fig. S1b)**, and include barren soils, heathlands (dominated by evergreen and deciduous shrubs), meadows (dominated by graminoids and forbs), and fens **(Suppl. Fig. S1c)**. In previous studies, we have established in the area a systematic fine-scale sampling of microclimate, soil conditions, and vegetation in topographically distinct environments [30–32]. Local variation in topography and soil properties creates a mosaic of habitats characterized by contrasting ecological conditions. This makes the study setting ideal to investigate species-environment relationships and ecosystem functioning in the tundra [32–34].

97 Sampling for metagenomic analysis was performed across 43 sites (barren soils, n = 2;
98 heathlands, n = 18; meadows, n = 7; fens, n = 16) in July 2017 and July 2018, during the growing
99 season in the northern hemisphere. Samples were obtained with a soil corer sterilized with 70%
100 ethanol and, when possible, cores were split into organic and mineral samples using a sterilized
101 spatula. In total, 41 organic and 28 mineral samples were obtained for metagenomic analysis
102 **(Suppl. Table S1)**. Samples were transferred to a whirl-pack bag and immediately frozen in
103 dry ice. Samples were transported frozen to the laboratory at the University of Helsinki and
104 kept at –80°C until analyses.

105 Soil physicochemical characterization and *in situ* measurement of GHG fluxes

106 Soil physicochemical characterization was done using an extended set of 228 sites distributed
107 across the study area. Soil pH, moisture, and soil organic matter (SOM) content were measured
108 according to Finnish (SFS) and international (ISO) standards (SFS 300, ISO 10390, and SFS
109 3008). Carbon (C) and N content were measured using a Vario Micro Cube machine (Elementar,
110 Langenselbold, Germany).

111 *In situ* ecosystem-level N_2O and methane (CH_4) fluxes were measured from 101 sites using a
112 static, non-steady state, non-flow-through system composed of a darkened acrylic chamber
113 (20 cm diameter, 25 cm height) [4, 35]. Measurements were conducted between 2nd July and 2nd
114 August 2018, between 10 am and 5 pm. At each site, five 25 mL gas samples were taken during
115 a 50-minute chamber closure and transferred to evacuated Exetainer vials (Labco, Lampeter,
116 UK). Gas samples were analysed using an Agilent 7890B gas chromatograph (Agilent
117 Technologies, Santa Clara, CA, USA) equipped with an autosampler (Gilson, Middleton, WI,
118 USA) and a flame ionization detector for CH_4 and an electron capture detector for N_2O. Gas
119 concentrations were calculated from the gas chromatograph peak areas based on standard
120 curves with a CH_4 concentration of 0–100 ppm and a N_2O concentration of 0–5000 ppb.

121 Metagenome sequencing and processing of raw data

122 Total DNA and RNA were co-extracted as previously described [32]. Briefly, extraction was
123 performed on 0.5 g of soil using a hexadecyltrimethyl ammonium bromide (CTAB), phenol-
124 chloroform, and bead-beating protocol. DNA was purified using the AllPrep DNA Mini Kit
125 (QIAGEN, Hilden, Germany) and quantified using the Qubit dsDNA BR Assay Kit
126 (ThermoFisher Scientific, Waltham, MA, USA). Library preparation for Illumina metagenome
127 sequencing was performed using the Nextera XT DNA Library Preparation Kit (Illumina, San
128 Diego, CA, USA). Metagenomes were obtained for 69 samples across two paired-end NextSeq

129 (132–170 bp) and one NovaSeq (2 x 151 bp) runs **(Suppl. Table S1)**. Two samples were
130 additionally sequenced with Nanopore MinION. For this, libraries were prepared using the
131 SQK-LSK109 Ligation Sequencing Kit with the long fragment buffer (Oxford Nanopore
132 Technologies, Oxford, UK) and the NEBNext Companion Module for Oxford Nanopore
133 Technologies Ligation Sequencing Kit (New England Biolabs). Each sample was sequenced for
134 48 hours on one R9.4 flow cell.

135 The quality of the raw Illumina data was verified with fastQC v0.11.9 [36] and multiQC v1.8
136 [37]. Cutadapt v1.16 [38] was then used to trim 3' adapters and low-quality base calls (q < 20)
137 and to filter out short reads (< 50 bp). Nanopore data were basecalled with GPU guppy v4.0.11
138 using the high-accuracy model and applying a minimum quality score of 7. The quality of the
139 basecalled Nanopore data was assessed with pycoQC v2.5.0.21 [39] and adapters were trimmed
140 with Porechop v0.2.4 [40].

141 Read-based analyses

142 Taxonomic and functional profiles of the microbial communities were obtained using a read-
143 based approach, i.e. based on unassembled Illumina data. Due to differences in sequencing
144 depth across the samples, the dataset was resampled to 2,000,000 reads per sample with seqtk
145 v1.3 [41]. Taxonomic profiles were obtained by annotating SSU rRNA gene sequences against
146 the SILVA database release 111 [42] with METAXA v2.2 [43]. Functional profiles were obtained
147 by blastx searches against the KEGG database release 86 [44] with DIAMOND v0.9.14 [45].
148 Only matches with a maximum e-value of 10^{-5} and minimum identity of 60% were considered.
149 The KEGG Orthology (KO) identifier of the best hit was assigned to each read and KEGG
150 modules were summarised with the package keggR v0.9.1 [46] in R v3.6.3 [47]. Differences in
151 taxonomic and functional community structure were assessed using non-metric
152 multidimensional scaling (NMDS) and permutational ANOVA (PERMANOVA) with the
153 package vegan v2.5.6 [48] in R v3.6.3 [47] (functions *metaMDS* and *adonis*, respectively).
154 Differences in richness (i.e. the number of genera and functional genes) between the ecosystems
155 were assessed using one-way analysis of variance (ANOVA) followed by Tukey's HSD test with
156 the *lm*, *aov*, and *TukeyHSD* functions in R v3.6.3 [47]. The same tests were performed to assess
157 differences in the abundance of individual genes across soil ecosystems, with p-values adjusted
158 to control for false discovery rates. Relationships between the abundance of denitrification genes
159 and N_2O flux rates were assessed using linear regression in R v3.6.3[47] (functions *lm* and *aov*).

Metagenome assembling and binning

Metagenome assembling of the Illumina data was performed as two co-assemblies. One co-assembly comprised the upland soils (barren, heathland, and meadow; n = 47) and the other the fen samples (n = 22). For each co-assembly, reads from the respective samples were pooled and assembled with MEGAHIT v1.1.1.2 [49]. Assembling of the Nanopore data was done for each sample individually with metaFlye v2.7.1 [50], and contigs were corrected based on Illumina data from the respective sample with bowtie v2.3.5 [51], SAMtools v1.9 [52], and pilon v1.23 [53]. Quality assessment of the (co-)assemblies was obtained with metaQUAST v5.0.2 [54].

Binning of metagenome-assembled genomes (MAGs) was performed separately for each Illumina and Nanopore (co-)assembly with anvi'o v6.2 [55] after discarding contigs shorter than 2500 bp. Gene calls were predicted with prodigal v2.6.3 [56]. Single-copy genes were identified with HMMER v.3.2.1 [57] and classified with DIAMOND v0.9.14 [45] against the Genome Taxonomy Database (GTDB) release 04-RS89 [58, 59]. Illumina reads were mapped to the contigs with bowtie v2.3.5 [51] and SAM files were sorted and indexed using SAMtools v1.9 [52]. Due to their large sizes, Illumina co-assemblies were split into 100 smaller clusters based on differential coverage and tetranucleotide frequency with CONCOCT v1.0.0 [60]. Contigs were then manually sorted into bins based on the same composition and coverage metrics using the *anvi-interactive* interface in anvi'o v6.2 [55]. Nanopore contigs were binned directly without pre-clustering. Bins that were ≥ 50% complete according to the presence of single-copy genes were further refined using the *anvi-refine* interface in anvi'o v6.2 [55]. In addition to taxonomic signal (based on single-copy genes classified against GTDB), either differential coverage or tetranucleotide frequency was used to identify and remove outlying contigs. The former was used for bins with a large variation in contig coverage across samples, and the latter for those with marked differences in GC content across contigs. Bins ≥ 50% complete and ≤ 10% redundant – hereafter referred as MAGs – were kept for downstream analyses.

Phylogenomic analyses of MAGs and metabolic reconstruction

Phylogenetic placement of MAGs was done based on 122 archaeal and 120 bacterial single-copy genes with GTDB-Tk v1.3.0 [61] and the GTDB release 05-RS95 [58, 59]. Acidobacteriota MAGs containing denitrification genes were submitted to further phylogenomic analyses alongside all genomes of Acidobacteriota strains and candidate taxa available on GenBank (n = 69; accessed on 9 October 2020). For this, the amino acid sequence of 23 ribosomal proteins was retrieved for each genome with anvi'o v6.2 [55] and aligned with MUSCLE v3.8.1551 [62]. A maximum

192 likelihood tree was then computed based on the concatenated alignments with FastTree v2.1.11
193 [63]. *Escherichia coli* ATCC 11775 was used to root the tree.

194 For metabolic reconstruction, MAGs were annotated against the KOfam hidden Markov model
195 (HMM) database [64] with HMMER v.3.2.1 [57] using the pre-computed score thresholds of each
196 HMM profile. The *anvi-estimate-metabolism* program in anvi'o v6.2 [55] was then used to
197 predict the metabolic capabilities of the MAGs. A metabolic pathway was considered present in
198 MAGs containing at least 75% of the genes involved in the pathway. Carbohydrate-active
199 enzymes (CAZymes) were annotated with dbCAN v.2.0 based on the dbCAN v7 HMM database
200 [65]. Only hits with an e-value < 1×10^{-14} and coverage > 0.35 were considered.

201 Annotation of genes involved in denitrification

202 MAGs were screened for genes involved in denitrification (*nirK*, *nirS*, *norB*, and *nosZ*) using a
203 three-step approach. First, predicted amino acid sequences were searched against the KOfam
204 HMM database with KofamScan v1.3.0 [64]. Only matches with scores above the pre-computed
205 family-specific thresholds were kept. Sequences were then inspected for the presence of
206 conserved residues at the following positions: *nirK*, Cu-binding and active sites [66]; *nirS*,
207 c-heme and d_1-heme binding sites [67]; *norB*, binding of the catalytic centres cyt b, b_3, and Fe_b
208 [68]; *nosZ*: binding of the Cu_Z and Cu_A centres [68]. For this, amino acid sequences were aligned
209 with MAFFT v7.429 [69] and alignments were visualized with Unipro UGENE v38.1 [70].
210 Sequences which did not contain the correct amino acid at the aforementioned positions were
211 removed. Finally, resulting amino acid sequences were aligned with MAFFT v7.429 [69] along
212 with reference sequences from Graf et al. [14] and a maximum-likelihood tree was computed
213 with FastTree v2.1.11 [63] using the LG+GAMMA model. Annotation of denitrification genes
214 was also performed for previously published genomes retrieved from GenBank. These included
215 a set of 1529 metagenome-assembled genomes (MAGs) obtained from soils in Stordalen Mire,
216 northern Sweden [28] and the GenBank Acidobacteriota genomes described above.

217 MAG dereplication and read recruitment analysis

218 Prior to read recruitment analyses, Illumina and Nanopore MAGs were dereplicated based on
219 a 90% average nucleotide identity (ANI) threshold with fastANI v1.3 [71] to remove redundancy
220 (i.e. MAGs that were recovered multiple times across the different assemblies). Read
221 recruitment analyses were then then performed with CoverM [72]. For this, Illumina reads were
222 mapped to the set of non-redundant MAGs with minimap v2.17 [73] and relative abundances
223 were calculated as a proportion of the reads mapping to each MAG.

224 # Results

225 ## Environmental characterization and *in situ* GHG fluxes

226 Our sampling design in Kilpisjärvi included two soil depths across four ecosystems that are
227 characteristic of polar and alpine environments (barren soils, heathlands, meadows, and fens)
228 **(Suppl. Figure S1a,b,c)**. Soil ecosystems differed in vegetation cover and physicochemical
229 composition, with fens in particular being characterized by higher pH, moisture, and N content
230 (one-way ANOVA, $R^2 = 0.16–0.64$, $p < 0.001$) and, together with the meadows, lower C:N ratio
231 (one-way ANOVA, $R^2 = 0.49$, $p < 0.001$) **(Suppl. Fig. S1d)**. *In situ* measurements of GHG fluxes
232 showed a high sink-source variability in net N_2O fluxes across the ecosystems **(Suppl. Fig.**
233 **S1e)**. Although the average N_2O flux across all sites was small (net consumption of
234 6 µg N_2O m^{-2} day^{-1}), high N_2O emission at rates of up to 660 µg N_2O m^{-2} day^{-1} was observed at
235 some sites. Net CH_4 emissions were observed exclusively at the fen sites **(Suppl. Fig. S1e)**.

236 ## Differences in microbial community structure across soils ecosystems

237 We obtained more than 9 billion Illumina (1.4 Tb) and 7 million Nanopore (21.5 Gb) reads from
238 the 69 soil metagenomes (mean: 19.9 Gb, minimum: 0.7 Gb, maximum: 82.9 Gb) **(Suppl. Table**
239 **S1)**. Read-based analyses of unassembled SSU rRNA gene sequences showed that microbial
240 community composition differed across the ecosystems, with fen soils harbouring contrasting
241 microbial communities compared to the other ecosystems (PERMANOVA, $R^2 = 0.24–0.51$,
242 $p < 0.001$) **(Suppl. Fig. S2a,b)**. No differences in community structure were observed between
243 soil depths or the interaction between soil ecosystem and depth (PERMANOVA, $p > 0.05$).

244 Among previously described (i.e. not unclassified) taxa, microbial communities in barren,
245 heathland, and meadow soils were dominated by aerobic and facultative anaerobic heterotrophs
246 such as *Acidobacterium*, *Actinoallomurus*, *Conexibacter*, *Mycobacterium*, *Bradyrhizobium*,
247 *Rhodoplanes*, *Ktedonobacter*, *Gemmata*, and *Chthoniobacter* **(Suppl. Fig. S2c)**. On the other
248 hand, fen soils were dominated by methanogenic archaea from the genera *Methanobacterium*
249 and *Methanosarcina* and anaerobic bacteria such as *Geobacter*, but also the putative aerobic
250 heterotroph *Candidatus* Koribacter. Communities from different ecosystems also differed in the
251 abundance of denitrification genes (*nirK*, *nirS*, *norB*, and *nosZ*), which were more abundant in
252 the meadows and fens (one-way ANOVA, $R^2 = 0.47–0.72$, $p < 0.001$) **(Suppl. Fig. S2d)**. Fen soils
253 also had a higher abundance of genes involved in sulfate reduction (*dsrA* and *dsrB*) and
254 methanogenesis (*mcrA* and *mcrB)* (one-way ANOVA, $R^2 = 0.61–0.89$, $p < 0.001$). We did not
255 observe a significant relationship between N_2O flux rates and neither the abundance of

individual N cycle genes nor the ratio between *nirK*/*nirS* and *nosZ* abundances (linear regression, $p > 0.05$).

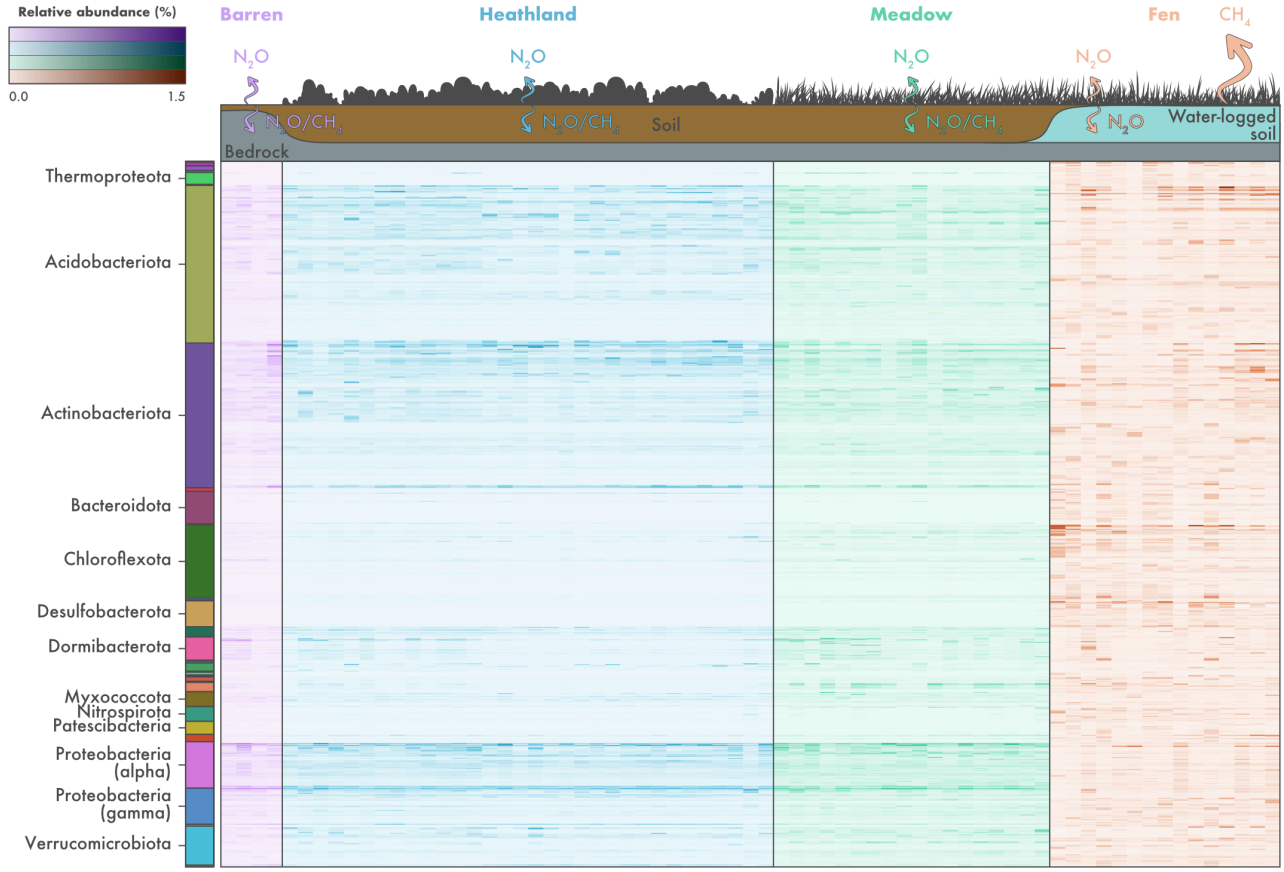
Fig. 1 | Genome-resolved metagenomics of tundra soils. Phylogenomic placement and assembly statistics of 796 metagenome-assembled genomes (MAGs) recovered from soils in Kilpisjärvi, northern Finland. Unrooted maximum likelihood tree based on concatenated alignments of amino acid sequences from 122 archaeal and 120 bacterial single-copy genes. Additional information about the MAGs can be found in **Suppl. Table S2**.

A manually curated genomic database from tundra soil metagenomes

Two Illumina co-assemblies and two individual Nanopore assemblies yielded more than 4 million contigs longer than 2,500 bp, with a total assembly size of 21.1 Gb. The co-assemblies covered a significant fraction of the original metagenomic data, with an average read recruitment rate of 54.6% across samples (minimum: 22.9%, maximum: 75.8%). Using anvi'o [55], we obtained 8,043 genomic bins and manually curated these to a set of 796 MAGs that were at least 50% complete and no more than 10% redundant **(Fig. 1, Suppl. Table S2)**. According to estimates based on domain-specific single-copy genes, the obtained MAGs were on average 65.4% complete (minimum: 50.0%, maximum: 100.0%) and 2.7% redundant (minimum: 0.0%, maximum: 9.9%) **(Suppl. Table S2)**. Phylogenomic analyses based on 122 archaeal and 120 bacterial single-copy genes placed the MAGs across 35 bacterial and archaeal phyla according to the GTDB classification [58, 59] **(Fig. 1).** The most represented phyla were Acidobacteriota (n = 172), Actinobacteriota (n = 163), Proteobacteria (Alphaproteobacteria, n = 54; Gammaproteobacteria, n = 39), Chloroflexota (n = 84), and Verrucomicrobiota (n = 43). Most MAGs (n = 703) belonged to genera that do not comprise formally described species, including 303 MAGs that were placed outside genus-level lineages currently described in GTDB and thus likely represent novel genera **(Suppl. Table S2)**.

To investigate their distribution across the different soil ecosystems, MAGs were dereplicated based on a 90% ANI threshold, yielding a set of 761 non-redundant MAGs **(Fig. 2)**. On average, 15.8% of the reads from each sample were recruited by the set of non-redundant MAGs (minimum: 7.6%, maximum: 30.5%). In agreement with the read-based assessment, we observed differences in MAG composition across the soil ecosystems, with only 50 MAGs shared between the heathland, meadow, and fen soils **(Suppl. Fig. S3a)**. Fen soils harboured the highest number of MAGs, with an average of 155 MAGs per sample **(Suppl. Fig. S3b)**. Although barren and fen soils had similar taxonomic richness according to the read-based estimates, only a small number of MAGs was detected in the barren soils (average of four MAGs per sample). This is likely a result of limited sampling and sequencing of this ecosystem, which consisted of four samples and a total of 7.9 Gb of metagenomic data **(Suppl. Table S1)**. The number of MAGs in heathland and meadow soils was similar (average of 47 and 63 MAGs per sample, respectively) **(Suppl. Fig. S3b)**. In general, barren, heathland, and meadow soils were dominated by the same set of MAGs **(Suppl. Fig. S3c**). These included members of the Acidobacteriota (*Sulfotelmatobacter* and unclassified genera in the class Acidobacteriae), Actinobacteriota (*Mycobacterium* and unclassified genera in the family Streptosporangiaceae), and Proteobacteria (Alphaproteobacteria: *Reyranella*, *Bradyrhizobium*, and unclassified

297 Xanthobacteraceae; Gammaproteobacteria: unclassified Steroidobacteraceae). On the other
298 hand, fen soils were dominated by MAGs that were not assigned to formally described genera,
299 including lineages of Acidobacteriota (family Koribacteraceae), Actinobacteriota (family
300 Solirubrobacteraceae), Chloroflexota (class Ellin6529), Desulfobacterota (order
301 Desulfobaccales), and Halobacterota.



302 **Fig. 2 | Microbial community composition across different soil ecosystems in the**
303 **tundra.** Relative abundance of 761 non-redundant metagenome-assembled genomes (MAGs)
304 recovered from soils in Kilpisjärvi, northern Finland. Relative abundances were computed as a
305 proportion of the reads mapping to each MAG. Phylum-level taxonomic assignments are shown
306 for the major groups found. The scheme on the top of the figure represents ecosystem-level
307 methane (CH_4) and nitrous oxide (N_2O) fluxes based on *in situ* measurements **(Suppl. Fig. S1)**.

308 **Microorganisms from tundra soils have truncated denitrification pathways**

309 To gain insights into the microorganisms involved with the cycling of N_2O in tundra soils, we
310 searched the MAGs for genes involved in denitrification, a key process controlling N_2O
311 production in soils [11]. We focused on the *nirK*/*nirS*, *norB*, and *nosZ* genes which encode the
312 nitrite (Nir), nitric oxide (Nor), and nitrous oxide (Nos) reductases, respectively [12]. Gene

313 annotation was performed using a three-step approach to avoid false positives consisting of
314 distantly related homologues that are not involved in denitrification. Putative genes were first
315 identified by searching predicted amino acid sequences against curated HMMs from the KOfam
316 database [64]. Positive matches were then aligned and manually inspected for the presence of
317 conserved residues at specific positions associated with the binding of co-factors and active sites
318 [66–68]. Sequences containing the correct amino acid at these core positions were further
319 submitted to phylogenetic analyses along with sequences from a comprehensive database of
320 archaeal and bacterial genomes [14] **(Suppl. Fig. S4, Suppl. Fig. S5, Suppl. Fig. S6, Suppl.**
321 **Fig. S7)**.

322 We identified denitrification genes in 110 of the 796 MAGs (13.8%) **(Suppl. Table S2)**. These
323 were affiliated with the archaeal phylum Thermoproteota and many bacterial phyla such as
324 Proteobacteria (Gamma- and Alphaproteobacteria), Acidobacteriota, Bacteroidota,
325 Actinobacteriota, Chloroflexota, and Verrucomicrobiota **(Fig. 3a)**. However, only 17 MAGs were
326 assigned to a validly described genera **(Suppl. Table S2)**. These included members of the
327 Acidobacteriota (*Solibacter*, *Sulfotelmatobacter*, *Terracidiphilus*, and *Gaiella*), Myxococcota
328 (*Anaeromyxobacter*), Planctomycetota (*Singulisphaera*), Proteobacteria (Alphaproteobacteria:
329 *Bauldia*, *Bradyrhizobium*, *Methylocella*, and *Reyranella*; Gammaproteobacteria: *Gallionella*
330 and *Rhizobacter*), and Verrucomicrobiota (*Lacunisphaera* and *Opitutus*). On average, 1.8% of
331 the reads in each sample were recruited by all denitrifiers combined (minimum: 0.4%,
332 maximum: 6.1%). In general, denitrifiers were most abundant in the fens (1.0–6.1%) and least
333 abundant in the heathlands (0.4–2.1%). When taking into account only reads mapping to the
334 796 MAGs, denitrifiers represented on average 11.1% of the reads in each sample
335 (minimum: 3.0%, maximum: 26.8%).

336 Genes involved in denitrification were found exclusively in MAGs with truncated denitrification
337 pathways, i.e. MAGs missing one or more genes involved in the complete denitrification process
338 **(Fig. 3a)**. Of the 110 MAGs harbouring denitrification genes, the vast majority (n = 104)
339 encoded only one of the Nir, Nor, and Nos enzymes and no MAG encoded all the three enzymes
340 required for the reduction of NO_2^- to N_2. Unsurprisingly, co-occurrence of genes encoding the
341 three enzymes was also not observed in any of the other genomic bins of lower quality that were
342 discarded from the final MAG dataset (i.e. bins that were < 50% complete and/or > 10%
343 redundant). To verify if microorganisms with truncated denitrification pathways are common
344 in other tundra systems, we expanded our analysis to 1529 MAGs recovered from permafrost
345 peatland, bog, and fen soils in Stordalen Mire, northern Sweden [28]. Among these, 225 MAGs
346 (14.7%) contained denitrification genes **(Suppl. Fig. S8)**. MAGs encompassed a similar
347 taxonomic profile as observed in the Kilpisjärvi dataset, and MAGs with truncated

denitrification pathways were also the norm in Stordalen Mire soils. Only one MAG, assigned to the Gammaproteobacteria genus *Janthinobacterium*, encoded all the Nir, Nor, and Nos enzymes required for the reduction of NO_2^- to N_2.

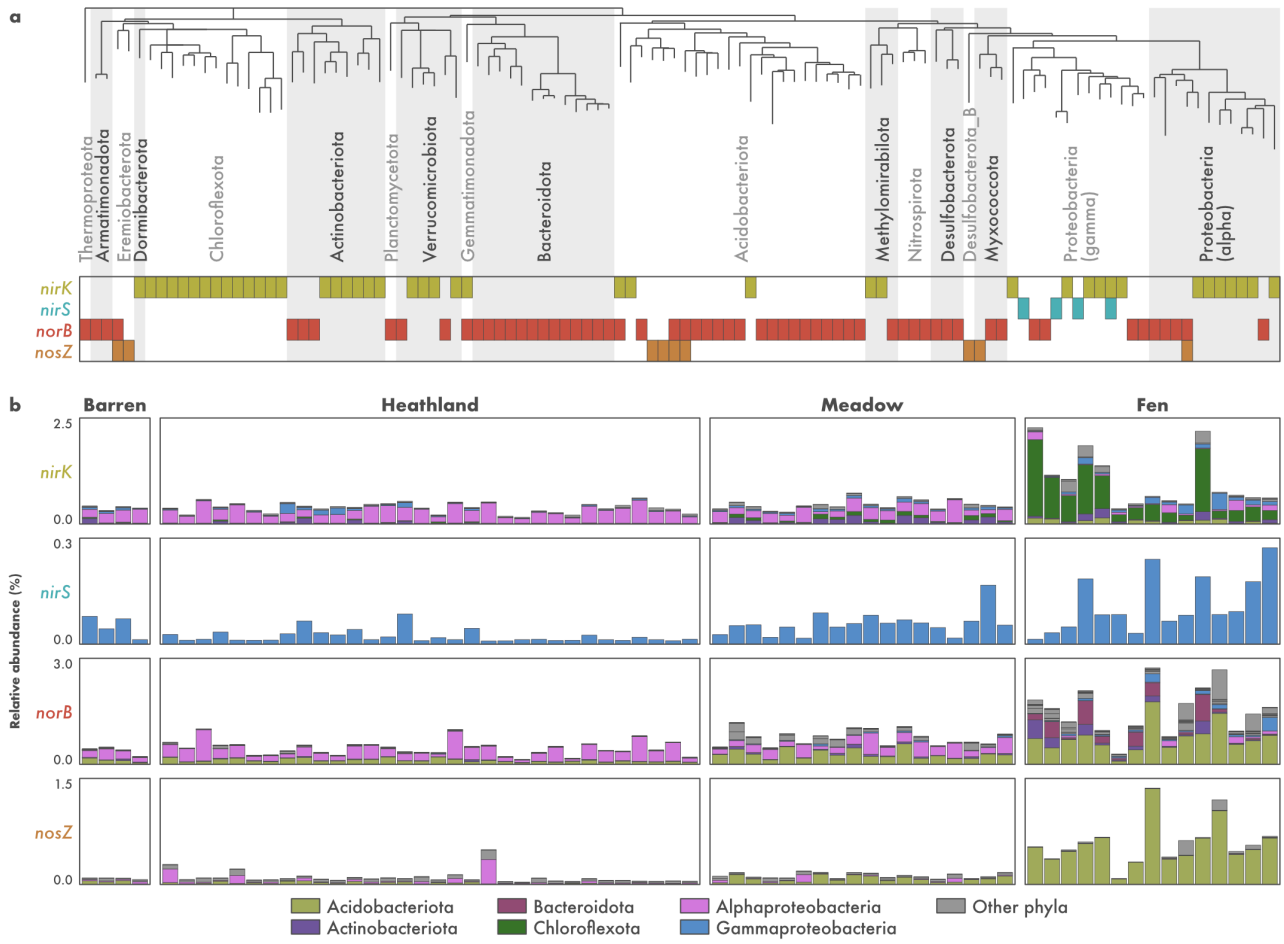


Fig. 3| Metabolic potential for denitrification in tundra soils. a) Distribution of denitrification genes across 110 metagenome-assembled genomes (MAGs) recovered from tundra soils in Kilpisjärvi, northern Finland. Genes encoding the nitrite (*nirK*/*nirS*), nitric oxide (*norB*), and nitrous oxide (*nosZ*) reductases were annotated using a three-step approach including (1) identification using hidden Markov models from the KOfam database, (2) manual inspection for the presence of conserved residues at positions associated with the binding of co-factors and active sites, and (3) phylogenetic analyses along with sequences from archaeal and bacterial genomes **(Suppl. Fig. S4–S7)**. **b)** Phylum-level relative abundance of microorganisms harbouring denitrification genes across the different soil ecosystems, computed as a proportion of reads mapping to each MAG.

361 **Microorganisms affiliated with the Chloroflexota lineage Ellin6529 are the**
362 **main denitrifiers *stricto sensu* in fen soils**

363 The reduction of NO_2^- to NO, performed by microorganisms harbouring the *nirK* or *nirS* genes,
364 is the hallmark step of denitrification and is often referred to as denitrification *stricto sensu* as
365 it involves the conversion of a soluble substrate to a gaseous product thus leading to the removal
366 of N from the system [12]. Within the 110 Kilpisjärvi MAGs harbouring genes involved in
367 denitrification, 46 contained *nirK*/*nirS* genes and are thus potential denitrifiers *stricto sensu*
368 **(Fig. 3a)**. These belonged mainly to the bacterial phyla Chloroflexota, Actinobacteriota, and
369 Proteobacteria (Alpha- and Gammaproteobacteria). Most MAGs (n = 43) contained the *nirK*
370 gene, which encodes the copper-containing form of Nir **(Suppl. Fig. S4)**. The *nirS* gene encoding
371 the cytochrome cd_1-containing form of Nir was present in four Gammaproteobacteria MAGs
372 **(Suppl. Fig. S5)**, including one MAG that contained both genes.

373 The composition of potential denitrifier *stricto sensu* communities differed across the ecosystems
374 **(Fig. 3b)**. MAGs belonging to the Alphaproteobacteria class of the Proteobacteria were the most
375 abundant in the barren, heathland, and meadow soils, particularly the MAG KUL-0154
376 assigned to the genera *Bradyrhizobium* **(Suppl. Fig. S9)**. Two other Alphaproteobacteria
377 MAGs that do not correspond to formally described genera in the families Acetobacteraceae and
378 Beijerinckiaceae (KUL-0057 and KUL-0056, respectively) were also found at lower relative
379 abundances. In addition, one Actinobacteriota MAG assigned to an uncharacterized genus in
380 the family Gaiellaceae (KWL-0073), was abundant in the meadow soils. On the other hand, fen
381 communities were dominated by MAGs belonging to the phylum Chloroflexota **(Fig. 3b)**, which
382 included seven MAGs assigned to the class-level lineage Ellin6529 **(Suppl. Fig. S9)**.

383 None of the Ellin6529 MAGs contained the key genes involved in autotrophic carbon fixation,
384 dissimilatory sulfate reduction, dissimilatory nitrate reduction to ammonia, and nitrogen
385 fixation **(Suppl. Table S3)**. Analysis of genes encoding terminal oxidases involved in the
386 aerobic respiratory electron chain revealed that all seven Ellin6529 MAGs harboured the
387 *coxABC* genes encoding the *aa3*-type cytochrome c oxidase**.** Four MAGs also contained the
388 *cydAB* genes encoding the cytochrome bd ubiquinol oxidase, a terminal oxidase with high
389 affinity for oxygen that also plays a role in protecting against harmful conditions such as
390 oxidative and nitrosative stress [74, 75]. The dominant MAGs in the barren, heathland, and
391 meadow soils contained a different set of genes encoding aerobic terminal oxidases. In addition
392 to the *cydAB* genes, the MAGs KUL-0057 and KUL-0154 also contained the *cyoABCD* genes
393 encoding the cytochrome o ubiquinol oxidase, which is the main terminal oxidase under highly
394 aerobic conditions [76], and KUL-0057 also contained genes encoding the *cbb3*-type cytochrome

395 c oxidase, a terminal oxidase with high affinity for oxygen [77]. Genes involved in the Calvin
396 cycle (e.g. *rbcL*, *rbcS*, and *prkB*) were found in the *Bradyrhizobium* MAG (KUL-0154), and none
397 of the key genes for autotrophic carbon fixation pathways were present in the other dominant
398 Alphaproteobacteria MAGs.

399 Acidobacteriota with the potential to reduce NO and N_2O are abundant in the fens

400 The stepwise reduction of NO to N_2O and N_2 carried out by microorganisms containing the *norB*
401 and *nosZ* genes, respectively, represents the final step of denitrification and the main biotic
402 control on N_2O emissions. Soil denitrification rates depend on multiple environmental
403 conditions such as adequate moisture and inorganic N availability, but whether it results in the
404 emission of N_2O or N_2 is ultimately linked to a balance between the activity of NO and N_2O
405 reducers [11, 15]. The *norB* and *nosZ* genes were identified in 62 and 9 Kilpisjärvi MAGs,
406 respectively, belonging mostly to the phyla Actinobacteriota, Bacteroidota, Acidobacteriota, and
407 Proteobacteria (class Alphaproteobacteria) **(Fig. 3a)**. With the exception of one
408 Gemmatimonadota and one Acidobacteriota MAG, *norB*- and *nosZ*-containing MAGs were
409 almost exclusively non-denitrifiers *stricto sensu*, i.e. they did not harboured the *nirK*/*nirS* genes
410 involved in the reduction of NO_2^- to NO. Most MAGs (n = 48) harboured a *norB* gene encoding
411 the monomeric, quinol-dependent form of Nor (qNor), while the remaining MAGs (n = 8) encoded
412 the cytochrome c-dependent Nor (cNor) **(Suppl. Fig. S6)**. In regards to the *nosZ* gene, most
413 MAGs (n = 6) contained sequences affiliated with the cluster II (also known as atypical) NosZ
414 [14, 15, 78] **(Suppl. Fig. S7)**. Only four MAGs contained both the *norB* and *nosZ* genes and
415 thus have the potential to reduce NO completely to N_2 **(Fig. 3a)**.

416 As observed for the denitrifier *stricto sensu* communities, the communities of potential NO and
417 N_2O reducers also differed between the ecosystems **(Fig. 3b)**. MAGs assigned to the
418 Alphaproteobacteria class of the Proteobacteria were the most abundant in the barren,
419 heathland, and meadow soils. In particular, the MAG KWL-0112 assigned to the genera
420 *Reyranella* was the dominant *norB*-containing MAG, while KUL-0116 (belonging to an
421 uncharacterized genus in the family Acetobacteraceae) was the dominant MAG harbouring the
422 *nosZ* gene **(Suppl. Fig. S9**). On the other hand, fen communities were dominated by
423 Acidobacteriota MAGs **(Fig. 3b)**, particularly the *norB*- and *nosZ*-containing MAG KWL-0326
424 affiliated with the class Thermoanaerobaculia **(Suppl. Fig. S9)**. This MAG contained the same
425 set of genes encoding aerobic terminal oxidases as found in the *nirK*-containing Ellin6529 MAGs
426 that were dominant in the fen sites, namely *coxABC* and *cydAB* **(Suppl. Table S3)**. Likewise,
427 the two MAGs that were dominant in the barren, heathland, and meadow soils (KWL-0112 and
428 KUL-0116) contained genes encoding the *cbb3*-type cytochrome c oxidase, which was also found

429 in one of the dominants *nirK*-containing Alphaproteobacteria MAG. No genes involved in carbon
430 fixation, dissimilatory sulfate reduction, dissimilatory nitrate reduction to ammonia, and
431 nitrogen fixation were found in any of the dominant *norB*- and *nosZ*-containing MAGs.

432 To elucidate the phylogenetic placement of the Acidobacteriota MAGs and to verify if the
433 potential for NO and N_2O reduction is present in other members of this phylum, we included in
434 our analysis all available genomes of Acidobacteriota strains and candidate taxa available on
435 GenBank. This revealed that genes encoding the Nir and Nos enzymes are widespread across
436 the phylum Acidobacteriota **(Fig. 4)**. Genes encoding the Nor enzyme were present in all but
437 one of the six Acidobacteriota subdivisions with genomes from cultured representatives. This
438 included the strains *Acidobacterium ailaaui* PMMR2 (subdivision Gp1), *Acidipila* sp. 4G-K13
439 (Gp1), *Silvibacterium bohemicum* DSM 103733 and *S. bohemicum* S15 (Gp1), Acidobacteriaceae
440 bacterium URHE0068 (Gp1), *Edaphobacter aggregans* DSM 19364 (Gp1), *Luteitalea pratensis*
441 DSM 100886 (Gp6), *Geothrix fermentans* DSM 14018 (Gp8), and *Thermoanaerobaculum*
442 *aquaticum* MP-01 (Gp23), as well as the candidate taxa *Candidatus* Koribacter versatilis
443 Ellin345 (Gp1), *Candidatus* Sulfotelmatomonas gaucii SbA5 (Gp1), and *Candidatus* Solibacter
444 usitatus Ellin6076 (Gp3). On the other hand, genes encoding the Nos enzyme were found only
445 in members of the subdivisions Gp6 and Gp23.

446 Discussion

447 The 796 MAGs obtained in the present study by a manual binning and curation effort represent
448 one of the largest genomic catalogues of microorganisms from tundra soils to date. Earlier gene-
449 centric investigations have demonstrated the potential for complete denitrification in tundra
450 soils [22, 79], however, these approaches fail to reveal the wider genomic context of the genes.
451 By applying the genome-resolved metagenomics approach, we traced denitrification genes to
452 specific microbial populations, thereby allowing a detailed investigation of the genomic makeup
453 of potential denitrifiers in tundra soils. This approach also enabled us to access the genomes of
454 uncultured, poorly characterized taxa, which comprise the majority of the microorganisms in
455 soils and other complex ecosystems [24, 25]. Our genome-resolved survey revealed that
456 denitrification in tundra soils is dominated by microorganisms with truncated denitrification
457 pathways, most of which represent poorly characterized taxa without cultured representatives.
458 With the support of *in situ* measurements of N_2O fluxes, we hypothesize that microorganisms
459 with truncated denitrification pathways are the main drivers of N_2O cycling in tundra soils.

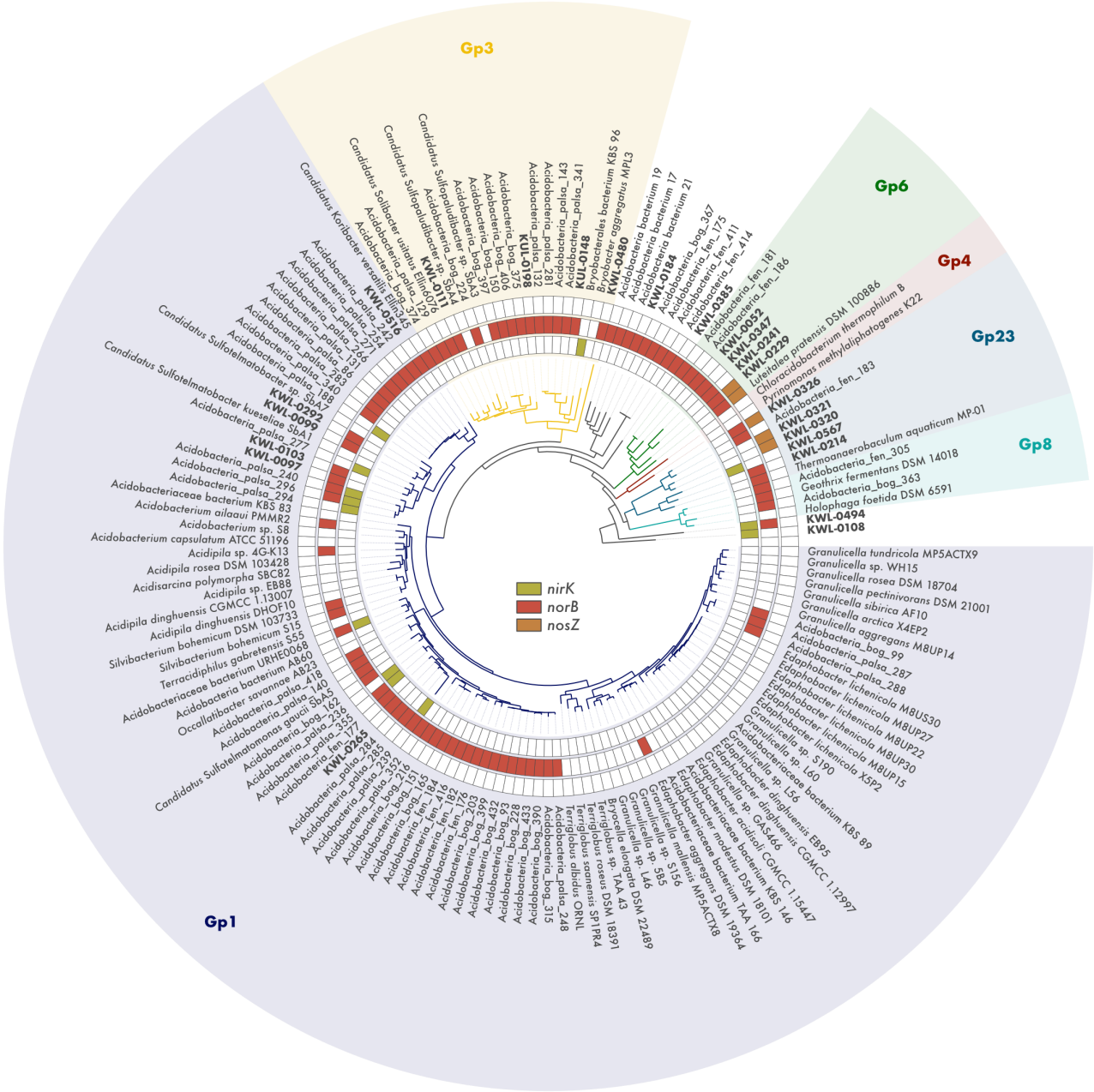


Fig. 4 | Metabolic potential for denitrification among members of the phylum Acidobacteriota. Phylogenomic analysis of 85 Acidobacteriota metagenome-assembled genomes (MAGs) containing denitrification genes recovered from tundra soils in Kilpisjärvi (northern Finland) and Stordalen Mire (northern Sweden), and 69 genomes of Acidobacteriota strains and candidate taxa. Maximum likelihood tree based on concatenated alignments of 23 ribosomal proteins and rooted with *Escherichia coli* ATCC 11775 (not shown). Genes encoding the nitrite (*nirK*), nitric oxide (*norB*), and nitrous oxide (*nosZ*) reductases were annotated using a three-step approach to avoid false positives (see methods).

468 The congruence of these findings in both our original analysis of northern Finland soils and our
469 re-analysis of a comprehensive metagenomic dataset from soils in northern Sweden [28]
470 suggests that truncated denitrification pathways are not a methodological artifact arising from
471 the metabolic reconstruction of fragmented genomes. Indeed, recent genome-resolved
472 investigations have shown that truncated metabolic pathways, also known as metabolic
473 handoffs, are the norm across a wide range of ecosystems such as grassland soil, aquifer
474 sediment, groundwater, and the ocean, and not only in relation to denitrification but other redox
475 transformations as well [80–82]. Although denitrification has traditionally been known as a
476 community effort performed by different microbial populations [12–15], these genome-resolved
477 metagenomic studies are beginning to reveal a more in-depth, ecosystem-centric representation
478 of the denitrification pathway. In addition to their predominance in genomic databases [14], it
479 appears that truncated denitrifiers are dominant within defined ecosystems across various
480 terrestrial and aquatic biomes, including the tundra. It has been suggested that the partitioning
481 of metabolic pathways across different populations via metabolic handoffs is advantageous as it
482 eliminates competition between enzymes accelerating substrate consumption [15, 83] and
483 provides flexibility and resilience to the communities in face of environmental disturbances [80].
484 We further hypothesize that the predominance of denitrification pathways characterized mostly
485 by metabolic handoffs in tundra soils could be related to N limitation. If metabolic handoffs
486 enable a more effective substrate consumption as previously suggested [15, 83], truncated
487 denitrification pathways would be favoured in tundra soils which are mostly N limited but
488 undergo rapid surges in N availability e.g. during the spring melting season [84].

489 Our results showed that tundra microbial communities differ between dryer upland ecosystems
490 (barren, heathland, and meadow soils) and the water-logged fens. This is likely explained by
491 large differences in soil moisture, which further affects oxygen availability in these ecosystems.
492 The dominant denitrifier populations in the mostly oxic dry upland soils, related to the genera
493 *Bradyrhizobium*, *Reyranella*, and other uncharacterized genera in the class
494 Alphaproteobacteria, encode aerobic terminal oxidases that are active under highly aerobic
495 conditions as well as oxidases with high oxygen affinity [76, 77]. This likely provides an adaptive
496 advantage in these soils by allowing rapid aerobic growth under standard conditions of high
497 oxygen availability as well as sustaining growth in microoxic niches within the soil matrix and
498 during periods of reduced oxygen availability (e.g. during the spring melting season). On the
499 other hand, fen soils are continuously inundated because the water table is at or near the soil
500 surface. The result is a mostly anoxic environment due to the slow rate at which oxygen diffuses
501 into the water-logged soil, favouring reduced rather than oxidized soil chemistry. In line with
502 this, we found a predominance of anaerobic processes in the fens, including a higher abundance

503 of genes involved in denitrification, sulfate reduction, and methanogenesis, the latter supported
504 by *in situ* observations of net CH_4 emission. Interestingly, the dominant populations in the fen
505 soils, assigned to the Chloroflexota lineage Ellin6529 and the subdivision Gp23 of the
506 Acidobacteriota, encode the cytochrome *bd* oxidase which plays a role in preventing the
507 inactivation of oxygen-sensitive enzymes and protecting against oxidative and nitrosative
508 stress, toxic compounds such as cyanide, and other stress conditions such as high temperature
509 and high pH [74, 75]. This cytochrome also has a high oxygen affinity, thus likely enabling these
510 microorganisms to grow aerobically later in the growing season when fen soils are drier and
511 thus less anoxic.

512 An interesting finding was that the communities of potential denitrifiers in the fen soils were
513 dominated by somewhat enigmatic taxa, namely potential NO_2^- reducers affiliated with the
514 class Ellin6529 of the Chloroflexota and NO/N_2O reducers assigned to the subdivision Gp23 of
515 the Acidobacteriota. Both groups are major members of microbial communities in soils
516 worldwide [85], and RNA-based investigations have shown that they are active in tundra soils
517 during both summer and winter seasons [32, 86]. Ellin6529 – formerly G04 – were first detected
518 by culture-independent methods in alpine tundra wet meadow soil in the Colorado Rocky
519 Mountains, USA [87], and later isolated in a study targeting slow-growing and mini-colony
520 forming bacteria from Australian agricultural soil [88]. However, their ecological, physiological,
521 and metabolic preferences remain largely unknown. *Thermoanaerobaculum aquaticum* MP-01,
522 the only cultivated member of the Acidobacteriota subdivision Gp23, is a strictly anaerobic
523 bacterium that has been shown to use Fe and Mn, but not NO_3^- nor NO_2^-, as electron acceptors
524 in anaerobic respiration [89]. However, studies investigating the use of nitrogen oxides in
525 anaerobic respiration usually provide soluble NO_3^- or NO_2^- as electron acceptors, not the gases
526 NO and N_2O, which bias against truncated denitrifiers that do not contain the *narG* and
527 *nirK*/*nirS* genes [90]. Their high abundance in the water-logged, anoxic fen soils suggest that
528 the Ellin6529 and Gp23 populations found in this study are able to grow anaerobically with the
529 use of NO and N_2O as electron acceptors. However, it is known that in addition to their role in
530 anaerobic respiration, NO and N_2O reduction can be used as a detoxification mechanism or even
531 as electron sink for metabolism. For example, the aerobe *Gemmatimonas aurantica* T-27 is not
532 able to grow on N_2O alone, but instead uses N_2O as electron acceptor transiently when oxygen
533 is depleted [91].

534 Despite differences in community structure, our results showed that both meadows and fens are
535 potential denitrification hotspots, with a higher abundance of denitrification genes than barren
536 and heathland soils. In addition, the low C:N ratio in the meadow and fen soils might indicate
537 higher rates of N mineralization and thus high N bioavailability, likely supporting

538 denitrification in these ecosystems [4, 92]. The largely anoxic conditions, high pH and SOM
539 content in fen soils, in particular, make them important sources of N_2O in the tundra [7].
540 Although the *in situ* N_2O fluxes measured in this study were highly variable across ecosystems,
541 most fen sites represented N_2O sources, and N_2O sinks in the fens were of smaller magnitude
542 compared to meadows and heathlands. However, no direct links could be observed between
543 denitrifier community composition and *in situ* N_2O fluxes. Indeed, modelling N_2O emissions
544 based on microbial community structure is challenging, as N_2O fluxes are characterized by a
545 high temporal and spatial variability driven by several environmental constraints related to soil
546 pH, N, moisture, and oxygen content [11]. Studies have suggested that incomplete denitrifiers
547 that contain Nir and Nor but lack Nos (thus having N_2O as the end product of denitrification)
548 can contribute substantially to soil N_2O emissions [93]. At the same time, non-denitrifying N_2O
549 reducers, i.e. microorganisms that contain Nos but lack Nir, can represent an important N_2O
550 sink [78, 94, 95]. However, while microorganisms with truncated denitrification pathways
551 appear to dominate the denitrifier communities investigated here and elsewhere [80–82], the
552 potential for complete denitrification is present at the ecosystem level. Our current knowledge
553 of the regulation of the denitrification process is largely based on the activity of model organisms
554 such as the complete denitrifier *Paracoccus denitrificans* [96], and so it remains unclear how a
555 community dominated by truncated denitrifiers interacts with each other and the environment
556 and impact N_2O emissions *in situ*.

557 Conclusions

558 A better understanding of denitrification is paramount for our ability to model N_2O emissions
559 and mitigate climate change [23, 96]. High-latitude environments in particular have
560 experienced amplified warming in recent decades, a trend that is likely to continue in the
561 coming centuries [10]. As mechanisms of GHG emissions are very climate sensitive, the
562 contribution of tundra soils to global GHG atmospheric levels is thus predicted to increase in
563 the future leading to a positive feedback loop [8, 9]. Compared with carbon dioxide and CH_4,
564 measurements of N_2O fluxes in tundra soils are sparse and are rarely coupled with a
565 characterization of the microorganisms involved, making the magnitude and drivers of N_2O
566 fluxes across the polar regions uncertain [7]. In addition to a better monitoring of N_2O emissions
567 throughout the tundra biome, our results suggest that a better understanding of the
568 contribution of tundra soil to global N_2O levels relies on the elucidation of the regulatory
569 mechanisms of metabolic handoffs in communities dominated by truncated denitrifiers.

Data availability

Raw metagenomic data and assembled MAGs have been submitted to the European Nucleotide Archive (ENA) under the project PRJEB41762.

Code availability

All the code used can be found in https://github.com/ArcticMicrobialEcology/Kilpisjarvi-MAGs.

Acknowledgements

This work was funded by the Academy of Finland (grants 314114 and 335354) and the University of Helsinki. SV was funded by the Microbiology and Biotechnology Doctoral Programme (MBDP). AMV was funded by the Academy of Finland (grant 286950), the Otto Malm Foundation, and the Gordon and Betty Moore Foundation (grant 8414). MEM was supported by the Academy of Finland (grants 314630 and 317054). We would like to acknowledge CSC – IT Centre for Science for providing the requisite computing resources and Kimmo Mattila for IT support; the staff from the Kilpisjärvi Biological Station, Tanja Orpana, Aino Rutanen, Anniina Sarekoski, Johanna Kerttula, and the members of the BioGeoClimate Modelling Lab for assistance with fieldwork and soil characterization; Jillian Banfield and Christina Biasi for helpful discussion; Laura Cappelatti for proof-reading the manuscript; and Murat Eren, Sebastian Lücker, Donovan Parks, and Antonios Kioukis for tips, recommendations, and troubleshooting.

Author contributions

JH and ML designed the research; SV and JH performed nucleic acid extraction and metagenomic library preparation; AMV and MEM designed and performed the GHG flux measurements and analyses; ISP analysed the data and wrote the manuscript; EER and TOD contributed with the analyses; all authors contributed to the final version of the manuscript.

Competing interests

The authors declare no competing interests.

References

1. IPCC (ed). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. 2013. Cambridge University Press, Cambridge.
2. Tian H, Xu R, Canadell JG, Thompson RL, Winiwarter W, Suntharalingam P, et al. A comprehensive quantification of global nitrous oxide sources and sinks. *Nature* 2020; **586**: 248–256.
3. Repo ME, Susiluoto S, Lind SE, Jokinen S, Elsakov V, Biasi C, et al. Large N_2O emissions from cryoturbated peat soil in tundra. *Nat Geosci* 2009; **2**: 189–192.
4. Marushchak ME, Pitkämäki A, Koponen H, Biasi C, Seppälä M, Martikainen PJ. Hot spots for nitrous oxide emissions found in different types of permafrost peatlands. *Glob Change Biol* 2011; **17**: 2601–2614.
5. Stewart KJ, Grogan P, Coxson DS, Siciliano SD. Topography as a key factor driving atmospheric nitrogen exchanges in arctic terrestrial ecosystems. *Soil Biol Biochem* 2014; **70**: 96–112.
6. Voigt C, Marushchak ME, Lamprecht RE, Jackowicz-Korczyński M, Lindgren A, Mastepanov M, et al. Increased nitrous oxide emissions from Arctic peatlands after permafrost thaw. *Proc Natl Acad Sci* 2017; **114**: 6238–6243.
7. Voigt C, Marushchak ME, Abbott BW, Biasi C, Elberling B, Siciliano SD, et al. Nitrous oxide emissions from permafrost-affected soils. *Nat Rev Earth Environ* 2020; **1**: 420–434.
8. Schuur EAG, McGuire AD, Schädel C, Grosse G, Harden JW, Hayes DJ, et al. Climate change and the permafrost carbon feedback. *Nature* 2015; **520**: 171–179.
9. Hugelius G, Loisel J, Chadburn S, Jackson RB, Jones M, MacDonald G, et al. Large stocks of peatland carbon and nitrogen are vulnerable to permafrost thaw. *Proc Natl Acad Sci* 2020; **117**: 20438–20446.
10. Post E, Alley RB, Christensen TR, Macias-Fauria M, Forbes BC, Gooseff MN, et al. The polar regions in a 2°C warmer world. *Sci Adv* 2019; **5**: eaaw9883.
11. Butterbach-Bahl K, Baggs EM, Dannenmann M, Kiese R, Zechmeister-Boltenstern S. Nitrous oxide emissions from soils: how well do we understand the processes and their controls? *Philos Trans R Soc B Biol Sci* 2013; **368**: 20130122.
12. Zumft WG. Cell biology and molecular basis of denitrification. *Microbiol Mol Biol Rev MMBR* 1997; **61**: 533–616.

13. Wallenstein MD, Myrold DD, Firestone M, Voytek M. Environmental controls on denitrifying communities and denitrification rates: insights from molecular methods. *Ecol Appl* 2006; **16**: 2143–2152.
14. Graf DRH, Jones CM, Hallin S. Intergenomic Comparisons Highlight Modularity of the Denitrification Pathway and Underpin the Importance of Community Structure for N_2O Emissions. *PLoS ONE* 2014; **9**: e114118.
15. Hallin S, Philippot L, Löffler FE, Sanford RA, Jones CM. Genomics and Ecology of Novel N_2O-Reducing Microorganisms. *Trends Microbiol* 2018; **26**: 43–55.
16. Liu X-Y, Koba K, Koyama LA, Hobbie SE, Weiss MS, Inagaki Y, et al. Nitrate is an important nitrogen source for Arctic tundra plants. *Proc Natl Acad Sci* 2018; **115**: 3398–3403.
17. Kou D, Yang G, Li F, Feng X, Zhang D, Mao C, et al. Progressive nitrogen limitation across the Tibetan alpine permafrost region. *Nat Commun* 2020; **11**: 3331.
18. Yergeau E, Kang S, He Z, Zhou J, Kowalchuk GA. Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitudinal transect. *ISME J* 2007; **1**: 163–179.
19. Yergeau E, Hogues H, Whyte LG, Greer CW. The functional potential of high Arctic permafrost revealed by metagenomic sequencing, qPCR and microarray analyses. *ISME J* 2010; **4**: 1206–1214.
20. Palmer K, Biasi C, Horn MA. Contrasting denitrifier communities relate to contrasting N_2O emission patterns from acidic peat soils in arctic tundra. *ISME J* 2012; **6**: 1058–1077.
21. Dai H-T, Zhu R-B, Sun B-W, Che C-S, Hou L-J. Effects of Sea Animal Activities on Tundra Soil Denitrification and *nirS*- and *nirK*-Encoding Denitrifier Community in Maritime Antarctica. *Front Microbiol* 2020; **11**: 573302.
22. Ortiz M, Bosch J, Coclet C, Johnson J, Lebre P, Salawu-Rotimi A, et al. Microbial Nitrogen Cycling in Antarctic Soils. *Microorganisms* 2020; **8**: 1442.
23. Yu T, Zhuang Q. Quantifying global N_2O emissions from natural ecosystem soils using trait-based biogeochemistry models. *Biogeosciences* 2019; **16**: 207–222.
24. Rappé MS, Giovannoni SJ. The Uncultured Microbial Majority. *Annu Rev Microbiol* 2003; **57**: 369–394.
25. Steen AD, Crits-Christoph A, Carini P, DeAngelis KM, Fierer N, Lloyd KG, et al. High proportions of bacteria and archaea across most biomes remain uncultured. *ISME J* 2019; **13**: 3126–3130.

26. Mackelprang R, Waldrop MP, DeAngelis KM, David MM, Chavarria KL, Blazewicz SJ, et al. Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. *Nature* 2011; **480**: 368–371.
27. Hultman J, Waldrop MP, Mackelprang R, David MM, McFarland J, Blazewicz SJ, et al. Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature* 2015; **521**: 208–212.
28. Woodcroft BJ, Singleton CM, Boyd JA, Evans PN, Emerson JB, Zayed AAF, et al. Genome-centric view of carbon processing in thawing permafrost. *Nature* 2018; **560**: 49–54.
29. Pirinen P, Simola H, Aalto J, Kaukoranta J-P, Karlsson P, Ruuhela R. Climatological statistics of Finland 1981–2010. 2012. Finnish Meteorological Institute, Helsinki.
30. le Roux PC, Aalto J, Luoto M. Soil moisture's underestimated role in climate change impact modelling in low-energy systems. *Glob Change Biol* 2013; **19**: 2965–2975.
31. Niittynen P, Heikkinen RK, Aalto J, Guisan A, Kemppinen J, Luoto M. Fine-scale tundra vegetation patterns are strongly related to winter thermal conditions. *Nat Clim Change* 2020.
32. Viitamäki S, Pessi IS, Virkkala A-M, Niittynen P, Kemppinen J, Eronen-Rasimus E, et al. The activity and functions of subarctic soil microbial communities vary across vegetation types. *bioRxiv* 2021.
33. le Roux PC, Pellissier L, Wisz MS, Luoto M. Incorporating dominant species as proxies for biotic interactions strengthens plant community models. *J Ecol* 2014; **102**: 767–775.
34. Kemppinen J, Niittynen P, Aalto J, le Roux PC, Luoto M. Water as a resource, stress and disturbance shaping tundra vegetation. *Oikos* 2019; **128**: 811–822.
35. Livingston GP, Hutchinson GL. Enclosure-based measurement of trace gas exchange: applications and sources of error. *Biogenic trace gases: Measuring emissions from soil and water*. Blackwell Science, Oxford, United Kingdom, pp 14–51.
36. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2019.
37. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 2016; **32**: 3047–3048.
38. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011; **17**: 10.
39. Leger A, Leonardi T. pycoQC, interactive quality control for Oxford Nanopore Sequencing. *J Open Source Softw* 2019; **4**: 1236.
40. Wick RR. Porechop. 2018.
41. Li H. seqtk. 2018.

694 42. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal
695 RNA gene database project: improved data processing and web-based tools. *Nucleic Acids*
696 *Res* 2012; **41**: D590–D596.
697 43. Bengtsson-Palme J, Hartmann M, Eriksson KM, Pal C, Thorell K, Larsson DGJ, et al.
698 METAXA 2: improved identification and taxonomic classification of small and large subunit
699 rRNA in metagenomic data. *Mol Ecol Resour* 2015; **15**: 1403–1414.
700 44. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*
701 2000; **28**: 27–30.
702 45. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat*
703 *Methods* 2015; **12**: 59–60.
704 46. Pessi IS. keggR: a tool to parse the results of BLAST/DIAMOND similarity searches made
705 against the KEGG GENES prokaryotes database. 2020.
706 47. R Core Team. R: A language and environment for statistical computing. 2020.
707 48. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. vegan:
708 Community Ecology Package. 2019.
709 49. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution
710 for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*
711 2015; **31**: 1674–1676.
712 50. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, et al. metaFlye:
713 scalable long-read metagenome assembly using repeat graphs. *Nat Methods* 2020.
714 51. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;
715 **9**: 357–359.
716 52. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence
717 Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.
718 53. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: An
719 Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly
720 Improvement. *PLoS ONE* 2014; **9**: e112963.
721 54. Mikheenko A, Saveliev V, Gurevich A. MetaQUAST: evaluation of metagenome assemblies.
722 *Bioinformatics* 2016; **32**: 1088–1090.
723 55. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, et al. Anvi'o: an advanced
724 analysis and visualization platform for 'omics data. *PeerJ* 2015; **3**: e1319.
725 56. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic
726 gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010;
727 **11**: 119.
728 57. Eddy SR. Accelerated Profile HMM Searches. *PLoS Comput Biol* 2011; **7**: e1002195.

729 58. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, et al. A
730 standardized bacterial taxonomy based on genome phylogeny substantially revises the tree
731 of life. *Nat Biotechnol* 2018; **36**: 996–1004.
732 59. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. A complete
733 domain-to-species taxonomy for Bacteria and Archaea. *Nat Biotechnol* 2020.
734 60. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning
735 metagenomic contigs by coverage and composition. *Nat Methods* 2014; **11**: 1144–1146.
736 61. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes
737 with the Genome Taxonomy Database. *Bioinformatics* 2019; btz848.
738 62. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
739 *Nucleic Acids Res* 2004; **32**: 1792–1797.
740 63. Price MN, Dehal PS, Arkin AP. FastTree 2 – Approximately Maximum-Likelihood Trees for
741 Large Alignments. *PLoS ONE* 2010; **5**: e9490.
742 64. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, et al. KofamKOALA:
743 KEGG Ortholog assignment based on profile HMM and adaptive score threshold.
744 *Bioinformatics* 2020; **36**: 2251–2252.
745 65. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, et al. dbCAN2: a meta server for
746 automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 2018; **46**: W95–
747 W101.
748 66. Decleyre H, Heylen K, Tytgat B, Willems A. Highly diverse *nirK* genes comprise two major
749 clades that harbour ammonium-producing denitrifiers. *BMC Genomics* 2016; **17**: 155.
750 67. Li Y, Bali S, Borg S, Katzmann E, Ferguson SJ, Schuler D. Cytochrome cd1 Nitrite
751 Reductase NirS Is Involved in Anaerobic Magnetite Biomineralization in Magnetospirillum
752 gryphiswaldense and Requires NirN for Proper d1 Heme Assembly. *J Bacteriol* 2013; **195**:
753 4297–4309.
754 68. Heylen K, Keltjens J. Redundancy and modularity in membrane-associated dissimilatory
755 nitrate reduction in *Bacillus*. *Front Microbiol* 2012; **3**.
756 69. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7:
757 Improvements in Performance and Usability. *Mol Biol Evol* 2013; **30**: 772–780.
758 70. Okonechnikov K, Golosova O, Fursov M. Unipro UGENE: a unified bioinformatics toolkit.
759 *Bioinformatics* 2012; **28**: 1166–1167.
760 71. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI
761 analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;
762 **9**: 5114.
763 72. Woodcroft BJ. CoverM.

764 73. Li H. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences.
765 *Bioinformatics* 2016; **32**: 2103–2110.
766 74. Borisov VB, Gennis RB, Hemp J, Verkhovsky MI. The cytochrome bd respiratory oxygen
767 reductases. *Biochim Biophys Acta BBA - Bioenerg* 2011; **1807**: 1398–1413.
768 75. Giuffrè A, Borisov VB, Arese M, Sarti P, Forte E. Cytochrome bd oxidase and bacterial
769 tolerance to oxidative and nitrosative stress. *Biochim Biophys Acta BBA - Bioenerg* 2014;
770 **1837**: 1178–1187.
771 76. Dinamarca MA, Ruiz-Manzano A, Rojo F. Inactivation of Cytochrome *o* Ubiquinol Oxidase
772 Relieves Catabolic Repression of the *Pseudomonas putida* GPo1 Alkane Degradation
773 Pathway. *J Bacteriol* 2002; **184**: 3785–3793.
774 77. Bueno E, Mesa S, Bedmar EJ, Richardson DJ, Delgado MJ. Bacterial Adaptation of
775 Respiration from Oxic to Microoxic and Anoxic Conditions: Redox Control. *Antioxid Redox*
776 *Signal* 2012; **16**: 819–852.
777 78. Sanford RA, Wagner DD, Wu Q, Chee-Sanford JC, Thomas SH, Cruz-Garcia C, et al.
778 Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils.
779 *Proc Natl Acad Sci* 2012; **109**: 19709–19714.
780 79. Makhalanyane TP, Van Goethem MW, Cowan DA. Microbial diversity and functional
781 capacity in polar soils. *Curr Opin Biotechnol* 2016; **38**: 159–166.
782 80. Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, et al. Thousands
783 of microbial genomes shed light on interconnected biogeochemical processes in an aquifer
784 system. *Nat Commun* 2016; **7**: 13219.
785 81. Diamond S, Andeer PF, Li Z, Crits-Christoph A, Burstein D, Anantharaman K, et al.
786 Mediterranean grassland soil C–N compound turnover is dependent on rainfall and depth,
787 and is mediated by genomically divergent microorganisms. *Nat Microbiol* 2019; **4**: 1356–
788 1367.
789 82. Sun X, Ward BB. Novel metagenome-assembled genomes involved in the nitrogen cycle from
790 a Pacific oxygen minimum zone. *ISME Commun* 2021; **1**: 26.
791 83. Lilja EE, Johnson DR. Segregating metabolic processes into different microbial cells
792 accelerates the consumption of inhibitory substrates. *ISME J* 2016; **10**: 1568–1578.
793 84. Westergaard-Nielsen A, Balstrøm T, Treier UA, Normand S, Elberling B. Estimating
794 meltwater retention and associated nitrate redistribution during snowmelt in an Arctic
795 tundra landscape. *Environ Res Lett* 2020; **15**: 034025.
796 85. Delgado-Baquerizo M, Oliverio AM, Brewer TE, Benavent-González A, Eldridge DJ,
797 Bardgett RD, et al. A global atlas of the dominant bacteria found in soil. *Science* 2018; **359**:
798 320–325.

86. Männistö MK, Kurhela E, Tiirola M, Häggblom MM. Acidobacteria dominate the active bacterial communities of Arctic tundra with widely divergent winter-time snow accumulation and soil temperatures. *FEMS Microbiol Ecol* 2013; **84**: 47–59.

87. Costello EK, Schmidt SK. Microbial diversity in alpine tundra wet meadow soil: novel Chloroflexi from a cold, water-saturated environment. *Environ Microbiol* 2006; **8**: 1471–1486.

88. Davis KER, Sangwan P, Janssen PH. Acidobacteria, Rubrobacteridae and Chloroflexi are abundant among very slow-growing and mini-colony-forming soil bacteria. *Environ Microbiol* 2011; **13**: 798–805.

89. Losey NA, Stevenson BS, Busse H-J, Damsté JSS, Rijpstra WIC, Rudd S, et al. *Thermoanaerobaculum aquaticum* gen. nov., sp. nov., the first cultivated member of Acidobacteria subdivision 23, isolated from a hot spring. *Int J Syst Evol Microbiol* 2013; **63**: 4149–4157.

90. Lycus P, Lovise Bøthun K, Bergaust L, Peele Shapleigh J, Reier Bakken L, Frostegård Å. Phenotypic and genotypic richness of denitrifiers revealed by a novel isolation strategy. *ISME J* 2017; **11**: 2219–2232.

91. Park D, Kim H, Yoon S. Nitrous Oxide Reduction by an Obligate Aerobic Bacterium, *Gemmatimonas aurantiaca* Strain T-27. *Appl Environ Microbiol* 2017; **83**.

92. Klemedtsson L, Von Arnold K, Weslien P, Gundersen P. Soil CN ratio as a scalar parameter to predict nitrous oxide emissions. *Glob Change Biol* 2005; **11**: 1142–1147.

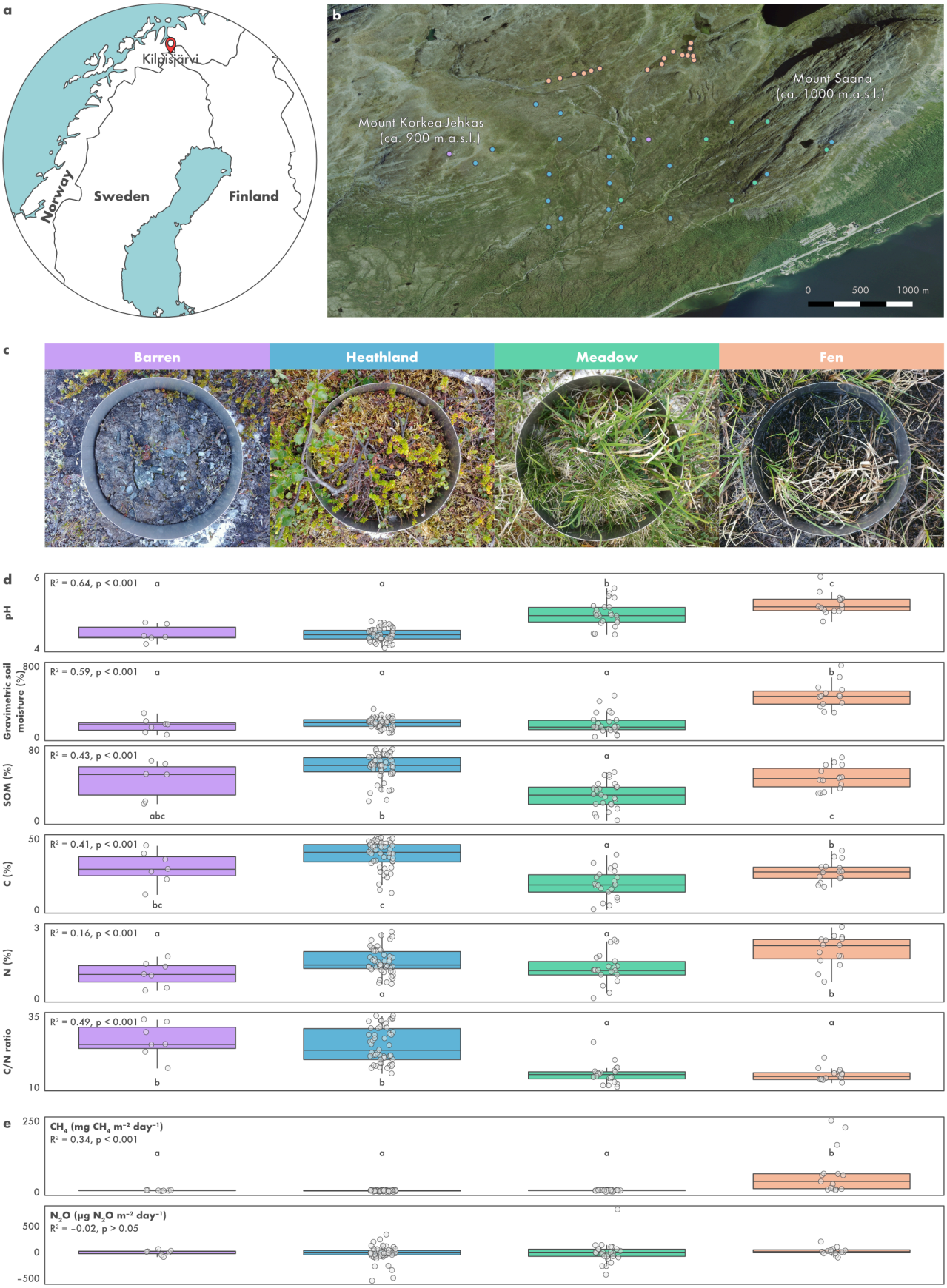
93. Philippot L, Andert J, Jones CM, Bru D, Hallin S. Importance of denitrifiers lacking the genes encoding the nitrous oxide reductase for N_2O emissions from soil: role of denitrifier diversity for N2O fluxes. *Glob Change Biol* 2011; **17**: 1497–1504.

94. Jones CM, Graf DR, Bru D, Philippot L, Hallin S. The unaccounted yet abundant nitrous oxide-reducing microbial community: a potential nitrous oxide sink. *ISME J* 2013; **7**: 417–426.

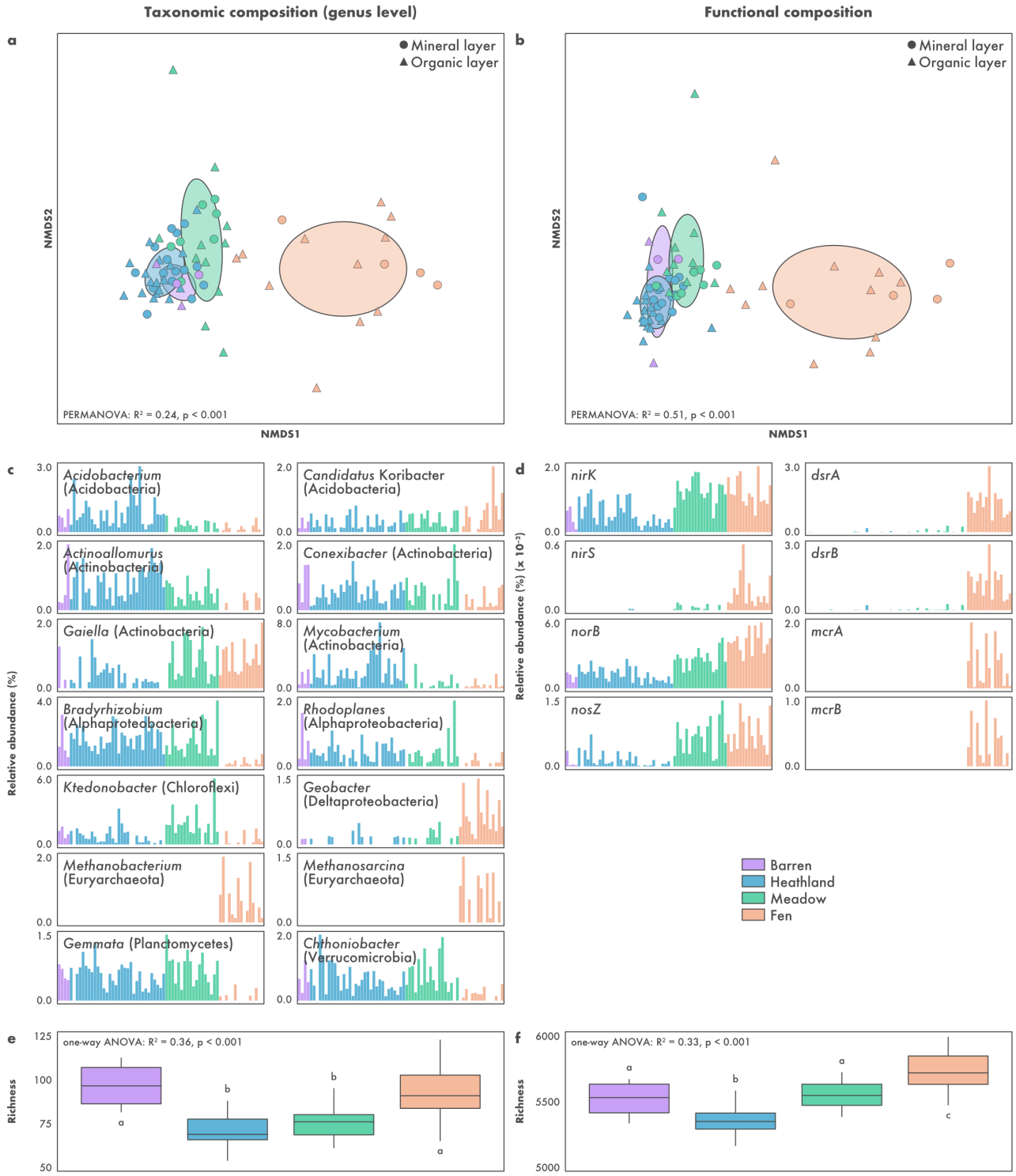
95. Jones CM, Spor A, Brennan FP, Breuil M-C, Bru D, Lemanceau P, et al. Recently identified microbial guild mediates soil N_2O sink capacity. *Nat Clim Change* 2014; **4**: 801–805.

96. Bakken LR, Bergaust L, Liu B, Frostegård Å. Regulation of denitrification at the cellular level: a clue to the understanding of N_2O emissions from soils. *Philos Trans R Soc B Biol Sci* 2012; **367**: 1226–1234.

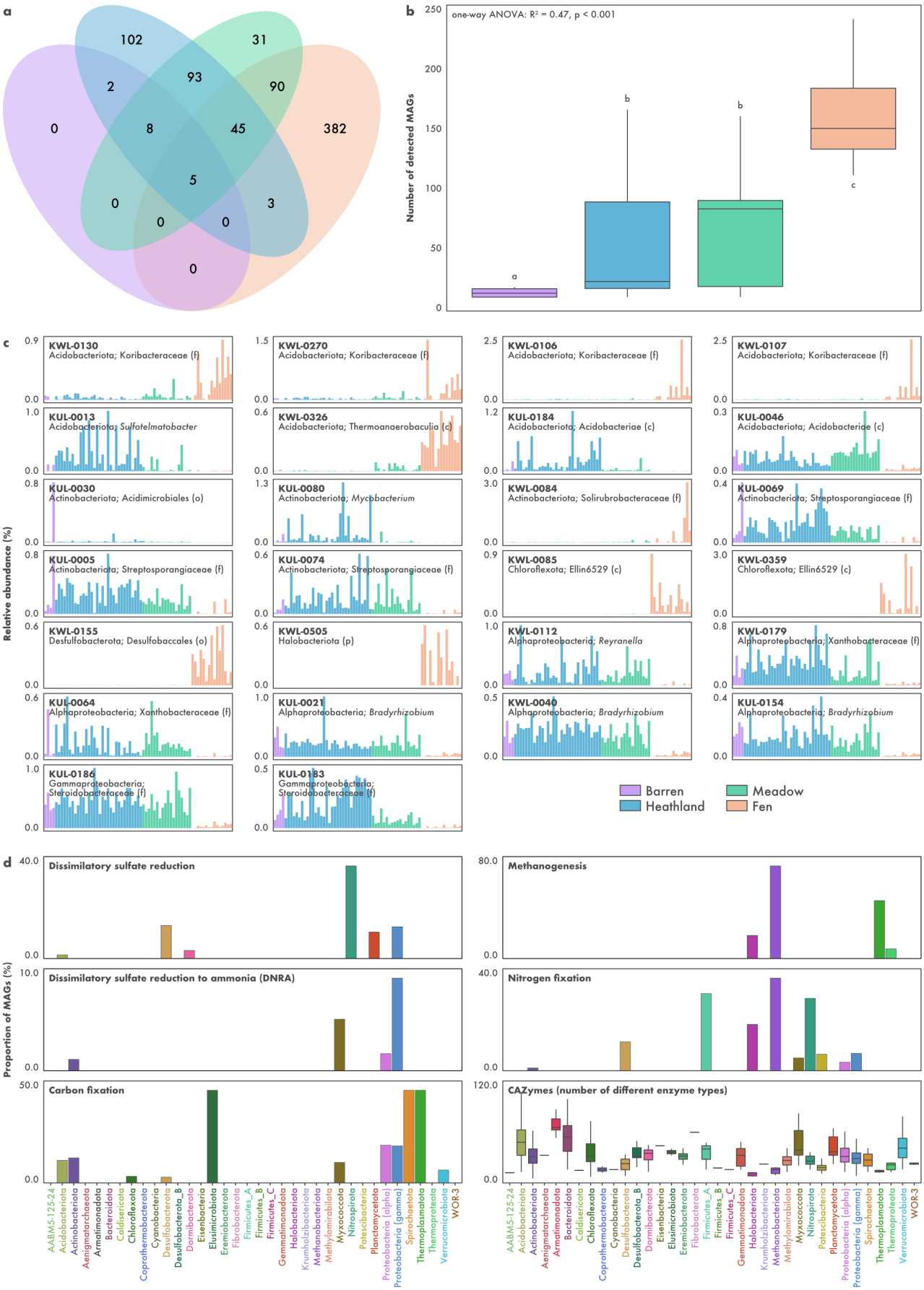
Supplementary Figures



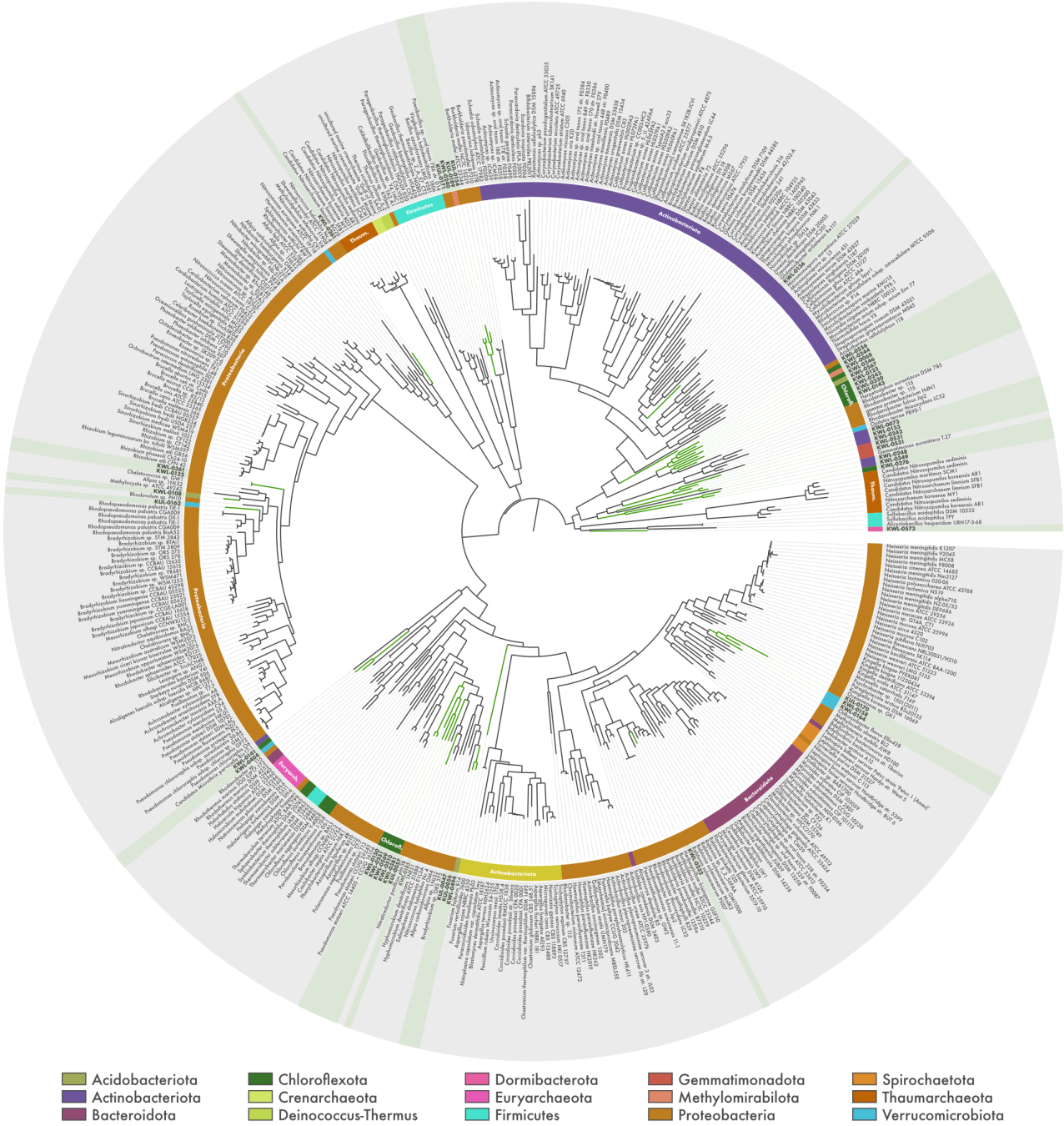
(Previous page) Suppl. Fig. S1 | Saana Nature Reserve, an area of mountain tundra in Kilpisjärvi, northern Finland. a) Map of Fennoscandia showing the location of Kilpisjärvi. **b)** Aerial overview of the study area showing the location of the 43 sites sampled for metagenomic analysis. Image provided by the National Land Survey of Finland under the Creative Commons CC BY 4.0 license. **c)** Pictures of the four types of soil ecosystems investigated. **d)** Physicochemical characterization of the soil ecosystems based on an extended set of 228 samples taken from the organic layer. **e)** *In situ* ecosystem-level nitrous oxide (N_2O) and methane (CH_4) fluxes measured from 101 sites using a static, non-steady state, non-flow-through system. Negative values represent net uptake and positive net emissions.


Taxonomic composition (genus level)
Functional composition
a
b
Mineral layer
Organic layer
NMDS2
NMDS1
PERMANOVA: R² = 0.24, p < 0.001
PERMANOVA: R² = 0.51, p < 0.001
c
d
Relative abundance (%)
Relative abundance (%) (x 10⁻²)
Acidobacterium (Acidobacteria)
Candidatus Koribacter (Acidobacteria)
Actinoallomurus (Actinobacteria)
Conexibacter (Actinobacteria)
Gaiella (Actinobacteria)
Mycobacterium (Actinobacteria)
Bradyrhizobium (Alphaproteobacteria)
Rhodoplanes (Alphaproteobacteria)
Ktedonobacter (Chloroflexi)
Geobacter (Deltaproteobacteria)
Methanobacterium (Euryarchaeota)
Methanosarcina (Euryarchaeota)
Gemmata (Planctomycetes)
Chthoniobacter (Verrucomicrobia)
nirK
dsrA
nirS
dsrB
norB
mcrA
nosZ
mcrB
Barren
Heathland
Meadow
Fen
e
f
Richness
one-way ANOVA: R² = 0.36, p < 0.001
one-way ANOVA: R² = 0.33, p < 0.001

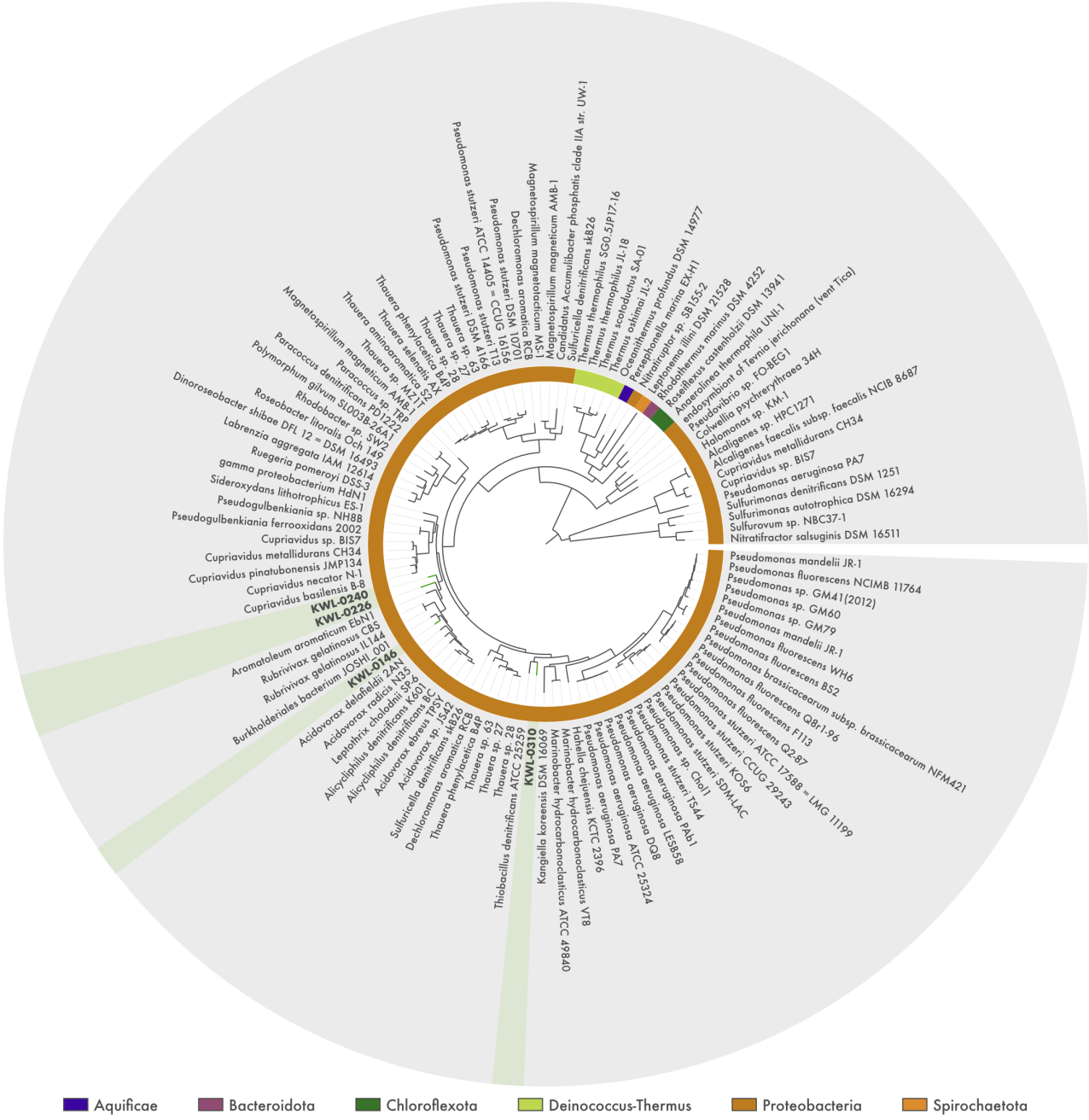
840 **(Previous page) Suppl. Fig. S2 | The microbial diversity of Kilpisjärvi soils as seen**
841 **using a read-based approach.** Read-based analyses were carried out using unassembled
842 Illumina data. Taxonomic composition was computed based on the annotation of SSU rRNA
843 gene sequences against the SILVA database with METAXA. Functional annotation was
844 performed by blastx searches against the KEGG database with DIAMOND. **a, b)** Non-metric
845 multidimensional scaling (NMDS) of taxonomic and functional community structure,
846 respectively. Differences between the ecosystems were assessed using permutational ANOVA
847 (PERMANOVA). **c)** Abundance profile of the five most abundant genera in each ecosystem.
848 **d)** Abundance profile of marker genes for denitrification, sulfate reduction, and
849 methanogenesis. **e, f)** Taxonomic (number of genera) and functional (number of functional
850 genes) richness. Differences in richness were assessed using one-way ANOVA followed by
851 Tukey's HSD test. Ecosystems that do not share the same letter are significantly different.


a
102
31
2
93
90
0
8
45
382
5
0
3
0
0
0
b
one-way ANOVA: R² = 0.47, p < 0.001
Number of detected MAGs
250
200
150
100
50
0
a
b
b
c
c
Relative abundance (%)
KWL-0130
Acidobacteriota; Koribacteraceae (f)
KWL-0270
Acidobacteriota; Koribacteraceae (f)
KWL-0106
Acidobacteriota; Koribacteraceae (f)
KWL-0107
Acidobacteriota; Koribacteraceae (f)
KUL-0013
Acidobacteriota; Sulfotelmatobacter
KWL-0326
Acidobacteriota; Thermoanaerobaculia (c)
KUL-0184
Acidobacteriota; Acidobacteriae (c)
KUL-0046
Acidobacteriota; Acidobacteriae (c)
KUL-0030
Actinobacteriota; Acidimicrobiales (o)
KUL-0080
Actinobacteriota; Mycobacterium
KWL-0084
Actinobacteriota; Solirubrobacteraceae (f)
KUL-0069
Actinobacteriota; Streptosporangiaceae (f)
KUL-0005
Actinobacteriota; Streptosporangiaceae (f)
KUL-0074
Actinobacteriota; Streptosporangiaceae (f)
KWL-0085
Chloroflexota; Ellin6529 (c)
KWL-0359
Chloroflexota; Ellin6529 (c)
KWL-0155
Desulfobacterota; Desulfobaccales (o)
KWL-0505
Halobacteriota (p)
KWL-0112
Alphaproteobacteria; Reyranella
KWL-0179
Alphaproteobacteria; Xanthobacteraceae (f)
KUL-0064
Alphaproteobacteria; Xanthobacteraceae (f)
KUL-0021
Alphaproteobacteria; Bradyrhizobium
KWL-0040
Alphaproteobacteria; Bradyrhizobium
KUL-0154
Alphaproteobacteria; Bradyrhizobium
KUL-0186
Gammaproteobacteria; Steroidobacteraceae (f)
KUL-0183
Gammaproteobacteria; Steroidobacteraceae (f)
Barren
Heathland
Meadow
Fen
d
Proportion of MAGs (%)
Dissimilatory sulfate reduction
Methanogenesis
Dissimilatory sulfate reduction to ammonia (DNRA)
Nitrogen fixation
Carbon fixation
CAZymes (number of different enzyme types)
AABM5-125-24
Acidobacteriota
Actinobacteriota
Aenigmatarchaeota
Armatimonadota
Bacteroidota
Caldisericota
Chloroflexota
Coprothermobacterota
Cyanobacteria
Desulfobacterota
Desulfobacterota_B
Dormibacterota
Eisenbacteria
Elusimicrobiota
Eremiobacterota
Fibrobacterota
Firmicutes_A
Firmicutes_B
Firmicutes_C
Gemmatimonadota
Halobacteriota
Krumholzibacteriota
Methanobacteriota
Methylomirabilota
Myxococcota
Nitrospirota
Patescibacteria
Planctomycetota
Proteobacteria (alpha)
Proteobacteria (gamma)
Spirochaetota
Thermoplasmatota
Thermoproteota
Verrucomicrobiota
WOR-3

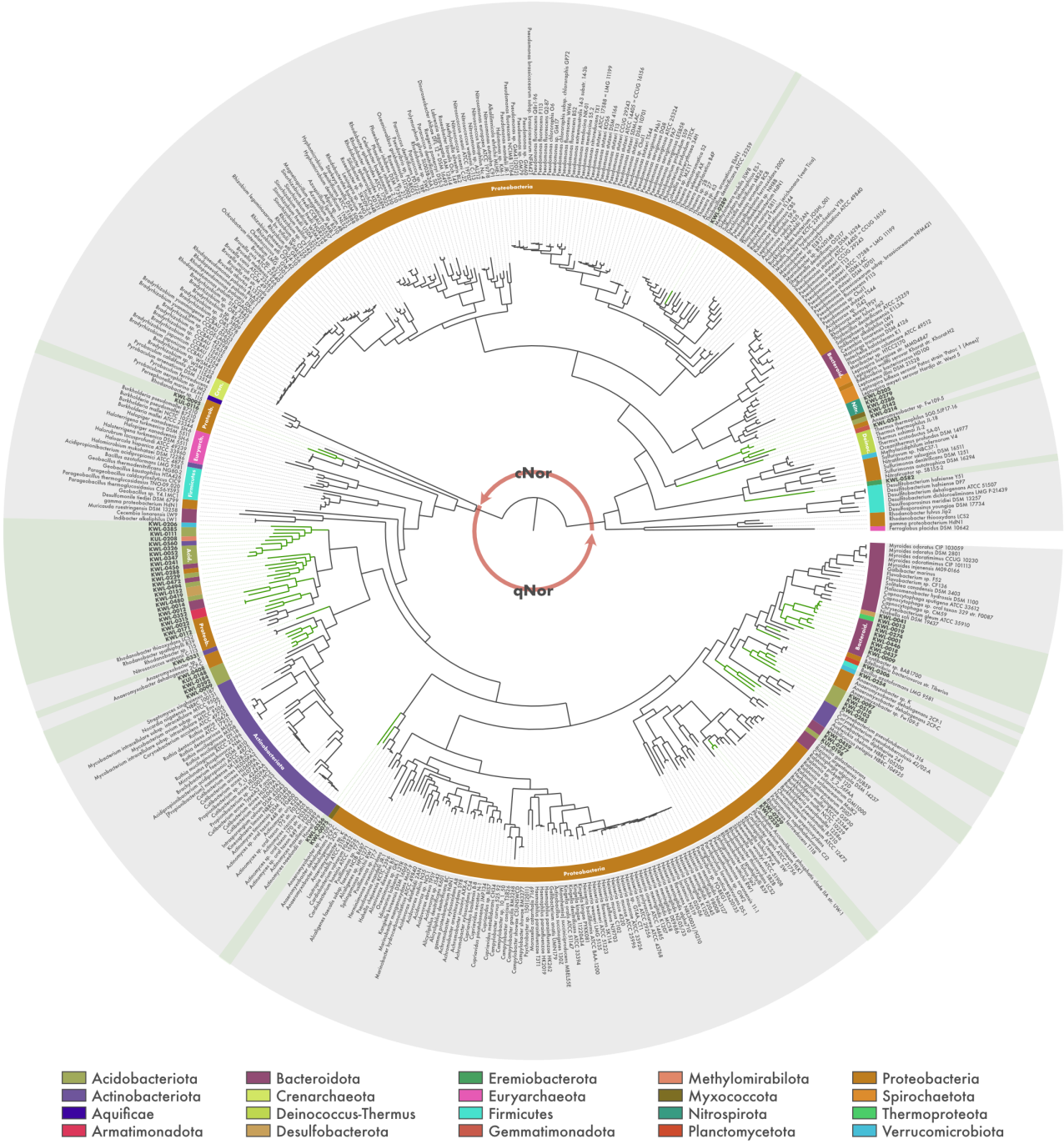
852 **(Previous page) Suppl. Fig. S3 | Overview of the microbial diversity in Kilpisjärvi**
853 **soils based on a genome-resolved approach. a)** Number of metagenome-assembled
854 genomes (MAGs) shared between the different ecosystems. **b)** Number of detected MAGs across
855 the ecosystems. **c)** Relative abundance of the ten most abundant MAGs in each ecosystem,
856 computed as a proportion of reads mapping to each MAG. **d)** Metabolic potential of the MAGs
857 based on the annotation of genes against the KOfam database. Barplots represent the
858 proportion of MAGs in each phylum with complete pathways, i.e. containing ≥ 75% of the genes
859 in the pathway. Boxplots of carbohydrate-active enzymes (CAZymes) show the number of
860 different enzyme types identified in each MAG.



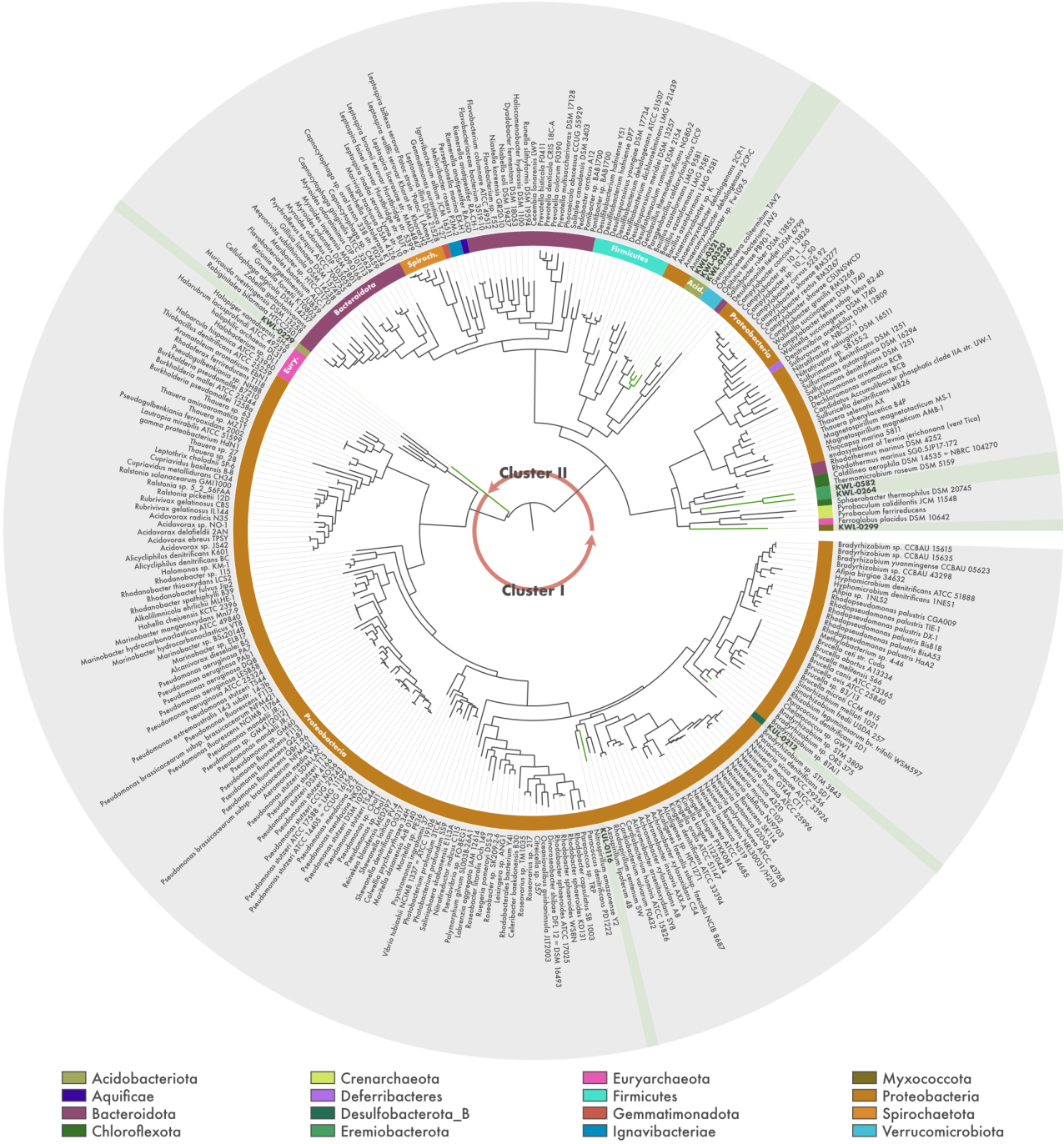
Suppl. Fig. S4 | Phylogeny of *nirK* sequences from metagenome-assembled genomes (MAGs) recovered from tundra soils in Kilpisjärvi, northern Finland. Midpoint-rooted maximum-likelihood tree of translated *nirK* sequences from Kilpisjärvi MAGs (highlighted) along with reference sequences from archaeal and bacterial genomes.



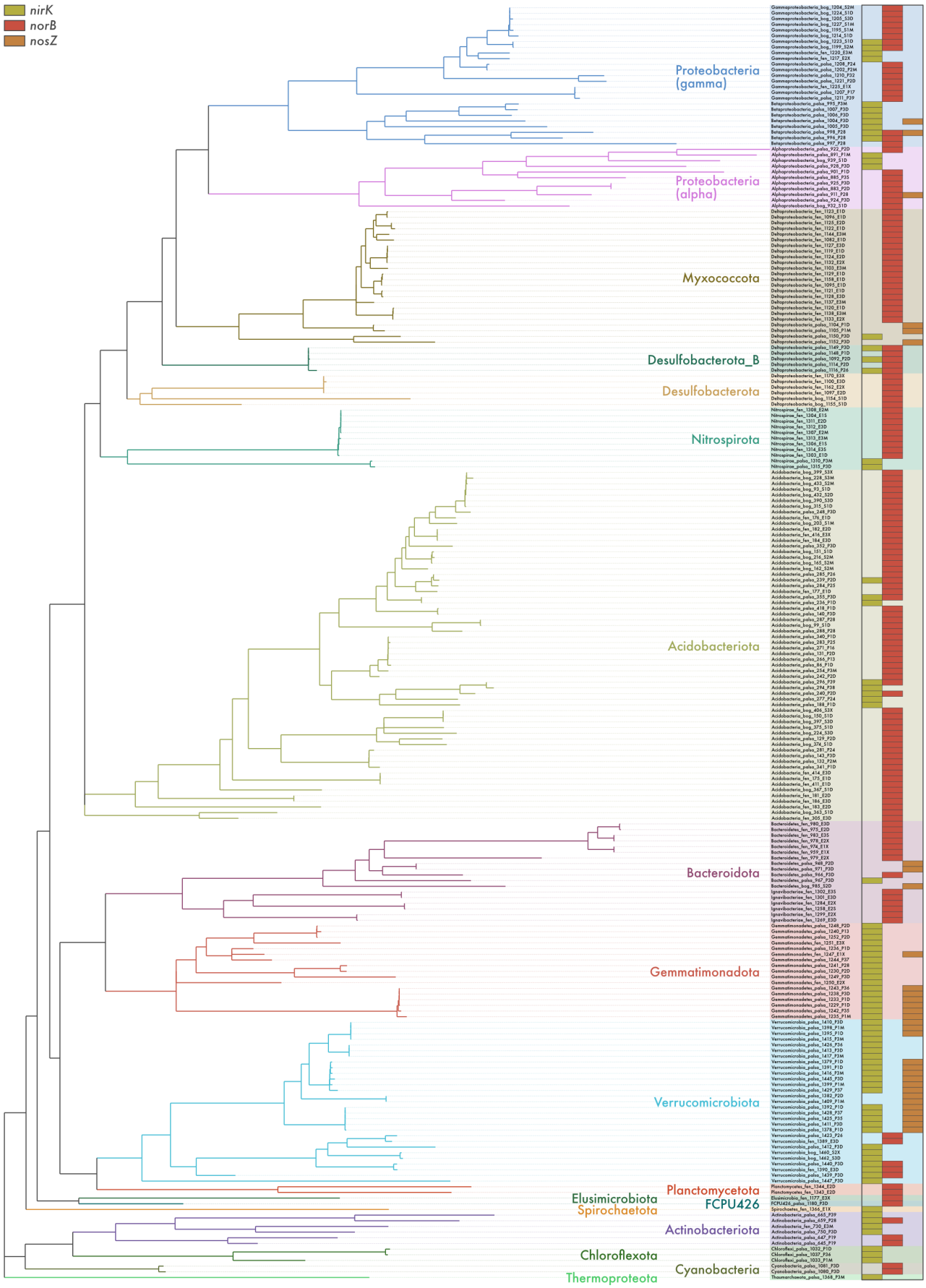
Suppl. Fig. S5 | Phylogeny of *nirS* sequences from metagenome-assembled genomes (MAGs) recovered from tundra soils in Kilpisjärvi, northern Finland. Midpoint-rooted maximum-likelihood tree of translated *nirS* sequences from Kilpisjärvi MAGs (highlighted) along with reference sequences from archaeal and bacterial genomes.



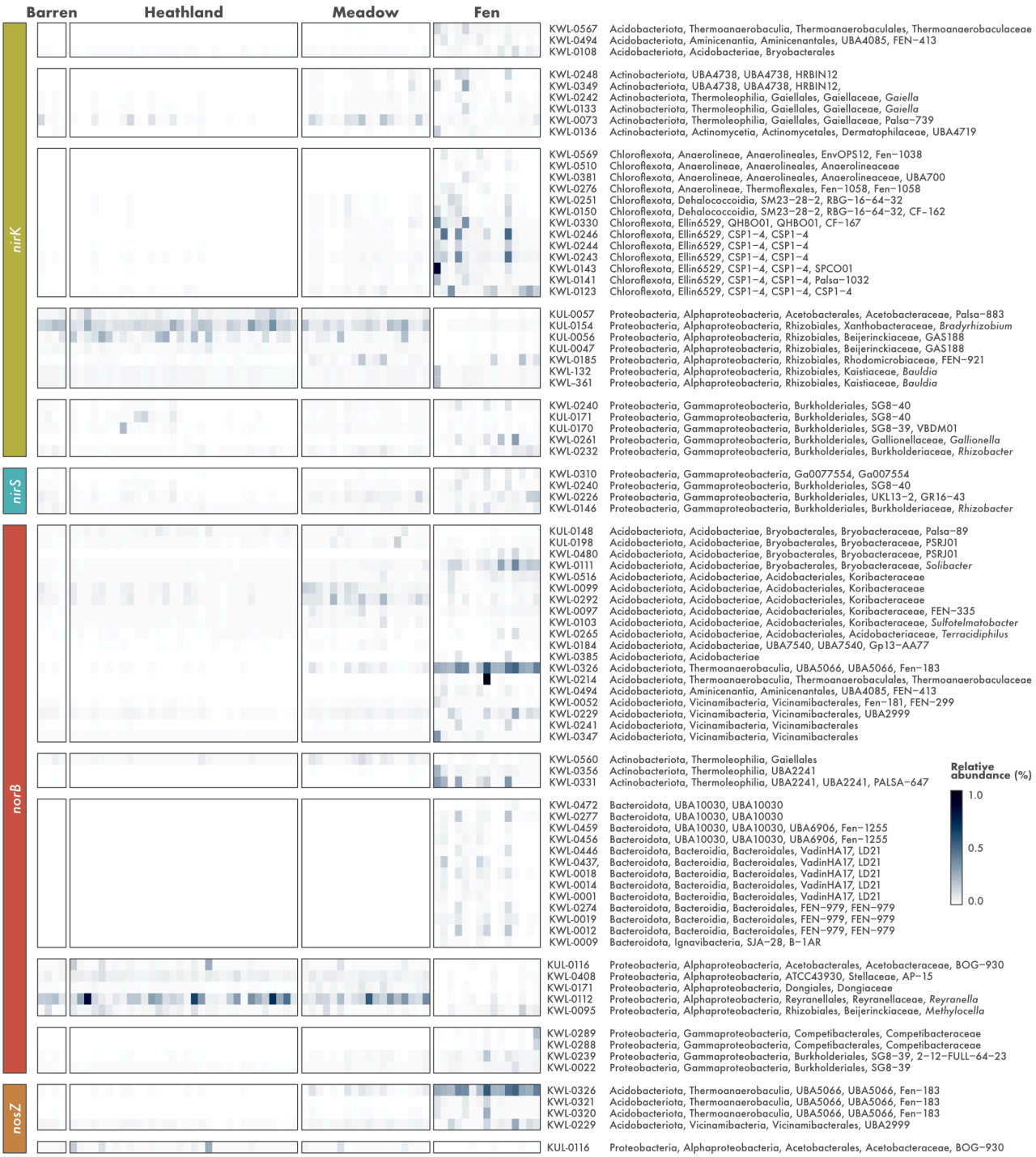
Suppl. Fig. S6 | Phylogeny of *norB* sequences from metagenome-assembled genomes (MAGs) recovered from tundra soils in Kilpisjärvi, northern Finland. Midpoint-rooted maximum-likelihood tree of translated *norB* sequences from Kilpisjärvi MAGs (highlighted) along with reference sequences from archaeal and bacterial genomes.



Suppl. Fig. S7 | Phylogeny of *nosZ* sequences from metagenome-assembled genomes (MAGs) recovered from tundra soils in Kilpisjärvi, northern Finland. Midpoint-rooted maximum-likelihood tree of translated *nosZ* sequences from Kilpisjärvi MAGs (highlighted) along with reference sequences from archaeal and bacterial genomes.


nirK
norB
nosZ
Proteobacteria (gamma)
Proteobacteria (alpha)
Myxococcota
Desulfobacterota_B
Desulfobacterota
Nitrospirota
Acidobacteriota
Bacteroidota
Gemmatimonadota
Verrucomicrobiota
Planctomycetota
Elusimicrobiota
FCPU426
Spirochaetota
Actinobacteriota
Chloroflexota
Cyanobacteria
Thermoproteota

(Previous page) Suppl. Fig. S8 | Metabolic potential for denitrification in Stordalen Mire soils. Distribution of denitrification genes across 225 metagenome-assembled genomes (MAGs) from permafrost peatland, bog, and fen soils in Stordalen Mire, northern Sweden. Genes encoding the nitrite (*nirK*), nitric oxide (*norB*), and nitrous oxide (*nosZ*) reductases were annotated using a three-step approach including (1) identification using hidden Markov models from the KOfam database, (2) manual inspection for the presence of conserved residues at positions associated with the binding of co-factors and active sites, and (3) phylogenetic analyses along with sequences from archaeal and bacterial genomes. Phylogenomic analysis of MAGs was done based on concatenated alignments of amino acid sequences from 122 archaeal and 120 bacterial single-copy genes.



Suppl. Fig. S9 | Relative abundance of metagenome-assembled genomes (MAGs) harbouring denitrification genes across different soil ecosystems in the tundra. MAGs were recovered from soils in Kilpisjärvi, northern Finland, and annotated for genes encoding the nitrite (*nirK*/*nirS*), nitric oxide (*norB*), and nitrous oxide (*nosZ*) reductases using a three-step approach. Relative abundances were computed as a proportion of reads mapping to each MAG.