Truncated denitrifiers dominate the denitrification pathway in tundra soil metagenomes

Igor S. Pessi[1,2], Sirja Viitamäki[1], Eeva Eronen-Rasimus[1,3], Tom O. Delmont[4], Miska Luoto[5] and Jenni Hultman[1,2,*]

[1]Department of Microbiology, University of Helsinki, Finland
[2]Helsinki Institute of Sustainability Science (HELSUS), Finland
[3]Finnish Environment Institute (SYKE), Helsinki, Finland
[4]Department of Bioinformatics, Genoscope, Paris, France
[5]Department of Geosciences and Geography, University of Helsinki, 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 these soils remain largely uncharacterized. In this study, we manually binned and curated 541 metagenome-assembled genomes (MAGs) from tundra soils in northern Finland. We then searched for MAGs encoding enzymes involved in denitrification, the main biotic process driving N_2O emissions. Denitrifying communities were dominated by poorly characterized taxa with truncated denitrification pathways, i.e. lacking one or more denitrification genes. Among these, MAGs with the metabolic potential to produce N_2O comprised the most diverse functional group. Re-analysis of a previously published metagenomic dataset from soils in northern Sweden supported these results, suggesting that truncated denitrifiers are dominant throughout the tundra biome.

Introduction

Nitrous oxide (N_2O) is a potent greenhouse gas (GHG) that has approximately 300 times the global warming potential of carbon dioxide on a 100-year scale[1]. 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[2–5]. The relative contribution of tundra soils

to global GHG emissions is predicted to increase in the future[6,7], as the warming rates experienced by high latitudes environments have been – and will likely continue to be – more than twice as high than in tropical and temperate regions[8].

The main biotic control on N_2O emissions is denitrification, which – like other biogeochemical cycles – is largely a microbial process[9]. Denitrification is a series of enzymatic steps in which nitrate (NO_3^-) is sequentially reduced to nitrite (NO_2), nitric oxide (NO), N_2O, and dinitrogen (N_2). The denitrification trait is found 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[10]. Approximately only one-third of the genomes from cultured denitrifiers sequenced to date encode the full set of enzymes needed for complete denitrification[11]. Thus, at the ecosystem level, denitrification is often a community effort performed by different microbial populations that each execute a part of the process[10,12].

Compared to high N_2O-emitting systems (e.g. agricultural and tropical soils), our knowledge of denitrifying communities in tundra soils is poor. As denitrification ultimately leads to the loss of N to the atmosphere, it enhances the N-limited status of tundra systems thus impacting microbial and plant productivity[13,14]. 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[15–18]. A better knowledge of the ecological, metabolic, and functional traits of denitrifiers is critical for improving current models and mitigating N_2O emissions[19]. This invariably relies on the characterization of the so-called uncultured majority, i.e. microorganisms that have not been cultured to date and comprise a high proportion of the microbial diversity in complex ecosystems[20,21]. 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[22–24]. However, this approach has not yet been applied to investigate the mechanisms driving N_2O cycling in the tundra.

Here, we used genome-resolved metagenomics to investigate the diversity and genomic composition of potential denitrifiers in tundra soils. By leveraging two comprehensive metagenomic datasets of Finnish and Swedish soils, we show that tundra denitrifying communities are dominated by microorganisms with truncated denitrification pathways, i.e. harbouring only a subset of the enzymes required for complete denitrification. Non-denitrifying N_2O-producers were the most diverse group and included many members of the phylum Acidobacteriota, a ubiquitous group in soils worldwide with a low representation in culture

collections. In contrast with earlier estimates based on genomes from cultured microorganisms[11], our genome-resolved metagenomics survey allowed access to the genomes of poorly characterized taxa and revealed that, within a defined ecosystem, the proportion of genomes encoding the complete denitrification pathway can be as low as 1%.

Results

A manually curated genomic database from tundra soil metagenomes

We analysed more than 9 billion Illumina (1.4 Tb) and 7 million Nanopore reads (21.5 Gb) from 69 soil metagenomes sampled in an area of mountain tundra biome in Kilpisjärvi, northern Finland **(Suppl. Fig. S1, Suppl. Table S1)**. In previous studies, we have established in the area a systematic fine-scale sampling of microclimate, soil conditions, and vegetation in topographically distinct habitats[25–27]. Local variation in topography and soil properties creates a mosaic of habitats characterized by contrasting ecological conditions (e.g. microclimate, vegetation productivity, cover, and biomass). This makes the study setting ideal to investigate biotic interactions and species-environment relationships in tundra ecosystems[27–29]. Our sampling design included two soil depths across four ecosystems that are characteristic of polar and alpine environments (barren soils, heathlands, meadows, and fens). Soil ecosystems differ in vegetation cover and physicochemical composition, with fens in particular being characterized by higher pH and moisture content **(Suppl. Fig. S1)**. Microbial community composition also differs, with fen soils harbouring contrasting microbial communities compared to the other ecosystems **(Suppl. Fig. S2)**.

Two Illumina co-assemblies and two individual Nanopore assemblies yielded nearly 20 million contigs longer than 1,000 bp, with a total assembly size of 50.0 Gb. Using anvi'o[30], we obtained 3,257 genomic bins and manually curated these to a set of 539 unique metagenome assembled genomes (MAGs) **(Fig. 1, Suppl. Fig. S3, Suppl. Fig. S4, Suppl. Table S2)**. On average, 8.6% of the reads from each sample were recruited by the MAGs (minimum: 2.7%, maximum: 22.4%). According to estimates based on domain-specific single-copy genes, the obtained MAGs are on average 67.7% complete (50.7–100.0%) and 2.5% redundant (0.0–9.9%) **(Suppl. Table S2)**. Phylogenomic analyses based on 122 archaeal and 120 bacterial single-copy genes in the context of the Genome Taxonomy Database (GTDB)[31,32] placed the Kilpisjärvi MAGs across a diverse range of bacterial and archaeal phyla, including Acidobacteriota (n = 127), Actinobacteriota (n = 126), Alphaproteobacteria (n = 44), Chloroflexota (n = 41), Gammaproteobacteria (n = 36), and Verrucomicrobiota (n = 31) **(Fig. 1, Suppl. Fig. S3)**. Of the 541 MAGs, only 78 were

assigned to a validly described genera **(Suppl. Table S2)**. Most MAGs (n = 463) belong to genera that do not comprise formally described species. These include 183 MAGs that were placed outside genus-level lineages currently described in GTDB and thus likely represent novel genera.

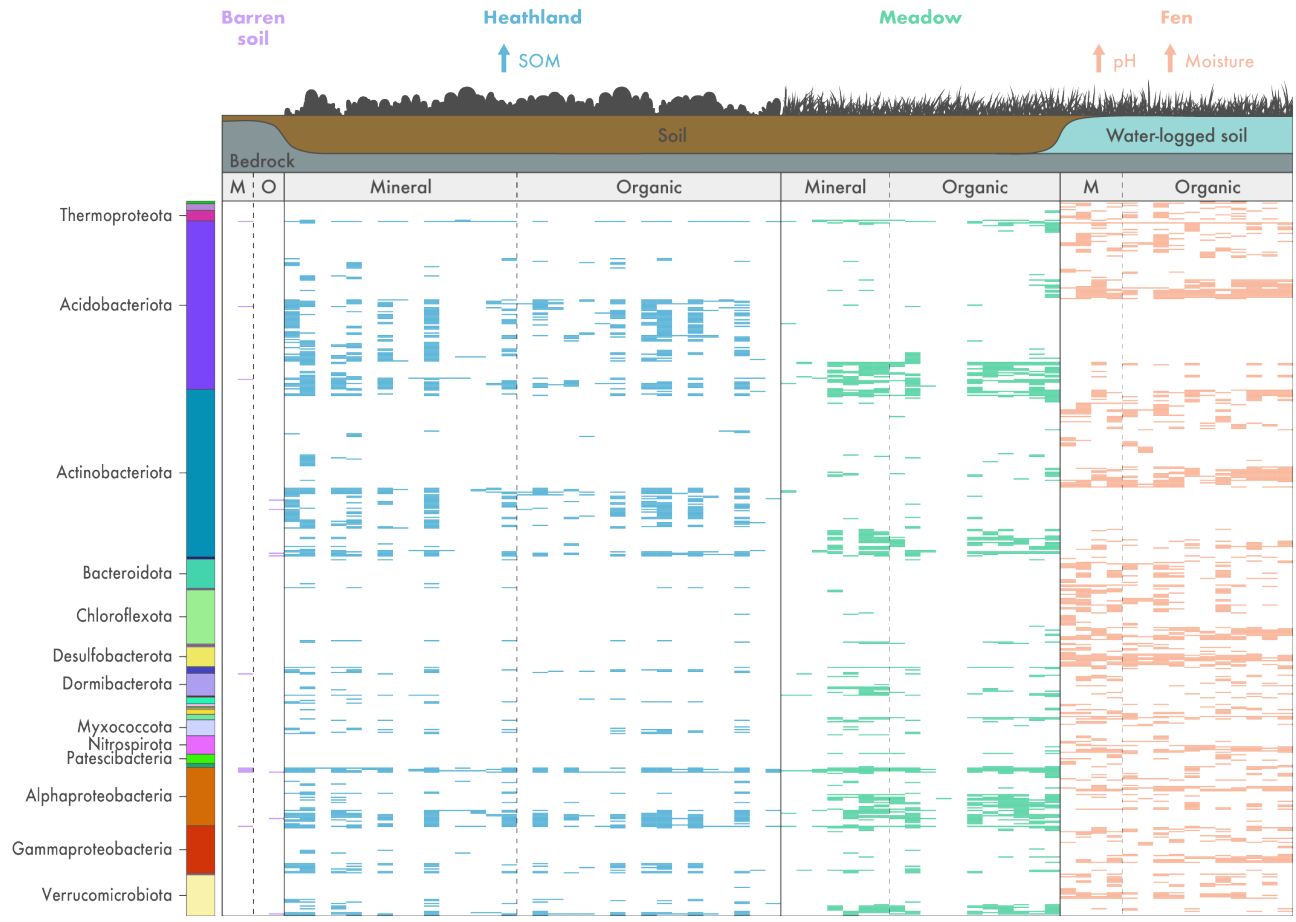


Fig. 1 | Genome-resolved metagenomics of microorganisms in tundra soils. Detection profile of 539 metagenome-assembled genomes (MAGs) recovered from mineral and organic soils across different ecosystems in Kilpisjärvi, northern Finland. A MAG was detected in a given sample if ≥ 50% of its nucleotides had ≥ 1x coverage. Phylum-level taxonomic assignments are shown only for the major groups found. A complete representation of all phyla can be found in **Suppl. Fig. S3** and additional information about the MAGs is provided in **Suppl. Table S2**.

The highest number of MAGs (n = 341) was detected in the fen soils **(Suppl. Fig. S4)**. Although barren and fen soils had similar taxonomic richness according to gene-centric estimates **(Suppl. Fig. S2)**, the genome-resolved approach recovered only a small fraction of the microbial diversity in barren soils (n = 15). This is likely a result of limited sampling and sequencing of this ecosystem **(Suppl. Table S1)**. The number of detected MAGs in heathland and meadow soils was similar (n = 223 and n = 217, respectively) **(Suppl. Fig. S4)**. In agreement with the

gene-centric assessment **(Suppl. Fig. S2)**, we observed differences in MAG composition across the soil ecosystems **(Suppl. Fig. S4)**. Only 41 MAGs (7.6%) were shared between the heathland and fen soils. On the other hand, meadow communities represent an intermediate state between heathland and fens, sharing 115 and 125 MAGs with these ecosystems, respectively. This is likely a reflection of edaphic similarities between the ecosystems **(Suppl. Fig. S1)**. Meadow soils have similar moisture levels to heathlands but, as fens, have graminoid plants as the dominant vegetation cover.

MAGs from tundra soils have truncated denitrification pathways

To gain insights into the microorganisms involved with the cycling of N_2O in tundra soils, we annotated the predicted amino acid sequences of the Kilpisjärvi MAGs across a range of protein databases (COG[33], KEGG[34], KOfam[35], RefSeq[36], and Swiss-Prot[37]). We then searched for MAGs with a metabolic potential for denitrification, the main process controlling N_2O production in soils[9]. The complete denitrification pathway is performed by microorganisms containing the *narG*/*napA*, *nirK*/*nirS*, *norB*, and *nosZ* genes, which encode the nitrate (Nar), nitrite (Nir), nitric oxide (Nor), and nitrous oxide (Nos) reductases, respectively[10]. In the Kilpisjärvi soils, the metabolic potential for denitrification was exclusively restricted to MAGs with truncated denitrification pathways, i.e. MAGs missing one or more denitrifying reductases **(Fig. 2)**. Of the 129 MAGs harbouring denitrifying reductases, 114 contain only one of the four enzymes and no MAG encode all the Nir, Nor, and Nos enzymes required for complete denitrification, i.e. the reduction of NO_2^- to N_2. MAGs encoding denitrifying reductases were detected mainly in the meadow and fen soils, where the abundance of denitrification genes was the highest **(Suppl. Fig. S2)**. MAGs belong to the archaeal phylum Thermoproteota and many bacterial phyla such as Gamma- and Alphaproteobacteria, Acidobacteriota, Bacteroidota, Actinobacteriota, and Chloroflexota **(Fig. 2)**. Only 19 MAGs were assigned to a validly described genera **(Suppl. Table S2)**.

Truncated pathways do not appear to be a methodological artifact arising from the metabolic reconstruction of fragmented genomes. Indeed, truncated denitrifiers include 23 high-quality MAGs (≥ 90% complete) and no relationship was observed between the number of denitrifying reductases encoded by the MAGs and their estimated completion values **(Suppl. Fig. S5)**. To verify if microorganisms with truncated denitrification pathways are common in other tundra systems, we expanded our analysis to 1529 MAGs recovered from permafrost peatland, bog, and fen soils in Stordalen Mire, northern Sweden[24]. Annotation and metabolic reconstruction of the Stordalen Mire MAGs corroborated the observations from Kilpisjärvi MAGs **(Suppl. Fig. S6)**. The distribution of denitrifying reductases was remarkably similar in MAGs from both systems,

and MAGs with truncated denitrification pathways were also the norm in Stordalen Mire soils. Of the 396 Stordalen Mire MAGs encoding denitrifying reductases, only six harbour all the Nir, Nor, and Nos enzymes required for complete denitrification. Altogether, the analysis of these two comprehensive datasets suggest that microorganisms with truncated denitrification pathways are widespread throughout tundra ecosystems.

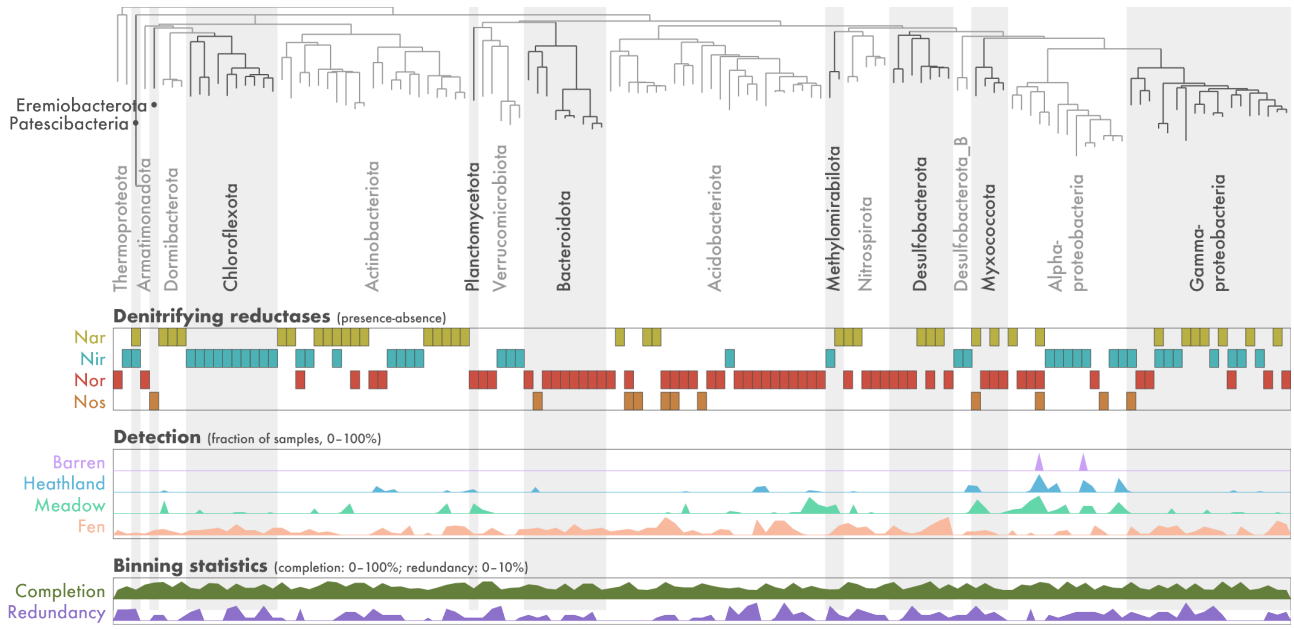


Fig. 2 | Metabolic potential for denitrification in metagenome-assembled genomes (MAGs) from tundra soils. Analysis of 129 MAGs encoding denitrifying reductases obtained from tundra soils in Kilpisjärvi, northern Finland. The presence of nitrate (Nar), nitrite (Nir), nitric oxide (Nor), and nitrous oxide (Nos) reductases was inferred based on the annotation of predicted amino acid sequences. Detection is shown as a fraction of the samples in each soil ecosystem. MAGs completion and redundancy were computed based on the presence of domain-specific single-copy genes. MAGs are ordered according to their phylogenetic placement, which was inferred based on concatenated alignments of 122 archaeal and 120 bacterial single-copy genes.

Fen soils harbour Chloroflexota MAGs with divergent Nir enzymes

The reduction of NO_2^- to NO, performed by microorganisms containing the Nir enzyme, is the hallmark step of denitrification and is often referred to as denitrification *stricto sensu*, as it involves the conversion of a soluble substrate to a gaseous product thus leading to the removal of N from the system[10]. Within the 129 Kilpisjärvi MAGs encoding denitrifying reductases, 41 are potential denitrifiers *stricto sensu* **(Fig. 2)**. Most MAGs (n = 38) harbour the copper-containing form of Nir encoded by the *nirK* gene. The cytochrome cd_1-containing form of Nir

167 encoded by the *nirS* gene was found in four Gammaproteobacteria MAGs, including one MAG
168 that encodes both forms of the enzyme. Most Nir-encoding MAGs belong to the bacterial phyla
169 Alpha- and Gammaproteobacteria, Chloroflexota, and Actinobacteriota **(Fig. 2)**. Denitrifying
170 *stricto sensu* communities differed between the meadow and fen soils **(Suppl. Figure S7a)**.
171 Alphaproteobacteria comprised half of the Nir-encoding MAGs in the meadow soils, while
172 Chloroflexota and Gammaproteobacteria were the majority in the fens. Interestingly, the amino
173 acid sequence composition of the Nir enzymes from Chloroflexota MAGs were quite divergent
174 **(Suppl. Figure S7b).** These had lower identity with reference sequences from RefSeq
175 (31.6–66.3%) than the sequences of MAGs from better characterized phyla such as the Alpha-
176 and Gammaproteobacteria (67.5–88.8% and 76.3–91.7%, respectively). Indeed, most Nir-
177 encoding Chloroflexota MAGs were assigned to the class-level lineage Ellin6529 **(Suppl. Table**
178 **S2)**, a major group in soils worldwide that at present does not include cultured
179 representatives[38].

180 The potential for N_2O production is widespread among the Acidobacteriota

181 The stepwise reduction of NO to N_2O and N_2 catalysed by the Nor and Nos enzymes represents
182 the final step of denitrification and the main biotic control on N_2O emissions. Soil denitrification
183 rates depend on multiple environmental conditions such as adequate moisture and inorganic N
184 availability, but whether it results in the emission of N_2O or N_2 is ultimately linked to a balance
185 between the activity of NO and N_2O reducers[9]. Kilpisjärvi MAGs with a metabolic potential for
186 NO and N_2O reduction (n = 56 and n = 11, respectively) are almost exclusively non-denitrifiers
187 *stricto sensu*, i.e. they do not encode the Nir enzyme involved in the reduction of NO_2^- to NO
188 **(Fig. 2).** Only four MAGs encode both Nor and Nos and thus have the potential to reduce NO
189 completely to N_2. Most MAGs with potential NO-reducing capability (n = 48) contain the qNor
190 enzyme encoded by the *norB* gene, which is a monomeric form of Nor found in denitrifying
191 bacteria and archaea and non-denitrifying pathogenic bacteria[39]. The remaining MAGs (n = 8)
192 harbour the cytochrome c-dependent form of Nor (cNor) encoded by *norBC*.

193 MAGs with the potential to reduce NO to N_2O were particularly prominent among the phylum
194 Acidobacteriota (n = 18 MAGs) **(Fig. 2)**, an ubiquitous group in tundra and other soil ecosystems
195 worldwide[38,40]. Acidobacteriota are very recalcitrant to culturing efforts. Of the 26 subdivisions
196 identified by 16S rRNA gene surveys, only seven contain genera that have been properly
197 described based on cultured representatives[41]. Our analysis of MAGs from tundra soils and all
198 available genomes of Acidobacteriota strains and candidate taxa suggests that the potential for
199 NO reduction to N_2O is widespread within members of this phylum **(Fig. 3)**. Potential N_2O
200 producers are particularly prominent in the subdivisions Gp1, Gp3, and Gp23, including the

genera *Sufotelmatomonas*, *Acidobacterium*, *Silvibacterium*, *Sulfotelmatobacter*, *Koribacter*, and *Thermoanaerobaculum*. On the other hand, genomes with potential to reduce NO completely to N_2 are mostly found in the subdivision Gp6 (including the strain *Luteitalea pratensis* DSM 100886) and in an early-branching subdivision without cultured representatives related to Gp23 **(Fig. 3)**. This clade includes the Kilpisjärvi MAG KWL-0326, an abundant member of the tundra communities detected in all fens and half of the meadow samples at relative abundances of up to nearly 1% **(Suppl. Fig. S4)**.

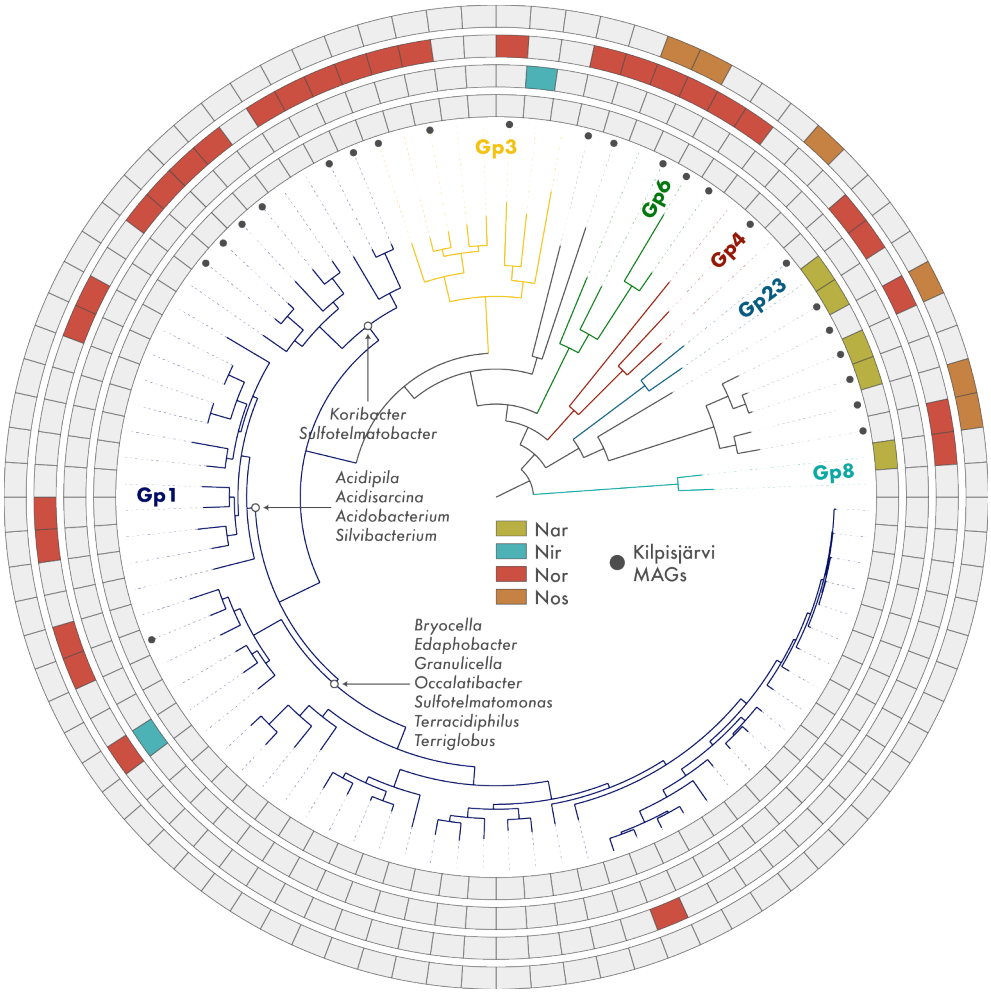


Fig. 3 | Metabolic potential for denitrification in members of the phylum Acidobacteriota. Phylogenomic analysis of 23 Acidobacteriota MAGs encoding denitrifying reductases recovered from tundra soils in Kilpisjärvi, northern Finland, and 64 published genomes of Acidobacteriota strains and candidate taxa. Maximum likelihood tree based on concatenated alignments of 23 ribosomal proteins and rooted with *Escherichia coli* ATCC 11775 (not shown). The presence of nitrate (Nar), nitrite (Nir), nitric oxide (Nor), and nitrous oxide (Nos) reductases was inferred based on the annotation of predicted amino acid sequences.

Discussion

The 539 MAGs obtained in the present study by a manual binning and curation effort represent one of the largest genomic catalogues of microorganisms from tundra soils to date **(Fig. 1, Suppl. Fig. S3, Suppl. Fig. S4, Suppl. Table S2)**. Our dataset is comparable to a previous metagenomic investigation of soils in Stordalen Mire, northern Sweden, which included 647 unique MAGs obtained from 1.7 Tb of metagenomic data[24]. Here, we leveraged these two comprehensive datasets to investigate the genomic makeup of microorganisms involved with denitrification in tundra soils, with the ultimate goal of gaining insights into mechanisms of N_2O emission from these systems. Our results strongly indicate that tundra soils are dominated by microorganisms with truncated denitrification pathways, most of which represent poorly characterized taxa without cultured representatives **(Fig. 2, Suppl. Fig. S5, Suppl. Fig. S6, Suppl. Fig. S7)**. Unlike well-known denitrifiers such as *Bradyrhizobium japonicum*[42,43], our genome-resolved metagenomics survey indicates that dominant microbial populations in tundra soils do not have the metabolic potential for complete denitrification. Instead, the potential for denitrification in tundra soils is highly modular, with each step of the pathway (NO_3^-, NO_2^-, NO, and N_2O reduction) encoded by different microbial populations.

It has been suggested that denitrification is a community process performed in synergy by different microbial taxa that execute only a subset of the complete denitrification pathway[10,12]. With the growing number of microbial genomes sequenced in recent decades, it became evident that only a fraction of the microorganisms involved in the denitrification pathway encode the enzymatic machinery needed for complete denitrification[11,42,43]. For instance, in a study that included 652 genomes of cultured microorganisms harbouring denitrifying reductases retrieved from GenBank, Graf et al.[11] found that only approximately 31% are complete denitrifiers. However, the magnitude of truncated denitrification pathways found in our dataset of tundra metagenomes was much more pronounced. Our genome-resolved metagenomics approach enabled us to access to the genomes of uncultured, poorly characterized taxa, which comprise the majority of the microorganisms in soils and other complex ecosystems[20,21]. By expanding the genomic catalogue of denitrifiers to include uncultured taxa, our results revealed a higher magnitude of truncated denitrification pathways than previous assessments based on genomes from cultured microorganisms. Moreover, from an ecosystem perspective, the contribution of microorganisms with truncated denitrification pathways has been largely overlooked. In tundra systems in particular, earlier investigations applying a gene-centric approach have shown that the potential for complete denitrification is present within defined microbial communities[44,45]. However, gene-centric approaches fail to reveal the wider genomic context in which these genes

249 are inserted. By applying the genome-resolved metagenomics approach, we traced
250 denitrification genes to specific microbial populations, thereby allowing a detailed investigation
251 of the genomic makeup of potential denitrifiers in tundra soils. In addition to their
252 predominance in genomic databases[11], our genome-resolved metagenomics survey revealed that
253 truncated denitrifiers are also dominant within a defined ecosystem.

254 Microorganisms harbouring the Nor enzyme – and thus with the potential to reduce NO to N_2O
255 – were the most diverse functional group and were particularly prominent among the phylum
256 Acidobacteriota **(Fig. 3)**. Members of this phylum, which comprise mostly oligotrophic (k-
257 strategist) taxa, make up an important fraction of the microbial communities in polar and alpine
258 soils[38,40]. Earlier genomic investigations have identified denitrifying reductases within
259 Acidobacteriota genomes, but have disregarded their potential role in denitrification due to the
260 presence of truncated pathways[46,47]. The potential contribution of Acidobacteriota to the soil
261 denitrification process in tundra soils is supported by other studies that have shown that,
262 despite not having a complete denitrification machinery, truncated denitrifiers can contribute
263 to *in situ* processes[43,48–51]. For example, Lycus et al.[48] applied an elegant cultivation approach
264 using different N oxides as sole electron acceptors and isolated a wide diversity of
265 microorganisms with different types of truncated denitrification pathways that are able to grow
266 on single N oxides. These included, for instance, microorganisms that can grow using NO as an
267 electron acceptor generating N_2O but are unable to perform any other step of the denitrification
268 process. Moreover, studies have suggested that incomplete denitrifiers that contain Nir and Nor
269 but lack Nos (thus having N_2O as the end product of denitrification) can contribute substantially
270 to soil N_2O emissions[49]. At the same time, non-denitrifying N_2O reducers, microorganisms that
271 contain Nos but lack Nir, can represent an important N_2O sink[43,50,51].

272 A great challenge in microbial ecology has been to link microbial community structure to
273 biogeochemical processes[52]. While microorganisms with truncated denitrification pathways
274 dominated the denitrifying communities investigated here, the potential for complete
275 denitrification was indeed present at the ecosystem level. Our current knowledge of the
276 regulation of the denitrification process is largely based on the activity of model organisms such
277 as the complete denitrifier *Paracoccus denitrificans*[53]. It thus remains unclear how such a
278 community dominated by truncated denitrifiers interacts with each other and the environment
279 and impact N_2O emissions *in situ*. A better understanding of the activity of truncated
280 denitrifiers is paramount for our ability to model N_2O emissions and mitigate climate
281 change[19,53]. High-latitude environments in particular have experienced amplified warming in
282 recent decades, a trend that is likely to continue in the coming centuries[8]. As mechanisms of
283 GHG emissions are very climate sensitive, the contribution of tundra soils to global GHG

284 atmospheric levels is thus predicted to increase in the future leading to a positive feedback
285 loop[6,7]. Compared with carbon dioxide and methane fluxes, measurements of N_2O emissions in
286 tundra soils are sparse, making the magnitude of N_2O fluxes across the polar regions uncertain[5].
287 In addition to a better monitoring of N_2O emissions throughout the tundra biome, our results
288 suggest that a better understanding of the contribution of tundra soil to global N_2O levels should
289 include the investigation of the mechanisms of metabolic regulation in communities dominated
290 by truncated denitrifiers.

291 Methods

292 Study area and sampling

293 The Saana Nature Reserve (69.04°N, 20.79°E) is located in Kilpisjärvi, northern Finland
294 **(Suppl. Fig. S1a)**. The area is part of the oroarctic mountain tundra biome, which is
295 characterized by a mean annual temperature of –1.9°C and annual precipitation of 487 mm[54].
296 Our study sites (n = 43) are distributed across Mount Saana and Mount Korkea-Jehkas and the
297 valley in between and include four soil ecosystems: barren soils (n = 2), heathlands (n = 18),
298 meadows (n = 7), and fens (n = 16) **(Suppl. Fig. S1b)**. Sampling was performed in July 2017
299 and July 2018, during the growing season in the northern hemisphere. Samples were obtained
300 with a soil corer sterilized with 70% ethanol and, when possible, soil cores were split into organic
301 and mineral samples using a sterilized spatula. In total, 41 organic and 28 mineral samples
302 were obtained. Samples were transferred to a whirl-pack bag and immediately frozen in dry ice.
303 Samples were transported frozen to the laboratory at the University of Helsinki and kept at
304 –80°C until analyses.

305 Metagenome sequencing

306 Total DNA and RNA were co-extracted as previously described[27]. Briefly, extraction was
307 performed on 0.5 g of soil using a hexadecyltrimethyl ammonium bromide (CTAB), phenol-
308 chloroform, and bead-beating protocol. DNA was purified using the AllPrep DNA Mini Kit
309 (QIAGEN, Hilden, Germany) and quantified using the Qubit dsDNA BR Assay Kit
310 (ThermoFisher Scientific, Waltham, MA, USA). Library preparation for Illumina metagenome
311 sequencing was performed using the Nextera XT DNA Library Preparation Kit (Illumina, San
312 Diego, CA, USA). Metagenomes were obtained for 69 samples across two paired-end NextSeq
313 (132–170 bp) and one NovaSeq (2 x 151 bp) runs. Two samples were additionally sequenced with
314 Nanopore MinION. For this, libraries were prepared using the SQK-LSK109 Ligation

315 Sequencing Kit with the long fragment buffer (Oxford Nanopore Technologies, Oxford, UK) and
316 the NEBNext Companion Module for Oxford Nanopore Technologies Ligation Sequencing Kit
317 (New England Biolabs). Each sample was sequenced for 48 hours on one R9.4 flow cell.

318 Processing of raw metagenomic data

319 The quality of the raw Illumina data was verified with fastQC[55] v0.11.9 and multiQC[56] v1.8.
320 Cutadapt[57] v1.16 was then used to trim 3' adapters and low-quality base calls (q < 20) and to
321 filter out short reads (< 50 bp). Nanopore data were basecalled with GPU guppy v4.0.11 using
322 the high-accuracy model and applying a minimum quality score of 7. The quality of the
323 basecalled Nanopore data was assessed with pycoQC[58] v2.5.0.21 and adapters were trimmed
324 with Porechop[59] v0.2.4.

325 Gene-centric analyses

326 Taxonomic and functional profiles of the microbial communities were obtained using a gene-
327 centric approach based on unassembled Illumina data. Due to differences in sequencing depth
328 across the samples, the dataset was resampled to 2,000,000 reads per sample with seqtk[60] v1.3.
329 Taxonomic profiles were obtained by annotating SSU rRNA gene sequences against the SILVA
330 database[61] release 111 with METAXA[62] v2.2. Functional profiles were obtained by blastx
331 searches against the KEGG database[34] release 86 with DIAMOND[63] v0.9.14. Only matches with
332 a maximum e-value of 10^{-5} and minimum identity of 60% were considered. The KEGG Orthology
333 (KO) identifier of the best hit was assigned to each read and KEGG modules were summarised
334 with the package keggR[64] v0.9.1 in R. Gene relative abundances were normalized to the
335 abundance of the *rpoB* gene. Differences in richness (i.e. the number of genera and functional
336 genes in each sample) between the ecosystems were assessed using one-way analysis of variance
337 (ANOVA) followed by Tukey's HSD test in R. The same test was performed to assess differences
338 in the abundance of individual N cycle genes. Differences in taxonomic and functional
339 community structure were assessed using non-metric multidimensional scaling (NMDS) and
340 permutational ANOVA (PERMANOVA) with the package vegan[65] v2.5.6 in R.

341 Metagenome assembling

342 Metagenome assembling of the Illumina data was performed as two co-assemblies. One co-
343 assembly was performed for the upland (barren, heathland, and meadow) and other for the fen
344 samples. For each co-assembly, reads from the respective samples were pooled and assembled
345 with MEGAHIT[66] v1.1.1.2. Assembling of the Nanopore data was done for each sample

346 individually with Flye[67] v2.7.1 in metagenome mode. Contigs were corrected based on Illumina
347 data from the respective sample with bowtie[68] v2.3.5, SAMtools[69] v1.9, and pilon[70] v1.23. Quality
348 assessment of the (co-)assemblies was obtained with metaQUAST[71] v5.0.2.

349 **Metagenome binning**

350 Binning was performed separately for each Illumina and Nanopore (co-)assembly with anvi'o[30]
351 v6.2 after discarding contigs shorter than 2500 bp. Gene calls were predicted with prodigal[72]
352 v2.6.3. Single-copy genes were identified with HMMER[73] v.3.2.1 and classified with
353 DIAMOND[63] v0.9.14 against the GTDB[31,32] release 04-RS89. Illumina reads were mapped to the
354 contigs with bowtie[68] v2.3.5 and SAM files were sorted and indexed using SAMtools[69] v1.9. Due
355 to their large sizes, Illumina co-assemblies were split into 100 smaller clusters based on
356 differential coverage and tetranucleotide frequency with CONCOCT[74] v1.0.0. Contigs were then
357 manually sorted into bins based on the same composition and coverage metrics using the 'anvi-
358 interactive' interface in anvi'o[30] v6.2. Nanopore contigs were binned directly without pre-
359 clustering with CONCOCT. Bins that were ≥ 50% complete according to the presence of single-
360 copy genes were further refined using the 'anvi-refine' interface in anvi'o[30] v6.2. In addition to
361 taxonomic signal (based on single-copy genes classified against GTDB), either differential
362 coverage or tetranucleotide frequency was used to identify and remove outlying contigs. The
363 former was used for bins presenting a large variation in contig coverage across samples, and
364 the latter for those that showed marked differences in GC content across contigs. Bins ≥ 50%
365 complete and ≤ 10% redundant were renamed as MAGs.

366 **Hybrid re-assembling**

367 Illumina and Nanopore bins were matched based on a 90% average nucleotide identity (ANI)
368 threshold with fastANI[75] v1.3. Pairs of bins sharing ≥ 90% ANI were then subjected to a hybrid
369 re-assembling approach with Unicycler[76] v0.4.8. For this, Illumina reads mapping to the
370 respective Illumina and Nanopore bins were retrieved with SAMtools[69] v1.9 and seqtk[60] v1.3
371 and assembled with SPAdes[77] v3.13.0. The SPAdes and Nanopore contigs were then assembled
372 with miniasm[78] v0.3 and the resulting assembly was polished based on the Illumina reads with
373 pilon[70] v1.23. Assemblies were processed with anvi'o as described above, with the exception that
374 resulting bins were refined based on tetranucleotide frequency only.

375 **MAG dereplication and read recruitment analysis**

376 To remove redundancy (i.e. MAGs that were recovered multiple times across the different
377 assemblies), Illumina and Nanopore MAGs were dereplicated based on a 90% ANI threshold
378 with fastANI[75] v1.3. Non-redundant MAGs were replaced by their respective 'hybrid' MAG if
379 the latter was of better quality (i.e. more complete, less redundant, with higher N_{50} value, and/or
380 included the 16S rRNA gene). Illumina reads were mapped to the set of non-redundant MAGs
381 with bowtie[68] v2.3.5 and SAM files were sorted and indexed using SAMtools[69] v1.9. MAG
382 detection and abundance was computed based on the nucleotide coverage across the Illumina
383 reads. A MAG was considered detected in a given sample if ≥ 50% of its nucleotides had ≥ 1x
384 coverage. Relative abundances were computed based on the number of reads recruited by each
385 MAG normalised by the number of quality-filtered reads in each sample.

386 **MAG annotation**

387 Genome annotation was performed based on predicted amino acid sequences. Sequences were
388 annotated against the COG database[33] release 12/2014 and the KEGG database[34] release 86
389 with DIAMOND[63] v0.9.14, and the KOfam database[35] release 11/2020 with HMMER[73] v.3.2.1.
390 A maximum e-value of 10^{-5} was used for the DIAMOND-based annotations, and pre-computed
391 family-specific thresholds were used for the HMMER-based annotation. Sequences assigned to
392 denitrifying reductases were confirmed by similarity searches against the RefSeq[36] and Swiss-
393 Prot databases using blastp[79] v2.9.0. Annotation of denitrifying reductases was also performed
394 for previously published genomes retrieved from GenBank. These included a set of 1529 MAGs
395 obtained from soils in Stordalen Mire, northern Sweden[24], and all genomes of Acidobacteriota
396 strains and candidate taxa.

397 **Phylogenomic analyses**

398 Phylogenetic assignment of MAGs was performed against the GTDB[31,32] release 05-RS95 based
399 on 122 archaeal and 120 bacterial single-copy genes with GTDB-Tk[80] v1.3.0. Acidobacteriota
400 MAGs encoding denitrifying reductases were submitted to further phylogenomic analyses
401 alongside the genomes of Acidobacteriota strains and candidate taxa retrieved from GenBank.
402 For this, the amino acid sequence of 23 ribosomal proteins was retrieved for each genome with
403 anvi'o[30] v6.2 and aligned with MUSCLE[81] v3.8.1551. A maximum likelihood tree was then
404 computed based on the concatenated alignments with FastTree[82] v2.1.10.

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.

References

1. Tian, H. *et al.* A comprehensive quantification of global nitrous oxide sources and sinks. *Nature* **586**, 248–256 (2020).
2. Repo, M. E. *et al.* Large N_2O emissions from cryoturbated peat soil in tundra. *Nat. Geosci.* **2**, 189–192 (2009).
3. Stewart, K. J., Grogan, P., Coxson, D. S. & Siciliano, S. D. Topography as a key factor driving atmospheric nitrogen exchanges in arctic terrestrial ecosystems. *Soil Biol. Biochem.* **70**, 96–112 (2014).
4. Voigt, C. *et al.* Increased nitrous oxide emissions from Arctic peatlands after permafrost thaw. *Proc. Natl. Acad. Sci.* **114**, 6238–6243 (2017).
5. Voigt, C. *et al.* Nitrous oxide emissions from permafrost-affected soils. *Nat. Rev. Earth Environ.* **1**, 420–434 (2020).
6. Hugelius, G. *et al.* Large stocks of peatland carbon and nitrogen are vulnerable to permafrost thaw. *Proc. Natl. Acad. Sci.* **117**, 20438–20446 (2020).
7. Schuur, E. A. G. *et al.* Climate change and the permafrost carbon feedback. *Nature* **520**, 171–179 (2015).
8. Post, E. *et al.* The polar regions in a 2°C warmer world. *Sci. Adv.* **5**, eaaw9883 (2019).
9. Butterbach-Bahl, K., Baggs, E. M., 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.* **368**, 20130122 (2013).
10. Zumft, W. G. Cell biology and molecular basis of denitrification. *Microbiol. Mol. Biol. Rev.* **61**, 533–616 (1997).

11. Graf, D. R. H., Jones, C. M. & Hallin, S. Intergenomic comparisons highlight modularity of the denitrification pathway and underpin the importance of community structure for N_2O emissions. *PLoS ONE* **9**, e114118 (2014).

12. Wallenstein, M. D., Myrold, D. D., Firestone, M. & Voytek, M. Environmental controls on denitrifying communities and denitrification rates: insights from molecular methods. *Ecol. Appl.* **16**, 2143–2152 (2006).

13. Kou, D. *et al.* Progressive nitrogen limitation across the Tibetan alpine permafrost region. *Nat. Commun.* **11**, 3331 (2020).

14. Liu, X.-Y. *et al.* Nitrate is an important nitrogen source for Arctic tundra plants. *Proc. Natl. Acad. Sci.* **115**, 3398–3403 (2018).

15. Yergeau, E., Kang, S., He, Z., Zhou, J. & Kowalchuk, G. A. Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitudinal transect. *ISME J.* **1**, 163–179 (2007).

16. Yergeau, E., Hogues, H., Whyte, L. G. & Greer, C. W. The functional potential of high Arctic permafrost revealed by metagenomic sequencing, qPCR and microarray analyses. *ISME J.* **4**, 1206–1214 (2010).

17. Palmer, K., Biasi, C. & Horn, M. A. Contrasting denitrifier communities relate to contrasting N_2O emission patterns from acidic peat soils in arctic tundra. *ISME J.* **6**, 1058–1077 (2012).

18. 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.* **11**, 573302 (2020).

19. Yu, T. & Zhuang, Q. Quantifying global N_2O emissions from natural ecosystem soils using trait-based biogeochemistry models. *Biogeosciences* **16**, 207–222 (2019).

20. Rappé, M. S. & Giovannoni, S. J. The uncultured microbial majority. *Annu. Rev. Microbiol.* **57**, 369–394 (2003).

21. Steen, A. D. *et al.* High proportions of bacteria and archaea across most biomes remain uncultured. *ISME J.* **13**, 3126–3130 (2019).

22. Mackelprang, R. *et al.* Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. *Nature* **480**, 368–371 (2011).

23. Hultman, J. *et al.* Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature* **521**, 208–212 (2015).

24. Woodcroft, B. J. *et al.* Genome-centric view of carbon processing in thawing permafrost. *Nature* **560**, 49–54 (2018).

25. le Roux, P. C., Aalto, J. & Luoto, M. Soil moisture's underestimated role in climate change impact modelling in low-energy systems. *Glob. Change Biol.* **19**, 2965–2975 (2013).

26. Niittynen, P. *et al.* Fine-scale tundra vegetation patterns are strongly related to winter thermal conditions. *Nat. Clim. Change* **10**, 1143–1148 (2020).

27. Viitamäki, S. The activity and functions of soil microbial communities across a climate gradient in Finnish subarctic. Master's thesis (University of Helsinki, 2019).

28. le Roux, P. C., Pellissier, L., Wisz, M. S. & Luoto, M. Incorporating dominant species as proxies for biotic interactions strengthens plant community models. *J. Ecol.* **102**, 767–775 (2014).

29. Kemppinen, J., Niittynen, P., Aalto, J., le Roux, P. C. & Luoto, M. Water as a resource, stress and disturbance shaping tundra vegetation. *Oikos* **128**, 811–822 (2019).

30. Eren, A. M. *et al.* Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**, e1319 (2015).

31. Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004 (2018).

32. Parks, D. H. *et al.* A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat. Biotechnol.* **38**, 1079–1086 (2020).

33. Galperin, M. Y., Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* **43**, D261–D269 (2015).

34. Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

35. Aramaki, T. *et al.* KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **36**, 2251–2252 (2020).

36. O'Leary, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–D745 (2016).

37. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).

38. Delgado-Baquerizo, M. *et al.* A global atlas of the dominant bacteria found in soil. *Science* **359**, 320–325 (2018).

39. Shiro, Y. Structure and function of bacterial nitric oxide reductases. *Biochim. Biophys. Acta - Bioenerg.* **1817**, 1907–1913 (2012).

40. Männistö, M. K., Kurhela, E., Tiirola, M. & Häggblom, M. M. Acidobacteria dominate the active bacterial communities of Arctic tundra with widely divergent winter-time snow accumulation and soil temperatures. *FEMS Microbiol. Ecol.* **84**, 47–59 (2013).

41. Kalam, S. *et al.* Recent understanding of soil Acidobacteria and their ecological significance: a critical review. *Front. Microbiol.* **11**, 580024 (2020).

42. Jones, C. M., Stres, B., Rosenquist, M. & Hallin, S. Phylogenetic analysis of nitrite, nitric oxide, and nitrous oxide respiratory enzymes reveal a complex evolutionary history for denitrification. *Mol. Biol. Evol.* **25**, 1955–1966 (2008).

43. Sanford, R. A. *et al.* Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils. *Proc. Natl. Acad. Sci.* **109**, 19709–19714 (2012).

44. Ortiz, M. *et al.* Microbial nitrogen cycling in Antarctic soils. *Microorganisms* **8**, 1442 (2020).

45. Makhalanyane, T. P., Van Goethem, M. W. & Cowan, D. A. Microbial diversity and functional capacity in polar soils. *Curr. Opin. Biotechnol.* **38**, 159–166 (2016).

46. Kielak, A. M., Barreto, C. C., Kowalchuk, G. A., van Veen, J. A. & Kuramae, E. E. The ecology of Acidobacteria: moving beyond genes and genomes. *Front. Microbiol.* **7**, (2016).

47. Eichorst, S. A. *et al.* Genomic insights into the Acidobacteria reveal strategies for their success in terrestrial environments. *Environ. Microbiol.* **20**, 1041–1063 (2018).

48. Lycus, P. *et al.* Phenotypic and genotypic richness of denitrifiers revealed by a novel isolation strategy. *ISME J.* **11**, 2219–2232 (2017).

49. Philippot, L., Andert, J., Jones, C. M., Bru, D. & Hallin, S. Importance of denitrifiers lacking the genes encoding the nitrous oxide reductase for N_2O emissions from soil. *Glob. Change Biol.* **17**, 1497–1504 (2011).

50. Jones, C. M. *et al.* Recently identified microbial guild mediates soil N_2O sink capacity. *Nat. Clim. Change* **4**, 801–805 (2014).

51. Jones, C. M., Graf, D. R., Bru, D., Philippot, L. & Hallin, S. The unaccounted yet abundant nitrous oxide-reducing microbial community: a potential nitrous oxide sink. *ISME J.* **7**, 417–426 (2013).

525 52. Schimel, J. Microbial ecology: Linking omics to biogeochemistry. *Nat. Microbiol.* **1**, 15028
526 (2016).

527 53. Bakken, L. R., Bergaust, L., Liu, B. & Frostegård, Å. Regulation of denitrification at the
528 cellular level: a clue to the understanding of N_2O emissions from soils. *Philos. Trans. R. Soc.*
529 *B Biol. Sci.* **367**, 1226–1234 (2012).

530 54. Pirinen, P. *et al. Climatological statistics of Finland 1981–2010.* (Finnish Meteorological
531 Institute, 2012).

532 55. Andrews, S. *FastQC: A Quality Control Tool for High Throughput Sequence Data.* (2019).
533 Available at https://www.bioinformatics.babraham.ac.uk/projects/fastqc.

534 56. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for
535 multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).

536 57. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
537 *EMBnet.journal* **17**, 10 (2011).

538 58. Leger, A. & Leonardi, T. pycoQC, interactive quality control for Oxford Nanopore
539 Sequencing. *J. Open Source Softw.* **4**, 1236 (2019).

540 59. Wick, R. R. *Porechop*. (2018). Available at https://github.com/rrwick/Porechop.

541 60. Li, H. *seqtk*. (2018). Available at https://github.com/lh3/seqtk.

542 61. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing
543 and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).

544 62. Bengtsson-Palme, J. *et al.* METAXA 2: improved identification and taxonomic classification
545 of small and large subunit rRNA in metagenomic data. *Mol. Ecol. Resour.* **15**, 1403–1414
546 (2015).

547 63. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND.
548 *Nat. Methods* **12**, 59–60 (2015).

549 64. Pessi, I. S. *keggR.* (2020). Available at https://github.com/igorspp/keggR.

550 65. Oksanen, J. *et al. vegan: community ecology package.* (2019). Available at https://cran.r-
551 project.org/web/packages/vegan.

552 66. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: an ultra-fast single-node
553 solution for large and complex metagenomics assembly via succinct de Bruijn graph.
554 *Bioinformatics* **31**, 1674–1676 (2015).

555 67. Kolmogorov, M. *et al.* metaFlye: scalable long-read metagenome assembly using repeat
556 graphs. *Nat. Methods* **17**, 1103–1110 (2020).

557 68. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*
558 **9**, 357–359 (2012).

559 69. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–
560 2079 (2009).

561 70. Walker, B. J. *et al.* Pilon: an integrated tool for comprehensive microbial variant detection
562 and genome assembly improvement. *PLoS ONE* **9**, e112963 (2014).

563 71. Mikheenko, A., Saveliev, V. & Gurevich, A. MetaQUAST: evaluation of metagenome
564 assemblies. *Bioinformatics* **32**, 1088–1090 (2016).

565 72. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site
566 identification. *BMC Bioinformatics* **11**, 119 (2010).

567 73. Eddy, S. R. Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).

568 74. Alneberg, J. *et al.* Binning metagenomic contigs by coverage and composition. *Nat. Methods*
569 **11**, 1144–1146 (2014).

570 75. Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High
571 throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat.*
572 *Commun.* **9**, 5114 (2018).

573 76. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: resolving bacterial genome
574 assemblies from short and long sequencing reads. *PLOS Comput. Biol.* **13**, e1005595 (2017).

575 77. Bankevich, A. *et al.* SPAdes: A new genome assembly algorithm and its applications to
576 single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).

577 78. Li, H. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences.
578 *Bioinformatics* **32**, 2103–2110 (2016).

579 79. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421
580 (2009).

581 80. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: a toolkit to classify
582 genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 1925–1927 (2019).

583 81. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high
584 throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

585 82. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 – approximately maximum-likelihood
586 trees for large alignments. *PLoS ONE* **5**, e9490 (2010).

587 Acknowledgements

588 This work was funded by the Academy of Finland (grants 314114 and 335354) and the
589 University of Helsinki. SV was funded by the Microbiology and Biotechnology Doctoral
590 Programme (MBDP). We would like to acknowledge the CSC – IT Centre for Science for
591 providing the requisite computing resources and Mr. Kimmo Mattila for IT support; the staff
592 from the Kilpisjärvi Biological Station, Tanja Orpana, Aino Rutanen, Anniina Sarekoski, Anna-
593 Maria Virkkala, and Miska Luoto's group members for assistance with fieldwork and soil
594 characterization; Christina Biasi and Maija Marushchak for helpful discussion; Laura
595 Cappelatti for revising the manuscript and assistance with figures; and Murat Eren, Sebastian
596 Lücker, Donovan Parks, and Antonios Kioukis for tips, recommendations, and troubleshooting
597 regarding the bioinformatics analyses.

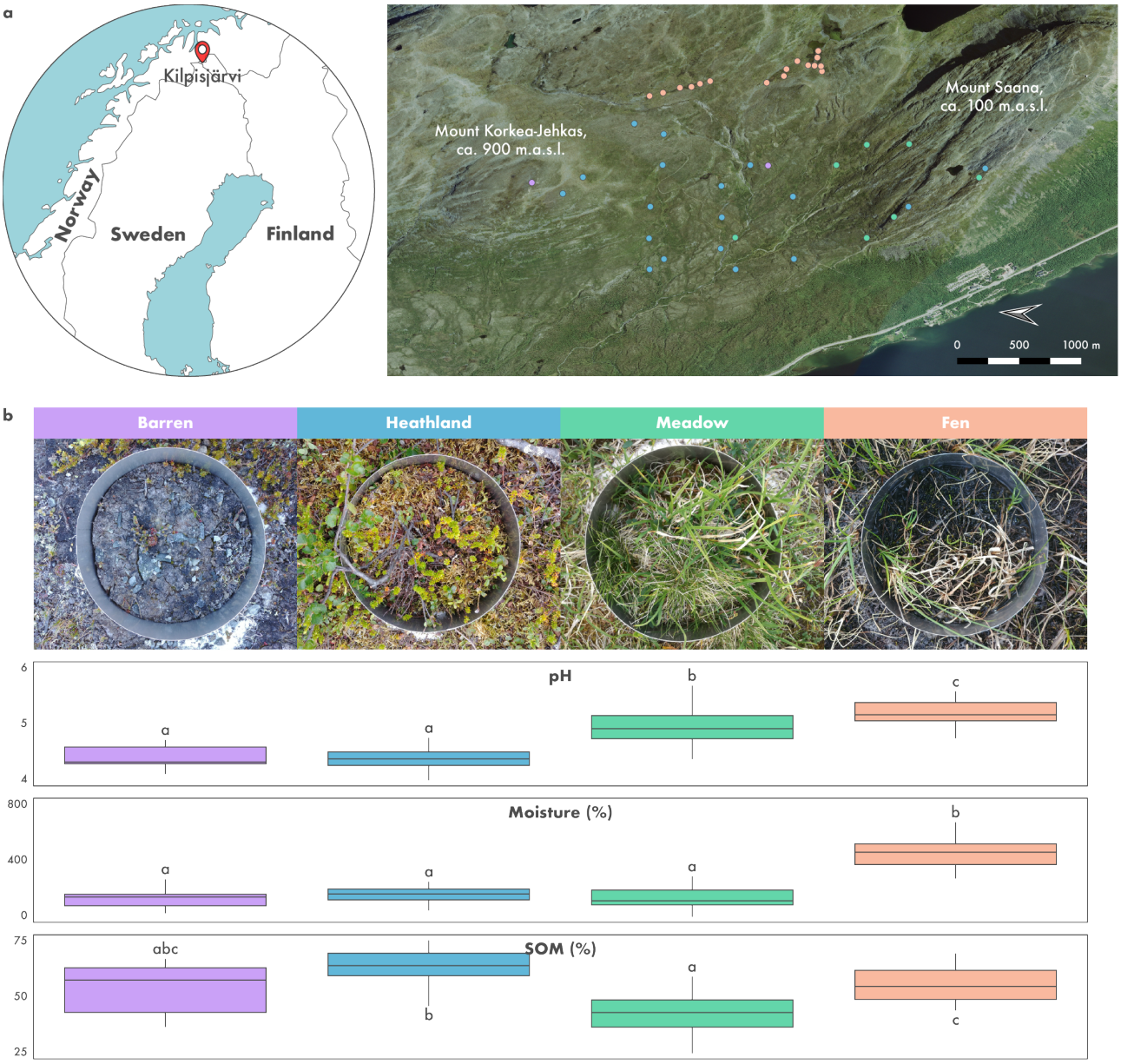
598 Author contributions

599 JH and ML designed the research; SV and JH performed nucleic acid extraction and library
600 preparation; ISP analysed the data and wrote the manuscript; EER and TOD contributed with
601 the analyses; all authors contributed to the final version of the manuscript.

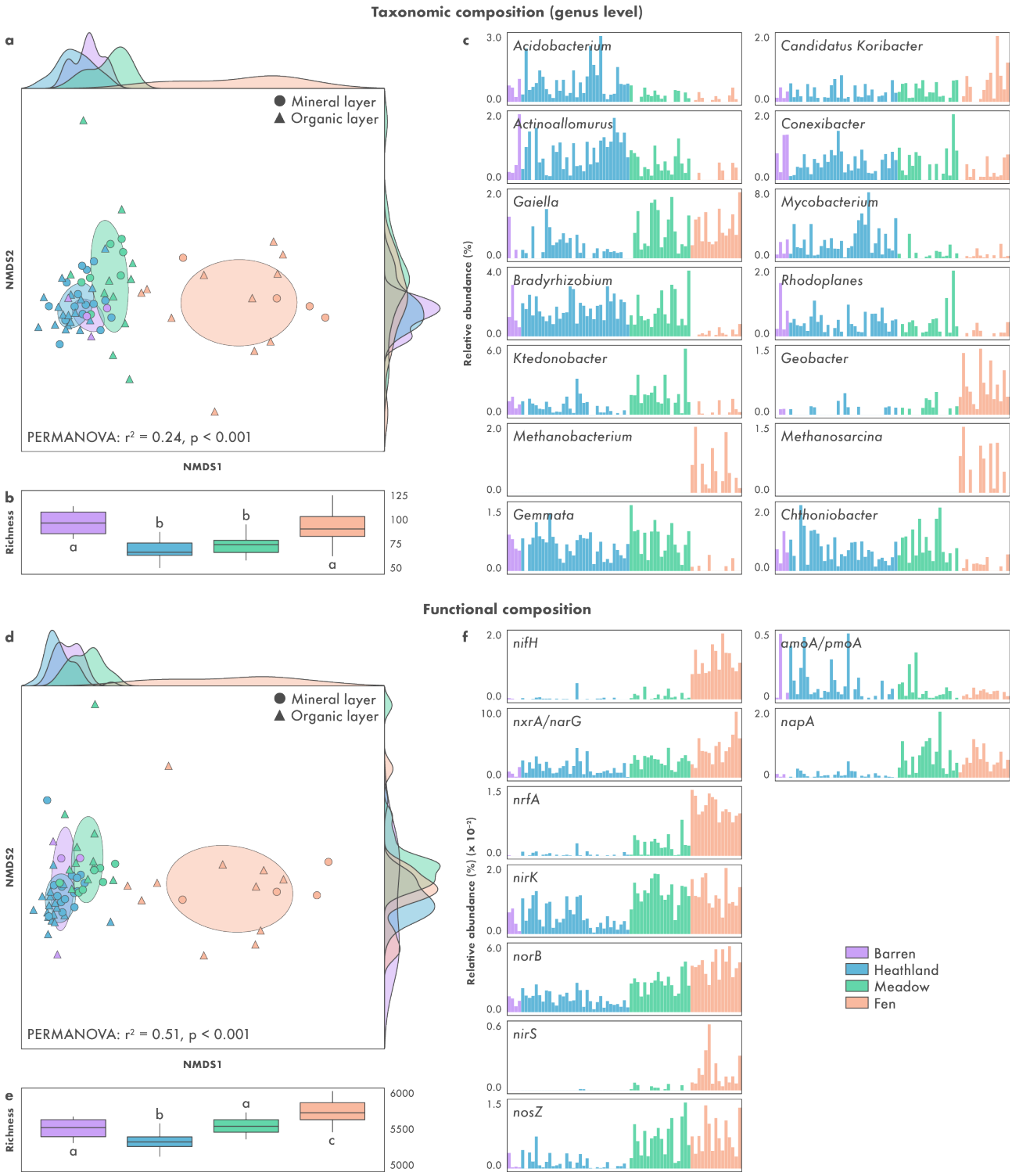
602 Competing interests

603 The authors declare no competing interests.

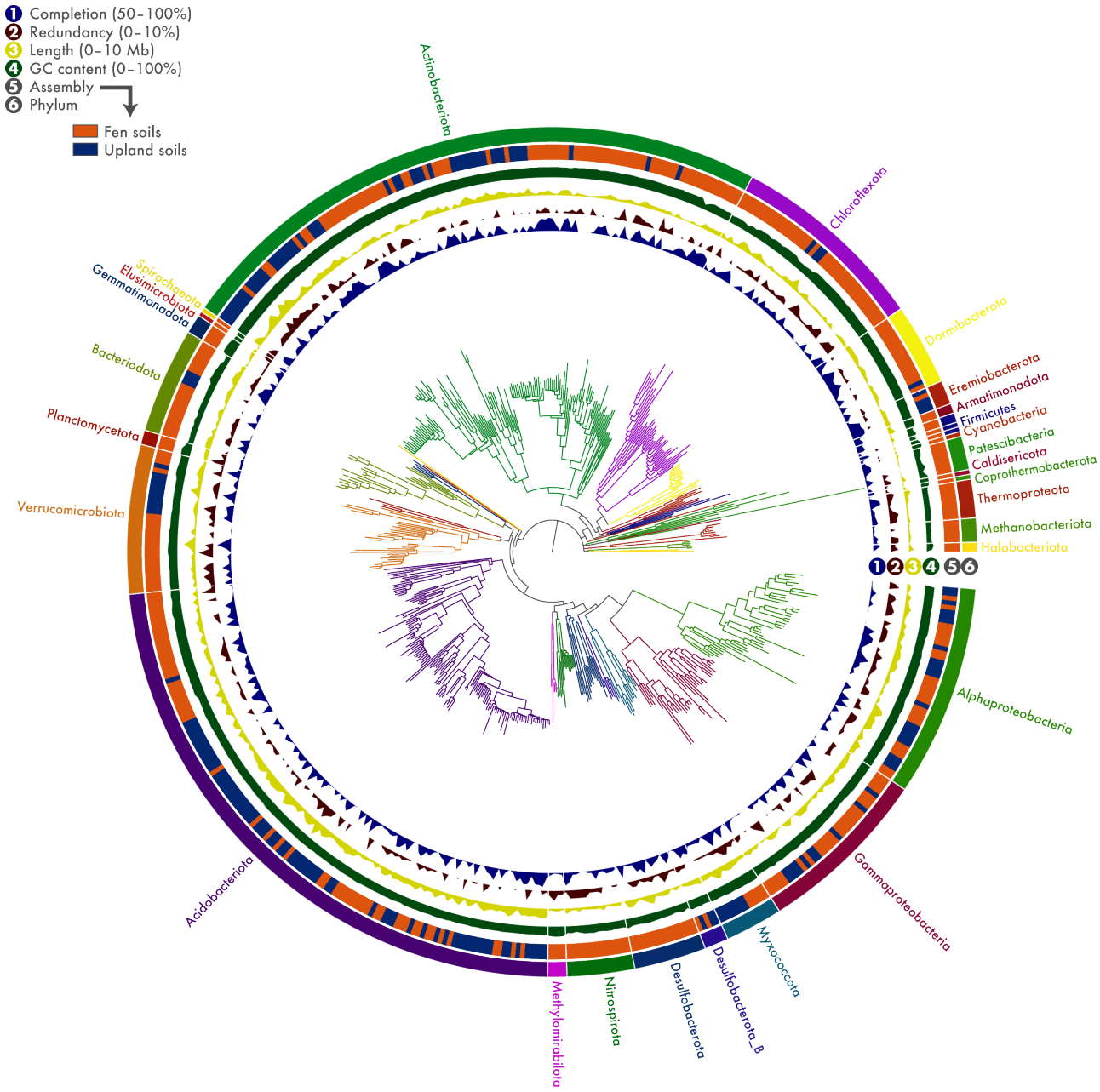
Supplementary Figures



Suppl. Fig. S1 | Saana Nature Reserve, an area of mountain tundra in Kilpisjärvi, northern Finland. a) Map of Fennoscandia showing the location of Kilpisjärvi and topographic overview of the study area. **b)** Characterization of the four types of soil ecosystems investigated. Ecosystems that do not share the same letter are significantly different (one-way ANOVA followed by Tukey's HSD test; $p < 0.001$).

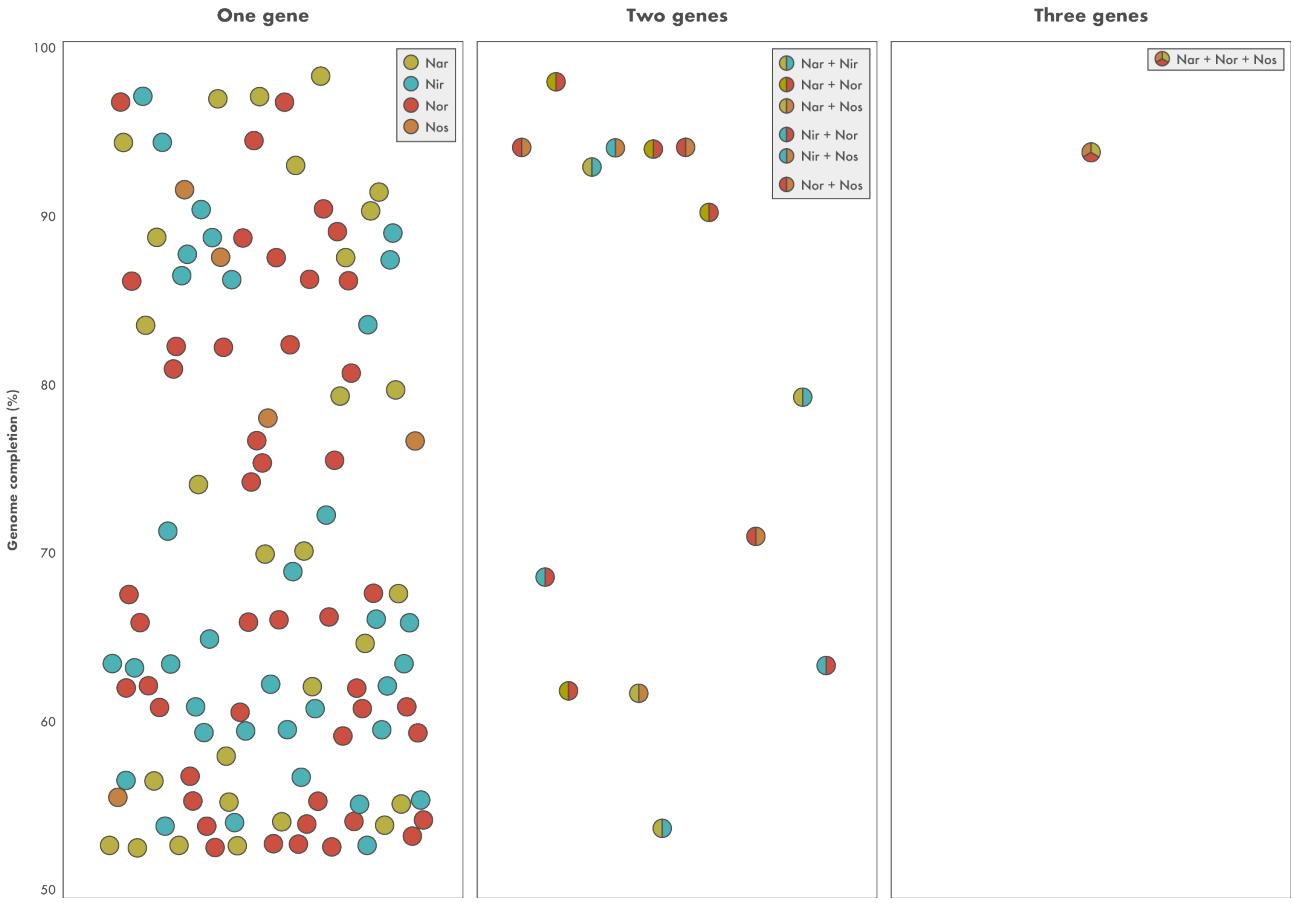

Taxonomic composition (genus level)
a
Mineral layer
Organic layer
NMDS2
NMDS1
PERMANOVA: r² = 0.24, p < 0.001
b
Richness
a
b
b
a
125
100
75
50
c
Relative abundance (%)
Acidobacterium
Candidatus Koribacter
Actinoallomurus
Conexibacter
Gaiella
Mycobacterium
Bradyrhizobium
Rhodoplanes
Ktedonobacter
Geobacter
Methanobacterium
Methanosarcina
Gemmata
Chthoniobacter
Functional composition
d
Mineral layer
Organic layer
NMDS2
NMDS1
PERMANOVA: r² = 0.51, p < 0.001
e
Richness
a
b
a
c
6000
5500
5000
f
Relative abundance (%) (x 10⁻²)
nifH
amoA/pmoA
nxrA/narG
napA
nrfA
nirK
norB
nirS
nosZ
Barren
Heathland
Meadow
Fen

610 **(Previous page) Suppl. Fig. S2 | The microbial diversity of Kilpisjärvi soils as seen**
611 **using a gene-centric approach.** Gene-centric analyses were carried out using unassembled
612 Illumina data. Taxonomic composition was computed based on the annotation of 16S rRNA gene
613 fragments against the SILVA database[61] with METAXA[62]. Functional annotation was
614 performed by blastx searches against the KEGG database[34] with DIAMOND[63]. **a, d)** Non-metric
615 multidimensional scaling (NMDS) of taxonomic and functional community structure.
616 Differences between the ecosystems were assessed using permutational ANOVA
617 (PERMANOVA). **b, e)** Taxonomic and functional richness (number of genera and functional
618 genes). Differences in richness were assessed using one-way ANOVA followed by Tukey's HSD
619 test. Ecosystems that do not share the same letter are significantly different ($p < 0.001$).
620 **c, f)** Abundance profile of the five most abundant genera in each ecosystem and marker genes
621 for the different steps of the N cycle.

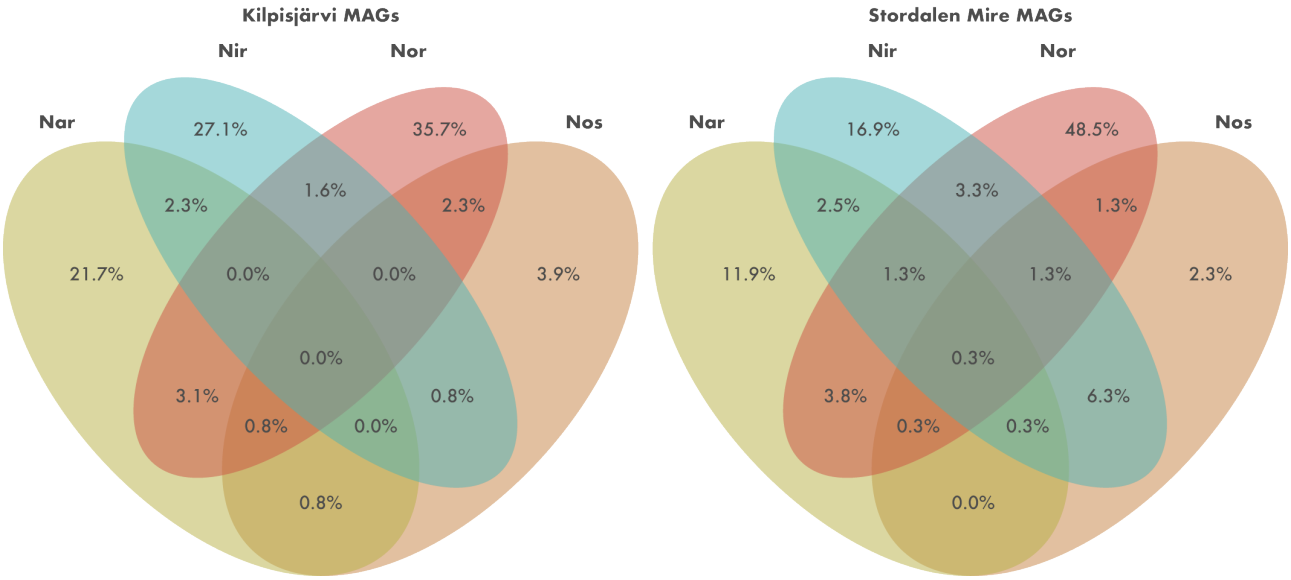


Suppl. Fig. S3 | Five hundred thirty-nine metagenome assembled genomes (MAGs) from tundra soils. Phylogenomic analysis of Kilpisjärvi MAGs based on concatenated alignments of amino acid sequences from 122 archaeal and 120 bacterial single-copy genes. MAGs completion and redundancy were computed based on the presence of single-copy genes.

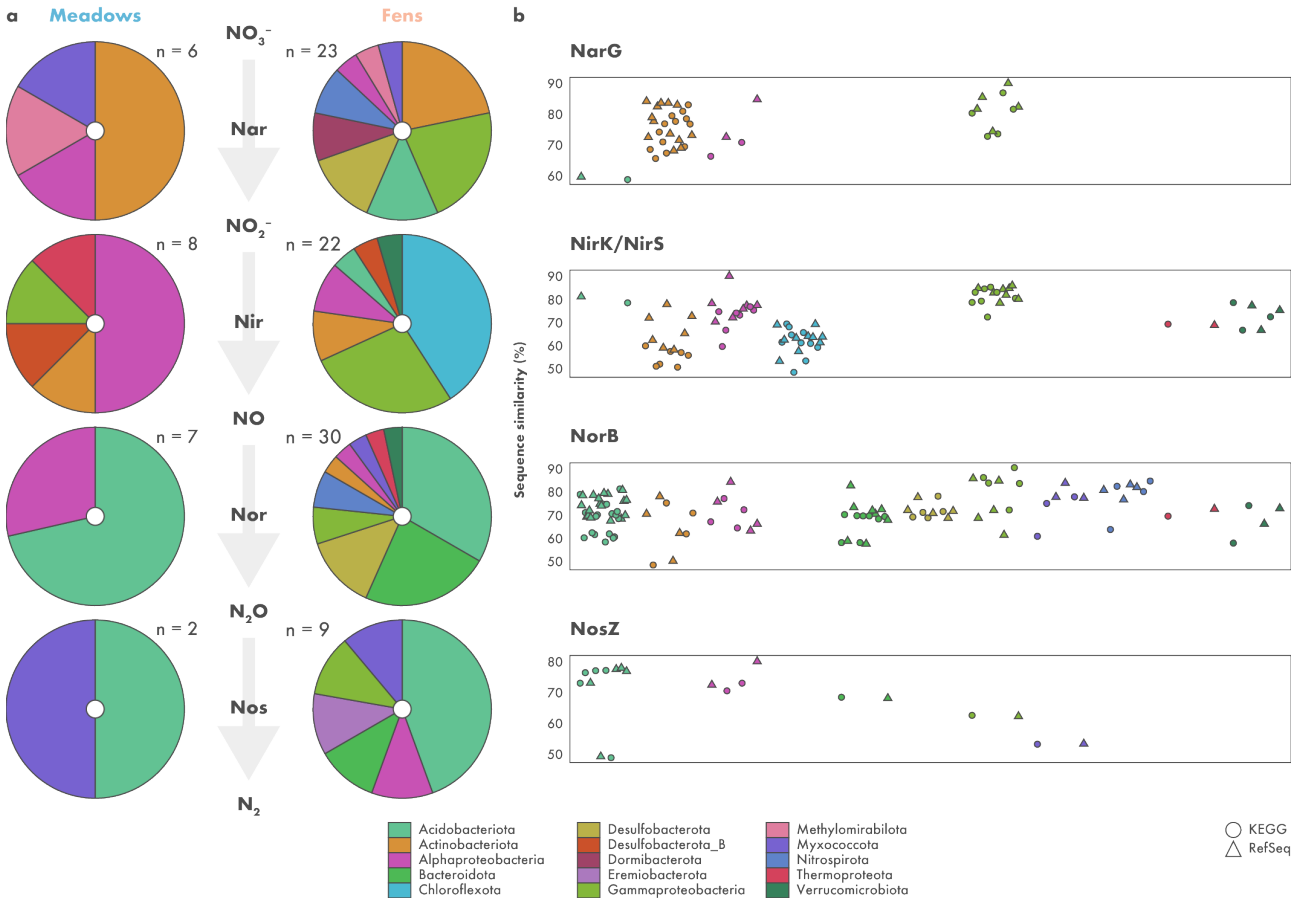
Suppl. Fig. S4 | Overview of the microbial diversity of Kilpisjärvi soils based on a genome- resolved approach. a) Number of detected MAGs across the ecosystems**.** A MAG was detected in a given sample if ≥ 50% of its nucleotides had ≥ 1x coverage. **b)** Detection profiles of MAGs shared between heathland, meadow, and fen soils. **c)** Relative abundance of the ten most abundant MAGs in each ecosystem.



Suppl. Fig. S5 | Relationship between MAG completion and number of denitrifying reductases. MAGs completion was computed based on the presence of single-copy genes. The presence of nitrate (Nar), nitrite (Nir), nitric oxide (Nor), and nitrous oxide (Nos) reductases was inferred based on the annotation of predicted amino acid sequences.



635 **Suppl. Fig. S6 | Percentage of MAGs encoding denitrifying reductases in Kilpisjärvi**
636 **and Stordalen Mire soils.** The presence of nitrate (Nar), nitrite (Nir), nitric oxide (Nor), and
637 nitrous oxide (Nos) reductases was inferred based on the annotation of predicted amino acid
638 sequences.



Suppl. Fig. S7 | Community structure and identity of NarG, NirK/NirS, NorB, and NosZ sequences. a) Taxonomic assignment of MAGs encoding each of the denitrifying reductases in meadow and fen soils. **b)** Sequence similarity of denitrifying reductases encoded by the MAGs with reference sequences in the KEGG and RefSeq databases.