Main environmental drivers of abundance, diversity, and community structure of comammox *Nitrospira* in paddy soils

Fuyun GAO[1,2,3], Yaying LI[1,2], Haoxin FAN[4], Jiantao XUE[4] and Huaiying YAO[1,2,4,*]

[1]*Key Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021 (China)*
[2]*Zhejiang Key Laboratory of Urban Environmental Processes and Pollution Control, CAS Haixi Industrial Technology Innovation Center in Beilun, Ningbo 315800 (China)*
[3]*University of Chinese Academy of Sciences, Beijing 100049 (China)*
[4]*Research Center for Environmental Ecology and Engineering, School of Environmental Ecology and Biological Engineering, Wuhan Institute of Technology, Wuhan 430073 (China)*

ABSTRACT

The recently discovered complete ammonia oxidizer (comammox *Nitrospira*) contains clades A and B that were established an independent one-step nitrification process; however, little is known about their environmental drivers or habitat distributions in agricultural soils. Previous studies on comammox *Nitrospira* in paddy soils have mainly focused on small-scale samples, and there is a lack of multi-site research on comammox *Nitrospira* in paddy soils. In this study, we conducted a survey of 36 paddy soil sites, aiming to understand the comammox *Nitrospira* community structure, abundance, and diversity and the degree to which they are affected by environmental factors at a large scale. Comammox *Nitrospira* was found to be widely distributed among the paddy soils. The abundance of comammox *Nitrospira* clade A was mostly lower than that of clade B, whereas its diversity was mostly higher than that of clade B. Correlation analysis showed that multiple factors affected the abundance of comammox *Nitrospira* including pH, soil organic matter, total carbon, total nitrogen, latitude, mean annual temperature and mean annual precipitation ($P < 0.05$). Moreover, there was a clear relationship between the comammox *Nitrospira* community and habitat, indicating that some amplicon sequence variants (ASVs) had a unique dominant position in specific habitat. Phylogenetic analysis showed that the ASVs of comammox *Nitrospira* clade A clustered with the known sequences in the paddy soils and were significantly different from the known sequences in other habitats; this may be related to the unique habitat of the paddy field. In contrast, comammox *Nitrospira* clade B showed no clear habitat dependence. These results support the wide distribution and high abundances of comammox *Nitrospira* in paddy soils, and provide novel insight into nitrogen cycling and nutrient management in agricultural ecosystems.

Key Words: abundance, comammox *Nitrospira*, community, diversity, paddy soil

INTRODUCTION

Paddy soils distribute widely in China, and paddy ecosystems are an important component of global agriculture (Hu *et al.*, 2015), with more than 158 million ha in use for arable land (Timilsina *et al.*, 2020). The drying-wetting stage of paddy soils provides a unique environment for microorganisms (Ishii *et al.*, 2011). Ammonia oxidation plays a vital role in the global nitrogen cycle. As the first rate-limiting step, it is executed by ammonia-oxidizing archaea (AOA) and ammonia-oxidizing bacteria (AOB). However, recent evidence has confirmed that complete ammonia oxidizers (comammox *Nitrospira*) can independently perform complete ammonia oxidation (Daims *et al.*, 2015; van Kessel *et al.*, 2015). According to the difference of ammonia monooxygenase, it was separated into two clades: clade A and clade B (Daims *et al.*, 2015; van Kessel *et al.*, 2015; Palomo *et al.*, 2018).

The abundances of comammox *Nitrospira* clade A and clade B responded differently to the various environmental factors that have been studied, the response of the community of ammonia oxidizers to environmental factors and habitat is still unknown (Liu *et al.*, 2018; Liu *et al.*, 2020; Shi *et al.*, 2020).

Previous studies indicated that comammox *Nitrospira* displayed high yield, high substrate affinity, relatively low maximum oxidation rate and slow growth under highly oligotrophic conditions (van Kessel *et al.*, 2015; Kits *et al.*, 2017). These results support early hypotheses that the slow growth in ammonia-depleted biofilms, soils, or sediments represents the fundamental niche of comammox *Nitrospira* (Costa *et al.*, 2006; Kits *et al.*, 2017). Although previous studies have indicated that the abundance of comammox *Nitrospira* is lower than the abundance of AOA and/or AOB in agricultural and forest soils (Li *et al.*, 2020; Shi *et al.*, 2020; Wang *et al.*, 2020) or approximate to the abundance of AOB in certain forest soils (Shi *et al.*, 2020). On the contrary, comammox *Nitrospira* clade A have been shown to outnumber clade B in alkaline paddy soils (Wang *et al.*, 2018) and AOA in acidic forest soils (Hu and He, 2017). Comammox *Nitrospira* preferred a weakly alkaline environment, because the optimal pH of key N-converting enzymes (for example, hydroxylamine dehydrogenase and ammonia monooxygenase) was approximately 7.0–8.0 (Blum *et al.*, 2018). Studies have examined the influence of various environmental factors on each ammonia oxidizer community, including soil temperature (Liu *et al.*, 2020), soil salinity (Liu *et al.*, 2020), fertilization (Wang *et al.*, 2018) and geographic factors (Liu *et al.*, 2020; Shi *et al.*, 2020; Zhang *et al.*, 2020). However, this is a relatively new field of study, and further research into the relationship between the comammox *Nitrospira* community and various environmental factors is still needed.

Paddy ecosystems are variable in time and space and the drying-wetting stage of paddy soils provides a unique environment for microorganisms (Ishii *et al.,* 2011). The Northeast Plain, the Yangtze River Basin and the southeast coast are the three most important rice-producing regions of China. The paddy fields therefore cover a wide range of climatic zones (Hou *et al.*, 2020). The particular dry-wet alternation in paddy soils causes regular fluctuations in its redox potential (Waqas *et al.*, 2019), with frequent aerobic-anoxic alternation. Although comammox *Nitrospira* has been studied in engineering systems, the driving factors for nitrification and the large-scale distribution of comammox *Nitrospira* in paddy soils remain poorly understood. Understanding of how the various environmental factors in soil affect the abundance, diversity and community of comammox *Nitrospira* is still essential. To understand the importance of these factors, we collected soil from 36 paddy fields across China to determine the soil characteristics and its molecular properties, study the main environmental drivers affecting the abundance, diversity and community of comammox *Nitrospira* clade A and clade B. Here, we applied quantitative real-time polymerase chain reaction (Q-PCR) and high-throughput sequencing combined with QIIME2 (Bolyen *et al.*, 2019) to comprehensively analyze comammox *Nitrospira* clade A and clade B communities. Through this study, we mainly hope to understand the following: (i) the distribution, abundance, diversity and community characteristics of comammox *Nitrospira* in Chinese paddy soils; (ii) the environmental driving factors of comammox *Nitrospira*; and (iii) the phylogenetic difference in the comammox *Nitrospira* community between paddy soil and other known habitats such as engineering systems, sediments, and forest soils. The results of this study could be applied to enhance the management strategy of nitrogen cycle in rice paddy soils as a result of improving understanding the main drivers on the environmental functions of comammox *Nitrospira* clade A and clade B.

MATERIALS AND METHODS

Sample collection and soil properties

In accordance with data from the National Bureau of Statistics of China (http://www.stats.gov.cn/), the sampling points set up in this study fully considered the dominant rice planting areas, planting years (> 15 years) and planting scale to ensure the typicality of the paddy soils. Thirty-six sites were selected in 12 provinces, and 108 soil samples were collected. The samples were taken from northeast, east, central, south and southwest China from July to September 2018. Soil samples were taken when the rice was harvested fallow, including three replicates for each site. Approximately 2 kg soils consisting of around eight soil cores (5 cm in diameter) were randomly collected from the upper 0–20 cm of soils across the whole field of each site, mixed well into three composite duplicate samples, placed into sampling bags, and transported on ice to the laboratory. Each sample was split into three parts. One part was air dried for chemical analysis, one part was stored at 4 °C for biochemical analysis and the third part was freeze-dried for DNA extraction. The information concerning the latitude and longitude, climate type, annual mean temperature, and annual mean precipitation of these 36 sites is shown in Supplementary Material (Tables SI and SII, see Supplementary Material for Tables SI and SII) of the supporting information.

The pH value of these paddy soils was measured in a soil-water volume ratio of 1:2.5 using a pH compound electrical level (PE-10, Germany). The total carbon (TC) and total nitrogen (TN) content were

determined by combustion decomposition using an element analyzer (Elementar Analysen Systeme GmbH, Germany). Microbial biomass carbon (MBC) and microbial biomass nitrogen (MBN) were calculated using the difference between samples fumigated with chloroform ($CHCl_3$) and those that were unfumigated, and then extracted with 0.5 mol L^{-1} potassium sulfate (K_2SO_4), as described by Meng *et al*. (2020). The soil ammonium (NH_4^+) and nitrate (NO_3^-) were extracted by shaking with potassium chloride (KCl) solution for 1 h. The volume ratio of soil to KCl was 1:10 and the KCl concentration in the final solution was 0.02 mol L^{-1}. Soil organic matter (SOM) content was measured by potassium permanganate ($K_2Cr_2O_7$) oxidation-volumetric, as described by Meng *et al*. (2021).

DNA extraction and Q-PCR

The FastDNA SPIN kit for soil (MP Biomedicals, Cleveland, OH, USA) was used to extract genomic DNA of each sample (~500 mg of soil) according to the manufacturer's manual. Before conducting the downstream analysis, a Nanodrop ND-2000 UV-vis spectrophotometer (NanoDrop Technologies, Wilmington, DE) was used to detect the concentration and quality of DNA.

We used Q-PCR to quantify functional genes (*amoA*) involved in the comammox *Nitrospira* process. The primers were equimolar mixtures of six pairs of primers of comaA-244F/comaA-659R and comaB-244F/comaB-659R, respectively (Pjevac *et al*., 2017). Q-PCR was executed to validate the functional gene abundance of *amoA* genes with SYBR Premix Ex TaqII (Promega, Madison, WI, USA) using a Light-Cycler Roche 480 instrument (Roche Molecular Systems). Each reaction was executed in a 20 μL volume containing about 5 ng of DNA, 0.25 μmol L^{-1} of each primer and 10 μL of SYBR Premix EX TaqII. The details of primers and thermal cycling conditions are shown in Table SIII (see Supplementary Material for Table SIII). Agarose gel electrophoresis was run with 2% concentration agarose gel for the target band, and the target band was cut down. After purification, the target band was loaded into PGM-T Easy Vector (Tian Gen, Beijing, China) to build the clone library. The positive clones were selected for sequencing and extracting the plasmids of *amoA*. The real-time PCR standard curves of all experiments in this study were operated by 10× gradient dilution of *amoA* positive plasmid DNA extracted from cloning. The negative control used the same amount of water as DNA as the template. The efficiency of all qPCR experiments was better than 88% and the error was less than 0.200.

Functional gene (amoA) sequencing

The comammox *Nitrospira* communities in paddy soils were analyzed with Illumina sequencing of comammox *Nitrospira* clade A and clade B functional genes. High-throughput paired end Illumina MiSeq sequencing (PE 300) was performed at Majorbio, Shanghai, China. After sequencing, Illumina MiSeq fastq reads of the comammox *Nitrospira* clade A and clade B *amoA* genes were analyzed using the QIIME2 platform (https://qiime2.org/) and were treated by the DADA2 plugin. The 8 bp barcode was collated to distribute to each sample and all reads < 375 bp were cut off and discarded. The trimmed sequences were chimera-detected and deleted using the vsearch algorithm, which can supposedly represent the real biological sequences present in the sample, known as amplicon sequence variants (ASVs). The depth was 1312 for comammox *Nitrospira* clade A and 1291 for comammox *Nitrospira* clade B. Clustering of high-quality sequences into ASVs was performed with vsearch pipeline at almost 100% similarity level. The neighbor-joining method was used to build a phylogenetic tree based on top1 (> 1% genus) representative sequences.

All the original sequences of comammox *Nitrospira amoA* gene have been submitted to GenBank. The accession numbers are SRR13149211-SRR13149411.

Statistical analysis

SPSS 22 for Windows was used to calculate the Spearman correlation coefficient, which was used to explain the significant correlation between pH, mean annual temperature (MAT), mean annual precipitation (MAP), geochemical properties and the abundances of comammox *Nitrospira* clade A and clade B. Redundancy analysis (RDA) was conducted to distinguish environmental factors driving the differences between comammox *Nitrospira* clade A and clade B. Before performing RDA, we used species-sample data to perform a detrended correspondence analysis to obtain the lengths of gradient in the analysis results. If this was less than 3, we initiated the RDA analysis. The representative sequences obtained in this study were compared with the reference sequences defined in NCBI database in BioEdit, and then a phylogenetic tree

was built by Molecular Evolutionary Genetic Analysis software (MEGA X), with a bootstrap of 1000 times (Kumar *et al.*, 2018). The phylogenetic tree was built by first translating nucleotide sequences into amino acid sequences. The sequence data were dealt with and analyzed with QIIME2. The vsearch pipeline was used to cluster high-quality sequences to amplicon variants at 100% similarity level. Aggregated boosted tree (ABT) models first converted the ASVs table into a beta diversity distance, and then converted all environmental factors into a matrix of differences between samples. The matrices were converted into one column of data, and cross-validation was performed. All other statistical tests were conducted in R software (v.4.0.2), involving the ggplot2, vegan, dismo, gbm and reshape2 packages.

RESULTS

Distribution and abundance of comammox Nitrospira clade A and clade B

Abundances and distribution of comammox *Nitrospira* clade A and clade B were compared among the sites by determining the relative abundance of their *amoA* gene (Fig. 1A). Comammox *Nitrospira* clade A and clade B *amoA* gene abundances ranged from 8.27×10^5 to 3.14×10^8 copies g^{-1} dry soil and from 4.94×10^6 to 9.67×10^8 copies g^{-1} dry soil, respectively. As a result, the highest copy number ratio of *amoA* gene between clade B and clade A was 68.96 in the ZJSX paddy field, and the lowest was 0.62 in the FJNP paddy field (Fig. 1A).

The abundances of comammox *Nitrospira* clade A and clade B genes were consistently strongly affected by pH. Regression analysis indicated that the *amoA* gene abundances of comammox *Nitrospira* clade A ($R^2 = 0.45$, $P < 0.001$) and clade B ($R^2 = 0.37$, $P < 0.001$) were positively related to soil pH (Fig. 1C, D). Principal component analysis (PCA) showed that 67.21% of total variation was explained by PC1, PC2 explained 21.53% of total variation, and the two axes together explained 88.74% of the variation (Fig. 1B).
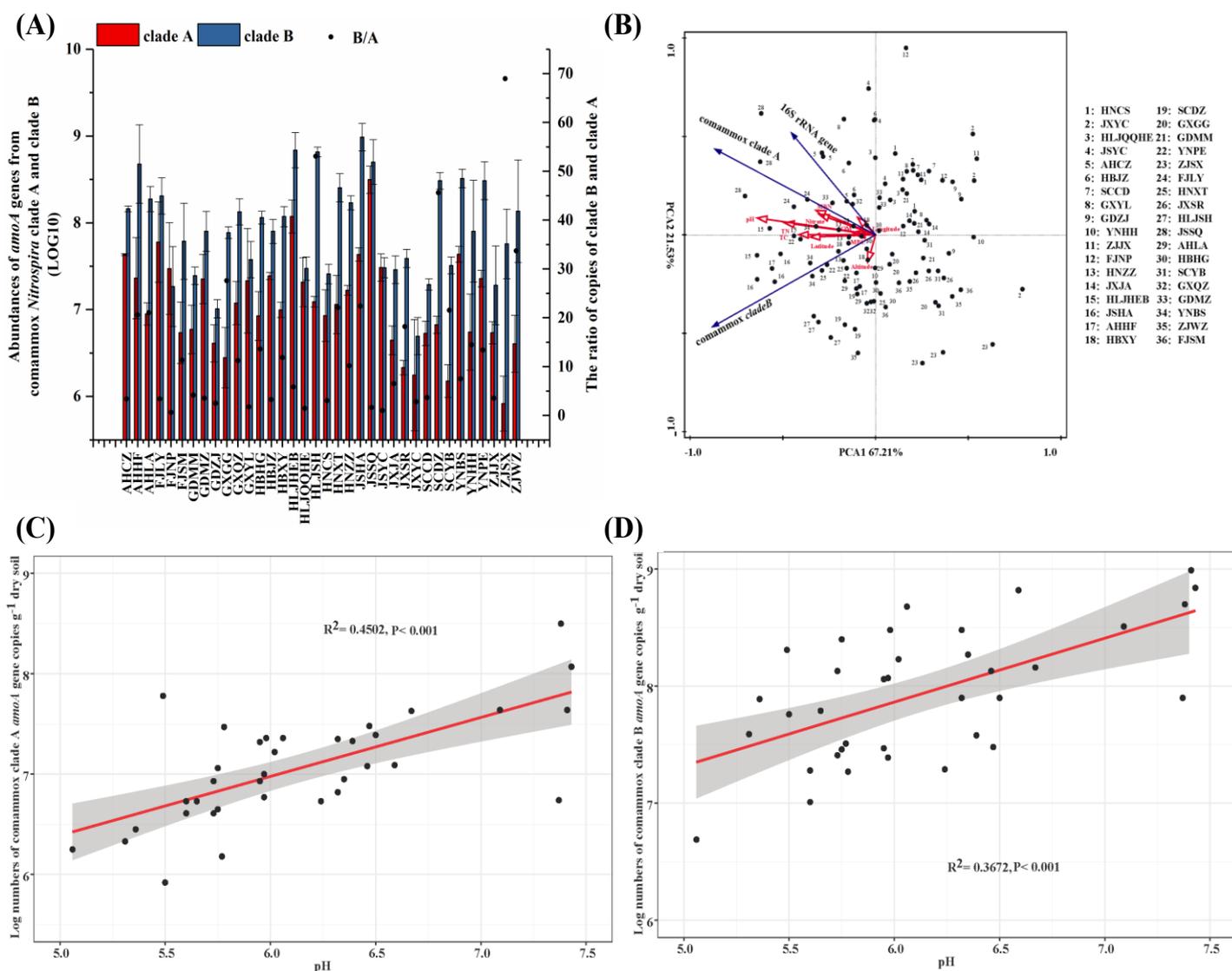
Fig. 1 Abundances of *amoA* genes from comammox *Nitrospira* clade A and B (A), principal component analysis (PCA) among comammox *Nitrospira* abundance and environmental variables (B), regression analysis between comammox *Nitrospira* clade A and clade B *amoA* gene copies (log numbers) and pH (C, D).

Moreover, according to Spearman's correlation analysis, the abundances of comammox *Nitrospira* clade A were significantly positively related to pH, SOM content, TC, TN and latitude ($P < 0.01$). The abundance of comammox *Nitrospira* clade A was positively related to MBN and longitude ($P < 0.05$) (Fig. S1, see Supplementary Material for Fig. S1). The abundance of comammox *Nitrospira* clade B was conspicuously positively related to pH, TC, TN and latitude ($P < 0.01$). Likewise, MBN and SOM content were positively related to comammox *Nitrospira* abundance ($P < 0.05$). Comammox *Nitrospira* clade A and clade B were significantly negatively related to MAT and MAP (Fig. S1). In addition, there was no significant correlation between comammox *Nitrospira* and other environmental factors (Fig. S1).

Alpha diversity of comammox Nitrospira clade A and clade B

Boxplots were used to exhibit the three diversity indexes of comammox *Nitrospira* clade A and clade B based on ASVs (Fig. 2). In this study, the mean Shannon index of comammox *Nitrospira* clade A and clade B was 3.46 and 1.49, respectively, suggesting that clade A has higher diversity than clade B in paddy soils. There was no significant difference among the Pielou's evenness values between the two clades. Furthermore, significant differences in Faith's phylogenetic diversity of comammox *Nitrospira* clade A and clade B communities were observed (Fig. 2). Spearman correlation analysis between the diversity of comammox *Nitrospira* clade A, clade B and environmental factors showed that comammox *Nitrospira* clade B diversity

was positively related to TN ($P < 0.05$) (Fig. S1).
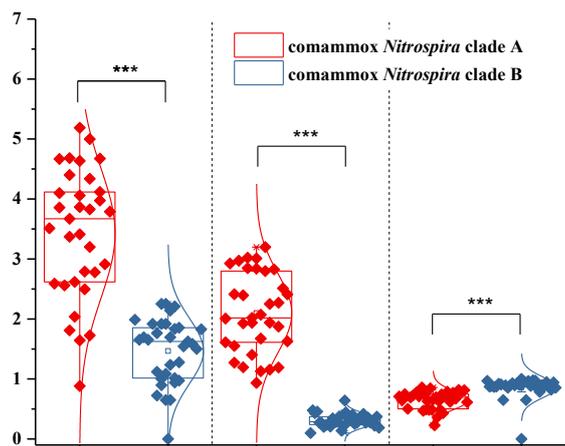
Fig. 2　Box plots shows the statistical results of α-diversity in paddy soils ($n = 36$). *** indicates p value is less than 0.001.

Community structure of comammox Nitrospira clade A and B

A total of 883521 and 843642 high-quality *amoA* sequences were collected from comammox *Nitrospira* clade A and clade B in all 36 samples, respectively. Among these sequences, 8430 (HLJHEB) to 84719 (JXYC) clade A *amoA* sequences (average 24542) were obtained for each sample. The read length of the sequence was between 375 to 415 bp, and 98% of the sequence length was 415 bp. Each sample obtained 398 (GDJZ) to 40630 (FJSM) clade B *amoA* sequences (23434 on average). The sequence read length was between 375 and 439 bp and 98% of the sequence length was 415 bp (Table SIV, see Supplementary Material for Table SIV).

RDA indicated that the community structure of comammox *Nitrospira* clade A and clade B were different among ASVs, and the influencing factors were different (Fig. 3). For comammox *Nitrospira* clade A, 9.73% of total variation was explained by RDA 1 and RDA 2 explained 7.46% of total variation, respectively (Fig. 3A). The community structure of clade A were related to longitude and SOM ($P < 0.05$). For comammox *Nitrospira* clade B, axis 1 explained 11.65% of the variation and axis 2 explained 6.55% of the variation, respectively (Fig. 3B). In comparison, SOM and MBN ($P < 0.05$) were related to the community structure of clade B (Fig. 3B). The differences in comammox *Nitrospira* clade A and clade B communities of paddy soils were displayed using the nonmetric multidimensional scaling analysis (NMDS) on the Bray-Curtis distance. This demonstrated no significant difference between comammox *Nitrospira* clade A and clade B communities (Fig. S2, see Supplementary Material for Fig. S2).
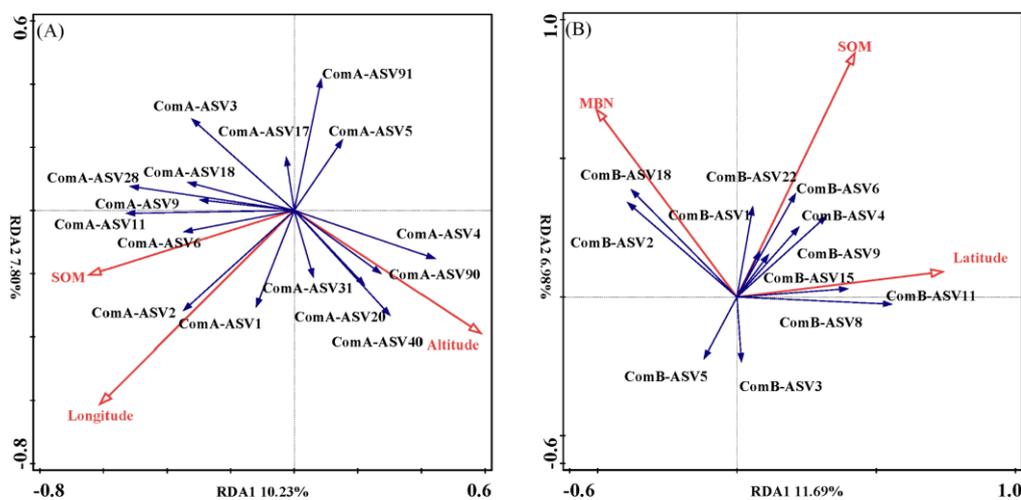
Fig. 3　Redundancy analysis (RDA) shows the relationship between the comammox *Nitrospira* clade A and clade B community composition and different environmental factors at sampling sites.

To explore the habitat relevance of these comammox *Nitrospira* clades, we searched the closely related *amoA* gene in the NCBI database. All these *amoA* genes belonged to *Nitrospira* II. Based on the existing sequences to construct phylogenetic trees, we found that they were not only closely related to paddy soil and paddy rhizosphere soil, but also closely related to the sequence of forests and engineering systems (Fig. 4A). In fact, our results indicated that the ASVs of comammox *Nitrospira* clade A were clustered with known sequences in paddy soils, which were significantly different from known sequences in other habitats (Fig. 4) and it may be caused by the unique habitats of the paddy fields. In addition, comA-ASV9 has a close phylogenetic distance with *Candidatus Nitrospira* nitrificans (Fig. 4B). Unlike comammox *Nitrospira* clade A, the ASVs of comammox *Nitrospira* clade B were not clustered with sequences of single habitat, and they were clustered with the sequences of paddy soils, forest soils and engineering systems (Fig. 4). comB-ASV6, comB-ASV11, comB-ASV26 and comB-ASV39 were clustered with the sequences of paddy soils. comB-ASV16 and comB-ASV59 were also related to acidic forest soil and neutral paddy soil in China (Fig. 4B). In addition, some ASVs were related to biofilters in groundwater wells and circulating aquarium systems (Fig. 4B). Surprisingly, comA-ASV96 and comA-ASV5 had unique advantages in paddy soils of HLJHEB and SCYB, respectively (Fig. 5A). Similarly, comB-ASV1, comB-ASV7, comB-ASV13 and comB-ASV39 were dominant in the comammox *Nitrospira* clade B communities of HLJSH, HLJHEB, SCYB and GDZJ, respectively (Fig. 5B).
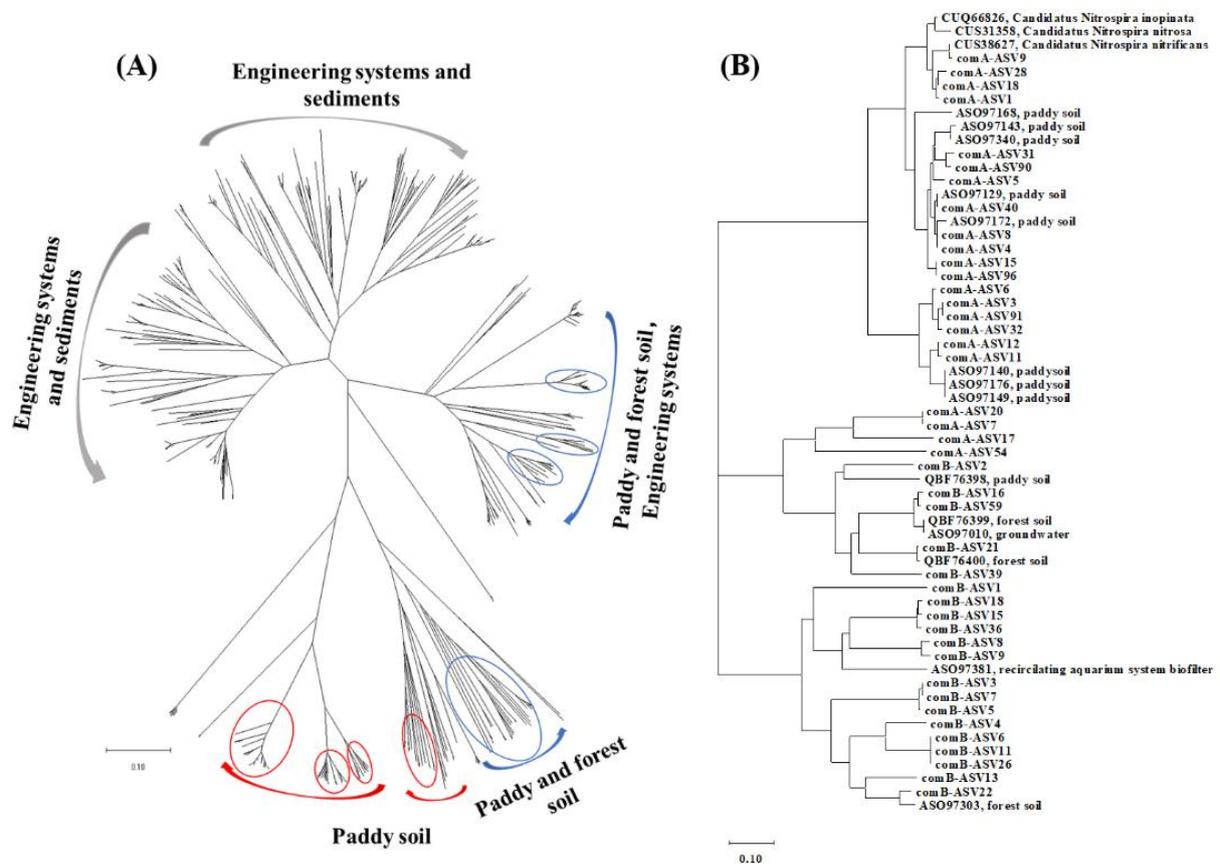
Fig. 4 (A) Neighbor-joining tree for dominant comammox *Nitrospira amoA* ASVs. The sequence in the circles indicates the clustering sequence in this study. Red represents comammox *Nitrospira* clade A and blue represents comammox *Nitrospira* clade B. (B) Neighbor-joining tree for dominant comammox *Nitrospira* clade A and clade B *amoA* ASVs (representatives with a relative abundance was top1). comammox *Nitrospira amoA* ASV representative sequences from this study were shown in comA-ASVXXX and comB-ASVXXX, respectively. The bootstrap values were 1000.
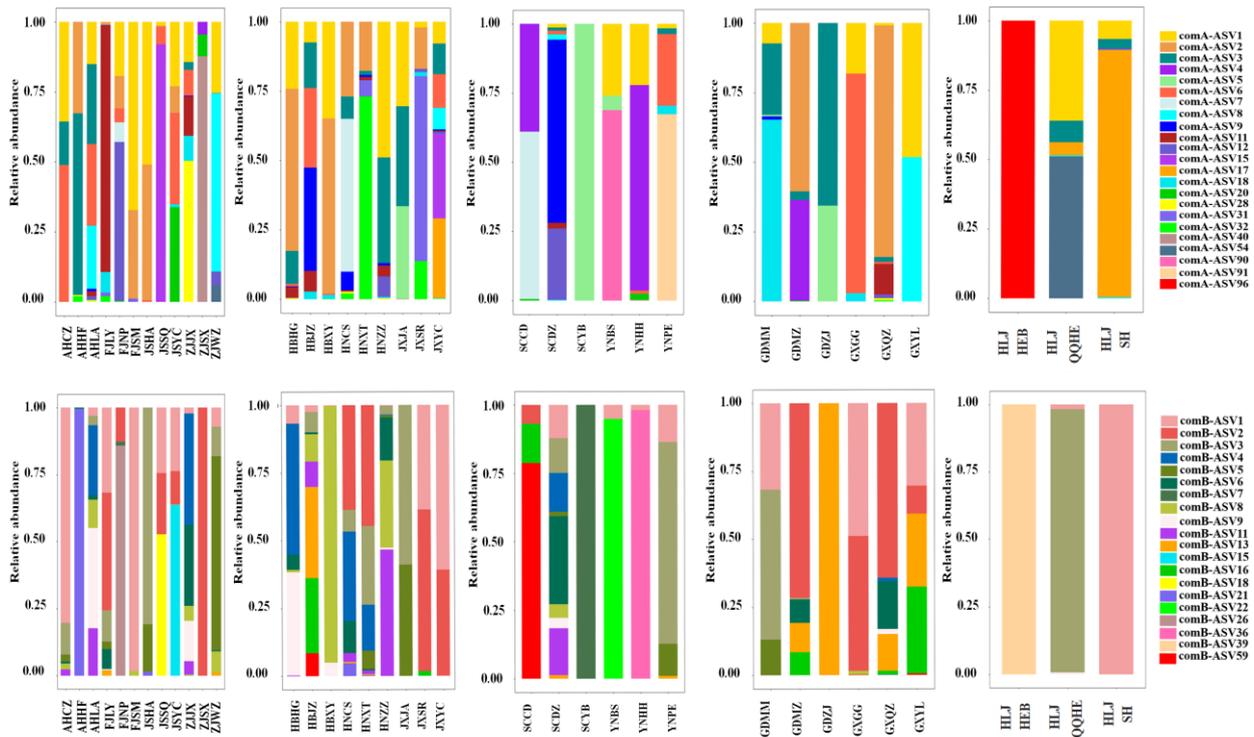
Fig. 5   (A) Relative abundance of phylogenetic clade A *amoA* ASVs representative sequences among different samples. (B) Relative abundance of phylogenetic clade B *amoA* ASVs representative sequences among different samples.

Relative influence of environmental factors on the abundance, diversity and community structure of comammox *Nitrospira*

The soil pH was the primary factor impacting abundance of clade A, giving rise to approximately 57.41% of the relative influence (Fig. 6A). Soil pH, latitude and MAT were the main factors affecting clade B abundance (76.89%, Fig. 6B). Soil nitrate, MBC and MBN content mainly affected clade A diversity (68.81%, Fig. 6C). Soil TC, nitrate content and altitude mainly affected clade B diversity (88.37%, Fig. 6D). The soil ammonium content, SOM, nitrate and TC were the major factors impacting clade A community structure, accounting for approximately 46.91% of the relative influence (Fig. 6E). Longitude, pH and MAT explained 8.19%–9.20% of the community structure, and thus were not considered main factors influencing the communities of comammox *Nitrospira* clade A (Fig. 6E). The ABT model also showed that longitude, latitude, TC, and SOM were the most important environmental factors influencing comammox *Nitrospira* clade B community structure, accounting for approximately 46.16% of the relative influence (Fig. 6F). Likewise, altitude, pH and nitrate only explained 8.34%–8.54% of the community structure, and thus were not considered major factors influencing the communities of comammox *Nitrospira* clade B (Fig. 6F). A clustered bar chart was used to display the result of Spearman correlation analysis, which was consistent with the significant correlations observed in the ABT analysis (Fig. S1).
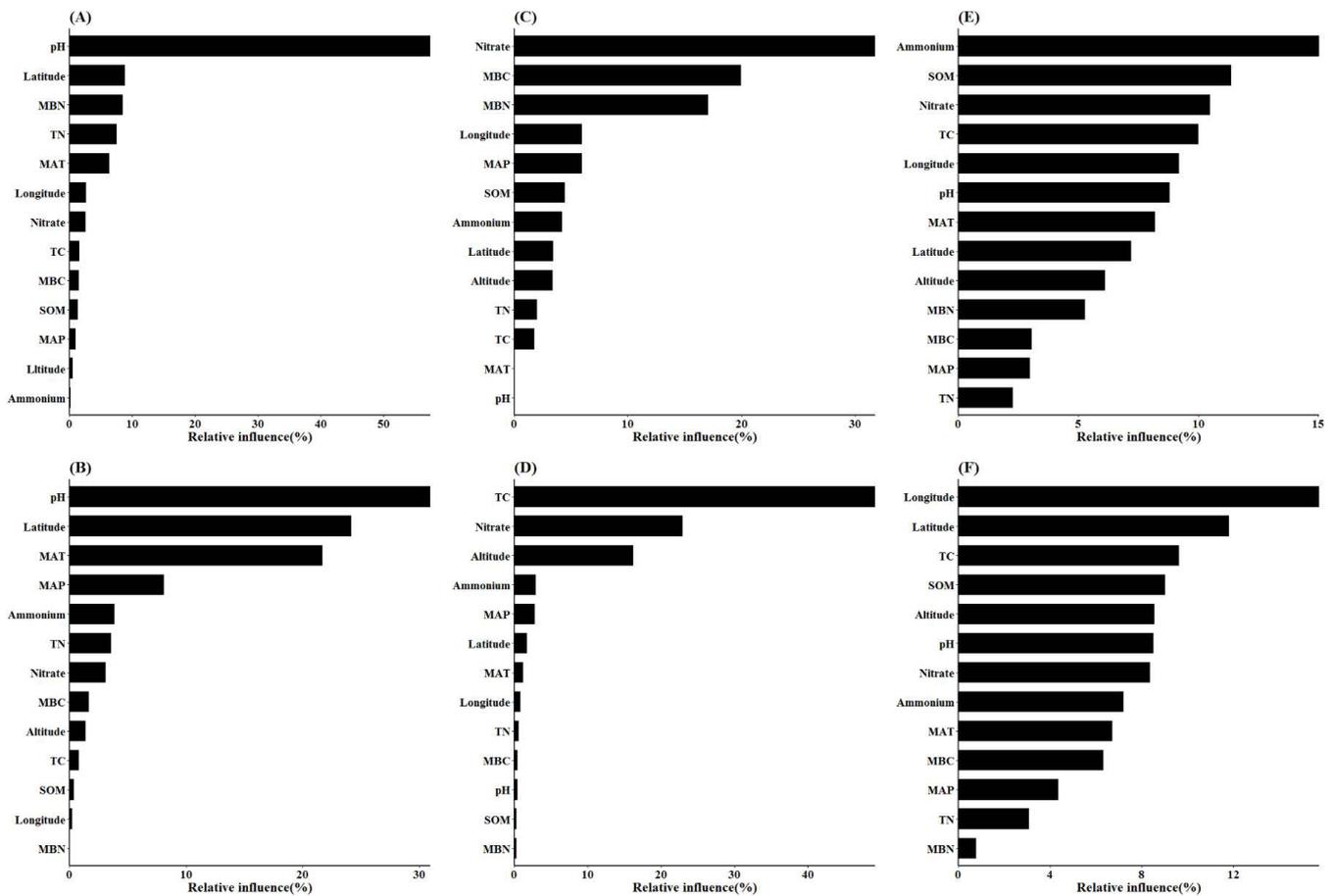
Fig. 6   Relative variable importance plot (%) of environmental drivers for (A, B) abundance, (C, D) diversity and (E, F) composition of comammox *Nitrospira* clade A and clade B by aggregated boosted tree (ABT) models. (A) comammox *Nitrospira* clade A abundance, (B) comammox *Nitrospira* clade B abundance, (C) Shannon index of comammox *Nitrospira* clade A, (D) Shannon index of comammox *Nitrospira* clade B, (E) Bray-Curtis distance of comammox *Nitrospira* clade A and (F) Bray-Curtis distance of comammox *Nitrospira* clade B.

DISCUSSION

To our knowledge, this is the first study to combine comammox *Nitrospira* community structure and abundance together with diversity at a large scale in Chinese paddy soils. The qPCR results indicated that both comammox *Nitrospira* clade A and clade B were prevalent in the paddy soils (Fig. 1A). Quantitative studies by other researchers have shown that in different habitats, AOA, AOB and comammox *Nitrospira* coexisted in diverse abundance patterns (Bartelme *et al.*, 2017; Fowler *et al.*, 2018; Li *et al.*, 2020). Studies have shown that comammox *Nitrospira* in paddy soil accounted for 22.9%–48.09% of the total detected autotrophic ammonia oxidizing microorganisms (Pjevac *et al.*, 2017). In paddy soils, the Shannon index of comammox *Nitrospira* clade A was 2.21–5.39 (Fig. 2), similar to findings for previous habitats, such as forest soil (2.54), plateau soil (5.33), sediment (2.87–5.03), and water (2.44–4.87) (Xia *et al.*, 2018). Moreover, the ratio of clade B/A *amoA* gene abundance ranged from 0.62 to 68.96, indicating much greater gene abundance of clade B than clade A, except for JSYC and FJNP (Fig. 1A). Some studies have also shown that the abundance of comammox *Nitrospira* in agricultural and forest soils was lower than that of AOA and/or AOB (Pjevac *et al.*, 2017; Kataoka *et al.*, 2018; Li *et al.*, 2020), higher than the abundance of AOA or AOB in acidic forest soils and grassland soils (Hu and He, 2017; Kataoka *et al.*, 2018) or similar to AOA in alkaline forest soils (Hu and He, 2017). However, no positive results for comammox *Nitrospira* clade B were obtained in terrestrial ecosystems of Australia (Li *et al.*, 2020; Li *et al.*, 2021). In addition, comammox *Nitrospira* clade A outnumbered clade B in alkaline paddy soils, with a ratio of clade A/B of 6.1, 14.4 and 43.1 in the three treatments (Wang *et al.*, 2018). These results indicate that pH may play an important role in controlling comammox *Nitrospira* abundance.

Even though research on comammox *Nitrospira* is accumulating rapidly, especially data on comammox *Nitrospira* abundance, pH had a good effect in controlling comammox *Nitrospira* abundance in paddy soils.

Comammox *Nitrospira* clade A and clade B were highly abundant in forest soils with the gradient of soil pH (Hu and He, 2017), and a significant positive correlation between pH and the abundance of comammox *Nitrospira* was noticed (Li *et al*., 2020; Lu *et al*., 2020; Xue *et al*., 2020). The relationship between the abundance of comammox *Nitrospira* and pH may be that the optimal pH of key N-converting enzymes (for example, HAO and AMO) was approximately 7.0–8.0 (Blum *et al*., 2018). We found that a variety of factors jointly drove comammox *Nitrospira*, and pH may be the most important factor affecting the comammox *Nitrospira* process. Low pH reduces the availability of ammonia, reduces the potential toxic effects of free ammonia (Gubry-Rangin *et al*., 2011) and also reduces the availability of nitrite and increases the toxic effects of nitrous acid (Boer and Kowalchuk, 2001). By sequencing the global soil AOA *amoA* gene, it was found that the AOA sequences attributable to Cluster 14 and Cluster 4 were obtained from the environment of pH < 5.50 and pH > 7.20, which indicates that AOA have different dominant communities under different pH conditions (Silva *et al*., 2006). The ASVs of comammox *Nitrospira* obtained in certain pH ranges may correspond to the adaptability of comammox *Nitrospira* to soil. However, certain environments (for example, high temperature, halophilic, heavy metal contamination or oligotrophic conditions) may contain other undetected comammox *Nitrospira*. In addition, the adaptability of ammonia oxidizing microorganisms to soil conditions also varies with factors such as pH, water content, substrate concentration and temperature (Shen *et al*., 2008; Sho *et al*., 2011; Stevens *et al*., 2020).

The community structure of both clade A and B was affected by both longitude and latitude (Fig. 3). Geographical factors have previously been shown to have a strong influence on AOA and AOB communities (Hu *et al*., 2015; Chen *et al*., 2017; Liu *et al*., 2018; Liu *et al*., 2019). Therefore, it is not difficult to understand the relationship between geographic factors (longitude and latitude explaining 16.40% of clade A and 27.53% of clade B variation, respectively) and comammox *Nitrospira* communities. In addition, SOM affected the community diversity of AOA (Chen *et al*., 2019), and SOM also influenced the comammox *Nitrospira* community (explaining 11.37% of clade A and 9.02% of clade B variation) (Fig. 6E, F). Under different ammonium concentrations, both AOA and AOB showed clear community changes, and the changes were different (Hu *et al*., 2012). Ammonium showed a strong relationship with the community of comammox *Nitrospira* clade A (15.04%), higher than that with comammox *Nitrospira* clade B (7.20%) (Fig. 6E, F). These relationships between environmental factors and comammox *Nitrospira* may be due to the fact that comammox *Nitrospira* clade A and clade B contain two different types of ammonia transporters (Palomo *et al*., 2018), resulting in different ammonia absorption characteristics. The Rh-type ammonia transporter (AOB, comammox *Nitrospira* clade A) has ammonia affinity and high absorption in the millimolar range capacity, while MEP-type ammonia transfer protein (AOA, comammox *Nitrospira* clade B) has higher affinity (micromolar range) and absorption capacity (Palomo *et al*., 2018). For comammox *Nitrospira* clade A, ammonium is the primary influencing factor of the community. This may be due to the long-term dry-wet alternation of the rice field, which makes ammonium occur at the millimolar level. At this level, comammox *Nitrospira* clade A has a higher ammonium affinity and absorption capacity.

Ammonium is the primary influencing factor of the comammox *Nitrospira* clade A community, but for comammox *Nitrospira* clade B, the effect of ammonium is weaker. NMDS showed that there was no clear separation between the comammox *Nitrospira* clade A communities and no obvious separation between the comammox *Nitrospira* clade B communities (Fig. S2). Although its higher affinity for the substrate ammonia may enhance its ability to compete with other microorganisms in soil with an extremely low ammonia concentration, this does not mean that the comammox *Nitrospira* community has obvious environmental separation. As the combined effect of soil pH value, temperature and fertilization also affect the ammonia oxidation process, there is no significant difference in the community composition of comammox (Li et al., 2018; Wang et al., 2017). Soil pH was closely related to the overall ammonia-oxidizer community (Hu *et al*., 2013; Hu *et al*., 2015), and the large-scale biogeographic pattern was mainly controlled by pH value, geographic distance and climatic factors (Hu *et al*., 2015). Comparative genomics analysis has shown that there were significant differences between comammox *Nitrospira* clade A and clade B (Palomo *et al*., 2018). These differences mean that environmental factors such as pH, temperature and TN may affect the functional importance of comammox *Nitrospira* in natural habitats (Ai *et al*., 2013). This may be because the factors that adjust microbial communities are usually intricate and touch upon multiple environmental variables (Hu *et al*., 2015). Therefore, to better understand the molecular mechanism and relative influence of comammox *Nitrospira*, it is necessary to conduct more research, especially regarding pure culture and metagenomics.

CONCLUSIONS

Our results provide new evidence that comammox *Nitrospira* are widely distributed in the paddy fields

of Chinese agricultural soil, suggesting their potential functions of soil nitrification in these environments, which can be influenced by a range of factors. The abundance of comammox *Nitrospira* clade A was mostly lower than that of clade B, while its diversity was mostly higher than that of clade B. Phylogenetic analysis showed the ASVs were clustered with the three types of comammox *Nitrospira* that are currently purely cultured. There was clear relationship between the comammox *Nitrospira* community and habitat conditions. Some ASVs had a unique dominant position in some habitats. These findings can help to explain the process of ammonia oxidation by ammonia-oxidizing microorganisms, and how their distinct preferences for environmental factors and habitat define the relative contributions of comammox *Nitrospira* clade A and clade B to soil nitrification.

ACKNOWLEDGEMENTS

SUPPLEMENTARY MATERIAL

Supplementary material for this article can be found in the online version.

REFERENCES

Ai C, Liang G Q, Sun J W, Wang X B, He P, Zhou W. 2013. Different roles of rhizosphere effect and long-term fertilization in the activity and community structure of ammonia oxidizers in a calcareous fluvo-aquic soil. *Soil Biol Biochem.* **57**: 30-42.

Bartelme R P, McLellan S L, Newton R J. 2017. Freshwater recirculating aquaculture system operations drive biofilter bacterial community shifts around a stable nitrifying consortium of ammonia-oxidizing archaea and comammox *Nitrospira. Front Microbiol.* **8**.

Blum J M, Su Q X, Ma Y J, Valverde Pérez B, Domingo Félez C, Jensen M M, Smets B F. 2018. The pH dependency of N-converting enzymatic processes, pathways and microbes: effect on net N_2O production. *Environ Microbiol.* **20**: 1623-1640.

Boer W D, Kowalchuk G A. 2001. Nitrification in acid soils: micro-organisms and mechanisms. *Soil Biol Biochem.* **33**: 853-866.

Bolyen E, Rideout J R, Dillon M R, Bokulich N A, Abnet C C, Al-Ghalith G A, Alexander H, Alm E J, Arumugam M, Asnicar F, Bai Y, Bisanz J E, Bittinger K, Brejnrod A, Brislawn C J, Brown C T, Callahan B J, Caraballo-Rodriguez A M, Chase J, Cope E K, Da Silva R, Diener C, Dorrestein P C, Douglas G M, Durall D M, Duvallet C, Edwardson C F, Ernst M, Estaki M, Fouquier J, Gauglitz J M, Gibbons S M, Gibson D L, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley G A, Janssen S, Jarmusch A K, Jiang L, Kaehler B D, Kang K B, Keefe C R, Keim P, Kelley S T, Knights D, Koester I, Kosciolek T, Kreps J, Langille M G I, Lee J, Ley R, Liu Y X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin B D, McDonald D, McIver L J, Melnik A V, Metcalf J L, Morgan S C, Morton J T, Naimey A T, Navas-Molina J A, Nothias L F, Orchanian S B, Pearson T, Peoples S L, Petras D, Preuss M L, Pruesse E, Rasmussen L B, Rivers A, Robeson M S, 2nd, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song S J, Spear J R, Swafford A D, Thompson L R, Torres P J, Trinh P, Tripathi A, Turnbaugh P J, Ul-Hasan S, van der Hooft J J J, Vargas F, Vazquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber K C, Williamson C H D, Willis A D, Xu Z Z, Zaneveld J R, Zhang Y, Zhu Q, Knight R, Caporaso J G. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol.* **37**: 852-857.

Chen C H, Liang H, Gao D W. 2019. Community diversity and distribution of ammonia-oxidizing archaea in marsh wetlands in the black soil zone in North-east China. *Front Environ Sci Eng.* **13**.

Chen R R, Zhong L H, Jing Z W, Guo Z Y, Li Z P, Lin X G, Feng Y Z. 2017. Fertilization decreases compositional variation of paddy bacterial community across geographical gradient. *Soil Biol Biochem.* **114**: 181-188.

Costa E, Perez J, Kreft J U. 2006. Why is metabolic labour divided in nitrification? *Trends Microbiol.* **14**: 213-219.

Daims H, Lebedeva E V, Pjevac P, Han P, Herbold C, Albertsen M, Jehmlich N, Palatinszky M, Vierheilig J, Bulaev A, Kirkegaard R H, von Bergen M, Rattei T, Bendinger B, Nielsen P H, Wagner M. 2015. Complete nitrification by *Nitrospira* bacteria. *Nature.* **528**: 504-519.

Fowler S J, Palomo A, Dechesne A, Mines P D, Smets B F. 2018. Comammox *Nitrospira* are abundant ammonia oxidizers in diverse groundwater-fed rapid sand filter communities. *Environ Microbiol.* **20**: 1002-1015.

Gubry-Rangin C, Hai B, Quince C, Engel M, Thomson B C, Phillip J, Michael S, I G R, I. P J, W. N G. 2011. Niche specialization of terrestrial archaeal ammonia oxidizers. *P Natl Acad Sci USA.* **108**: 21206-21211.

Hou J Y, Wu L H, Liu W X, Ge Y Y, Mu T T, Zhou T, Li Z, Zhou J W, Sun X, Luo Y M, Christie P. 2020. Biogeography and diversity patterns of abundant and rare bacterial communities in rice paddy soils across China. *Sci Total Environ.* **730**: 139116.

Hu B L, Liu S, Shen L D, Zheng P, Xu X Y, Lou L P. 2012. Effect of different ammonia concentrations on community succession of ammonia-oxidizing microorganisms in a simulated paddy soil column. *PLOS ONE.* **7**: e44122.

Hu H W, He J Z. 2017. Comammox—a newly discovered nitrification process in the terrestrial nitrogen cycle. *J Soil Sediment.* **17**: 2709-2717.

Hu H W, Zhang L M, Dai Y, Di H J, He J Z. 2013. pH-dependent distribution of soil ammonia oxidizers across a large geographical scale as revealed by high-throughput pyrosequencing. *J Soil Sediment.* **13**: 1439-1449.

Hu H W, Zhang L M, Yuan C L, Zheng Y, Wang J T, Chen D L, He J Z. 2015. The large-scale distribution of ammonia oxidizers in paddy soils is driven by soil pH, geographic distance, and climatic factors. *Front Microbiol.* **6**.

Ishii S, Ikeda S, Minamisawa K, Senoo K. 2011. Nitrogen cycling in rice paddy environments: past achievements and future challenges. *Microbes Environ*. 26: 282-292.

Kataoka T, Suzuki K, Irino T, Yamamoto M, Higashi S, Liu H. 2018. Phylogenetic diversity and distribution of bacterial and archaeal *amoA* genes in the East China Sea during spring. *Arch Microbiol*. 200: 329-342.

Kits K D, Sedlacek C J, Lebedeva E V, Han P, Bulaev A, Pjevac P, Daebeler A, Romano S, Albertsen M, Stein L Y, Daims H, Wagner M. 2017. Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle. *Nature*. 549: 269 - 272.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 35: 1547-1549.

Li C Y, Hu H W, Chen Q L, Chen D L, He J Z. 2020. Niche differentiation of clade A comammox *Nitrospira* and canonical ammonia oxidizers in selected forest soils. *Soil Biol Biochem*. 149: 107925.

Li C Y, Hu H W, Chen Q L, Yan Z Z, Thi Nguyen B A, Chen D L, He J Z. 2021. Niche specialization of comammox *Nitrospira* clade A in terrestrial ecosystems. *Soil Biol Biochem*. 156: 108231.

Li Y Y, Chapman S J, Nicol G W, Yao H Y. 2018. Nitrification and nitrifiers in acidic soils. *Soil Biol Biochem*. 116: 290-301.

Liu J J, Yu Z H, Yao Q, Sui Y Y, Shi Y, Chu H Y, Tang C X, Franks A E, Jin J, Liu X B, Wang G H. 2018. Ammonia-oxidizing archaea show more distinct biogeographic distribution patterns than ammonia-oxidizing bacteria across the black soil zone of northeast China. *Front Microbiol*. 9: 00171.

Liu J J, Yu Z H, Yao Q, Sui Y Y, Shi Y, Chu H Y, Tang C X, Franks A E, Jin J, Liu X B, Wang G H. 2019. Biogeographic distribution patterns of the archaeal communities across the black soil zone of northeast China. *Front Microbiol*. 10: 00023.

Liu S F, Wang H Y, Chen L M, Wang J W, Zheng M S, Liu S T, Chen Q, Ni J R. 2020. Comammox *Nitrospira* within the Yangtze River continuum: community, biogeography, and ecological drivers. *ISME J*. 14: 2488-2504.

Liu Z B, Zhang C J, Wei Q Y, Zhang S Y, Quan Z X, Li M. 2020. Temperature and salinity drive comammox *Nitrospira* community composition in mangrove ecosystems across southeastern China. *Sci Total Environ*. 742: 140456.

Lu S D, Sun Y J, Lu B Y, Zheng D Y, Xu S W. 2020. Change of abundance and correlation of *Nitrospira inopinata*-like comammox and populations in nitrogen cycle during different seasons. *Chemosphere*. 241: 125098.

Meng X T, Li Y Y, Yao H Y, Wang J, Dai F, Wu Y P, Chapman S. 2020. Nitrification and urease inhibitors improve rice nitrogen uptake and prevent denitrification in alkaline paddy soil. *Appl Soil Ecol*. 154: 103665.

Meng X T, Liao H K, Fan H X, Zhang X C, Li Y Y, Yao H Y, Razavi B S. 2021. The geographical scale dependence of diazotroph assembly and activity: Effect of a decade fertilization. *Geoderma*. 386.

Palomo A, Pedersen A G, Fowler S J, Dechesne A, Sicheritz-Ponten T, Smets B F. 2018. Comparative genomics sheds light on niche differentiation and the evolutionary history of comammox *Nitrospira*. *ISME J*. 12: 1779-1793.

Pinto A J, Marcus D N, Ijaz U Z, Bautista-de Lose Santos Q M, Dick G J, Raskin L. 2016. Metagenomic evidence for the presence of comammox *Nitrospira*-like bacteria in a drinking water system. *mSphere*. 1: 1-8.

Pjevac P, Schauberger C, Poghosyan L, Herbold C W, van Kessel M, Daebeler A, Steinberger M, Jetten M S M, Lucker S, Wagner M, Daims H. 2017. *AmoA*-targeted polymerase chain reaction primers for the specific detection and quantification of comammox *Nitrospira* in the environment. *Front Microbiol*. 8: 1508.

Shen J P, Zhang L M, Zhu Y G, Zhang J B, He J Z. 2008. Abundance and composition of ammonia-oxidizing bacteria and ammonia-oxidizing archaea communities of an alkaline sandy loam. *Environ Microbiol*. 10: 1601-1611.

Shi Y, Jiang Y Y, Wang S Y, Wang X M, Zhu G B. 2020. Biogeographic distribution of comammox bacteria in diverse terrestrial habitats. *Sci Total Environ*. 717: 137257.

Sho M, Masahito H, Takada H Y, Kazunari N, Masatsugu Y, Toshihiko K, Makoto T, Hiroko A. 2011. Quantitative analyses of ammonia-oxidizing archaea (AOA) and ammonia-oxidizing bacteria (AOB) in fields with different soil types. *Microbes Environ*. 26: 248-253.

Silva G d, Kennedy E M, Dlugogorski B Z. 2006. Ab initio procedure for aqueous-phase pKa calculation: the acidity of nitrous acid. *J Phys Chem A*. 110: 11371-11376.

Stevens B M, Sonderegger D L, Johnson N C. 2020. Microbial community structure across grazing treatments and environmental gradients in the Serengeti. *Soil Ecology Letters*. 4: 45-56.

Timilsina A, Bizimana F, Pandey B, Yadav R K P, Dong W, Hu C. 2020. Nitrous oxide emissions from paddies: understanding the role of rice plants. *Plants*. 9.

van Kessel M A, Speth D R, Albertsen M, Nielsen P H, Op den Camp H J, Kartal B, Jetten M S, Lucker S. 2015. Complete nitrification by a single microorganism. *Nature*. 528: 555-559.

Wang J C, Wang J L, Rhodes G, He J Z, Ge Y. 2019. Adaptive responses of comammox *Nitrospira* and canonical ammonia oxidizers to long-term fertilizations: Implications for the relative contributions of different ammonia oxidizers to soil nitrogen cycling. *Sci Total Environ*. 668: 224-233.

Wang J G, Xia F, Zeleke J, Zou B, Rhee S K, Quan Z X. 2017. An improved protocol with a highly degenerate primer targeting copper-containing membrane-bound monooxygenase genes for community analysis of methane- and ammonia-oxidizing bacteria. *FEMS Microbiol Ecol*. 93: fiw244.

Wang M Y, Huang G H, Zhao Z R, Dang C Y, Liu W, Zheng M S. 2018. Newly designed primer pair revealed dominant and diverse comammox *amoA* gene in full-scale wastewater treatment plants. *Bioresour Technol*. 270: 580-587.

Wang M, Wang Z H, Shi X J, Jiang X J. 2018. Long-term fertilization effects on the abundance of complete ammonia oxidizing bacteria(comammox *Nitrospira*) in a neutral paddy soil (in Chinese). *Environ Sci*. 39: 4727-4734.

Wang X M, Wang S Y, Jiang Y Y, Zhou J M, Han C, Zhu G B. 2020. Comammox bacterial abundance, activity, and contribution in agricultural rhizosphere soils. *Sci Total Environ*. 727: 138563.

Wang Z H, Cao Y Q, Zhu-Barker X, Nicol G W, Wright A L, Jia Z J, Jiang X J. 2019. Comammox *Nitrospira* clade B contributes to nitrification in soil. *Soil Biol Biochem*. 135: 392-395.

Waqas A, Huang J, Liu K, Muhammad Q, Numan K M, Chen J, Sun G, Huang Q H, Liu Y R, Liu G R, Sun M, Li C, Li D C, Sehrish A, Yodgar N, Sajid M, Zhang H M. 2019. Changes in phosphorus fractions associated with soil chemical properties under long-term organic and inorganic fertilization in paddy soils of southern China. *PLOS ONE*. 14: e0216881.

Xia F, Wang J G, Zhu T, Zou B, Rhee S K, Quan Z X. 2018. Ubiquity and diversity of complete ammonia oxidizers (comammox). *Appl Environ Microbiol*. 84: e01390-18.

Xue P, McBratney A B, Minasny B, O'Donnell T, Pino V, Fajardo M, Ng W, Wilson N, Deaker R. 2020. Soil bacterial depth

distribution controlled by soil orders and soil forms. *Soil Ecology Letters*. **4**: 57-68.

Xu S Y, Wang B Z, Li Y, Jiang D Q, Zhou Y T, Ding A Q, Zong Y X, Ling X, Zhang S Y, Lu H J. 2020. Ubiquity, diversity, and activity of comammox *Nitrospira* in agricultural soils. *Sci Total Environ*. **706**: 135684.1-135684.11.

Zhang S B, Qin W, Xia X H, Xia L Z, Li S L, Zhang L W, Bai Y B, Wang G Q. 2020. Ammonia oxidizers in river sediments of the Qinghai-Tibet Plateau and their adaptations to high-elevation conditions. *Water Res*. **173**: 115589.

SUPPLEMENTARY MATERIAL

Table SI Distribution of sampling points in the 36 sites.

Site	Coordinate	Site	Coordinate	Site	Coordinate
HNCS	28°26′30″N ,112°53′48″E	HNZZ	27°06′31″N, 113°12′40″E	HNXT	27°28′17″N, 112°41′1″E
JXYC	28°04′26″N, 115°09′2″E	JXJA	27°51′35″N, 115°26′43″E	JXSR	28°38′8″N, 118°06′42″E
HLJQQHE	47°29′10″N, 123°54′20″E	HLJHEB	45°40′10″N, 126°21′9″E	HLJSH	47°09′8″N, 126°17′32″E
JSYC	32°57′39″N, 120°29′25″E	JSHA	33°27′11″ N, 119°28′10″E	JSSQ	33°58′47″ N, 118°14′24″E
AHCZ	32°38′50″N, 118°27′23″E	AHHF	31°39′50″ N, 117°15′35″E	ANLA	31°44′47″N, 116°18′55″E
HBJZ	30°07′37″N, 112°33′15″E	HBXY	31°44′56″ N, 112°29′44E″	HBHG	31°04′0″N, 115°0′5″E
SCCD	30°31′4″N, 103°31′59″E	SCDZ	30°41′22″ N, 107°21′57″E	SCYB	29°08′13″N, 104°39′49″E
GXYL	22°29′56″ N, 110°19′32″E	GXGG	23°18′1″ N, 110°28′20″E	GXQZ	22°20′24″, 109°10′57″E
GDZJ	21°07′55″ N, 110°19′51″E	GDMM	21°44′35″N. 110°49′26″E	GDMZ	24°25′16″, 116°22′16″E
YNHH	23°24′31″N, 103°23′14″E	YNPE	23°42′41″ N, 101°09′28″E	YNBS	25°09′25″, 99°15′14″E
ZJJX	30°42′37″ N, 120°56′23″E	ZJSX	29°21′52.14″N, 120°55′45.73″E	ZJWZ	27°35′8″ N, 120°23′10″E
FJNP	27°22′36″ N, 117°52′47″E	FJLY	25°08′11″N, 116°52′45″E	FJSM	26°22′59″N, 118°12′27″E

Table SII Information about the latitude and longitude, climate type, annual mean temperature, and annual mean precipitation of these 36 sites.

Site	MAT	MAP	Climate type	Site	MAT	MAP	Climate type
HNCS	17.20	1361.60	Subtropical monsoon climate	SCDZ	16.15	1100.00	Subtropical monsoon climate
JXYC	17.20	1680.20	Subtropical monsoon climate	GXGG	21.50	1600.00	Subtropical monsoon climate
HLJQQHE	2.45	450.00	Temperate monsoon climate	GDMM	21.00	1264.00	Subtropical monsoon climate
JSYC	14.20	1014.70	Subtropical monsoon climate	YNPE	11.50	1500.00	Subtropical monsoon climate
AHCZ	15.40	1035.50	Subtropical monsoon climate	ZJSX	16.40	1300.00	Subtropical monsoon climate
HBJZ	16.25	1200.00	Subtropical monsoon climate	FJLY	19.00	1700.00	Subtropical monsoon climate
SCCD	16.00	1000.00	Subtropical monsoon climate	HNXT	17.05	1510.00	Subtropical monsoon climate
GXYL	21.80	1700.00	Subtropical monsoon climate	JXSR	17.50	1725.00	Subtropical monsoon climate
GDZJ	23.20	1750.00	Subtropical monsoon climate	HLJSH	2.80	536.00	Temperate monsoon climate
YNHH	18.80	1545.00	Subtropical monsoon climate	JSSQ	14.20	910.00	Subtropical monsoon climate
ZJJX	15.90	1168.60	Subtropical monsoon climate	AHLA	15.60	1100.00	Subtropical monsoon climate
FJNP	19.30	1660.00	Subtropical monsoon climate	HBHG	16.40	1358.00	Subtropical monsoon climate
HNZZ	17.00	1500.00	Subtropical monsoon climate	SCYB	17.90	1169.60	Subtropical monsoon climate
JXJA	18.30	1500.00	Subtropical monsoon climate	GXQZ	22.00	1600.00	Subtropical monsoon climate
HLJHEB	5.00	569.10	Temperate monsoon climate	GDMZ	21.45	1578.70	Subtropical monsoon climate
JSHA	14.45	1000.00	Subtropical monsoon climate	YNBS	16.00	1000.00	Subtropical monsoon climate
AHHF	15.50	995.30	Subtropical monsoon climate	ZJWZ	18.00	2150.00	Subtropical monsoon climate
HBXY	15.50	1000.00	Subtropical monsoon climate	FJSM	19.20	1700.00	Subtropical monsoon climate

Table SIII Quantitative real-time polymerase chain reaction (Q-PCR) and PCR primer sets and amplification conditions used in this study.

Target gene	Primer name	Primer sequence (5'-3')	Amplification conditions	Length	Reference
Comammox *Nitrospira* clade A *amoA*	comaA-244f_a	TACAACTGGGTGAACTA	95°C for 5 min, 45 cycles of 95°C for 30 s, 52°C for 45 s and 72°C for 1 min	415	Pjevac *et al.*, (2017)
	comaA-244f_b	TATAACTGGGTGAACTA			
	comaA-244f_c	TACAATTGGGTGAACTA			
	comaA-244f_d	TACAACTGGGTCAACTA			
	comaA-244f_e	TACAACTGGGTCAATTA			
	comaA-244f_f	TATAACTGGGTCAATTA			
	comaA-659r_a	AGATCATGGTGCTATG			
	comaA-659r_b	AAATCATGGTGCTATG			
	comaA-659r_c	AGATCATGGTGCTGTG			
	comaA-659r_d	AAATCATGGTGCTGTG			
	comaA-659r_e	AGATCATCGTGCTGTG			

	comaA-659r_f	AAATCATCGTGCTGTG	
	comaB-244f_a	TAYTTCTGGACGTTCTA	
	comaB-244f_b	TAYTTCTGGACATTCTA	
	comaB-244f_c	TACTTCTGGACTTTCTA	
	comaB-244f_d	TAYTTCTGGACGTTTTA	
	comaB-244f_e	TAYTTCTGGACATTTTA	
Comammox *Nitrospira* clade B *amoA*	comaB-244f_f	TACTTCTGGACCTTCTA	95°C for 5 min, 45 cycles of 95°C for 30 s, 52°C for 45 s and 72°C for 1 min
	comaB-659r_a	ARATCCAGACGGTGTG	415	Pjevac *et al.,* (2017)
	comaB-659r_b	ARATCCAAACGGTGTG	
	comaB-659r_c	ARATCCAGACAGTGTG	
	comaB-659r_d	ARATCCAAACAGTGTG	
	comaB-659r_e	AGATCCAGACTGTGTG	
	comaB-659r_f	AGATCCAAACAGTGTG	

Pjevac P, Schauberger C, Poghosyan L, Herbold C W, van Kessel M, Daebeler A, Steinberger M, Jetten M S M, Lucker S, Wagner M, Daims H. 2017. *AmoA*-targeted polymerase chain reaction primers for the specific detection and quantification of comammox *Nitrospira* in the environment. *Front Microbiol.* **8**: 1508.

Table SIV The statistical results of communities of comammox *Nitrospira* in paddy soil were showed with Illumina sequencing of comammox *Nitrospira* clade A and clade B *amoA* genes.

	Comammox Clade A	Comammox Clade B
Total	883521	843642
Min	8430	398
Max	84719	40630
Average	24542	23434
Read	375-415	375-439
ASVs	2700	1551
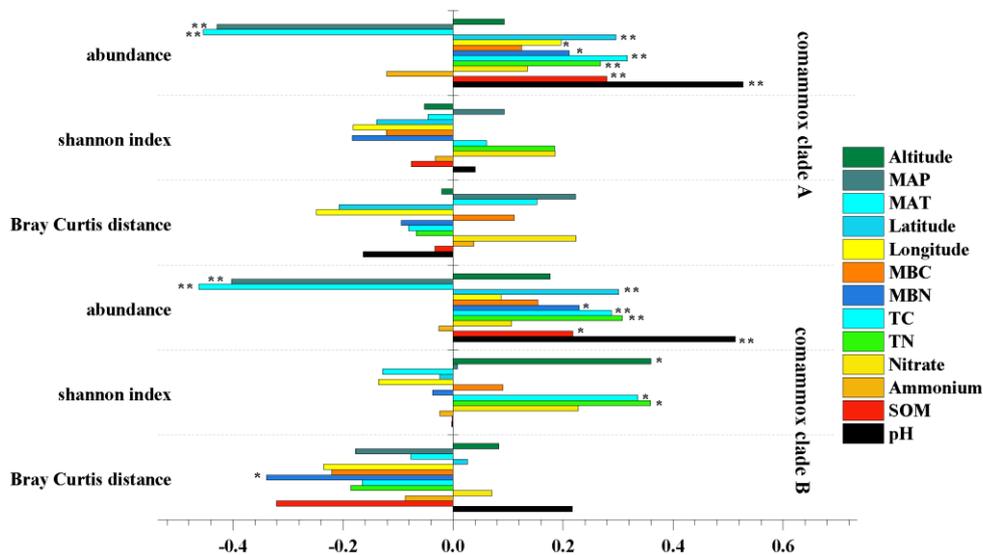
Fig. S1 Spearman correlation analyses of the abundance, diversity and community composition of comammox *Nitrospira* and ecological factors. The correlation coefficients are shown in the table. * *P* < 0.05, ** *P* <0.01.
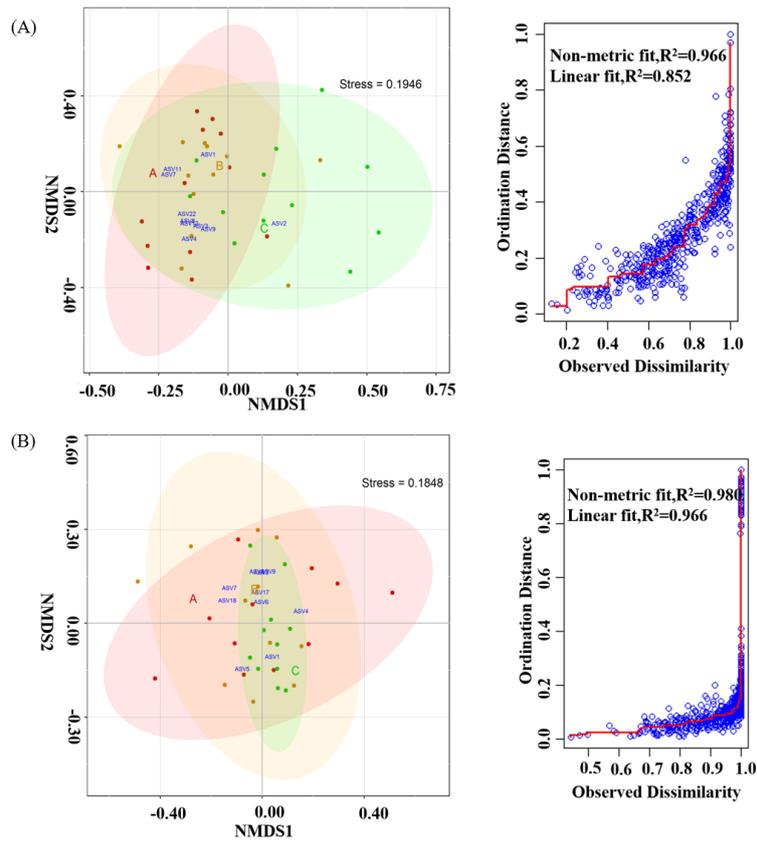
Fig. S2 Nonmetric multidimensional scaling (NMDS) ordination plot depicts the Bray–Curtis distance of (A) comammox *Nitrospira* clade A and (B) clade B communities. The two-dimensional stress value for the NMDS were 0.1946 and 0.1848 based on Bray Curtis distance. The R^2 was determined to evaluate the fit between ordination distances and observed dissimilarity. Comammox *Nitrospira amoA* ASVs representative sequences from this study are shown in ASVXX.