



## OPEN ACCESS

## EDITED BY

José M González,  
University of La Laguna, Spain

## REVIEWED BY

Matthew Kellom,  
Berkeley Lab (DOE), United States  
Sining Zhou,  
Chinese Academy of Sciences (CAS), China

## \*CORRESPONDENCE

Naomi Intrator  
✉ naomii@princeton.edu

RECEIVED 26 May 2025

ACCEPTED 01 September 2025

PUBLISHED 17 September 2025

## CITATION

Intrator N and Ward BB (2025) Aquatic copper-containing nitrite reductase gene (*nirK*) phylogeny and environmental distribution.

*Front. Microbiol.* 16:1635656.  
doi: 10.3389/fmicb.2025.1635656

## COPYRIGHT

© 2025 Intrator and Ward. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Aquatic copper-containing nitrite reductase gene (*nirK*) phylogeny and environmental distribution

Naomi Intrator\* and Bess B. Ward

Department of Geosciences, Princeton University, Princeton, NJ, United States

Nitrite reduction is an essential step in the oceanic Nitrogen cycle. Nitrite reductase genes, mainly *nirS* and *nirK*, are found in dozens of phyla, are often associated with denitrifiers, ammonia- and nitrite-oxidizing bacteria (AOB and NOB) as well as ammonia-oxidizing archaea (AOA). *nirK* is found throughout the ocean, including in oxygenated surface water as well as in oxygen minimum zones (OMZs). The diverse and complex evolutionary history of the *nirK* genes makes it challenging to study the population structure and distribution of *nirK* containing organisms in the environment. The organisms containing *nirK* play key roles in the global nitrogen cycle, including the loss of fixed N, and have the potential to influence nitrous oxide (N<sub>2</sub>O) emissions via multiple pathways. This study surveyed the phylogeny and environmental distribution of over 12,000 *nirK* genes, focusing on those originating from marine and aquatic sources. Sequences were clustered into OTUs based on DNA sequence identity and their phylogeny and environmental sources were examined. The distribution of the sequences showed habitat separation within taxonomic groups, i.e., the majority of the OTUs were associated with only one environmental source. Bacterial *nirK* is more diverse phylogenetically and has a wider distribution across environmental sources than archaeal *nirK*. Most of the bacterial sequences were obtained from marine sediments, but there was variation in the dominant environmental source across phyla and classes. Archaeal sequences demonstrated niche separation between phyla as sequences from the more phylogenetically diverse phylum, Euryarchaeota, were all isolated from hypersaline environments while Nitrososphaerota sequences came from a wider range of environmental sources. This study expands the known diversity of *nirK* genes and provides a clearer picture of how *nirK* organisms are distributed across diverse environments.

## KEYWORDS

nitrite reduction, *nirK*, oxygen minimum zone, nitrogen cycling, denitrification, AOA, ammonia oxidizing archaea, nitrification

## 1 Introduction

Nitrite (NO<sub>2</sub><sup>-</sup>) is an important source of biologically available fixed nitrogen—nitrogen that has been converted from atmospheric N<sub>2</sub> into usable forms—and a key intermediate in the global nitrogen cycle. In denitrification, the major route of fixed nitrogen loss from the biosphere to the atmosphere, nitrite is reduced to gaseous nitric oxide (NO) by one of two distinct nitrite reductases: the heme-coordinating cytochrome *cd*<sub>1</sub> type encoded by *nirS* genes and the copper-containing type encoded by *nirK* genes (Hochstein and Tomlinson, 1988; Glockner et al., 1993; Zumft, 1997). The complete pathway of denitrification, in which fixed nitrogen is successively reduced through anaerobic respiration, is as follows: nitrate (NO<sub>3</sub><sup>-</sup>), NO<sub>2</sub><sup>-</sup>, nitric oxide (NO), N<sub>2</sub>O, dinitrogen gas (N<sub>2</sub>) (Zumft, 1997). The intermediate N<sub>2</sub>O is a potent greenhouse gas, with approximately 300-fold greater global warming potential per molecule compared to carbon

dioxide on the 100-year timescale, and is a significant agent in ozone depletion (Felgate et al., 2012; Ravishankara et al., 2009). Co-occurrence patterns of denitrification genes in cultivated microbes showed *nosZ* (the gene encoding the enzyme that reduces N<sub>2</sub>O) co-occurred more frequently with *nirS* than with *nirK* (Graf et al., 2014). This suggested that *nirS*-containing denitrifiers are more likely to perform complete denitrification and, therefore contribute less to N<sub>2</sub>O emissions (Graf et al., 2014), compared to those containing *nirK* (Hallin et al., 2018; Pold et al., 2024). A positive correlation between N<sub>2</sub>O emissions and *nirK* abundances, and qPCR analyses, revealed that only 10–30% of the *nirK*-containing denitrifiers also possessed *nosZ* (Clark et al., 2012). Therefore, the distribution and prevalence of *nirK* may influence net N<sub>2</sub>O emissions. Most previous research examining *nirK* focused on terrestrial sources, predominately soils, leaving *nirK* understudied in marine and other aquatic environments including hydrothermal vent systems, salt marshes, wastewaters, freshwater and estuary systems. Additionally, one-third of atmospheric N<sub>2</sub>O originates in aquatic environments, predominantly from microbial metabolism (Ciais et al., 2013), and ~20% of global N<sub>2</sub>O emissions are estimated to come from natural processes in the global ocean (Gluschankoff et al., 2023). Therefore, this study focuses on *nirK* from those marine and aquatic environmental sources.

In addition to the denitrifiers, nitrifying organisms (ammonia- and nitrite-oxidizing bacteria: AOB and NOB), also contain *nirK* (Cantera and Stein, 2007). AOB obtain energy from the oxidation of ammonia (NH<sub>3</sub>) to NO<sub>2</sub><sup>-</sup>, while NOB gain energy from the oxidation of NO<sub>2</sub><sup>-</sup> to NO<sub>3</sub><sup>-</sup>; therefore, each plays a significant role in the global nitrogen cycle (Prosser, 1986). In AOB *nirK* is involved in the production of N<sub>2</sub>O through ‘nitrifier denitrification’ under low oxygen conditions (Dundee and Hopkins, 2001; Shaw et al., 2005) and by other, somewhat undefined, pathways under oxic conditions (Freitag et al., 2012). Nitrification has been found to contribute as much, or more, N<sub>2</sub>O to the atmosphere as heterotrophic denitrification in marine ecosystems (Dore et al., 1998; Ji et al., 2018; Buitenhuis et al., 2018). Lastly, *nirK* is also found in ammonia-oxidizing archaea (AOA) (Treusch et al., 2005); however, the function of *nirK* in AOA is still unclear. Both AOB and AOA also produce N<sub>2</sub>O under aerobic conditions, although the pathways are not clear and the involvement of *nirK* is not proven. N<sub>2</sub>O and N<sub>2</sub>O yield in AOB and AOA increase with increasing NO<sub>2</sub><sup>-</sup> concentrations and with decreasing oxygen levels (Santoro et al., 2011; Peng et al., 2015). However, culture experiments have shown AOA species were unable to produce N<sub>2</sub>O through the nitrifier denitrification pathway (Stieglmeier et al., 2014). A model utilizing archaeal *nirK* abundance could predict N<sub>2</sub>O production in oceanic oxygen minimum zones (OMZs), although not through the nitrifier denitrification pathway (Trimmer et al., 2016). A ‘hybrid formation’ of N<sub>2</sub>O has been proposed in which NO, produced by the enzyme encoded by *nirK*, reacts with hydroxylamine, an intermediate of ammonia oxidation (Vajjala et al., 2012; Wan et al., 2023), to produce N<sub>2</sub>O (Stieglmeier et al., 2014; Wan et al., 2023).

Although *nirK* has been known and studied in cultivated organisms for decades (Zumft, 1997), much is still unknown about the taxonomy, function, distribution, and biogeochemical impacts of *nirK*-containing organisms in the environment. This knowledge gap is partly attributed to the high taxonomic diversity of *nirK* that has been uncovered in the environment through PCR and sequence analyses (Helen et al., 2016; Liu et al., 2003; Oakley et al., 2006; Han et al., 2022; Zheng et al., 2024). Ongoing studies continue to expand our understanding of *nirK* diversity, suggesting that novel sequences

are likely yet to be discovered. The large *nirK* sequence divergence observed may be due to *nirK*'s complex evolutionary history (Helen et al., 2016). Due to the polyphyletic distribution of *nirK* and denitrification gene phylogenies being incongruent with 16S rRNA, horizontal gene transfer has been proposed to have a substantial role in its evolution (Heylen et al., 2006; Jang et al., 2019). However, analysis of G + C and codon usage showed no evidence for horizontal gene transfer and therefore other phenomena, such as gene divergence/duplication and lineage sorting, may better explain *nirK*'s evolutionary history (Jones et al., 2008). Recent work by Ming et al. (2024) found *nirK*-carrying MAGs were more taxonomically diverse and less phylogenetically cohesive than *nirS*-carrying MAGs. They suggested that the scattered taxonomic distribution of *nirK*-containing organisms may reflect horizontal gene transfer and ecological flexibility. It has been hypothesized that sequence variation in *nirK* may be linked to ecological differences rather than ancestry alone (Pold et al., 2024).

We analyzed more than 12,000 *nirK* sequences from published databases, focusing on sequences from marine and other aquatic environments. *nirK*-containing organisms play a key role in fixed N loss pathways and in the production of N<sub>2</sub>O, both processes that may be linked to N<sub>2</sub>O emissions. Therefore, characterization of *nirK* gene diversity and environmental distribution will help to understand their biogeochemical impacts. This study analyzes the diversity and distribution of bacterial and archaeal *nirK* sequences focusing on the environmental sources (primarily marine and aquatic) of the gene to provide further insights into what environments and taxa are most influential in nitrogen transformation processes.

## 2 Materials and methods

### 2.1 Database compilation

To encompass the full diversity range of the *nirK* gene all gene features and CDS (coding sequences) of annotated *nirK* nucleotide sequences were downloaded from the National Center for Biotechnology Information (NCBI) Genbank database on June 4th, 2020. These *nirK* gene sequences were extracted from whole genomes as well as fragments from environmental clones and metagenomes. Duplicate sequences were then removed by keeping only the first instance of each sequence and its accession number, resulting in a total of 25,925 *nirK* sequences. Taxonomy was determined for each sequence using a modified code developed from the R package, Taxonomizr (Sherrill-Mix, 2018). Sequences were separated into domains: Bacteria, Archaea and Eukaryota (24,776, 971 and 208 sequences respectively). For this study Eukaryota sequences were removed from further analysis. Sequences were filtered by length and only sequences with length ≥400 bp for bacteria (11,638 sequences) and ≥300 bp (894 sequences) for archaea were retained for analysis.

### 2.2 Database processing

The 12,532 bacterial and archaeal *nirK* sequences were clustered using CD-HIT (Li and Godzik, 2006; Fu et al., 2012) with a defined threshold of 87% sequence identity (Taroncher-Oldenburg et al., 2003; Ward et al., 2007; Bulow et al., 2008). This threshold was chosen to

cluster homologous genes at the level found to differentiate functional genes at the approximate level of species distinction initially reported for the *amoA* gene in AOB (Purkhold et al., 2000). The 87% threshold probably underestimates the number of “species” in the *nirK* database, compared to a genomic species definition of >95% Average Nucleotide Identity (ANI; Thompson et al., 2021). CD-HIT generated 2,185 bacterial and 86 archaeal clusters. The environmental sources of the sequences within each cluster were determined by searching NCBI and published literature. Metadata, including the environmental source and phylogenetic assignment (when available), was collected by searching the representative region using blastn optimized for highly similar sequences on NCBI (Camacho et al., 2009) (Supplementary Tables 1, 2).

Sequences were labeled with the following environmental sources: marine sediment, marine water column, marine OMZs, freshwater and estuary systems, hydrothermal vent systems, wastewater systems, aquatic other, terrestrial and animal (i.e., from the animal microbiome). Clusters that did not contain any marine or aquatic sequences were removed from further consideration. Approximately 30% of the 12,532 bacterial and archaeal *nirK* sequences were further analyzed after removing those outside the scope of this study, i.e., sequences that predominantly originated from terrestrial (mainly soil) sources and did not cluster with any aquatic sequences. The remaining clusters containing more than three sequences were each aligned with MAFFT v7.407 (Katoh and Standley, 2013), and a consensus sequence for each cluster was generated with emboss v6.6.0 (Stamatakis, 2014) (i.e., generating 180 bacterial and 19 archaeal consensus sequences). Clusters containing one or two sequences were removed from their cluster and treated as individual sequences that are identified by their accession numbers. This resulted in 180 marine and aquatic bacterial clusters (representing 2,327 sequences) and 655 unique bacterial sequences as well as 19 marine and aquatic archaeal clusters (representing 794 sequences) and 65 unique archaeal sequences.

Alignments of the consensus sequences and individual sequences for bacterial and archaeal sequences were made using MAFFT v7.407 (Katoh and Standley, 2013). Sequence lengths varied depending on how they were originally obtained (e.g., whole genomes vs. PCR fragments). To account for this, we extracted the longest continuous alignment region where the greatest number of clusters and individual sequences overlapped. This region was then used for downstream analyses. A 277 bp region was chosen for bacteria and a 161 bp region was used for archaea; however, 407 individual bacterial sequences did not align in the 277 bp region and were therefore not considered in further analyses (Supplementary Tables 1, 2). For the bacterial region, compared to the reference sequence of the *Nitrosomonas marina* strain Nm71's *nirK* gene (1,105 bp: NZ\_FOIA01000003.1), the 277-mer region is located around ~541–818 bp. The 161 bp archaeal region is around ~222–383 bp compared to the reference sequence of *Nitrosopumilus adriaticus* strain NF's *nirK* gene (1,425 bp: CP011070.1, Bayer et al., 2016).

The probe finding algorithm of Bulow et al. (2008) was used to further group the 277 and 161 bp regions into operational taxonomic units (OTUs) and assign a representative sequence for each OTU. An identity threshold of 87% was used to provide the optimal discrimination between related sequences at approximately the species level for previously studied functional genes (Ward et al., 2007; Bulow et al., 2008). The algorithm produced 353 bacterial and 50 archaeal OTUs. Phylogenetic trees were generated with the aligned bacterial and archaeal OTUs using RAXML v8.2.12 (Stamatakis, 2014) and viewed and edited in iTOL v6 (Letunic and Bork, 2006). Ten bacterial

clusters representing 40 sequences were removed from the phylogenetic tree (but can be found in Supplementary Table 1) because their extreme divergence prevented alignment. The phylogenetic trees represent a total of 2,552 bacterial sequences and 856 archaeal sequences.

## 3 Results

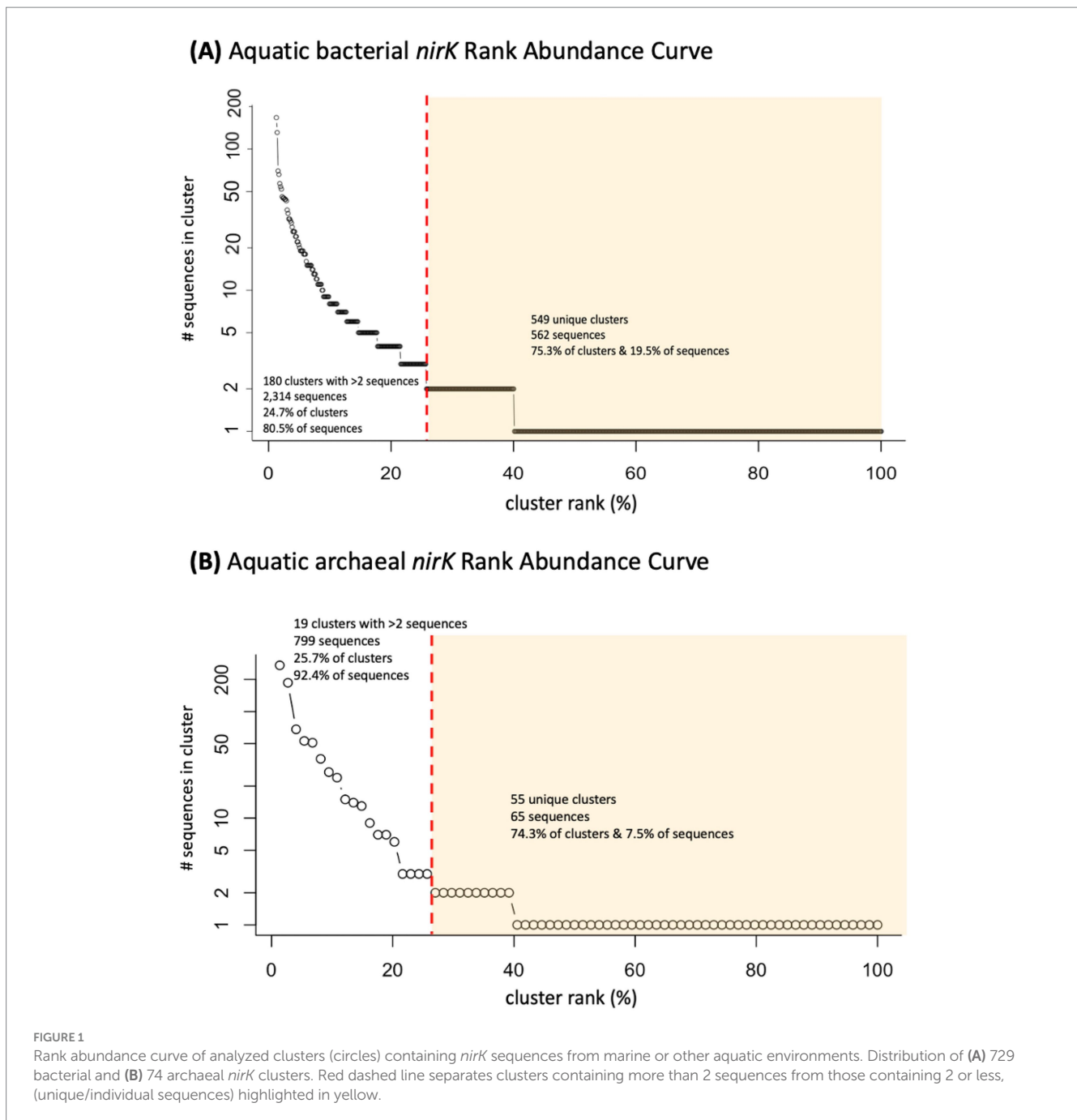
### 3.1 Aquatic *nirK* sequence compilation overview

A total of 12,532 *nirK* sequences were compiled and processed. Following clustering using CD-HIT and removal of clusters comprised of only non-aquatic sequences, 2,876 unique aquatic bacterial and 864 archaeal *nirK* sequences were used for further analyses. Most of the sequences were derived from PCR amplification but several sequences were pulled from full genomes. These aquatic sequences grouped into 729 bacterial and 74 archaeal clusters. Of those, 180 bacterial and 19 archaeal clusters contained more than 2 sequences. These 180 bacterial and 19 archaeal aquatic clusters represent only 24.7 and 25.7% of the clusters, but account for 80.5% (2,314 sequences) and 92.4% (799 sequences) of the bacterial and archaeal sequences, respectively. The two largest aquatic clusters contain 167 bacterial and 271 archaeal sequences (5.8% of aquatic bacterial and 31.4% of archaeal sequences). The remaining bacterial and archaeal clusters, i.e., the unique individual sequences, represent 75.3 and 74.3% of the clusters respectively, and 19.5 and 7.5% of the bacterial and archaeal sequences (562 and 65 sequences respectively). This distribution of sequences shows that while the majority of the clusters are unique, most of the sequences group into larger clusters (Figure 1), particularly for archaeal *nirK* sequences.

### 3.2 Bacterial *nirK* environmental distribution and phylogeny

The majority of the *nirK* sequences were bacterial. A total of 2,552 bacterial *nirK* sequences were grouped into 353 OTUs (based on the aligned fragment, 277 bp), which were further examined. Two hundred fourteen OTUs represent 254 unique individual sequences only (i.e., do not contain clusters). The remaining 139 OTUs contained the majority of the sequences (2,298 sequences). The largest OTU represents 167 sequences (Figure 1A). Most of the OTUs (267 OTUs, 75.6%) contain sequences obtained from only one of the environmental sources. However, there are three OTUs (one uncultured Alphaproteobacterium and two unclassified) represented by sequences from five different environmental sources, and four OTUs (three unclassified and one *Pseudomonas* sp.) representing four different environmental sources. The remainder of the OTUs represent three or fewer different environmental sources. Figure 2 depicts the phylogeny of the bacterial *nirK* sequences based on the aligned 277 bp region of the gene and shows the spread of environmental sources across OTUs. The environmental sources for each OTU (based on the 277 bp OTU analysis) were identified and the total number of OTUs representing an environmental source was determined. Bacterial OTUs were predominantly obtained from marine sediment (40.2%) followed by freshwater and estuary systems (28%) (Figure 3A).

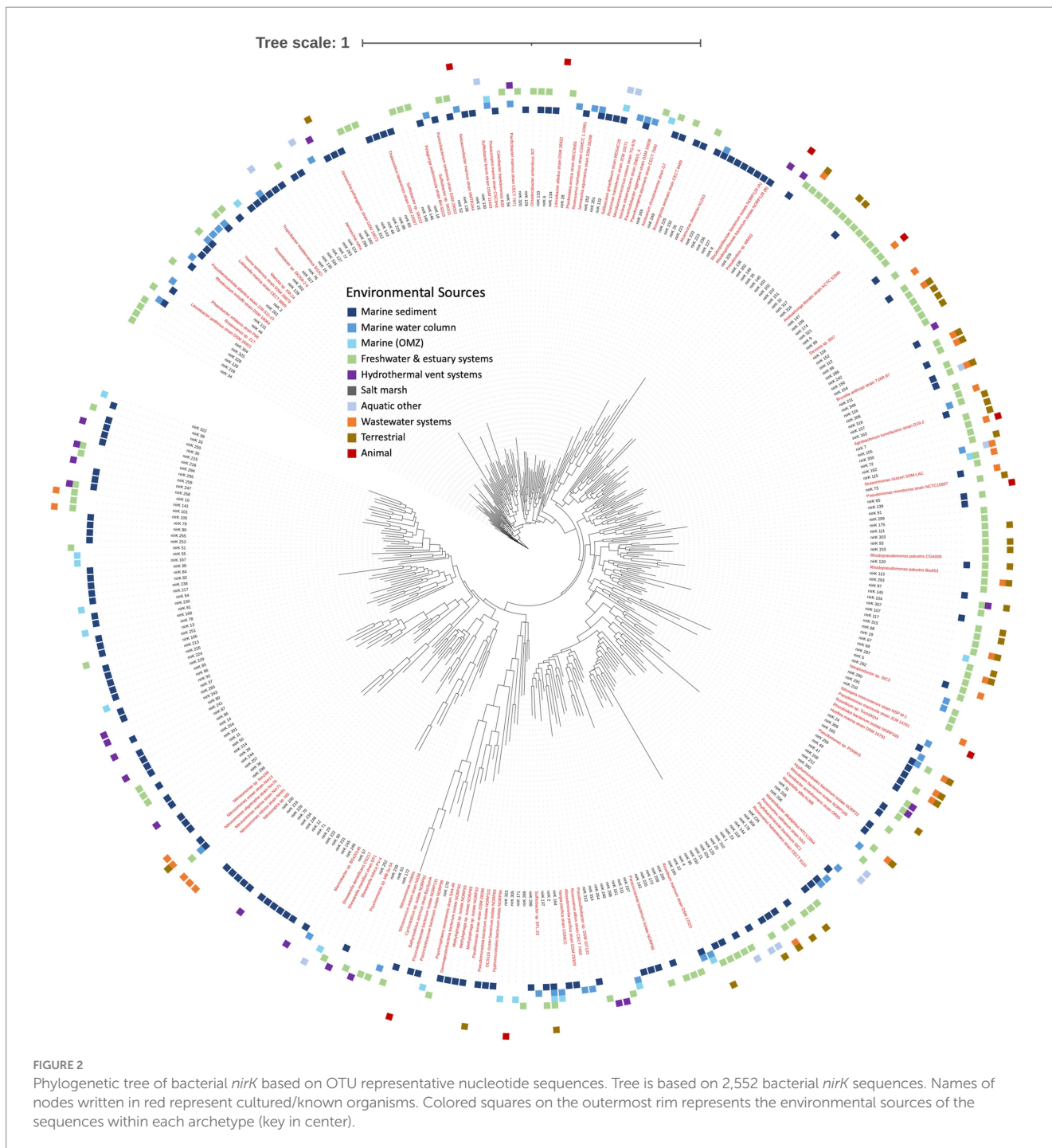
The taxonomy of each OTU's representative sequence (Figure 4A) was determined through phylogenetic inference based on the tree in



**Figure 2.** The phylum and class of the sequences within the bacterial *nirK* OTUs could be determined for 40 and 32.3% of OTUs respectively, while 60% of OTUs remained unclassified. Only 30.4, 28.2 and 18% could be assigned to the order, family, and genus, respectively, (Supplementary Table 1). The bacterial *nirK* genes that could be assigned phylogeny primarily consisted of sequences from the Pseudomonadota phylum (38.3%), with most classified as Alphaproteobacteria (20.9%), followed by Gammaproteobacteria (7.5%) and Betaproteobacteria (2.2%). The phylum Bacteroidota only made up 1.4% of OTUs and phylum Nitrospirota was the smallest represented phylum with only 1 OTU (0.3%). A few OTUs representing bacterial ammonia and nitrite oxidizers (AOB and NOB) were identified. Nine OTUs were classified as known AOB, encompassing a total of 15 sequences. Among these,

most were classified as Betaproteobacteria (8 OTUs, 14 sequences), with only one OTU/sequence belonging to the Gammaproteobacteria class.

The environmental sources of the bacterial *nirK* sequences are shown in the outer wheel of Figure 2 while Figure 5 summarizes the breakdown of environmental sources within the phyla and those which could not be classified, i.e., unknown. The dominant phylum, Pseudomonadota, was examined further by determining what percent of OTUs came from each environmental source (Figure 5). Pseudomonadota, which represented 139 OTUs (677 sequences), was primarily obtained from marine sediments (44.6%), followed by the marine water column (23.7%). The Pseudomonadota classes were further analyzed. Alphaproteobacteria (76 OTUs, 480 sequences) was predominantly retrieved from the marine water column (35.5%), closely

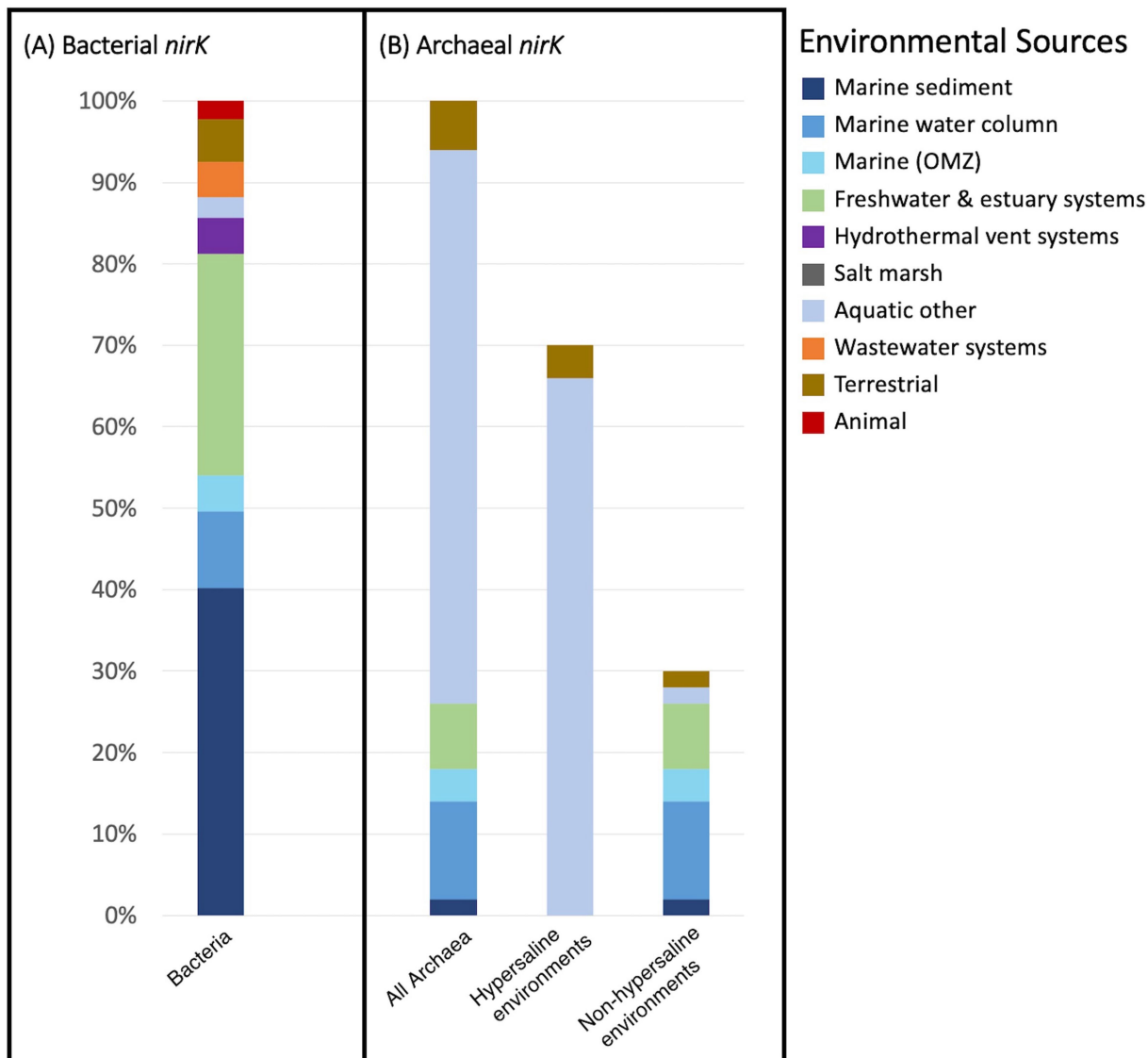


followed by marine sediment (32.9%). The main environmental source of Betaproteobacteria (8 OTUs, 58 sequences) was wastewater systems (62.5%). Lastly, Gammaproteobacteria sequences (27 OTUs, 78 sequences) were predominantly obtained from marine sediment (33.3%), followed by freshwater and estuarine systems (18.5%). On the other hand, the main environmental sources for Bacteroidota were terrestrial (60%); however, Bacteroidota only comprises 5 OTUs (15 sequences). The only OTU classified in the phylum Nitrospirota was obtained from a wastewater system. Most of the AOB OTUs were also obtained from wastewater systems. Those OTUs that could not be classified into phyla (218 OTUs, 1,899 sequences) were mostly obtained from marine sediment and freshwater and estuarine systems

(both 38.1%), followed by hydrothermal vent systems (6.9%). No sequences with unknown taxonomy were obtained from the marine water column.

### 3.3 Archaeal *nirK* environmental distribution and phylogeny

A total of 857 archaeal *nirK* sequences were grouped into 50 OTUs, which were further analyzed. Fifty-nine unique individual sequences made up 44 OTUs (containing 1–2 sequences each) while most of the sequences (805 sequences) were clustered into 14 OTUs.

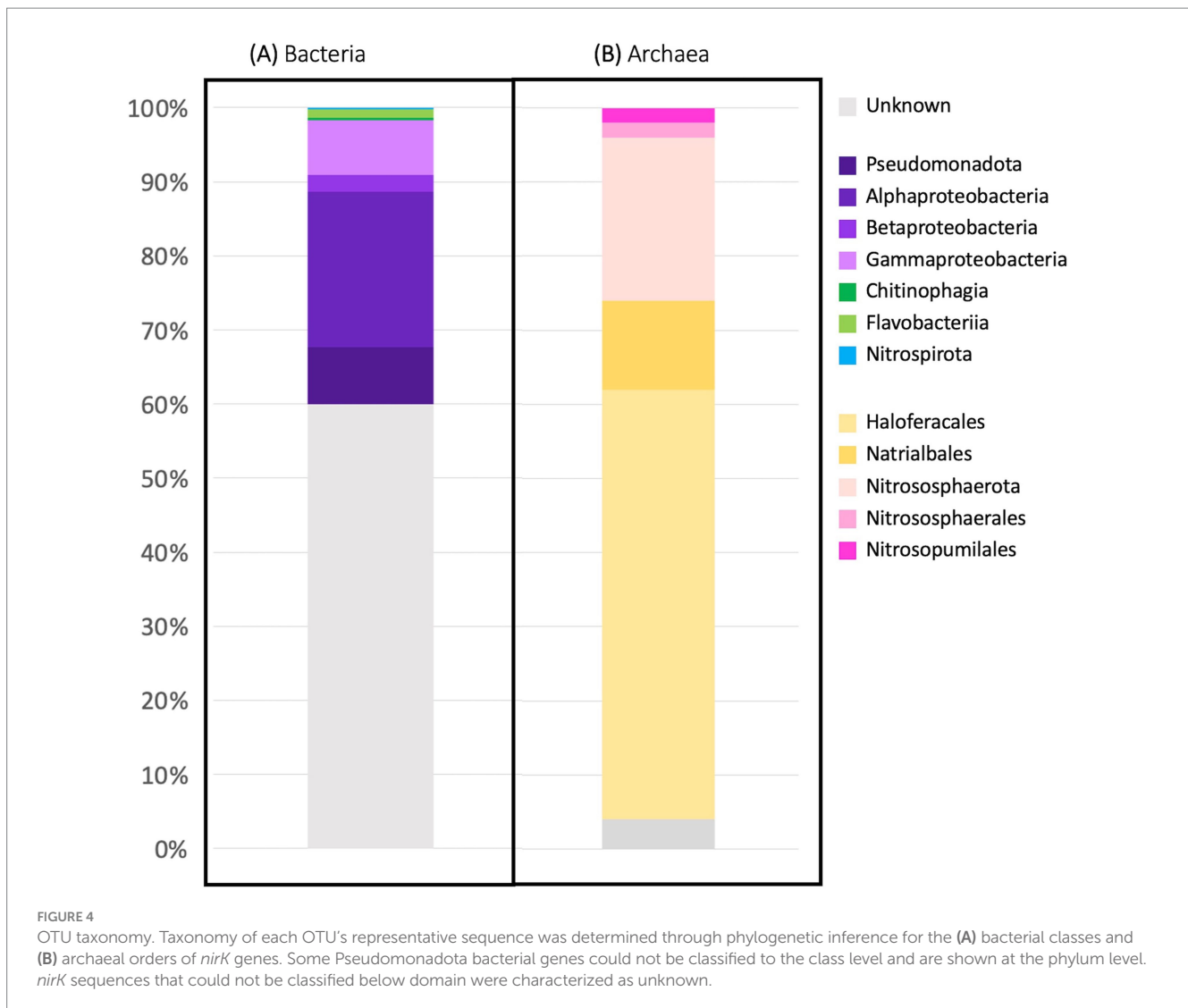


**FIGURE 3**  
Percent of representative archetype sequences found from each environmental source. Bar plots display the percent of OTU representative sequences found from each environmental source for (A) the 353 bacterial OTUs and (B) the 58 archaeal OTUs. Archaeal OTUs are further separated into those obtained from hypersaline environments and non-hypersaline environments.

The largest archaeal OTU represented 275 sequences (Figure 1B). Similar to the bacterial *nirK* OTUs, most of the OTUs from the archaea (37 OTUs, 74%) represent only one environmental source. Only one OTU (260 sequences) represents sequences from three different environmental sources and was classified as an uncultured Nitrososphaerota (formerly known as Thaumarchaeota). The phylogenetic tree of the representative archaeal *nirK* OTU sequences, based on the 161 bp region of the gene (Figure 6) portrays the spread of the environmental sources across OTUs. Figure 3B depicts the fraction of environmental sources for all OTU representative sequences. The majority of the archaeal *nirK* sequences were obtained from “other” aquatic environments (68%) such as salt/saline lakes, solar salterns, and brines, i.e., mostly hypersaline environments. Hypersaline environments were separated from the remaining representative environmental sources (Figure 3B). Of the archaeal OTUs obtained from non-hypersaline environments, 40% were

retrieved from the marine water column followed by 26.7% from freshwater and estuary systems.

Unlike the bacterial sequences, most of the archaeal *nirK* sequences could be classified into phylum/class (96%/74%) as well as 74% classified into order, family and genus (Figure 4B and Supplementary Table 2). Only two phyla were represented by the archaeal OTUs: Euryarchaeota (70%) and Nitrososphaerota (26%). All sequences originating from hypersaline environments were assigned to the phylum and class Euryarchaeota Halobacteria, while all those in non-hypersaline environments are within phylum Nitrososphaerota. Since Halobacteria within the phylum Euryarchaeota are exclusively found in hypersaline environments, the bar in Figure 3B labeled “hypersaline” reflects only Euryarchaeota, with environmental sources distributed as 97.1% aquatic (other) and 2.9% terrestrial. On the other hand, the bar showing non-hypersaline environments generally represents the environmental sources of the phylum Nitrososphaerota.



The two unknown OTUs were obtained from freshwater and estuary systems.

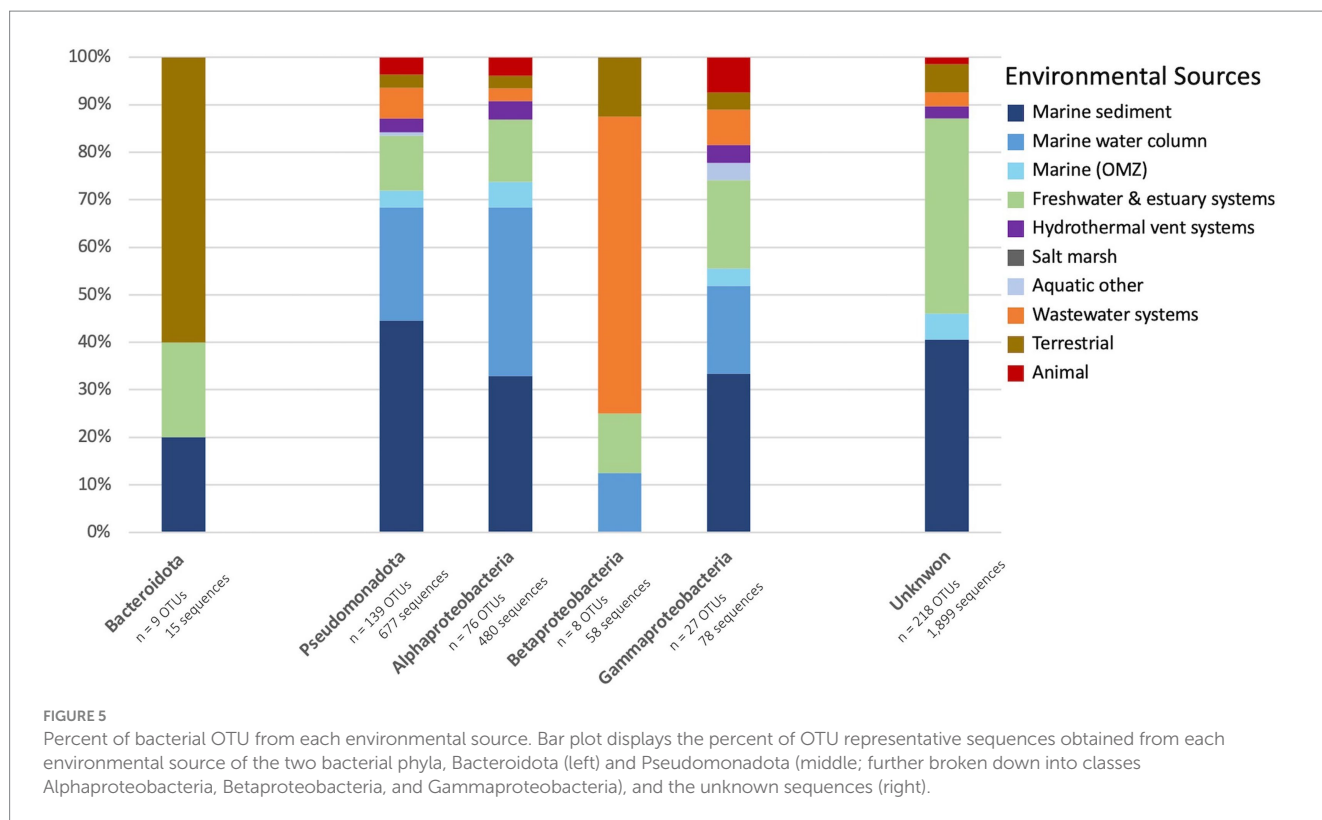
## 4 Discussion

### 4.1 *nirK* diversity

Despite decades of research on the *nirK* gene, our knowledge of the function, distribution, and taxonomy of *nirK*-containing organisms in the environment remains incomplete. This is partly attributed to the broad taxonomic diversity of organisms harboring *nirK* genes, as identified through phylogenetic analysis and high-throughput sequencing of these genes in natural ecosystems (Cantera and Stein, 2007; Helen et al., 2016; Sun and Jiang, 2022). Differences in the structure of *nirK* genes, functional domains, operons, and transcriptional regulatory signals suggest multiple functions (Cantera and Stein, 2007). This extensive diversity is illustrated in Figure 1, where approximately 75.3% of bacterial clusters and 74.3% of archaeal clusters are unique sequences. After further grouping into OTUs, ~50% of both the bacterial and archaeal OTUs represent individual

sequences, meaning approximately half of the branches on the phylogenetic trees (Figures 2, 6) represent unique sequences. Thus, although the OTU cutoff of 87% identity used here likely underestimates the number of different *nirK* sequences, the fact that the majority of the *nirK* sequences could not be assigned to phylogeny at the phylum level means that greater resolution in the “species” identity of the sequences would not be helpful at this time.

The high diversity of the *nirK* gene implies potential detection and analytical challenges, which suggests the size, distribution, and diversity of the microbial community of *nirK*-containing organisms is underestimated (Helen et al., 2016). A study by Cantera and Stein (2007) found previously published PCR primers used in various studies were unable to detect *nirK* genes from archaeal and several bacterial lineages from genomic DNA. In addition, Bonilla-Rosso et al. (2016) evaluated 79 *nirK* primers and determined there were no optimal universal primer pairs. Although this study examined a wide diversity of *nirK* sequences, we were limited to analyzing those publicly available from previous studies, and which met minimal length criteria, and therefore some diversity is likely missing due to the detection and analytical challenges recognized in environmental studies.



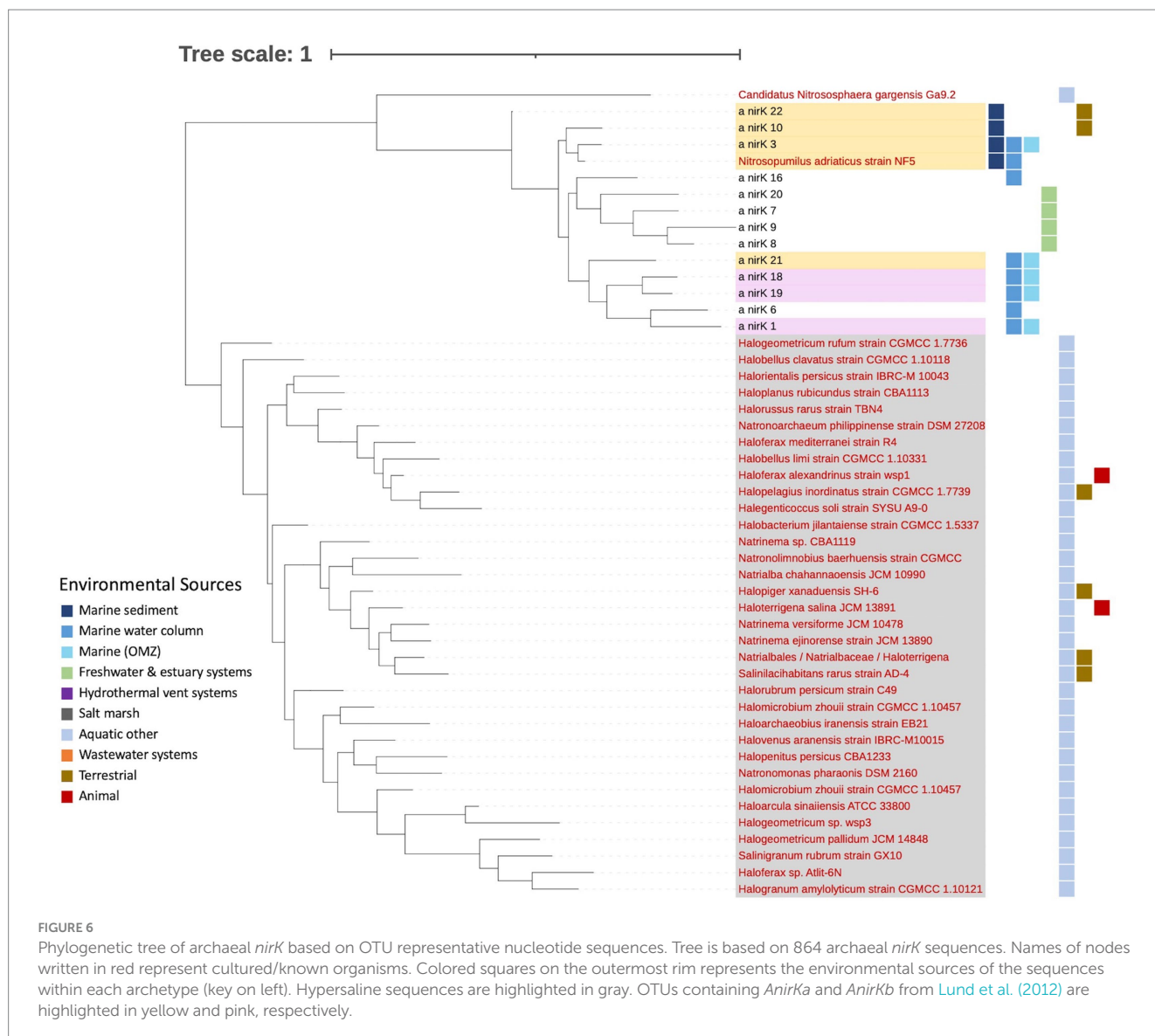
#### 4.1.1 Taxonomic diversity

The bacterial *nirK* represented the majority (95.5%) of the total aquatic *nirK* gene sequences analyzed here while archaeal sequences only made up 3.7%. These values are similar to a global prokaryotic census that found bacteria and archaea represented 93.3 and 6.7%, respectively, of over one billion prokaryotic reads (Louca et al., 2019). The smaller percentage of archaeal *nirK* (3.7% vs. 6.7%) compared to the global could be due to a research bias in microbial studies as more research has focused on bacteria.

The aquatic bacterial *nirK* sequences appear to be more phylogenetically diverse than the archaeal sequences (Figures 2, 6). The largest aquatic clusters contained 167 bacterial vs. 275 archaeal sequences, which represented 5.8% of aquatic bacterial and 32% archaeal sequences (Figure 1). Therefore, although more bacterial sequences were analyzed, they were more dissimilar to each other such that they grouped into smaller clusters. A higher diversity of bacterial vs. archaeal *nirK* sequences has been seen in other studies (Hansel et al., 2008; Helen et al., 2016; Ji et al., 2020) and is expected due to the significantly larger number of bacterial *nirK* sequences. In addition, a study by Hou et al. (2013), looking at *amoA*, which encodes the ammonia monooxygenase enzyme, and *nirK/nirS* genes, showed that despite a greater abundance of AOA vs. AOB, bacterial sequences still displayed greater phylogenetic diversity. In this study, the bacterial sequences came from a wider taxonomic range than observed for archaeal sequences (Figure 4). The bacterial *nirK* sequences in this study were identified as belonging to 3 distinct phyla, 6 classes and 53 genera, while archaeal sequences only included 2 phyla, 2 classes and 28 genera (Supplementary Tables 1, 2). Although they contain much unidentified and unexplored diversity within those phyla, the aquatic sequences analyzed here represent only a small fraction of the total global diversity of *nirK* described by Pold et al.

(2024), who found *nirK* in at least 24 bacterial phyla, as well as the two archaeal phyla also represented here. Pold et al. (2024) divided the world into 28 biomes, only eight of which were the subject of the current study. Thus, a large portion of the bacterial *nirK* diversity found by Pold et al. (2024) came from terrestrial sources. Nevertheless, metagenomes from the eight aquatic biomes investigated by Pold et al. (2024), particularly the mid and high latitude ocean biomes, contained some of the highest *nirK* abundances and diversity per gigabase sequenced, which is consistent with the breadth of diversity we report here for aquatic systems.

Although Pseudomonadota was the largest phylum found among the bacterial sequences, most of the OTUs and sequences were unknown and could not be identified at the phylum level (60 and 73.3% respectively). Pseudomonadota (formerly known as the phylum Proteobacteria) has been identified as the dominant phylum in several environmental studies including those from marine sediments, marine OMZs, freshwater and estuary systems (respectively: Lee and Francis, 2017; Muck et al., 2019; Chen et al., 2017; Li et al., 2017). This is in contrast to the archaeal *nirK* sequences where the majority of the representative sequences could be identified and were classified in either the Nitrososphaerota or Euryarchaeota. Although the Euryarchaeota phylum represents the majority of the OTUs (70%), it only represents only a small portion of sequences 8.2% (i.e., the majority of the Euryarchaeota OTUs came from unique individual sequences). On the other hand, the Nitrososphaerota phylum represented 26% of OTUs but 87.2% of all archaeal sequences. This implies the Euryarchaeota phylum is more diverse than the Nitrososphaerota phylum in aquatic environments, which may be related to the environmental conditions in which these organisms are found (Castro-Fernandez et al., 2017; see section 4.2).



## 4.2 Ecological distribution (environmental sources) and niche differentiation

Due to the extensive and continually expanding understanding of *nirK* gene diversity, it is challenging to fully understand and study the ecological distribution of *nirK*-containing organisms. In this study we analyzed the environmental sources and environmental distribution of the *nirK* gene. Starting with a database of over 12,000 sequences, we clustered and identified *nirK* sequences through phylogenetic inference, subsequently identifying their environmental origins (Figures 2, 3, 6). Although our data compilation also included *nirK* sequences from published whole genomes, it is crucial to acknowledge that many sequences in this study were obtained via PCR amplification and clone libraries. Consequently, the patterns observed in the environmental sources and distribution discussed below are subject to PCR bias. With that caveat in mind, it is intriguing to explore the biogeographical and ecological patterns present in our extensive *nirK* database.

Most OTUs contained sequences from a single environmental source (75.6 and 74% of OTUs for bacterial and archaeal sequences),

indicating they are unique to their specific environments (i.e., display biogeography). This supports previous findings of niche separation found between and within environmental sources (Oakley et al., 2006; Wang et al., 2014; Lee and Francis, 2017; Saarenheimo et al., 2017; Muck et al., 2019; Hatzepichler, 2012 and references therein). In this study, bacterial OTUs have a wider distribution across our defined environmental sources than archaeal OTUs (Figures 2, 3, 6). This implies bacterial *nirK* microbes are adapted to a broader range of environments than archaeal *nirK* organisms, consistent with the greater gene diversity observed. Pold et al. (2024) found that *nirK* diversity and abundance usually far exceeded that of *nirS* (across 28 different environments), and that specific clades of both genes were associated with specific environments. *nirS* and *nirK* are associated with different biomes and different environmental conditions, suggesting different functionality.

Bacterial *nirK* OTUs were predominantly retrieved from marine sediment while the majority of the archaeal OTUs were obtained from other aquatic environments. These other aquatic environments are primarily extreme conditions, particularly those with high salinity (e.g., salt/saline lakes, solar salterns, and brines). Niche differentiation

was evident, as all sequences originating from hypersaline environments were assigned to the Euryarchaeota phylum. The Euryarchaeota phylum exhibited greater phylogenetic diversity, containing more unique individual sequences than the other archaeal phylum found in our study, Nitrososphaerota. This diversity in Euryarchaeota is underpinned by their broad range of metabolic capabilities and environmental adaptations, allowing them to thrive in various extreme environments (Castro-Fernandez et al., 2017; Martinez-Espinosa, 2020). On the other hand, Nitrososphaerota, while being phylogenetically less diverse, exhibits greater ecological diversity as it is found in a wider range of environmental sources (Figure 3B). This broad ecological distribution reflects its critical role in the global nitrogen cycle, particularly in ammonia oxidation (i.e., AOA), which is a key component of the N cycle in all environments (Zheng et al., 2024). However, the functional role of *nirK* in these AOA remains uncertain (Lund et al., 2012; Kop et al., 2024). Thus, this niche differentiation implies archaeal *nirK*-containing organisms contribute to the nitrogen cycle in distinct ways: Euryarchaeota likely engaging in denitrification in hypersaline environments (Torregrosa-Crespo et al., 2018) and Nitrososphaerota in ammonia oxidation in a wider range of environments.

A similar study (Biller et al., 2012) focused on the archaeal *amoA* gene, which encodes the  $\alpha$ -subunit of the ammonia monooxygenase enzyme required for ammonia oxidation. That study aimed to understand the ecological factors influencing the distribution and diversity of AOA. Examining over 8,000 *amoA* sequences, they found, similar to our study, that the majority of sequences tend to cluster with others from the same environment. The findings by Biller et al. (2012), like ours, support the existence of multiple distinct aquatic AOA clades in the environment. They suggested that factors such as salinity, depth, and temperature could be selective pressures driving the niche partitioning of AOA and influencing *amoA* diversity. Their study found ~75% of their total *amoA* sequences to be unique, very similar to the 74.3% we found for archaeal *nirK*. However, unlike our study, the majority of their *amoA* sequences came from coastal sediments, whereas our study found that the *nirK* sequences from Nitrososphaerota, likely representing AOA, were predominantly obtained from the marine water column. This discrepancy maybe due to a sampling bias between research conducted on *amoA* versus *nirK*. Additionally, 16S rRNA gene abundances of Thaumarchaeota (Nitrososphaerota) were greater than AOA *amoA* suggesting not all Thaumarchaeota contain *amoA* or were missed due to prior primer bias (Tang et al., 2023). Biller et al. (2012) also revealed, similar to our study, hydrothermal vents and wastewater systems were underrepresented environments in terms of total numbers of sequences (Figure 3B). Notably, 75% of the AOB OTUs were obtained from wastewater systems. This distribution is consistent with previous studies that have highlighted the prevalence of AOB and Betaproteobacteria in wastewater treatment systems (Yin et al., 2018; Wu et al., 2016) while AOA dominate marine systems.

The diversity of Nitrososphaerota (Thaumarchaeal) *nirK* genes in coastal and marine environments was previously explored by Lund et al. (2012), who identified two distinct groups, AnirKa and AnirKb, based on primer sets. The two variants showed contrasting distributions in the water column of Monterey Bay and the California Current. AnirKa was more prevalent in the epi- to mesopelagic Monterey Bay, while AnirKb was dominant in the meso- to bathypelagic California

Current. In addition, only AnirKa could be detected in sediments. This separation persists in our analyses, as the AnirKa and AnirKb sequences of Lund et al. (2012) fell into separate clusters (Figure 6; yellow represents AnirKa, pink represents AnirKb). This separation further demonstrates niche differentiation between *nirK* sequences. Metadata such as depth, as shown by Lund et al. (2012), could reveal further unseen separations in the rest of our sequences and provide deeper insights into the ecological roles and adaptations of these Nitrososphaerota groups in different marine environments. *nirK* resolved the diversity of AOA better than *amoA*, distinguishing three pelagic ecotypes in the marine water column (Reij et al., 2019).

The widespread distribution and overall large diversity of *nirK*-containing organisms suggest that the ability to reduce nitrite is an ecologically significant and favorable trait. While not solely linked to canonical denitrification, this ability may partially regulate net N<sub>2</sub>O emissions (Hallin et al., 2018; Clark et al., 2012; Shaw et al., 2005). Such large gene diversity and distribution suggest differing *nirK*-containing organisms will have differing nitrite affinities and reduction efficiencies, potentially influenced by genetic diversity and ecological adaptations. Several factors such as oxygen concentration, nutrient availability, redox conditions, etc., could influence the observed environmental distribution and should be examined further. For instance, in a study in the Yellow River Estuary phylogenetic analysis revealed differing *nirK* clusters between sites of high versus low dissolved oxygen concentrations (Li et al., 2017). Principal-component analysis of biogeochemical data and denitrifying organisms conducted along the Pacific coast of Mexico revealed nitrate concentrations and oxygen levels were key factors controlling the structure of denitrifying communities (Liu et al., 2003). Similarly, in rice paddies changes in redox chemistry in the water led to shifts in the active microbial community and expression of functional genes including *nirK* (Yoshida et al., 2009). Phylogenetically distinct organisms exhibit unique ranges and tolerances to environmental conditions, often enabling them to occupy specific ecological niches. The spread of these diverse taxa across environments has implications for global N-cycling and is an active area of study.

Niche differentiation was also observed between and within the bacterial phyla, Pseudomonadota and Bacteroidota (Figure 5). Bacteroidota sequences were primarily obtained from terrestrial environments and Pseudomonadota from marine sediment. The Pseudomonadota classes were further analyzed to resolve finer scale patterns in environmental distribution. The largest class, Alphaproteobacteria, was mainly derived from the marine water column and marine sediment (35.5 and 32.9% respectively). Similarly, Gammaproteobacteria sequences were predominantly obtained from marine sediment (33.3%), followed by the marine water column and freshwater and estuary systems (both 18.5%). However, the smallest class, Betaproteobacteria, was dominated by sequences from wastewater systems (62.5%), including the AOB. The variation in environmental sources between classes suggests that environmental selection operates at a finer scale than just the phylum level. This indicates the need for further studies to determine the precise extent of biogeography and niche differentiation at these finer taxonomic levels. Niche partitioning has been found in relation to particulate concentrations, salinity, nitrogen concentrations, organic matter, temperature, pH, oxygen levels, etc. (Muck et al., 2019; Jones and Hallin, 2010; Lee and Francis, 2017; Hatzenpichler, 2012); likely, there is no single driving factor that controls *nirK* niche separation in the

wider context of microbial genomes. Lastly, the majority of the bacterial *nirK* sequences could not be classified into specific phyla. These unknown sequences were predominantly obtained from marine sediment and freshwater and estuary systems. This indicates that these environments contain phylogenetic and metabolic diversity that should be explored to elucidate the potential roles for *nirK* in nitrogen cycle and respiratory functions.

### 4.3 Summary and future directions

This study expands our understanding of the diversity and ecological distribution of *nirK*-containing organisms in marine and other aquatic environments. The *nirK* gene is found in organisms in a wide range of genera inhabiting diverse environments. The organisms containing *nirK* play essential roles in the global nitrogen cycle through a variety of pathways and have the potential to influence N<sub>2</sub>O emissions. Although *nirK* was characterized over two decades ago (Zumft, 1997), much remains unknown about the distribution, taxonomy, and biogeochemical impacts on the environment. This study analyzed over 12,000 *nirK* sequences and found large diversity of the gene: based on clustering at 87% identity, the majority of the OTUs were unique. This extensive sequence diversity introduces detection and analytical challenges. PCR primers can only detect a limited range of the phylogenetically diverse *nirK* and therefore universal PCR primers for *nirK* are not possible. In turn, it is likely that *nirK* diversity and environmental abundance are greatly underestimated. The ecological significance of this underestimation should be further examined, preferably using metagenomics to avoid PCR primer bias.

We found bacterial *nirK* to have a wider taxonomic range and distribution across environmental sources compared to archaeal *nirK*. Differences in community composition and environmental distribution among *nirK*-containing organisms are likely to impact environmental nitrogen cycling causing various ecological and environmental consequences, because the function of *nirK* varies among organisms (e.g., denitrifiers, NOB, AOA, and AOB). However, the exploration of these consequences is beyond the scope of this study and additional field studies are required to understand the biogeochemical impacts of this variation. The majority of the OTUs analyzed in this study were each obtained from only one environmental source, suggesting most *nirK* organisms are specific to an environment. The spread of the diverse taxa across environments has implications for nitrogen cycling and its connections with other diverse metabolisms.

This study examined the environmental distribution of the *nirK* gene with a focus on marine and other aquatic environments to explore the breadth of environmental sources. The majority of the sequences were bacterial, which were predominately obtained from marine sediments. In addition, most bacterial OTUs could not be taxonomically classified even to the phylum level. Pseudomonadota was the largest identified phylum of bacterial *nirK*, most of which were Alphaproteobacteria. Niche differentiation was observed between and within the bacterial phyla. Bacteroidota sequences were primarily obtained from terrestrial environments and Pseudomonadota from marine sediment. Finer scale patterns in environmental distribution were observed by looking at the Pseudomonadota classes, where

Alpha- and Gammaproteobacteria were primarily isolated from marine sources (water column and sediment), while Betaproteobacteria was dominated by sequences from wastewater systems. Archaeal sequences also demonstrated niche differentiation as the more phylogenetically diverse phylum, Euryarchaeota, were all isolated from extreme hypersaline environments while Nitrososphaerota sequences came from a wider range of environmental sources. The mechanisms responsible for the environmental distribution and niche separation of *nirK*, such as nutrient availability, oxygen concentration, redox state, etc., are not well understood and should be examined further as it impacts nitrogen and biogeochemical cycling.

A major challenge in establishing *nirK* sequence diversity in an environmental context was the lack of standardized environmental metadata such as temperature, salinity, latitude and longitude, dissolved oxygen concentrations, depth, etc. Understanding the factors that influence the environmental distribution and niche separation of *nirK* in different environments requires such data. These limitations highlight the necessity of additional studies on cultivated organisms to better understand the specific factors driving *nirK* niche differentiation.

### Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author/s.

### Author contributions

NI: Formal analysis, Methodology, Investigation, Writing – original draft, Visualization, Data curation, Conceptualization. BW: Conceptualization, Supervision, Writing – review & editing, Project administration, Funding acquisition, Resources.

### Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by National Science Foundation grants to BBW (1657663) (1946516).

### Acknowledgments

We would like to acknowledge Maggie C. Y. Lau (University of Chinese Academy of Sciences) for her suggestions on the best method approaches. We appreciate Mandy Lin for her assistance determining environmental sources of sequences.

### Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations,

## References

- Bayer, B., Vojvoda, J., Offre, P., Alves, R. J. E., Elisabeth, N. H., Garcia, J. A. L., et al. (2016). Physiological and genomic characterization of two novel marine thaumarchaeal strains indicates niche differentiation. *ISME J.* 10, 1051–1063. doi: 10.1038/ismej.2015.200
- Biller, S. J., Mosier, A. C., Wells, G. F., and Francis, C. A. (2012). Global biodiversity of aquatic ammonia-oxidizing archaea is partitioned by habitat. *Front. Microbiol.* 3:252. doi: 10.3389/fmicb.2012.00252
- Bonilla-Rosso, G., Wittorf, L., Jones, C. M., and Hallin, S. (2016). Design and evaluation of primers targeting genes encoding no-forming nitrite reductases: implications for ecological inference of denitrifying communities. *Sci. Rep.* 6. doi: 10.1038/srep39208
- Buitenhuis, E. T., Suntharalingam, P., and Le Querre, C. (2018). Constraints on global oceanic emissions of N<sub>2</sub>O from observations and models. *Biogeosci.* 15, 2161–2175. doi: 10.5194/bg-15-2161-2018
- Bulow, S. E., Francis, C. A., Jackson, G. A., and Ward, B. B. (2008). Sediment denitrifier community composition and *nirS* gene expression investigated with functional gene microarrays. *Environ. Microbiol.* 10, 3057–3069. doi: 10.1111/j.1462-2920.2008.01765.x
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST architecture and applications. *BMC Bioinform.* 10, 1–9. doi: 10.1186/1471-2105-10-421
- Cantera, J. J. L., and Stein, L. Y. (2007). Molecular diversity of nitrite reductase genes (*nirK*) in nitrifying bacteria. *Environ. Microbiol.* 9, 765–776. doi: 10.1111/j.1462-2920.2006.01198.x
- Castro-Fernandez, V., Zamora, R., Morande, A. H., Vallejos, G., Gonzalez-Ordenes, F., and Guix, V. (2017). “Evolution, metabolism and molecular mechanisms underlying extreme adaptation of Euryarchaeota and its biotechnological potential” in *Archaea – new biocatalysts, novel pharmaceuticals and various biotechnological applications need full citation*. eds. H. Sghaier, A. Najjari and K. Ghedira (Rijeka, Croatia: InTech).
- Chen, R., Deng, M., He, X., and Hou, J. (2017). Enhancing nitrate removal from freshwater pond by regulating carbon/nitrogen ratio. *Front. Microbiol.* 8:1712. doi: 10.3389/fmicb.2017.01712
- Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., et al. (2013). “Carbon and other biogeochemical cycles supplementary material,” in *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. eds. T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, et al. Cambridge, UK: Cambridge University Press.
- Clark, I. M., Buchkina, N., Jhurrea, D., Goulding, K. W. T., and Hirsch, P. R. (2012). Impacts of nitrogen application rates on the activity and diversity of denitrifying bacteria in the Broadbalk wheat. *Experiment. Phil. Trans. R. Soc. B* 367, 1235–1244. doi: 10.1098/rstb.2011.0314
- Dore, J. E., Popp, B. N., Karl, D. M., and Sansone, F. J. (1998). A large source of atmospheric nitrous oxide from subtropical North Pacific surface waters. *Nature* 396, 63–66. doi: 10.1038/23921
- Dundee, L., and Hopkins, D. W. (2001). Different sensitivities to oxygen of nitrous oxide production by *Nitrosomonas europaea* and *Nitrosolobus multififormis*. *Soil Biol. Biochem.* 33, 1563–1565. doi: 10.1016/s0038-0717(01)00059-1
- Felgate, H., Giannopoulos, G., Sullivan, M. J., Gates, A. J., Clarke, T. A., Baggs, E., et al. (2012). The impact of copper, nitrate and carbon status on the emission of nitrous oxide by two species of bacteria with biochemically distinct denitrification pathways. *Environ. Microbiol.* 14, 1788–1800. doi: 10.1111/j.1462-2920.2012.02789.x
- Freng, A., Wallace, D. W. R., and Bange, H. W. (2012). Global oceanic production of nitrous oxide. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 367, 1245–1255. doi: 10.1098/rstb.2011.0360
- Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565
- Glockner, A. B., Jünger, A., and Zumft, W. G. (1993). Copper-containing nitrite reductase from *Pseudomonas aureofaciens* is functional in a mutationally cytochrome cd 1-free background (NirS<sup>-</sup>) of *Pseudomonas stutzeri*. *Arch. Microbiol.* 160, 18–26. doi: 10.1007/bf00258141
- Gluschankoff, N., Santoro, A. E., Buchwald, C., and Casciotti, K. L. (2023). Shifts in the isotopic composition of nitrous oxide between El Niño and La Niña in the eastern tropical South Pacific. *Glob. Biogeochem. Cycles* 37, 1–19. doi: 10.1029/2023gb007959
- Graf, D. R. H., Jones, C. M., and Hallin, S. (2014). Intergenomic comparisons highlight modularity of the denitrification pathway and underpin the importance of community structure for N<sub>2</sub>O emissions. *PLoS One* 9:e114118. doi: 10.1371/journal.pone.0114118
- Hallin, S., Philippot, L., Löffler, F. E., Sanford, R. A., and Jones, C. M. (2018). Genomics and ecology of novel N<sub>2</sub>O-reducing microorganisms. *Trends Microbiol.* 26, 43–55. doi: 10.1016/j.tim.2017.07.003
- Han, B., Liu, S.-Y., Han, L.-L., Song, X.-T., Mo, L.-Y., and Zhang, L.-M. (2022). The contribution of *nirK* gene-containing Thaumarchaea to denitrification and N<sub>2</sub>O production across coastal sediment and terrestrial ecosystems. *J. Soils Sediments* 22, 2246–2261. doi: 10.1007/s11368-022-03219-7
- Hansel, C. M., Fendorf, S., Jardine, P. M., and Francis, C. A. (2008). Changes in bacterial and archaeal community structure and functional diversity along a geochemically variable soil profile. *Appl. Environ. Microbiol.* 74, 1620–1633. doi: 10.1128/AEM.01787-07
- Hatzenpichler, R. (2012). Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. *Appl. Environ. Microbiol.* 78, 7501–7510. doi: 10.1128/AEM.01960-12
- Helen, D., Kim, H., Tytgat, B., and Anne, W. (2016). Highly diverse *nirK* genes comprise two major clades that harbour ammonium-producing denitrifiers. *BMC Genomics* 17:155. doi: 10.1186/s12864-016-2465-0
- Heylen, K., Gevers, D., Vanparys, B., Wittebolle, L., Geets, J., Boon, N., et al. (2006). The incidence of *nirS* and *nirK* and their genetic heterogeneity in cultivated denitrifiers. *Environ. Microbiol.* 8, 2012–2021. doi: 10.1111/j.1462-2920.2006.01081.x
- Hochstein, L. I., and Tomlinson, G. A. (1988). The enzymes associated with denitrification. *Ann. Rev. Microbiol.* 42, 231–261. doi: 10.1146/annurev.mi.42.100188.001311
- Hou, J., Cao, X., Song, C., and Zhou, Y. (2013). Predominance of ammonia-oxidizing archaea and *nirK* gene-bearing denitrifiers among ammonia-oxidizing and denitrifying populations in sediments of a large urban eutrophic Lake (Lake Donghu). *Can. J. Microbiol.* 59, 456–464. doi: 10.1139/cjm-2013-0083
- Jang, J., Sakai, Y., Senoo, K., and Ishii, S. (2019). Potentially mobile denitrification genes identified in *Azospirillum* sp. strain TSH58. *Appl. Environ. Microbiol.* 85, 1–12. doi: 10.1128/aem.02474-18
- Ji, Q., Buitenhuis, E., Suntharalingam, P., Sarmiento, J. L., and Ward, B. B. (2018). Global nitrous oxide production determined by oxygen sensitivity of nitrification and denitrification. *Glob. Biogeochem. Cycles* 32, 1790–1802. doi: 10.1029/2018GB005887
- Ji, Y., Conrad, R., and Xu, H. (2020). Responses of archaeal, bacterial, and functional microbial communities to growth season and nitrogen fertilization in rice fields. *Biol. Fertil. Soils* 56, 81–95. doi: 10.1007/s00374-019-01404-4
- Jones, C. M., and Hallin, S. (2010). Ecological and evolutionary factors underlying global and local assembly of denitrifier communities. *ISME J.* 4, 633–641. doi: 10.1038/ismej.2009.152

or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2025.1635656/full#supplementary-material>

### SUPPLEMENTARY TABLE 1

Archetype sequence information and metadata for bacterial *nirK* OTUs.

### SUPPLEMENTARY TABLE 2

Archetype sequence information and metadata for archaeal *nirK* OTUs.

- Jones, C. M., Stres, B., Rosenquist, M., and Hallin, S. (2008). Phylogenetic analysis of nitrite, nitric oxide, and nitrous oxide respiratory enzymes reveal a complex evolutionary history for denitrification. *Mol. Biol. Evol.* 25, 1955–1966. doi: 10.1093/molbev/msn146
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kop, L. F. M., Koch, H., Jetten, M. S. M., Daims, H., and Lückner, S. (2024). Metabolic and phylogenetic diversity in the phylum Nitrospina revealed by comparative genome analyses. *ISME Comms.* 4:4. doi: 10.1093/ismeco/ycad017
- Lee, J. A., and Francis, C. A. (2017). Spatiotemporal characterization of San Francisco Bay denitrifying communities: a comparison of *nirK* and *nirS* diversity and abundance. *Microb. Ecol.* 73, 271–284. doi: 10.1007/s00248-016-0865-y
- Letunic, I., and Bork, P. (2006). Interactive tree of life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23, 127–128. doi: 10.1093/bioinformatics/btl529
- Li, F., Li, M., Shi, W., Li, H., Sun, Z., and Gao, Z. (2017). Distinct distribution patterns of proteobacterial *nirK*- and *nirS*-type denitrifiers in the Yellow River estuary, China. *Can. J. Microbiol.* 63, 708–718. doi: 10.1139/cjm-2017-0053
- Liu, X., Tiquia, S. M., Holguin, G., Wu, L., Nold, S. C., Devol, A. H., et al. (2003). Molecular diversity of denitrifying genes in continental margin sediments within the oxygen-deficient zone off the Pacific coast of Mexico. *Appl. Environ. Microbiol.* 69, 3549–3560. doi: 10.1128/aem.69.6.3549-3560.2003
- Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158
- Louca, S., Mazel, F., Doebeli, M., and Parfrey, L. W. (2019). A census-based estimate of earth's bacterial and archaeal diversity. *PLoS Biol.* 17:17. doi: 10.1371/journal.pbio.3000106
- Lund, M. B., Smith, J. M., and Francis, C. A. (2012). Diversity, abundance and expression of nitrite reductase (*nirK*)-like genes in marine thaumarchaea. *ISME J.* 6, 1966–1977. doi: 10.1038/ismej.2012.40
- Martinez-Espinoza, R. M. (2020). Microorganisms and their metabolic capabilities in the context of the biogeochemical nitrogen cycle at extreme environments. *Int. J. Mol. Sci.* 21:4228. doi: 10.3390/ijms21124228
- Ming, Y., Abdullah Al, M., Zhang, D., Zhu, W., Liu, H., Cai, L., et al. (2024). Insights into the evolutionary and ecological adaptation strategies of *nirS*- and *nirK*-type denitrifying communities. *Mol. Ecol.* 33:e17507. doi: 10.1111/mec.17507
- Muck, S., De Corte, D., Clifford, E. L., Bayer, B., Herndl, G. J., and Sintés, E. (2019). Niche differentiation of aerobic and anaerobic ammonia oxidizers in a high latitude deep oxygen minimum zone. *Front. Microbiol.* 10:2141. doi: 10.3389/fmicb.2019.02141
- Oakley, B. B., Francis, C. A., Roberts, K. J., Fuchsman, C. A., Srinivasan, S., and Staley, J. T. (2006). Analysis of nitrite reductase (*nirK* and *nirS*) genes and cultivation reveal depauperate community of denitrifying bacteria in the Black Sea suboxic zone. *Environ. Microbiol.* 9, 118–130. doi: 10.1111/j.1462-2920.2006.01121.x
- Peng, L., Ni, B. J., Ye, L., and Yuan, Z. (2015). The combined effect of dissolved oxygen and nitrite on N<sub>2</sub>O production by ammonia oxidizing bacteria in an enriched nitrifying sludge. *Water Res.* 73, 29–36. doi: 10.1016/j.watres.2015.01.021
- Pold, G., Bonilla-Rosso, G., Saghai, A., Strous, M., Jones, C. M., and Hallin, S. (2024). Phylogenetics and environmental distribution of nitric oxide forming nitrite reductases reveals their distinct functional and ecological roles. *ISME Commun.* 4, 1–15. doi: 10.1093/ismeco/ycac020
- Prosser, J. I. (Ed.) (1986). *Nitrification*. Washington, DC, USA: IRL Press.
- Purkhold, U., Pommerening-Roser, A., Juretschko, S., Schmid, M. C., Koops, H. P., and Wagner, M. (2000). Phylogeny of all recognized species of ammonia oxidizers based on comparative 16S rRNA and *amoA* sequence analysis: implications for molecular diversity surveys. *Appl. Environ. Microbiol.* 66, 5368–5382. doi: 10.1128/AEM.66.12.5368-5382.2000
- Ravishankara, A. R., Daniel, J. S., and Portmann, R. W. (2009). Nitrous oxide (N<sub>2</sub>O): the dominant ozone-depleting substance emitted in the 21st century. *Science* 326, 123–125. doi: 10.1126/science.1176985
- Reij, L., Tolar, B. B., Smith, J. M., Chavez, F. P., and Francis, C. A. (2019). Depth distributions of nitrite reductase (*nirK*) gene variants reveal spatial dynamics of thaumarchaeal ecotype populations in coastal Monterey Bay. *Environ. Microbiol.* 21, 4032–4045. doi: 10.1111/1462-2920.14753
- Saarenheimo, J., Aalto, S. L., Rissanen, A. J., and Tirola, M. (2017). Microbial community response on wastewater discharge in boreal Lake sediments. *Front. Microbiol.* 8:750. doi: 10.3389/fmicb.2017.00750
- Santoro, A. E., Buchwald, C., McIlvin, M. R., and Casciotti, K. L. (2011). Isotopic signature of N<sub>2</sub>O produced by marine ammonia-oxidizing archaea. *Science* 333, 1282–1285. doi: 10.1126/science.1208239
- Shaw, L. J., Nicol, G. W., Smith, Z., Fear, J., Prosser, J. I., and Baggs, E. M. (2005). *Nitrosospira* spp. can produce nitrous oxide via a nitrifier denitrification pathway. *Environ. Microbiol.* 8, 214–222. doi: 10.1111/j.1462-2920.2005.00882.x
- Sherrill-Mix, S. (2018). Taxonomizr: Functions to work with NCBI accessions and taxonomy. Available online at: <https://cran.r-project.org/package=taxonomizr> (Accessed June 4, 2020).
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Stieglmeier, M., Mooshammer, M., Kitzler, B., Wanek, W., Zechmeister-Boltenstern, S., Richter, A., et al. (2014). Aerobic nitrous oxide production through N-nitrosating hybrid formation in ammonia-oxidizing archaea. *ISME J.* 8, 1135–1146. doi: 10.1038/ismej.2013.220
- Sun, H., and Jiang, S. (2022). A review on *nirS*-type and *nirK*-type denitrifiers via a scientometric approach coupled with case studies. *Environ. Sci.: Processes Impacts* 24, 221–232. doi: 10.1039/d1em00518a
- Tang, W., Ward, B. B., Beman, M., Bristow, L., Clark, D., Fawcett, S., et al. (2023). Database of nitrification and nitrifiers in the global ocean. *Earth Syst. Sci. Data* 15, 5039–5077. doi: 10.5194/essd-15-5039-2023
- Taroncher-Oldenburg, G., Griner, E. M., Francis, C. A., and Ward, B. B. (2003). Oligonucleotide microarray for the study of functional gene diversity in the nitrogen cycle in the environment. *Appl. Environ. Microbiol.* 69, 1159–1171. doi: 10.1128/AEM.69.2.1159-1171.2003
- Thompson, C. C., Vidal, L., Salazar, V., and Swings, J., and Thompson, F. L. (2021). Microbial genomic taxonomy. *Trends in the systematics of bacteria and fungi*. 168–178. doi: 10.1079/9781789244984.0168
- Torregrosa-Crespo, J., Bergaust, L., Pire, C., and Martínez-Espinoza, R. M. (2018). Denitrifying haloarchaea: sources and sinks of nitrogenous gases. *FEMS Microbiol. Lett.* 365, 1–6. doi: 10.1093/femsle/fnx270
- Treusch, A. H., Leininger, S., Kletzin, A., Schuster, S. C., Klenk, H., and Schleper, C. (2005). Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. *Environ. Microbiol.* 7, 1985–1995. doi: 10.1111/j.1462-2920.2005.00906.x
- Trimmer, M., Chronopoulou, P.-M., Maanoja, S. T., Upstill-Goddard, R. C., Kitidis, V., and Purdy, K. J. (2016). Nitrous oxide as a function of oxygen and archaeal gene abundance in the North Pacific. *Nat. Commun.* 7, 1–10. doi: 10.1038/ncomms13451
- Vajjala, N., Martens-Habbena, W., Sayavedra-Soto, L. A., Schauer, A., Bottomley, P. J., Stahl, D. A., et al. (2012). Hydroxylamine as an intermediate in ammonia oxidation by globally abundant marine archaea. *Proc. Natl. Acad. Sci. USA* 110, 1006–1011. doi: 10.1073/pnas.1214272110
- Wang, L., Zheng, B., Nan, B., and Hu, P. (2014). Diversity of bacterial community and detection of *nirS* and *nirK*-encoding denitrifying bacteria in sandy intertidal sediments along Laizhou Bay of Bohai Sea, China. *Mar. Pollut. Bull.* 88, 215–223. doi: 10.1016/j.marpolbul.2014.09.002
- Wan, X. S., Hou, L., Kao, S. J., Zhang, Y., Sheng, H. X., Shen, H., et al. (2023). Pathways of N<sub>2</sub>O production by marine ammonia-oxidizing archaea determined from dual-isotope labeling. *Proc. Natl. Acad. Sci.* 120:120. doi: 10.1073/pnas.2220697120
- Ward, B. B., Eveillard, D., Kirshtein, J. D., Nelson, J. D., Voytek, M. A., and Jackson, G. A. (2007). Ammonia-oxidizing bacterial community composition in estuarine and oceanic environments assessed using a functional gene microarray. *Environ. Microbiol.* 9, 2522–2538. doi: 10.1111/j.1462-2920.2007.01371.x
- Wu, J., He, C., van Loosdrecht, M. C., and Pérez, J. (2016). Selection of ammonium oxidizing bacteria (AOB) over nitrite oxidizing bacteria (NOB) based on conversion rates. *Chem. Eng. J.* 304, 953–961. doi: 10.1016/j.cej.2016.07.019
- Yin, Z., Bi, X., and Xu, C. (2018). Ammonia-oxidizing archaea (AOA) play with ammonia-oxidizing bacteria (AOB) in nitrogen removal from wastewater. *Archaea*, 8429145. doi: 10.1155/2018/8429145
- Yoshida, M., Ishii, S., Otsuka, S., and Senoo, K. (2009). Temporal shifts in diversity and quantity of *nirS* and *nirK* in a rice paddy field soil. *Soil Biol. Biochem.* 41, 2044–2051. doi: 10.1016/j.soilbio.2009.07.012
- Zheng, Y., Wang, B., Gao, P., Yang, Y., Xu, B., Su, X., et al. (2024). Novel order-level lineage of ammonia-oxidizing archaea widespread in marine and terrestrial environments. *ISME J.* 18, 1–13. doi: 10.1093/ismej/wrad002
- Zumft, W. G. (1997). Cell biology and molecular basis of denitrification. *Microbiol. Mol. Biol. Rev.* 61, 533–616. doi: 10.1128/mmb.61.4.533-616.1997