Candidatus Nitrosopolaris, a genus of putative ammonia-oxidizing archaea with a polar/alpine distribution

Igor S. Pessi[1,2], Aino Rutanen[1], and Jenni Hultman[1,2,3*]

[1]Department of Microbiology, University of Helsinki, Helsinki, Finland
[2]Helsinki Institute of Sustainability Science (HELSUS), Helsinki, Finland
[3]Natural Resources Institute Finland (LUKE), Helsinki, Finland
[*]Corresponding author: jenni.hultman@helsinki.fi

Abstract

Ammonia-oxidizing archaea (AOA) are key players in the nitrogen cycle. Here, we report four novel metagenome-assembled genomes (MAGs) assigned to the genus "UBA10452", an uncultured lineage of putative AOA in the family Nitrososphaeraceae. Analysis of other eight previously reported MAGs and publicly available amplicon sequencing data revealed that the UBA10452 lineage is predominantly found in acidic polar and alpine soils. We propose a novel *Candidatus* genus, *Ca.* Nitrosopolaris, with four species representing clear biogeographical/habitat clusters. In addition to the presence of genes involved in cold adaptation, we hypothesize that the distribution of *Ca.* Nitrosopolaris across the cold biosphere might also be due to the combination of a polar origin and limited dispersal capabilities throughout geological time.

Introduction

Nitrification – the oxidation of ammonia to nitrite and further oxidation to nitrate – is a crucial part of the nitrogen (N) cycle providing a link between reduced and oxidized forms of N. The first step of nitrification, ammonia oxidation, is carried out mainly by aerobic chemolithoautotrophic microorganisms that grow by coupling the energy obtained from the oxidation of ammonia with carbon dioxide (CO_2) fixation (Lehtovirta-Morley, 2018). For more than a century after the isolation of *Nitrosomonas europaea* (Winogradsky, 1890), this process was thought to be carried out exclusively by ammonia-oxidizing bacteria (AOB). However, metagenomic investigations in the early 21st century revealed the presence of ammonia oxidation genes within uncultivated archaea (Treusch *et al.*, 2004; Venter *et al.*, 2004). The first cultured representative of ammonia-oxidizing archaea (AOA), *Nitrosopumilus maritimus*, was

obtained shortly after (Könneke *et al.*, 2005). It is now well established that AOA outnumber AOB by orders of magnitude in many terrestrial and aquatic environments, particularly in oligotrophic environments with low N input (Leininger *et al.*, 2006; Schleper and Nicol, 2010; Lehtovirta-Morley, 2018). Among the reasons for their ecological success is an enzymatic machinery with higher affinity for ammonia and a more efficient CO_2 fixation pathway than their bacterial counterparts (Martens-Habbena *et al.*, 2009; Könneke *et al.*, 2014; Kerou *et al.*, 2016).

AOA were originally described as members of the phylum Crenarchaeota and later accommodated into a new phylum, the Thaumarchaeota (Brochier-Armanet *et al.*, 2008; Spang *et al.*, 2010). In the Genome Taxonomy Database (GTDB), Thaumarchaeota are represented as the class Nitrososphaeria in the phylum Thermoproteota (Rinke *et al.*, 2021). AOA are notoriously difficult to cultivate and so far only three genera have been formally described based on axenic cultures: *Nitrosopumilus* (Qin *et al.*, 2017) and *Nitrosarchaeum* (Jung *et al.*, 2018) in the order Nitrosopumilales ("marine group I.1a") and *Nitrososphaera* (Stieglmeier *et al.*, 2014) in the order Nitrososphaerales ("terrestrial group I.1b"). Several provisional *Candidatus* genera have also been proposed based on non-axenic enrichments, e.g. *Ca.* Nitrosocaldus (de la Torre *et al.*, 2008), *Ca.* Nitrosotalea (Lehtovirta-Morley *et al.*, 2011), and *Ca.* Nitrosocosmicus (Lehtovirta-Morley *et al.*, 2016). Moreover, the growing use of genome-resolved metagenomics in recent years has resulted in the identification of tens of novel, currently uncultured lineages in the class Nitrososphaeria (Rinke *et al.*, 2021). These lineages are phylogenetically distinct from both formally described and *Candidatus* taxa and are identified with so-called placeholder names (e.g., the genus "UBA10452" in the family Nitrososphaeraceae). The identification of these novel lineages by metagenomics greatly expands our knowledge of the diversity of AOA but detailed descriptions of their metabolic and ecological features are generally lacking.

Ammonia oxidation is an important process in polar soils despite commonly N limited and cold conditions (Alves *et al.*, 2013; Siljanen *et al.*, 2019; Hayashi *et al.*, 2020). AOA generally outnumber AOB in oligotrophic polar soils and are often represented by a very limited number of species (Alves *et al.*, 2013; Magalhães *et al.*, 2014; Richter *et al.*, 2014; Siljanen *et al.*, 2019; Ortiz *et al.*, 2020; Pessi *et al.*, 2021). Due to their predominance, AOA are important contributors to the N cycle in polar soils and are thus key players in the cycling of the potent greenhouse gas nitrous oxide (N_2O). Contrary to earlier assumptions, polar soils are increasingly recognized as important sources of N_2O (Voigt *et al.*, 2020). Both the nitrite originated from the oxidation of ammonia as well as the nitrate produced in the second step of nitrification are the substrates of denitrification, an anaerobic process that has N_2O as a gaseous intermediate (Butterbach-Bahl *et al.*, 2013). Moreover, AOA have been directly implicated in the production of N_2O under oxic

conditions via several mechanisms such as hydroxylalamine oxidation and nitrifier denitrification (Wu *et al.*, 2020). However, both the direct and indirect role of AOA in the cycling of N_2O is much less understood compared to their bacterial counterparts.

Recently, we have applied a genome-resolved metagenomics approach to gain insights into the microorganisms involved with the cycling of greenhouse gases in tundra soils from Kilpisjärvi, Finland (Pessi *et al.*, 2021). Analysis of *amoA* genes encoding the alpha subunit of the enzyme ammonia monooxygenase (Amo) revealed a very low diversity of ammonia oxidizers, with only four genes annotated as *amoA* out of 23.5 million assembled genes. Three of these were most closely related to the *amoA* gene of the comammox bacterium *Ca.* Nitrospira inopinata (Daims *et al.*, 2015). The remaining *amoA* gene was binned into a metagenome-assembled genome (MAG) assigned to the genus "UBA10452", an uncharacterized archaeal lineage in the class Nitrososphaeria (former phylum Thaumarchaeota) (Rinke *et al.*, 2021). Here, we i) report four novel UBA10452 MAGs obtained from tundra soils in Rásttigáisá, Norway; ii) characterize the phylogenetic, genomic, metabolic, and biogeographic properties of the UBA10452 lineage; and iii) propose the creation of a new *Candidatus* genus, *Ca.* Nitrosopolaris.

Methods

Genomic dataset

We analysed 12 MAGs belonging to the genus "UBA10452" in the order Nitrososphaerales ("terrestrial group I.1b"), class Nitrososphaeria (Rinke *et al.*, 2021) **(Table 1, Suppl. Table S1)**. Of these, six MAGs from permafrost soils in Canada (Chauhan *et al.*, 2014; Parks *et al.*, 2017) and one MAG from polar desert soils in Antarctica (Ji *et al.*, 2017) were retrieved from GTDB release 95 (Parks *et al.*, 2018, 2020), and one MAG from tundra soils in Finland (Pessi *et al.*, 2021) was retrieved from GenBank. The remaining four MAGs were binned from tundra soils in Rásttigáisá, Norway, as described in Rutanen (2020). Briefly, metagenomes were obtained with Illumina NextSeq from 10 soil samples taken from a 10–15 cm depth in July 2017. Metagenomes were assembled individually and as one co-assembly with MEGAHIT v1.1.1 (Li *et al.*, 2015) and manually binned with anvi'o v6.0 (Eren *et al.*, 2015) according to Pessi *et al.* (2021).

Table 1. List of metagenome-assembled genomes (MAGs) belonging to the UBA10452 lineage (*Candidatus* Nitrosopolaris).

MAG	Isolation source	Accession	Ref.
COA_Bin_4_1	Tundra soils, Rásttigáisá, Norway	PRJEB49283 *	[1]
S89_Bin_2	Tundra soils, Rásttigáisá, Norway	PRJEB49283 *	[1]
S100_Bin_4	Tundra soils, Rásttigáisá, Norway	PRJEB49283 *	[1]
S1130_Bin_3	Tundra soils, Rásttigáisá, Norway	PRJEB49283 *	[1]
KWL-0179	Tundra soils, Kilpisjärvi, Finland	PRJEB41762 *	[2]
UBA272	Permafrost soils, Nunavut, Canada	GCA_002504425.1	[3]
UBA273	Permafrost soils, Nunavut, Canada	GCA_002501935.1	[3]
UBA347	Permafrost soils (active layer), Nunavut, Canada	GCA_002495965.1	[3]
UBA348	Permafrost soils (active layer), Nunavut, Canada	GCA_002501855.1	[3]
UBA466	Permafrost soils (active layer), Nunavut, Canada	GCA_002498345.1	[3]
UBA536	Permafrost soils (active layer), Nunavut, Canada	GCA_002496625.1	[3]
RRmetagenome_bin19	Polar desert soils, Wilkes Land, Antarctica	GCA_003176995.1	[4]

1. This study.
2. Pessi et al. (2021).
3. Parks et al. (2017), based on data originally published by Chauhan et al. (2014).
4. Ji et al. (2017).

* BioProject accessions in GenBank/ENA. Submissions to GenBank/ENA are being processed and thereafter this table will be updated with the accession numbers for the individual MAGs.

Genome annotation

We used anvi'o v7.0 (Eren *et al.*, 2015) to predict gene calls with Prodigal v2.6.3 (Hyatt *et al.*, 2010), identify ribosomal genes and a set of 76 archaeal single-copy genes with HMMER v.3.3 (Eddy, 2011), and compute genome completion and redundancy levels based on the presence of the 76 single-copy genes. We also employed anvi'o v7.0 to annotate the gene calls against the KOfam (Aramaki *et al.*, 2020) and the Pfam (Mistry *et al.*, 2021) databases with HMMER v.3.3 (Eddy, 2011) and the COG database (Galperin *et al.*, 2021) with DIAMOND v0.9.14 (Buchfink *et al.*, 2015). Additionally, we used BLASTP v2.11.0 (Camacho *et al.*, 2009) to annotate the gene calls against the arCOG database (Makarova *et al.*, 2015) and to compare the amino acid sequences of genes identified as *amoA*, *amoB*, or *amoC* to the RefSeq (O'Leary *et al.*, 2016) and Swiss-Prot (The UniProt Consortium, 2019) databases. Finally, we used BLASTN v2.11.0 (Camacho *et al.*, 2009) to compare *amoA* genes against the nt database and the curated *amoA* database of Alves *et al.* (2018).

127 ## Phylogenomic and phylogenetic analyses

128 For phylogenomic analysis, we used a set of 59 archaeal single-copy genes that were present in
129 at least 80% of the genomes. In addition to the 12 UBA10452 MAGs, we retrieved from GenBank
130 other 33 genomes belonging to the family Nitrososphaeraceae and the genome of
131 *Nitrosopumilus maritimus* SCM1 to be used as an outgroup. We used anvi'o v7.0 (Eren *et al.*,
132 2015) to recover the predicted amino acid sequence for each of the 59 genes, align them
133 individually with MUSCLE v3.8.1551 (Edgar, 2004), and generate a concatenated alignment.
134 We then computed a maximum likelihood tree with IQ-TREE v2.1.4 employing the automatic
135 model selection and 1000 bootstraps (Nguyen *et al.*, 2015). Finally, we also used anvi'o v7.0 to
136 compute pairwise average nucleotide identities (ANI) with pyani v0.2.10 (Pritchard *et al.*, 2016)
137 and to carry out phylogenetic analysis of the *amoA* and 16S rRNA genes as described for the
138 phylogenomic analysis (i.e. alignment with MUSCLE and tree building with IQ-TREE).

139 ## Abundance and geographic distribution

140 We employed read recruitment to compute the relative abundance of UBA10452 lineage across
141 the metagenomics datasets from which the MAGs were originally recovered. These datasets
142 consisted of 10 samples from tundra soils in Rásttigáisá, Norway (this study); 69 samples from
143 tundra soils in Kilpisjärvi, Finland (Pessi *et al.*, 2021); 13 samples from permafrost soils in
144 Nunavut, Canada (Chauhan *et al.*, 2014; Stackhouse *et al.*, 2015); and three samples from polar
145 desert soils in Wilkes Land, Antarctica (Ji *et al.*, 2017). We used fasterq-dump v2.10.8
146 (https://github.com/ncbi/sra-tools) to retrieve the raw metagenomic data from the Sequence
147 Read Archive (SRA). We then used CoverM v0.6.1 (https://github.com/wwood/CoverM) to map
148 the reads to the MAGs with minimap v2.17 (Li, 2016) and to compute relative abundances based
149 on the proportion of reads recruited by the MAGs. In addition, we used IMNGS (Lagkouvardos
150 *et al.*, 2016) to further investigate the geographic distribution of the UBA10452 lineage. For
151 this, we used the 16S rRNA gene sequence of the MAG RRmetagenome_bin19 as query to screen
152 422,877 amplicon sequencing datasets in SRA with UBLAST (Edgar, 2010). We considered only
153 matches with ≥ 99.0% similarity.

Results

Genomic characteristics of the UBA10452 lineage

UBA10452 is an uncultured genus-level lineage in the order Nitrososphaerales ("terrestrial group I.1b), class Nitrososphaeria (former phylum Thaumarchaeota) (Rinke *et al.*, 2021). The UBA10452 lineage is currently represented by eight MAGs in GTDB and GenBank in addition to the four MAGs obtained in the present study **(Table 1, Fig. 1a)**. Genome completion and redundancy estimated with anvi'o v7.0 (Eren *et al.*, 2015) based on the presence of 76 single-copy genes range from 50.0–90.8% and 2.6–9.2%, respectively **(Fig. 1b, Suppl. Table S1)**. The MAG RRmetagenome_bin19, with 90.8% completion, 3.9% redundancy, and a nearly complete (1462 bp) 16S rRNA gene, represents a high-quality MAG according to the MIMAG standard (Bowers *et al.*, 2017). The remaining 11 MAGs are of medium quality (≥ 50% complete, < 10% redundant), four of which also include the 16S rRNA gene. The genome size of UBA10452 MAGs ranges from 0.8 Mb (MAG S1130_Bin_3, 60.5% complete) to 4.0 Mb (MAG RRmetagenome_bin19, 90.8% complete). G+C content ranges from 38.1 to 41.5%.

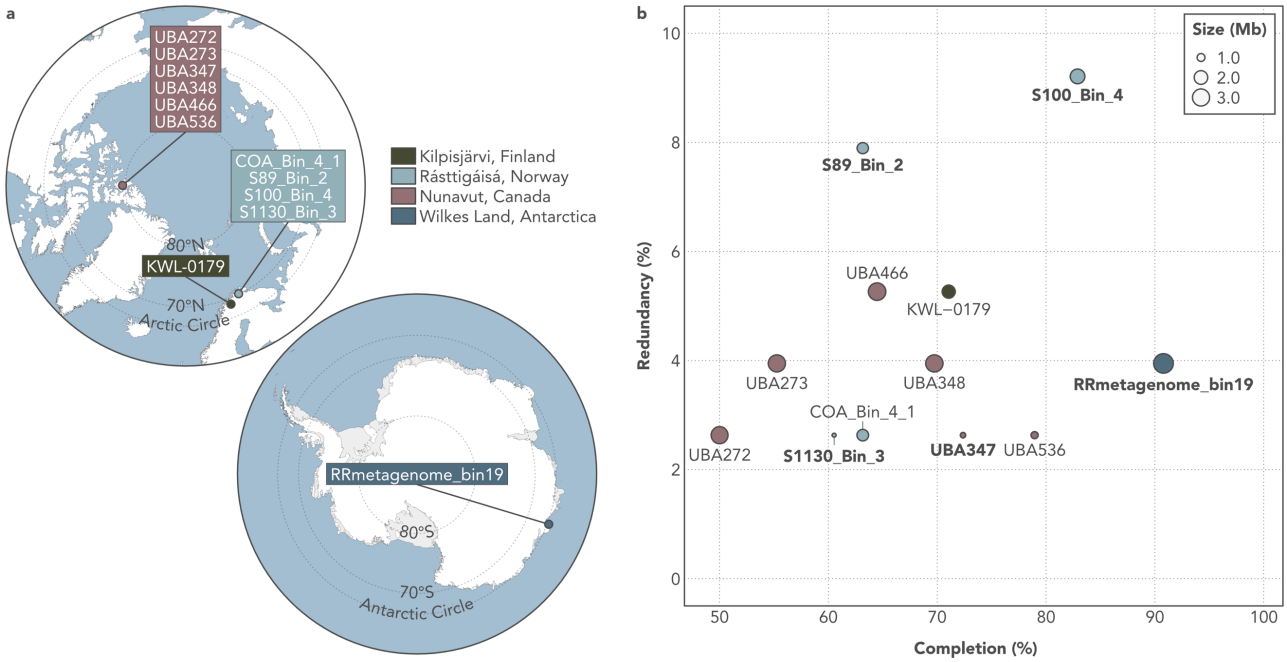


Figure 1. Geographic origin and assembly statistics of metagenome-assembled genomes (MAGs) assigned to the UBA10452 lineage (*Candidatus* Nitrosopolaris). a) Maps showing the geographic origin of the 12 UBA10452 MAGs. **b)** Genome completion, redundancy, and size of the UBA10452 MAGs. Completion and redundancy levels were computed based on the presence of 76 single-copy genes. MAGs in bold include the 16S rRNA gene.

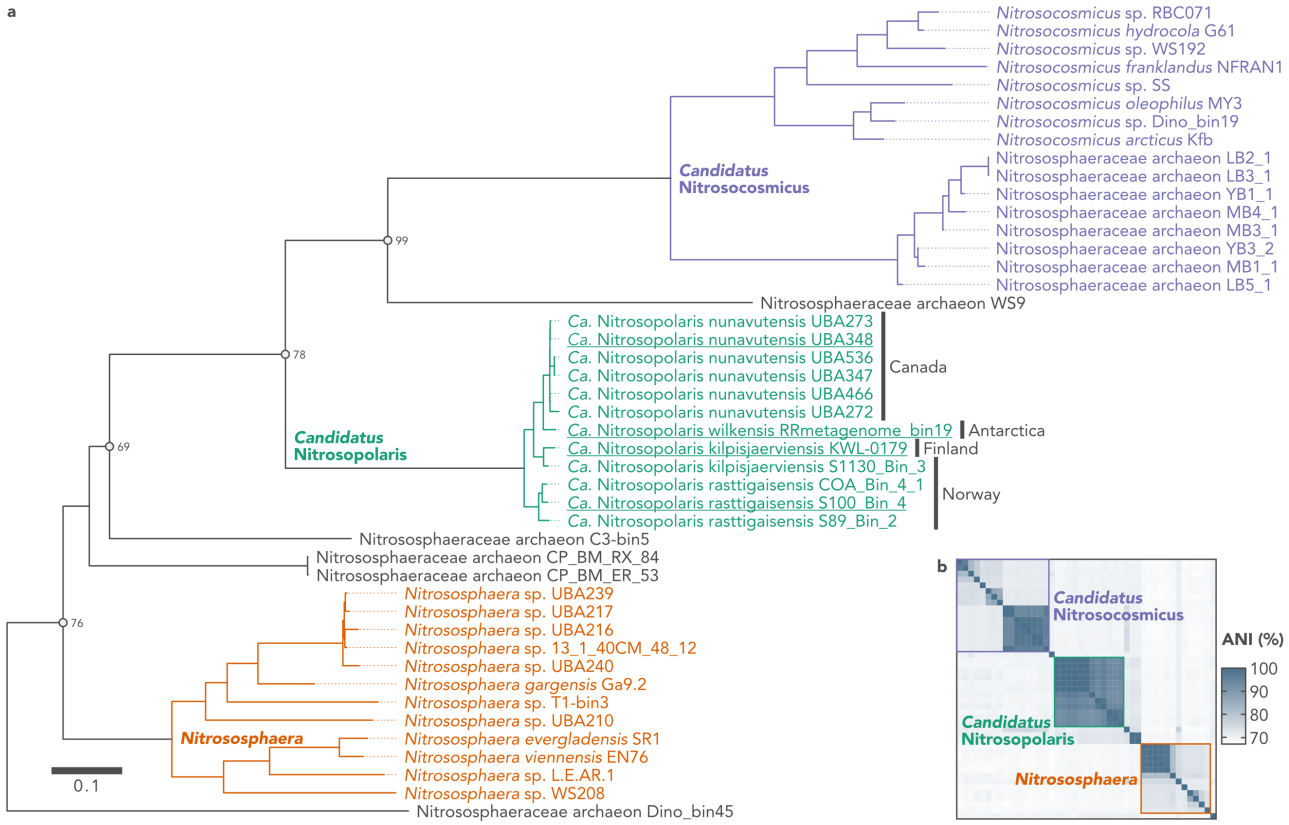


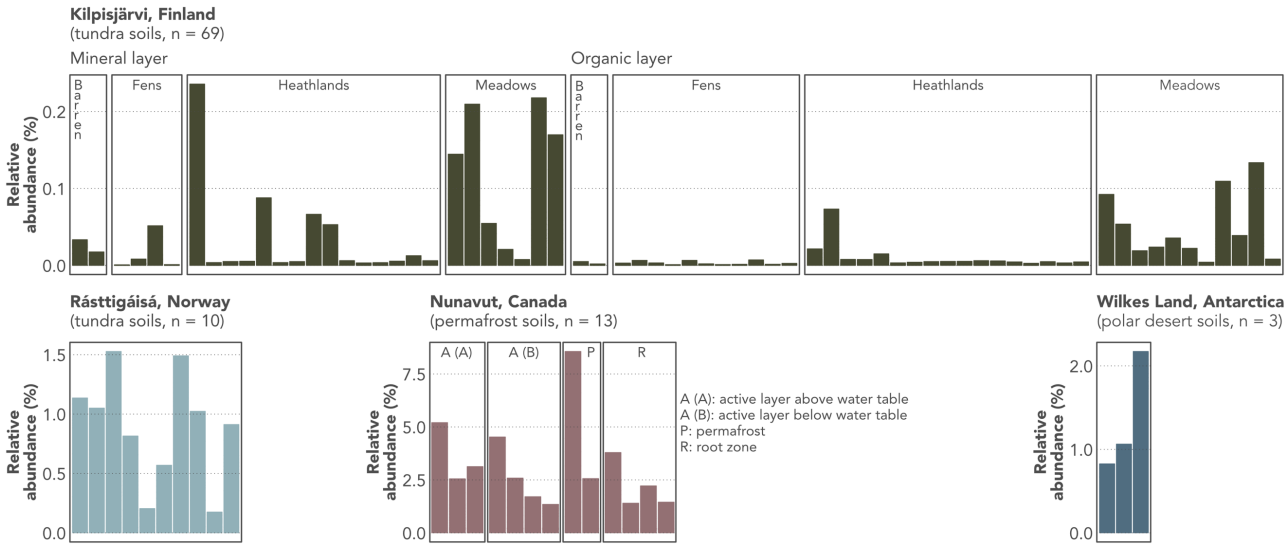
Figure 2. Phylogenomic analysis of the UBA10452 lineage (*Candidatus* Nitrosopolaris). a) Maximum likelihood tree based on 59 single-copy genes from 12 metagenome-assembled genomes (MAGs) assigned to the UBA10452 lineage and 33 other Nitrososphaeraceae genomes available on GenBank. The tree was rooted with *Nitrosopumilus maritimus* SCM1 (not shown). Nodes are supported by bootstrap values of 100% unless shown otherwise. Representatives for the four proposed species are indicated in underscore. **b)** Pairwise average nucleotide identity (ANI) between Nitrososphaeraceae genomes. Rows and columns are ordered from top to bottom and left to right, respectively, according to the top-bottom order of leaves in panel a.

UBA10452 is a distinct lineage in the family Nitrososphaeraceae

Phylogenomic analysis based on 59 single-copy genes confirmed the placement of UBA10452 as a distinct lineage outside *Nitrososphaera* and *Ca.* Nitrosocosmicus, the two described genera in the family Nitrososphaeraceae (Rinke *et al.*, 2021) **(Fig. 2a)**. Separation of UBA10452 from *Nitrososphaera* and *Ca.* Nitrosocosmicus is also supported by ANI and 16S rRNA gene analyses. UBA10452 MAGs share 68.9% ± 0.5 and 69.5% ± 0.3 ANI with *Nitrososphaera* and *Ca.* Nitrosocosmicus, respectively **(Fig. 2b)**. At the 16S rRNA gene level, UBA10452 MAGs are 94.8% ± 1.2 and 95.4% ± 0.2 similar to *Nitrososphaera* and *Ca.* Nitrosocosmicus, respectively **(Suppl. Fig. S1)**. Given that UBA10452 represents a clear, distinct lineage in the family

192 Nitrososphaeraceae, we consider that UBA10452 should be recognized as a *Candidatus* genus
193 and propose the name *Ca.* Nitrosopolaris.

194 Pairwise ANI values between *Ca.* Nitrosopolaris MAGs range from 90.9 to 99.9% **(Fig. 2b)**.
195 Based on a 96% ANI threshold (Ciufo *et al.*, 2018), the 12 *Ca.* Nitrosopolaris MAGs represent
196 four distinct species **(Fig. 2a)**. Two of these, one comprising the six Canadian MAGs (Chauhan
197 *et al.*, 2014; Parks *et al.*, 2017) and the other consisting solely of the Antarctic MAG (Ji *et al.*,
198 2017), correspond to the two existing species in GTDB release 95 ("UBA10452 sp002501855"
199 and "UBA10452 sp003176995", respectively). Here we suggest renaming these species as *Ca.*
200 Nitrosopolaris nunavutensis and *Ca.* Nitrosopolaris wilkensis, respectively, according to the
201 geographic origin of the MAGs. The Finnish MAG (Pessi *et al.*, 2021) plus one of the Norwegian
202 MAGs obtained in the present study (S1130_Bin_3) represent a novel species, for which we
203 suggest the name *Ca.* Nitrosopolaris kilpisjaerviensis. Finally, the three remaining MAGs
204 obtained in the present study (COA_Bin_4_1, S89_Bin_2, and S100_Bin_4) correspond to
205 another novel species, which we propose to be named as *Ca.* Nitrosopolaris rasttigaisensis.
206 However, the separation of *Ca.* Nitrosopolaris into four species is not supported by the analysis
207 of the 16S rRNA gene **(Suppl. Fig. S1)**. The pairwise similarity between 16S rRNA gene
208 sequences across the four ANI clusters range from 99.5 to 99.9%, which is above the proposed
209 threshold of 98.7% for species delineation (Kim *et al.*, 2014).



210 **Figure 3. Relative abundance of metagenome-assembled genomes (MAGs) assigned to**
211 **the UBA10452 lineage (*Candidatus* Nitrosopolaris).** Relative abundances were
212 investigated in the four datasets originally used to bin the UBA10452 MAGs. Abundances were
213 computed as the proportion of the reads recruited by the MAGs across the datasets.

Ca. Nitrosopolaris has a predominantly polar distribution

All 12 UBA10452/*Ca.* Nitrosopolaris MAGs were obtained from tundra, permafrost, and polar desert soils **(Table 1, Fig. 1a)**. These soils are typical of polar and alpine environments, being characterized by low pH (4.8–5.1), carbon (C) (1.0–7.3%), and N (0.1–0.3%) content (Stackhouse *et al.*, 2015; Ji *et al.*, 2017; Pessi *et al.*, 2021). To gain insights into the ecology of *Ca.* Nitrosopolaris, we used read recruitment to quantify its abundance across the datasets which were used to assemble the MAGs. *Ca.* Nitrosopolaris was most abundant in the permafrost from nutrient-poor (C: 1.0%, N: 0.1%) mineral cryosoils in Nunavut, Canadian Arctic, where it reached relative abundances of up to 8.6% **(Fig. 3)**. On the other hand, *Ca.* Nitrosopolaris was least abundant in the more nutrient-rich (C: 7.3%, N: 0.3%) tundra soils from Kilpisjärvi, Finland, where it was found particularly in samples taken from the mineral layer of heathland and meadow soils.

In order to investigate further the geographic distribution of *Ca.* Nitrosopolaris, we used IMNGS (Lagkouvardos *et al.*, 2016) to screen 422,877 16S rRNA gene amplicon sequencing datasets in SRA. Sequences matching the 16S rRNA gene of *Ca.* Nitrosopolaris (≥ 99.0% similarity) were found across 1281 datasets, mostly consisting of soil (n = 750), freshwater (n = 104), and rhizosphere samples (n = 100). Matched reads accounted for 6.0% of the total number of reads in these datasets (8.9 out of 149.1 million sequences). Of these, the overwhelming majority (8.7 million reads, 97.9%) come from Antarctic soil datasets, particularly from 149 sites in the vicinity of Davis Station, Princess Elizabeth Land (Bissett *et al.*, 2016) **(Suppl. Fig. S2a)**. The relative abundance of *Ca.* Nitrosopolaris was above 50% of the total archaeal community in 70 of these sites and reached values as high as 88.8% **(Suppl. Fig. S2b)**.

Ca. Nitrosopolaris harbours genes for ammonia oxidation, autotrophy, and cold adaptation

Annotation of protein-coding genes revealed that *Ca.* Nitrosopolaris harbours the *amoA*, *amoB*, *amoC*, and *amoX* genes encoding the enzyme ammonia monooxygenase (Amo) which catalyses the oxidation of ammonia to hydroxylamine **(Fig. 4a, Suppl. Table S2)**. As for other AOA, homologues of the *hao* gene were not found in the *Ca.* Nitrosopolaris MAGs. In AOB, this gene encodes the enzyme hydroxylamine dehydrogenase (Hao) which takes part in the oxidation of hydroxylamine to nitrite, a mechanism that remains unknown in AOA (Lehtovirta-Morley, 2018). The *amo* genes in all *Ca.* Nitrosopolaris MAGs are distributed across two separate contigs **(Fig. 4a)**. One of the contigs contains the *amoC*, *amoX*, and *amoA* genes; however, the *amoC* gene is truncated and found in only two MAGs. In some MAGs, the other contig contains a

247 second, full-length copy of the *amoC* gene followed by *amoB*. Not all MAGs contain all *amoABCX*
248 genes. However, considering that the MAGs present varying levels of completion and since the
249 putative localization of the genes corresponds to start or end of contigs, it is likely that missing
250 genes are an artifact of truncated assemblies rather than due to gene loss. In addition to *amo*
251 genes, *Ca.* Nitrosopolaris also encodes an ammonia transporter of the Amt family involved in
252 the uptake of extracellular ammonia and genes involved in the biosynthesis of the molybdenum
253 cofactor (Moco) **(Suppl. Table S2)**.

254 The *amoA* gene of *Ca.* Nitrosopolaris has a length of 651 bp and belongs to the NS-γ-2.3.2 cluster
255 **(Fig. 4b)**. One exception is the MAG UBA272 which contains a longer *amoA* gene (873 bp) with
256 a long insert of ambiguous base calls, most likely an artifact from assembling and/or scaffolding.
257 Sequences belonging to the NS-γ-2.3.2 cluster are found majorly in acidic soils (Alves *et al.*,
258 2018). Moreover, analysis of sequences from GenBank showed that the *amoA* gene of *Ca.*
259 Nitrosopolaris is related (≥ 96% nucleotide similarity) to sequences recovered mostly from cold
260 environments, i.e., the Artic, Antarctica, and alpine regions such as the Tibetan Plateau **(Fig.**
261 **4c)**. Among these, the *amoA* sequences of *Ca.* Nitrosopolaris MAGs are most closely related
262 (98.9–99.7% nucleotide similarity) to uncultured sequences from Antarctic soil (MH318339 and
263 MH318307), grassland soil in Iceland (JQ403917 and JQ403913), and the Tibetan Plateau
264 (GQ143258, GQ143220, GQ143135, KF004058, and MK987234) (Daebeler *et al.*, 2012; Xie *et al.*,
265 2014; Wang *et al.*, 2019; Zhang *et al.*, 2019) **(Fig. 4b)**.

266 Similarly to other AOA, *Ca.* Nitrosopolaris is likely autotrophic as it harbours genes for the
267 hydroxypropionate-hydroxybutylate pathway of CO_2 fixation, complexes I–V of the electron
268 transfer chain, the citric acid cycle, and gluconeogenesis **(Suppl. Table S2)**. *Ca.* Nitrosopolaris
269 also harbours urease and urea transporter genes, thus indicating the ability to generate
270 ammonia from urea. Like other AOA, the gene content of *Ca.* Nitrosopolaris indicates a potential
271 for mixo- or organotrophic metabolism, with multiple copies of genes encoding proteins involved
272 in carbohydrate metabolism and transport such as glucose/sorbosone dehydrogenases,
273 permeases of the major facilitator superfamily (MFS), and pyruvate oxidases. In contrast to
274 other AOA, we did not detect genes encoding the enzyme cyanate hydratase enzyme that is
275 involved in the production of ammonia from cyanate. Similarly, we did not detect genes involved
276 in the assembly of pili, flagellar apparatus, and chemotaxis, which are common in other AOA.

277 To investigate possible mechanisms underlying the distribution of *Ca.* Nitrosopolaris, we looked
278 for genes with known or predicted roles in cold adaptation and growth (Raymond-Bouchard *et*
279 *al.*, 2018). *Ca.* Nitrosopolaris harbours several copies of genes encoding proteins involved in cold
280 shock, DNA replication and repair, membrane and peptidoglycan alteration, osmotic and

oxidative stress, and translation/transcription such as different molecular chaperones, universal stress proteins, RecA/RadA recombinases, glycosyltransferases, sodium-hydrogen antiporters, thioredoxin reductases, peroxiredoxins, and DNA/RNA helicases **(Suppl. Table S2).**

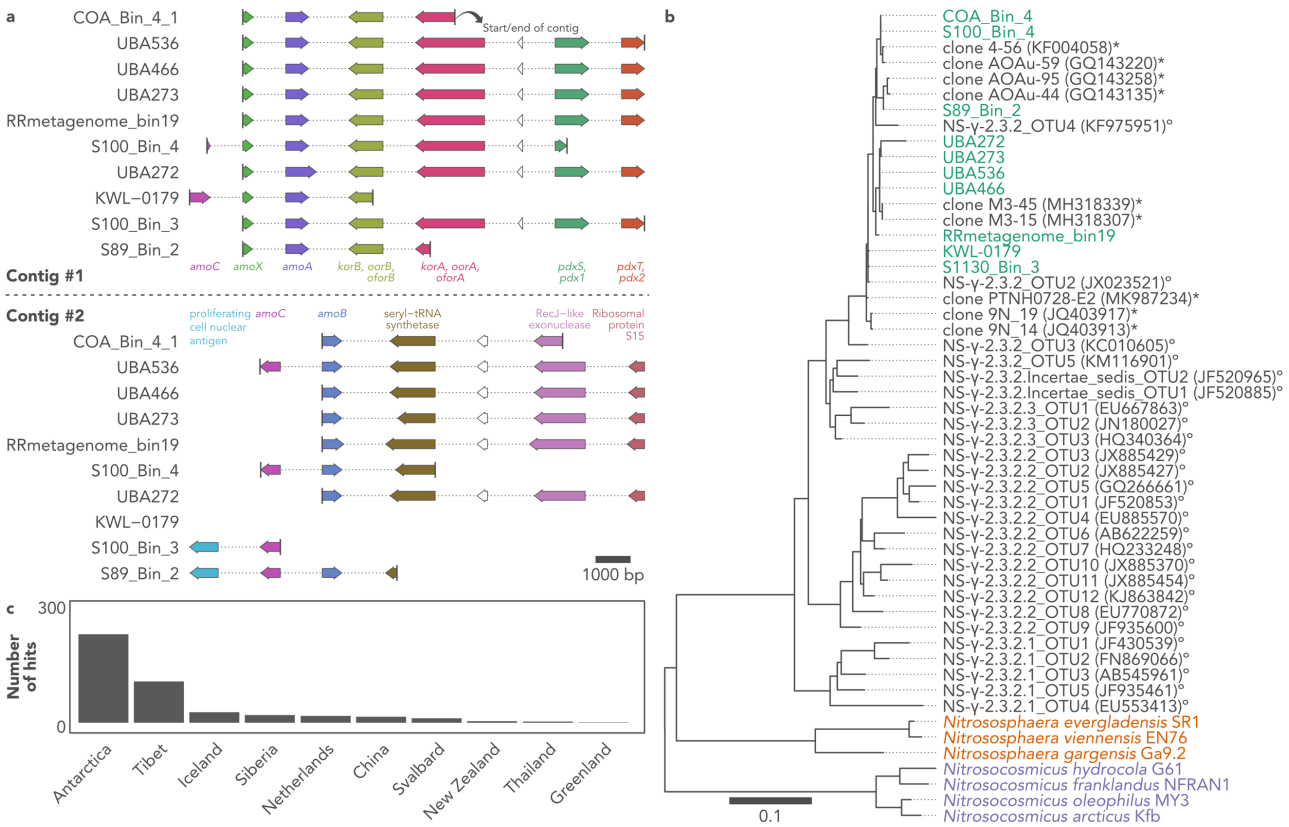


Figure 4. The ammonia monooxygenase (*amo*) genes of UBA10452 (*Candidatus* Nitrosopolaris). a) Representation of two contigs containing *amo* genes in metagenome-assembled genomes (MAGs) assigned to the UBA10452 lineage. Two MAGs which do not contain the *amoA* gene are omitted (UBA347 and UBA348). **b)** Maximum likelihood tree of the *amoA* sequence of UBA10452 MAGs and related sequences from GenBank (asterisks) and Alves *et al.* (2018) (circles). **c)** Geographic origin of sequences from GenBank with ≥ 96.0% nucleotide similarity to the *amoA* sequence of UBA10452 MAGs.

Discussion

Genome-resolved metagenomics has revolutionized our knowledge of archaeal diversity by giving us access to the genome of uncultured microorganisms at an unprecedented rate (Tahon *et al.*, 2021). In a recent metagenomic investigation of tundra soils in northern Finland (Pessi *et al.*, 2021), we have manually binned and curated a MAG belonging to the genus "UBA10452",

298 an uncultured and largely uncharacterized lineage in the order Nitrososphaerales ("terrestrial
299 group I.1b") of the class Nitrososphaeria (former phylum Thaumarchaeota) (Rinke *et al.*, 2021).
300 Here, we binned four other UBA10452 MAGs from tundra soils in Rásttigáisá, Norway, and
301 carried out a phylogenomic, metabolic, and biogeographic characterization of this lineage. Our
302 results indicate that the UBA10452 lineage consists of putative AOA with a geographic
303 distribution mostly restricted to polar ecosystems. We suggest the recognition of UBA10452 as
304 a *Candidatus* genus, for which we propose the name *Ca.* Nitrosopolaris (*nitrosus*: Latin
305 adjective meaning nitrous; *polaris*: Latin adjective meaning of or pertaining to the poles).

306 The findings from our polyphasic analysis consisting of phylogenomic, ANI, and 16S rRNA gene
307 analyses support the placement of *Ca.* Nitrosopolaris outside *Nitrososphaera* and *Ca.*
308 Nitrosocosmicus in the family Nitrososphaeraceae, as previously suggested (Rinke *et al.*, 2021).
309 Although there is no accepted universal ANI threshold for genus-level delineation, the ANI
310 between *Ca.* Nitrosopolaris and other members of the family Nitrososphaeraceae investigated
311 in this study is below the median ANI value of 73.1% observed between different microbial
312 genera belonging the same family/order (Barco *et al.*, 2020). Furthermore, pairwise similarity
313 values between *Ca.* Nitrosopolaris and other Nitrososphaeraceae are in the same range of the
314 ANI values observed between *Nitrososphaera* and *Ca.* Nitrosocosmicus genomes (68.5% ± 0.7).
315 Analysis of 16S rRNA gene sequences also support *Ca.* Nitrosopolaris as an independent
316 lineage, given that pairwise 16S rRNA gene similarities between *Ca.* Nitrosopolaris and other
317 Nitrososphaeraceae are below the median similarity value of 96.4% observed between related
318 genera across different microbial phyla (Yarza *et al.*, 2014).

319 Our results further indicate that the 12 *Ca.* Nitrosopolaris MAGs represent four different
320 species based on a 96% ANI threshold (Ciufo *et al.*, 2018). In addition to the two current species
321 in GTDB release 95 established based on rank normalization (Parks *et al.*, 2018, 2020), the
322 inclusion of the four MAGs obtained in the present study resulted in the identification of two
323 novel species. Here, we propose to name these species according to the geographic origin of the
324 representative MAG (*Ca.* N. nunavutensis, *Ca.* N. wilkensis, *Ca.* N. kilpisjaerviensis, and *Ca.*
325 N. rasttigaisensis). It is important to note that the separation of *Ca.* Nitrosopolaris into four
326 species based on ANI values is not readily supported by the analysis of 16S rRNA gene
327 sequences, which are ≥ 99.5% similar across the four species. Although a 98.7% threshold has
328 been proposed, species delineation based solely on the 16S rRNA gene can be problematic given
329 that microorganisms belonging to different species can share identical 16S rRNA gene
330 sequences (Kim *et al.*, 2014). One example of this is *Ca.* Nitrosocosmicus arcticus and *Ca.*
331 Nitrosocosmicus oleophilus, two species of AOA which share an identical 16S rRNA gene
332 sequence despite having divergent genomes with only 83.0% ANI (Alves *et al.*, 2019). It thus

333 appears reasonable to conclude that the 12 *Ca.* Nitrosopolaris MAGs indeed represent four
334 different species as suggested by the ANI analysis. If cultured representatives become available
335 in the future, phenotypic and ecophysiological characterization of these isolates could help
336 resolve the taxonomy of *Ca.* Nitrosopolaris.

337 *Ca.* Nitrosopolaris harbours the complete set of *amoA* genes responsible for chemolithotrophic
338 growth via ammonia oxidation (Lehtovirta-Morley, 2018). Although *in silico* analyses provide
339 valuable predictions, metabolic capabilities inferred by genomic annotation need to be
340 confirmed based on the analysis of isolated/enriched cultures or with the help of other indirect
341 methods such as stable isotope probing (SIP) (Gadkari *et al.*, 2020). Indeed, studies have
342 suggested that some members of the class Nitrososphaeria do not oxidize ammonia or are not
343 obligate chemolithoautotrophs, being able to grow on organic compounds as alternative energy
344 and/or C sources (Mussmann *et al.*, 2011; Pester *et al.*, 2011). Nevertheless, the putative
345 ammonia oxidation capability of *Ca.* Nitrosopolaris is supported by its close phylogenetic
346 relationship to *Nitrososphaera* and *Ca.* Nitrosocosmicus, two genera which have been
347 demonstrated to grow by oxidizing ammonia (Stieglmeier *et al.*, 2014; Lehtovirta-Morley *et al.*,
348 2016). In addition, the presence of several genes involved in carbohydrate transport and
349 metabolism suggest that *Ca.* Nitrosopolaris, like other AOA, might be able to grow myxo- or
350 organotrophically (Pester *et al.*, 2011).

351 In addition to the geographical origin of the MAGs, large-scale screening of 16S rRNA gene and
352 *amoA* sequences from SRA and GenBank indicate that *Ca.* Nitrosopolaris is restricted to soils
353 in the cold biosphere. Furthermore, the four proposed *Ca.* Nitrosopolaris species form coherent
354 biogeographical/habitat clusters: *Ca.* N. nunavutensis, comprising MAGs obtained from
355 permafrost soils in Nunavut, Canada; *Ca.* N. wilkensis, corresponding to one MAG from polar
356 desert soils in Wilkes Land, Antarctica; and *Ca.* N. kilpisjaerviensis and *Ca.* N. rasttigaisensis,
357 comprising MAGs obtained from mineral tundra soils in two relatively close regions in northern
358 Fennoscandia (Kilpisjärvi and Rásttigáisá, respectively). The discovery of *Ca.* Nitrosopolaris
359 complements the list of microbial taxa that appear to be adapted to life in cold environments,
360 such as the mat-forming cyanobacteria *Phormidesmis priestleyi* (Komárek *et al.*, 2009) and
361 *Shackeltoniella antarctica* (Strunecky *et al.*, 2020) and the sea-ice bacteria *Polaribacter* and
362 *Psychrobacter* (Bowman, 2013). Furthermore, the abundance profile of *Ca.* Nitrosopolaris
363 observed in this study, which was characterized by a higher relative abundance in mineral
364 cryosoil permafrost and polar desert soils compared to tundra soils, indicates that *Ca.*
365 Nitrosopolaris is particularly adapted to the highly oligotrophic conditions found in some of the
366 most extreme environments in the cryosphere.

367 Investigation of the genome of *Ca.* Nitrosopolaris provided insights on possible adaptations to
368 cold and oligotrophic environments. For instance, *Ca.* Nitrosopolaris harbour multiple copies of
369 several genes that have been implicated in tolerance to cold, such as genes encoding molecular
370 chaperones, DNA/RNA helicases, and universal stress proteins (Raymond-Bouchard *et al.*,
371 2018). In fact, it has been suggested that the environmental characteristics of polar soils favour
372 AOA in general (Alves *et al.*, 2013; Siljanen *et al.*, 2019). AOA commonly predominate over AOB
373 in soils with low N content, and acidic conditions such as those found in polar soils further limit
374 the availability of ammonia (Schleper and Nicol, 2010; Lehtovirta-Morley, 2018). The ecological
375 success of AOA in oligotrophic and acidic soils has been linked to the higher affinity of their
376 ammonia oxidation machinery compared to their bacterial counterparts (Martens-Habbena *et*
377 *al.*, 2009; Kerou *et al.*, 2016). Furthermore, the hydroxypropionate-hydroxybutylate pathway of
378 CO_2 fixation encoded by *Ca.* Nitrosopolaris and other AOA appears to be more energy efficient
379 than the Calvin cycle employed by AOB (Könneke *et al.*, 2014). However, these traits are shared
380 between *Ca.* Nitrosopolaris and other AOA and thus do not readily explain the apparent
381 ecological success of *Ca.* Nitrosopolaris in cold environments. Indeed, mechanisms of cold
382 adaptation are evolutionary and functionally complex and involve many features that cannot
383 be observed by metagenomics alone (e.g. gene regulation and membrane modifications)
384 (Cavicchioli, 2006). Structural, transcriptomics, and proteomics analysis of cultured isolates
385 could help shed further light on possible adaptations to cold in *Ca.* Nitrosopolaris.

386 In addition to possible mechanisms of cold adaptation, we hypothesize that the distribution of
387 *Ca.* Nitrosopolaris could be related to historical factors. A recent molecular dating study has
388 suggested that the origin of the AOA clade group I.1b (order Nitrososphaerales) coincides with
389 severe glaciation events that happened during the Neoproterozoic (Yang *et al.*, 2021). If these
390 estimates are accurate, it would imply that *Ca.* Nitrosopolaris and all other lineages in group
391 I.1b share a common ancestor that appeared when the global climate was characterized by sub-
392 zero temperatures, having likely evolved at glacial refugia such as nunataks or regions with
393 geothermal activity. We further hypothesize that, given the apparent lack of motility and
394 chemotaxis machineries, *Ca.* Nitrosopolaris remained restricted to polar and alpine
395 environments following deglaciation, while its sister lineages (*Nitrososphaera* and *Ca.*
396 Nitrosocosmicus) were able to expand their distribution to other climates and are now found in
397 meso- and even thermophilic environments (Schleper and Nicol, 2010; Lehtovirta-Morley,
398 2018).

399 Due to low temperatures throughout the year, polar soils store a large amount of organic matter
400 and have thus served as important carbon sinks. At present, polar soils are considered minor
401 yet significant sources of N_2O (Voigt *et al.*, 2020) but, if warming trends continue at the levels

402 observed currently, polar soils might become major contributors to the global N_2O budget. For
403 instance, the AOA *Ca.* Nitrosocosmicus arcticus isolated from Arctic soil has an ammonia
404 oxidation optimum at temperatures well above those found *in situ* (Alves *et al.*, 2019). Given
405 that both the direct and indirect roles of AOA in the cycling of N_2O in polar soils remain largely
406 undetermined, a better understanding of polar microbial communities is paramount to model
407 current and future N_2O fluxes from this biome.

408 Acknowledgments

409 This study was funded by the Academy of Finland (project 335354) and the University of
410 Helsinki. We would like to acknowledge the CSC – IT Centre for Science, Finland, for providing
411 computing resources; Miska Luoto and the BioGeoClimate Modelling Lab at the University of
412 Helsinki for providing the Rásttigáisá samples; Eeva Eronen-Rasimus, Laura Cappelatti, and
413 Benoit Durieu for helpful suggestions and discussions; and Jouko Rikkinen and Heino Vänskä
414 for nomenclatural advice.

415 Data availability

416 Genomic assemblies can be found in GenBank/ENA under the accession numbers listed in
417 **Table 1**.

418 Author contributions

419 ISP, AR, and JH designed the study. ISP performed the analyses and wrote the manuscript. AR
420 and JH revised the manuscript.

421 Competing interests

422 The authors declare no conflict of interests.

423 References

424 Alves, R.J.E., Kerou, M., Zappe, A., Bittner, R., Abby, S.S., Schmidt, H.A., et al. (2019) Ammonia
425 Oxidation by the Arctic Terrestrial Thaumarchaeote Candidatus Nitrosocosmicus arcticus Is
426 Stimulated by Increasing Temperatures. *Front Microbiol* **10**: 1571.

427 Alves, R.J.E., Minh, B.Q., Urich, T., von Haeseler, A., and Schleper, C. (2018) Unifying the
428 global phylogeny and environmental distribution of ammonia-oxidising archaea based on amoA
429 genes. *Nat Commun* **9**: 1517.

430 Alves, R.J.E., Wanek, W., Zappe, A., Richter, A., Svenning, M.M., Schleper, C., and Urich, T.
431 (2013) Nitrification rates in Arctic soils are associated with functionally distinct populations of
432 ammonia-oxidizing archaea. *ISME J* **7**: 1620–1631.

433 Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., and Ogata, H.
434 (2020) KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score
435 threshold. *Bioinformatics* **36**: 2251–2252.

436 Barco, R.A., Garrity, G.M., Scott, J.J., Amend, J.P., Nealson, K.H., and Emerson, D. (2020) A
437 Genus Definition for *Bacteria* and *Archaea* Based on a Standard Genome Relatedness Index.
438 *mBio* **11**:.

439 Bissett, A., Fitzgerald, A., Meintjes, T., Mele, P.M., Reith, F., Dennis, P.G., et al. (2016)
440 Introducing BASE: the Biomes of Australian Soil Environments soil microbial diversity
441 database. *GigaScience* **5**: 21.

442 Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K.,
443 et al. (2017) Minimum information about a single amplified genome (MISAG) and a
444 metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* **35**: 725–731.

445 Bowman, J.P. (2013) Sea-Ice Microbial Communities. In *The Prokaryotes*. Rosenberg, E.,
446 DeLong, E.F., Lory, S., Stackebrandt, E., and Thompson, F. (eds). Berlin, Heidelberg: Springer
447 Berlin Heidelberg, pp. 139–161.

448 Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008) Mesophilic
449 crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol*
450 **6**: 245–252.

451 Buchfink, B., Xie, C., and Huson, D.H. (2015) Fast and sensitive protein alignment using
452 DIAMOND. *Nat Methods* **12**: 59–60.

453 Butterbach-Bahl, K., Baggs, E.M., Dannenmann, M., Kiese, R., and Zechmeister-Boltenstern,
454 S. (2013) Nitrous oxide emissions from soils: how well do we understand the processes and their
455 controls? *Philos Trans R Soc B Biol Sci* **368**: 20130122.

456 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden,
457 T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.

458 Cavicchioli, R. (2006) Cold-adapted archaea. *Nat Rev Microbiol* **4**: 331–343.

459 Chauhan, A., Layton, A.C., Vishnivetskaya, T.A., Williams, D., Pfiffner, S.M., Rekepalli, B., et
460 al. (2014) Metagenomes from Thawing Low-Soil-Organic-Carbon Mineral Cryosols and
461 Permafrost of the Canadian High Arctic. *Genome Announc* **2**:.

462 Ciufo, S., Kannan, S., Sharma, S., Badretdin, A., Clark, K., Turner, S., et al. (2018) Using
463 average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the
464 NCBI. *Int J Syst Evol Microbiol* **68**: 2386–2392.

465 Daebeler, A., Abell, G.C.J., Bodelier, P.L.E., Bodrossy, L., Frampton, D.M.F., Hefting, M.M.,
466 and Laanbroek, H.J. (2012) Archaeal dominated ammonia-oxidizing communities in Icelandic
467 grassland soils are moderately affected by long-term N fertilization and geothermal heating.
468 *Front Microbiol* **3**:.

469 Daims, H., Lebedeva, E.V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., et al. (2015) Complete
470 nitrification by Nitrospira bacteria. *Nature* **528**: 504–509.

471 Eddy, S.R. (2011) Accelerated Profile HMM Searches. *PLoS Comput Biol* **7**: e1002195.

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

474 Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST.
475 *Bioinformatics* **26**: 2460–2461.

476 Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., and Delmont, T.O.
477 (2015) Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**: e1319.

478 Gadkari, P.S., McGuinness, L.R., Männistö, M.K., Kerkhof, L.J., and Häggblom, M.M. (2020)
479 Arctic tundra soil bacterial communities active at subzero temperatures detected by stable
480 isotope probing. *FEMS Microbiol Ecol* **96**: fiz192.

481 Galperin, M.Y., Wolf, Y.I., Makarova, K.S., Vera Alvarez, R., Landsman, D., and Koonin, E.V.
482 (2021) COG database update: focus on microbial diversity, model organisms, and widespread
483 pathogens. *Nucleic Acids Res* **49**: D274–D281.

484 Hayashi, K., Tanabe, Y., Fujitake, N., Kida, M., Wang, Y., Hayatsu, M., and Kudoh, S. (2020)
485 Ammonia Oxidation Potentials and Ammonia Oxidizers of Lichen–Moss Vegetated Soils at Two
486 Ice-free Areas in East Antarctica. *Microbes Environ* **35**: n/a.

487 Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010)
488 Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC*
489 *Bioinformatics* **11**: 119.

490 Ji, M., Greening, C., Vanwonterghem, I., Carere, C.R., Bay, S.K., Steen, J.A., et al. (2017)
491 Atmospheric trace gases support primary production in Antarctic desert surface soil. *Nature*
492 **552**: 400–403.

493 Jung, M.-Y., Islam, Md.A., Gwak, J.-H., Kim, J.-G., and Rhee, S.-K. (2018) Nitrosarchaeum
494 koreense gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon member
495 of the phylum Thaumarchaeota isolated from agricultural soil. *Int J Syst Evol Microbiol* **68**:
496 3084–3095.

497 Kerou, M., Offre, P., Valledor, L., Abby, S.S., Melcher, M., Nagler, M., et al. (2016) Proteomics
498 and comparative genomics of *Nitrososphaera viennensis* reveal the core genome and adaptations
499 of archaeal ammonia oxidizers. *Proc Natl Acad Sci* **113**: E7937–E7946.

500 Kim, M., Oh, H.-S., Park, S.-C., and Chun, J. (2014) Towards a taxonomic coherence between
501 average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of
502 prokaryotes. *Int J Syst Evol Microbiol* **64**: 346–351.

503 Komárek, J., Ka□tovský, J., Ventura, S., and Turicchia, S. □marda (2009) The cyanobacterial
504 genus Phormidesmis. *Algol Stud* **129**: 41–59.

505 Könneke, M., Bernhard, A.E., de la Torre, J.R., Walker, C.B., Waterbury, J.B., and Stahl, D.A.
506 (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**: 543–546.

507 Könneke, M., Schubert, D.M., Brown, P.C., Hugler, M., Standfest, S., Schwander, T., et al.
508 (2014) Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2
509 fixation. *Proc Natl Acad Sci* **111**: 8239–8244.

510 Lagkouvardos, I., Joseph, D., Kapfhammer, M., Giritli, S., Horn, M., Haller, D., and Clavel, T.
511 (2016) IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for
512 ecology and diversity studies. *Sci Rep* **6**: 33721.

513 Lehtovirta-Morley, L.E. (2018) Ammonia oxidation: Ecology, physiology, biochemistry and why
514 they must all come together. *FEMS Microbiol Lett* **365**:.

515 Lehtovirta-Morley, L.E., Ross, J., Hink, L., Weber, E.B., Gubry-Rangin, C., Thion, C., et al.
516 (2016) Isolation of ' *Candidatus* Nitrosocosmicus franklandus', a novel ureolytic soil archaeal
517 ammonia oxidiser with tolerance to high ammonia concentration. *FEMS Microbiol Ecol* **92**:
518 fiw057.

519 Lehtovirta-Morley, L.E., Stoecker, K., Vilcinskas, A., Prosser, J.I., and Nicol, G.W. (2011)
520 Cultivation of an obligate acidophilic ammonia oxidizer from a nitrifying acid soil. *Proc Natl*
521 *Acad Sci* **108**: 15892–15897.

522 Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G.W., et al. (2006) Archaea
523 predominate among ammonia-oxidizing prokaryotes in soils. *Nature* **442**: 806–809.

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

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

529 Magalhães, C.M., Machado, A., Frank-Fahle, B., Lee, C.K., and Cary, S.C. (2014) The ecological
530 dichotomy of ammonia-oxidizing archaea and bacteria in the hyper-arid soils of the Antarctic
531 Dry Valleys. *Front Microbiol* **5**:.

532 Makarova, K., Wolf, Y., and Koonin, E. (2015) Archaeal Clusters of Orthologous Genes
533 (arCOGs): An Update and Application for Analysis of Shared Features between Thermococcales,
534 Methanococcales, and Methanobacteriales. *Life* **5**: 818–840.

535 Martens-Habbena, W., Berube, P.M., Urakawa, H., de la Torre, J.R., and Stahl, D.A. (2009)
536 Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria.
537 *Nature* **461**: 976–979.

538 Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., et
539 al. (2021) Pfam: The protein families database in 2021. *Nucleic Acids Res* **49**: D412–D419.

540 Mussmann, M., Brito, I., Pitcher, A., Sinninghe Damste, J.S., Hatzenpichler, R., Richter, A., et
541 al. (2011) Thaumarchaeotes abundant in refinery nitrifying sludges express amoA but are not
542 obligate autotrophic ammonia oxidizers. *Proc Natl Acad Sci* **108**: 16771–16776.

543 Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015) IQ-TREE: A Fast and
544 Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol*
545 **32**: 268–274.

546 O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., et al. (2016)
547 Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and
548 functional annotation. *Nucleic Acids Res* **44**: D733–D745.

549 Ortiz, M., Bosch, J., Coclet, C., Johnson, J., Lebre, P., Salawu-Rotimi, A., et al. (2020) Microbial
550 Nitrogen Cycling in Antarctic Soils. *Microorganisms* **8**: 1442.

551 Parks, D.H., Chuvochina, M., Chaumeil, P.-A., Rinke, C., Mussig, A.J., and Hugenholtz, P.
552 (2020) A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat Biotechnol*.

553 Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., and
554 Hugenholtz, P. (2018) A standardized bacterial taxonomy based on genome phylogeny
555 substantially revises the tree of life. *Nat Biotechnol* **36**: 996–1004.

556 Parks, D.H., Rinke, C., Chuvochina, M., Chaumeil, P.-A., Woodcroft, B.J., Evans, P.N., et al.
557 (2017) Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree
558 of life. *Nat Microbiol* **2**: 1533–1542.

559 Pessi, I.S., Viitamäki, S., Virkkala, A.-M., Eronen-Rasimus, E., Delmont, T.O., Marushchak,
560 M.E., et al. (2021) Truncated denitrifiers dominate the denitrification pathway in tundra soil
561 metagenomes. *bioRxiv*. Preprint (https://doi.org/10.1101/2020.12.21.419267).

562 Pester, M., Schleper, C., and Wagner, M. (2011) The Thaumarchaeota: an emerging view of their
563 phylogeny and ecophysiology. *Curr Opin Microbiol* **14**: 300–306.

564 Pritchard, L., Glover, R.H., Humphris, S., Elphinstone, J.G., and Toth, I.K. (2016) Genomics
565 and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal*
566 *Methods* **8**: 12–24.

567 Qin, W., Heal, K.R., Ramdasi, R., Kobelt, J.N., Martens-Habbena, W., Bertagnolli, A.D., et al.
568 (2017) Nitrosopumilus maritimus gen. nov., sp. nov., Nitrosopumilus cobalaminigenes sp. nov.,
569 Nitrosopumilus oxyclinae sp. nov., and Nitrosopumilus ureiphilus sp. nov., four marine
570 ammonia-oxidizing archaea of the phylum Thaumarchaeota. *Int J Syst Evol Microbiol* **67**: 5067–
571 5079.

572 Raymond-Bouchard, I., Goordial, J., Zolotarov, Y., Ronholm, J., Stromvik, M., Bakermans, C.,
573 and Whyte, L.G. (2018) Conserved genomic and amino acid traits of cold adaptation in subzero-
574 growing Arctic permafrost bacteria. *FEMS Microbiol Ecol* **94**:.

575 Richter, I., Herbold, C.W., Lee, C.K., McDonald, I.R., Barrett, J.E., and Cary, S.C. (2014)
576 Influence of soil properties on archaeal diversity and distribution in the McMurdo Dry Valleys,
577 Antarctica. *FEMS Microbiol Ecol* **89**: 347–359.

578 Rinke, C., Chuvochina, M., Mussig, A.J., Chaumeil, P.-A., Davín, A.A., Waite, D.W., et al. (2021)
579 A standardized archaeal taxonomy for the Genome Taxonomy Database. *Nat Microbiol* **6**: 946–
580 959.

581 Rutanen, A. (2020) Genome-resolved metagenomics provides insights into polar microbial
582 community composition, diversity and functional potential related to greenhouse gas cycling.
583 MSc thesis (http://hdl.handle.net/10138/320320).

Schleper, C. and Nicol, G.W. (2010) Ammonia-Oxidising Archaea – Physiology, Ecology and Evolution. In *Advances in Microbial Physiology*. Elsevier, pp. 1–41.

Siljanen, H.M.P., Alves, R.J.E., Ronkainen, J.G., Lamprecht, R.E., Bhattarai, H.R., Bagnoud, A., et al. (2019) Archaeal nitrification is a key driver of high nitrous oxide emissions from arctic peatlands. *Soil Biol Biochem* **137**: 107539.

Spang, A., Hatzenpichler, R., Brochier-Armanet, C., Rattei, T., Tischler, P., Spieck, E., et al. (2010) Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. *Trends Microbiol* **18**: 331–340.

Stackhouse, B.T., Vishnivetskaya, T.A., Layton, A., Chauhan, A., Pfiffner, S., Mykytczuk, N.C., et al. (2015) Effects of simulated spring thaw of permafrost from mineral cryosol on CO_2 emissions and atmospheric CH_4 uptake. *J Geophys Res Biogeosciences* **120**: 1764–1784.

Stieglmeier, M., Klingl, A., Alves, R.J.E., Rittmann, S.K.-M.R., Melcher, M., Leisch, N., and Schleper, C. (2014) Nitrososphaera viennensis gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *Int J Syst Evol Microbiol* **64**: 2738–2752.

Strunecky, O., Raabova, L., Bernardova, A., Ivanova, A.P., Semanova, A., Crossley, J., and Kaftan, D. (2020) Diversity of cyanobacteria at the Alaska North Slope with description of two new genera: Gibliniella and Shackletoniella. *FEMS Microbiol Ecol* **96**: fiz189.

Tahon, G., Geesink, P., and Ettema, T.J.G. (2021) Expanding Archaeal Diversity and Phylogeny: Past, Present, and Future. *Annu Rev Microbiol* **75**: 359–381.

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

de la Torre, J.R., Walker, C.B., Ingalls, A.E., Könneke, M., and Stahl, D.A. (2008) Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environ Microbiol* **10**: 810–818.

Treusch, A.H., Kletzin, A., Raddatz, G., Ochsenreiter, T., Quaiser, A., Meurer, G., et al. (2004) Characterization of large-insert DNA libraries from soil for environmental genomic studies of Archaea. *Environ Microbiol* **6**: 970–980.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., et al. (2004) Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science* **304**: 66–74.

Voigt, C., Marushchak, M.E., Abbott, B.W., Biasi, C., Elberling, B., Siciliano, S.D., et al. (2020) Nitrous oxide emissions from permafrost-affected soils. *Nat Rev Earth Environ* **1**: 420–434.

Wang, Q., Zhu, R., Zheng, Y., Bao, T., and Hou, L. (2019) Effects of sea animal colonization on the coupling between dynamics and activity of soil ammonia-oxidizing bacteria and archaea in maritime Antarctica. *Biogeosciences* **16**: 4113–4128.

Winogradsky, S. (1890) Recherches sur les organisms de la nitrification. *Ann Inst Pasteur* **4**: 215–231.

Wu, L., Chen, X., Wei, W., Liu, Y., Wang, D., and Ni, B.-J. (2020) A Critical Review on Nitrous Oxide Production by Ammonia-Oxidizing Archaea. *Environ Sci Technol* **54**: 9175–9190.

Xie, Z., Le Roux, X., Wang, C., Gu, Z., An, M., Nan, H., et al. (2014) Identifying response groups of soil nitrifiers and denitrifiers to grazing and associated soil environmental drivers in Tibetan alpine meadows. *Soil Biol Biochem* **77**: 89–99.

Yang, Y., Zhang, C., Lenton, T.M., Yan, X., Zhu, M., Zhou, M., et al. (2021) The Evolution Pathway of Ammonia-Oxidizing Archaea Shaped by Major Geological Events. *Mol Biol Evol* **38**: 3637–3648.

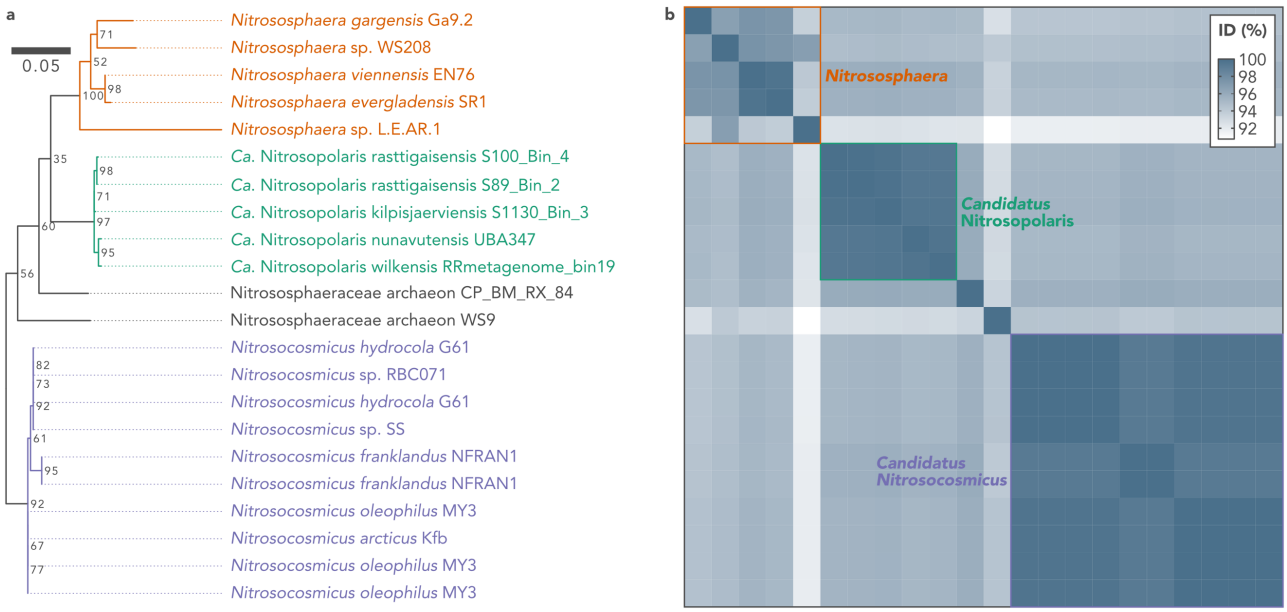
Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.-H., et al. (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* **12**: 635–645.

Zhang, S., Xia, X., Li, S., Zhang, L., Wang, G., Li, M., et al. (2019) Ammonia Oxidizers in High-Elevation Rivers of the Qinghai-Tibet Plateau Display Distinctive Distribution Patterns. *Appl Environ Microbiol* **85**:.

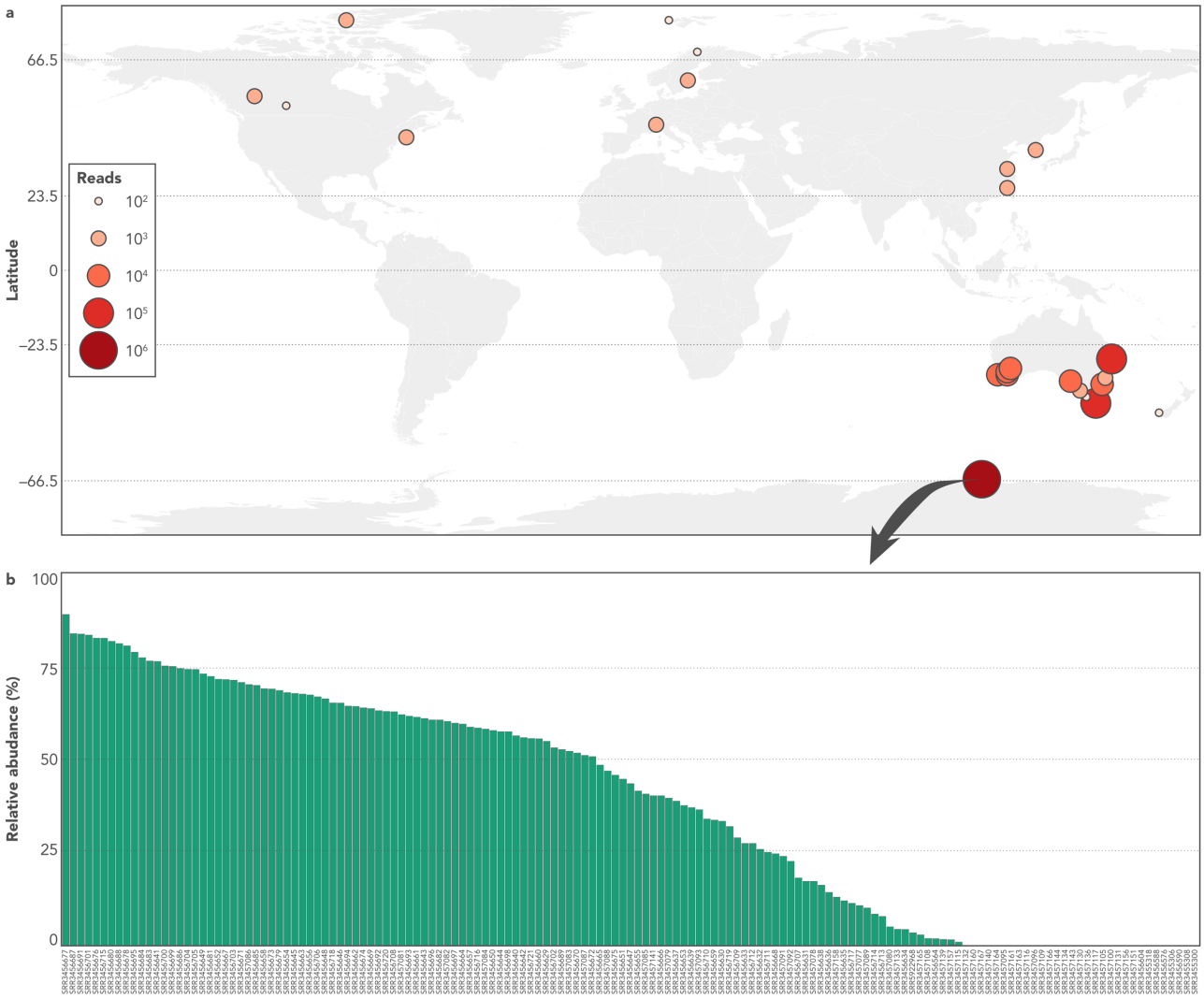
Supplementary Information

Suppl. Table S1 (separate .xlsx file). Additional information on metagenome-assembled genomes (MAGs) belonging to the UBA10452 lineage (*Candidatus* Nitrosopolaris).

Suppl. Table S2 (separate .xlsx file). List of selected genes from metagenome-assembled genomes (MAGs) belonging to the UBA10452 lineage (*Candidatus* Nitrosopolaris).



Suppl. Figure S1. The 16S rRNA gene of UBA10452 (*Candidatus* Nitrosopolaris). a) Phylogenetic analysis of the 16S rRNA gene sequence of five metagenome-assembled genomes (MAGs) assigned to the UBA10452 lineage and other Nitrososphaeraceae genomes available on GenBank. Maximum likelihood tree rooted with *Nitrosopumilus maritimus* SCM1 (not shown). Bootstrap values for node support are indicated. **b)** Pairwise similarity between 16S rRNA gene sequences. Note that some genomes contain multiple copies of the 16S rRNA gene.



Suppl. Figure S2. Geographic distribution of the UBA10452 lineage (*Candidatus* Nitrosopolaris). a) Distributional patterns of *Ca.* Nitrosopolaris based on the screening of 422,877 16S rRNA gene amplicon sequencing datasets in the Sequence Read Archive (SRA). Datasets with few matches (< 0.1% or < 100 reads) are not shown. **b)** Abundance of *Ca.* Nitrosopolaris across 149 16S rRNA gene amplicon sequencing datasets from soils in the vicinity of Davis Station, Princess Elizabeth Land, Antarctica (BioProject PRJNA317932). Relative abundances were computed as the proportion of reads matching the sequence of *Ca.* Nitrosopolaris in each sample.