

# Rothamsted Repository Download

## A - Papers appearing in refereed journals

Clark, I. M., Hughes, D. J., Fu, Q., Abadie, M. and Hirsch, P. R. 2021. Metagenomic approaches reveal differences in genetic diversity and relative abundance of nitrifying bacteria and archaea in contrasting soils . *Scientific Reports*. 11. <https://doi.org/10.1038/s41598-021-95100-9>

The publisher's version can be accessed at:

- <https://doi.org/10.1038/s41598-021-95100-9>

The output can be accessed at:

<https://repository.rothamsted.ac.uk/item/985q7/metagenomic-approaches-reveal-differences-in-genetic-diversity-and-relative-abundance-of-nitrifying-bacteria-and-archaea-in-contrasting-soils>.

© 5 August 2021, Please contact [library@rothamsted.ac.uk](mailto:library@rothamsted.ac.uk) for copyright queries.

**Metagenomic approaches reveal differences in genetic diversity and relative abundance of nitrifying bacteria and archaea in contrasting soils**

Ian M Clark<sup>a</sup>, David J Hughes<sup>b</sup>, Qingling Fu<sup>a,c</sup>, Maider Abadie<sup>a</sup>, Penny R Hirsch<sup>a\*</sup>

<sup>a</sup> Sustainable Agriculture Sciences Department, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK

<sup>b</sup> Computational and Analytical Sciences, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK

<sup>c</sup> College of Resources and Environment, Huazhong Agricultural University, Wuhan 430070, Hubei Province, PR China

\*corresponding author

[penny.hirsch@rothamsted.ac.uk](mailto:penny.hirsch@rothamsted.ac.uk)

tel: 0044 (0)1582 938669

fax: 0044 (0)1582 760981

**Keywords**

soil metagenome, soil metatranscriptome, nitrification potential, ammonia oxidizing archaea, ammonia oxidizing bacteria, nitrite oxidizing bacteria, comammox, anammox

## Abstract

The abundance and phylogenetic diversity of functional genes involved in nitrification were assessed in Rothamsted soils under contrasting management regimes - permanent bare fallow, grassland and arable (wheat) cultivation maintained for more than 50 years. Metagenome and metatranscriptome analysis indicated that nitrite oxidizing bacteria (NOB) were more abundant than ammonia oxidizing archaea (AOA) and bacteria (AOB) in all soils. The most abundant AOA and AOB in the metagenomes were, respectively, *Nitrososphaera* and *Ca. Nitrososcosmicus* (family Nitrososphaeraceae) and *Nitrospira* and *Nitrosomonas* (family Nitrosomonadaceae). The most abundant NOB were *Nitrospira* including the comammox species *Nitrospira inopinata*, *Ca. N. nitrificans* and *Ca. N. nitrosa*. Anammox bacteria were also detected. *Nitrospira* and the AOA Nitrososphaeraceae showed most transcriptional activity in arable soil. Similar numbers of sequences were assigned to the *amoA* genes of AOA and AOB, highest in the arable soil metagenome and metatranscriptome. The AOB *amoA* reads included those from comammox *Nitrospira* clades A and B, in addition to Nitrosomonadaceae. Nitrification potential assessed in soils amended or not with DCD (an inhibitor of AOB but not AOA) was highest in arable soils and lower in all soil microcosm assays containing DCD, indicating AOB were responsible for oxidizing ammonium fertilizer added to these soils.

## Introduction

Soil microorganisms are essential for maintaining supplies of bioavailable N for plants, whether in natural ecosystems or agricultural systems <sup>1</sup>. A key step in the nitrogen cycle is nitrification, the oxidation of ammonia derived from breakdown of organic matter, animal wastes or man-made fertilizer to nitrate via hydroxylamine and nitrite <sup>2</sup>. The positively-charged ammonium cation (the predominant form of ammonia in soil) is relatively immobile, associating with negatively charged surfaces. The more mobile nitrate anion is more likely to reach roots but is liable to be lost by leaching from soil and is also the substrate for denitrification, an anaerobic process which results in losses of the greenhouse gas nitrous oxide <sup>3</sup>. Bacteria and archaea containing the enzyme ammonia monooxygenase (AMO) catalyse production of hydroxylamine from ammonia, usually considered to be the rate-limiting step <sup>2,4</sup> in nitrification. A variety of compounds such as dicyandiamide (DCD) inhibit AMO in bacteria and are applied to soil to slow conversion of fertilizer-N, limiting losses and improving the efficiency of agricultural production <sup>5</sup>. The relative susceptibility of AOA and AOB to commercial nitrification inhibitors in soil is not fully known but in lab culture, a representative of the most abundant AOA soil genus was unaffected by DCD at a dose five-fold higher than that needed to completely inhibit the most abundant AOB genus <sup>6</sup>.

Ammonia oxidizing bacteria (AOB) belong to the betaproteobacterial family Nitrosomonadaceae and the gammaproteobacterial genera *Candidatus Nitrosoglobulus terrae* and *Nitrosococcus* <sup>7,8</sup>. They are obligate chemoautotrophs that use type I RuBisCO to fix CO<sub>2</sub> via the Calvin cycle and grow very slowly in laboratory culture as a consequence of the low energy yield from oxidation of ammonia <sup>2,7</sup>. The Nitrosomonadaceae are the dominant AOB in soil and two genera are currently recognised: *Nitrospira* (which now includes *Nitrosovibrio* and *Nitrosolobus*, previously classified as separate genera) and *Nitrosomonas* <sup>7</sup>.

The ammonia oxidizing archaea (AOA) are considered to be oligotrophic chemoautotrophs. Currently, five major AOA lineages are recognised: Nitrososphaerales (NS); Nitrosopumilales (NP);

Candidatus Nitrosocaldales (NC); Candidatus Nitrosotaleales (NT); and a separate deeply-branching group related to NP and NT (Aves et al., 2018). In soil, the dominant lineage is NS which contains the genus *Nitrososphaera* and a related *Candidatus* genus, Nitrosocosmicus. *Nitrososphaera* isolates are difficult to grow in the lab but culture experiments have shown that some isolates are not obligate chemoautotrophs<sup>9-11</sup>. Complete ammonia oxidation to nitrate (comammox) has been demonstrated in bacteria belonging to sublineage II of the monophyletic clade Nitrospira, previously thought to comprise only nitrite oxidizers<sup>12,13</sup>. The enzyme AMO is distinct in AOB, AOA and comammox bacteria, the genes readily identified using metagenomics, qPCR and amplicon sequencing, providing evidence that AOA are often more abundant than AOB and respond differently to land management<sup>14-17</sup>. The comammox *amoA* genes fall into two clades, A and B, the former found in *Nitrospira* that include the named comammox species *N. inopinata*, *Ca. N. nitrificans* and *Ca. N. nitrosa*; the latter in a range of environmental *Nitrosospira* metagenomes<sup>18,19</sup>. Gene exchange between the betaproteobacterial AOB and Nitrospira may explain similarities in *amoA* sequences from the two phyla<sup>20</sup>. Other groups that oxidize ammonia using different mechanisms are the anaerobic ammonia oxidizing (anammox) Planctomycetes in marine environments and wastewater<sup>21</sup> and certain heterotrophic bacteria and fungi in acid soils<sup>22</sup>.

The AMO in some archaea have lower ammonia saturation and inhibition constants than those in bacteria, indicating that they would benefit less from the relatively high levels of ammonia in agricultural soils<sup>9</sup>, although the cultured soil archaeon *Nitrososphaera viennensis* grows at levels similar to those optimal for soil AOB<sup>23</sup>. The mechanism by which hydroxylamine is oxidized to nitrite is also different, with no evidence in AOA for the hydroxylamine oxidoreductase (HAO) present in AOB<sup>24-26</sup>. The comammox *N. inopinata* also has a high ammonia affinity and low maximum ammonia oxidation rate<sup>27</sup>. Comammox and AOA *amoA* abundance is reported to co-vary in soil, indicating both have an advantage in soils with low available ammonia, in contrast to the non-comammox AOB<sup>28</sup>. However, to date, there are few studies focussing on comammox ecology in soil

and it is possible that they include *Nitrospira* that are adapted to benefit from relatively high levels of available ammonia<sup>29</sup>.

Chemoautotrophic bacteria that oxidize nitrite to nitrate (NOB), generating energy for aerobic growth, belong to the phyla Proteobacteria, Chloroflexi, Nitrospina, and Nitrospira (which includes the commamox group)<sup>30</sup>. The nitrite oxidoreductases (NXR), molybdopterin-containing members of the large DMSO reductase family<sup>31</sup>, are located in the cytoplasm of proteobacterial NOB and the periplasm of the other phyla. The distinction between nitrate reductase encoded by *narG* and the cytoplasmic NXR encoded by *nxrA* is defined by their cellular environments; however, the genes also fall into distinct groups in a phylogenetic analysis of the DMSO reductase superfamily<sup>31</sup>. Periplasmic NXR, as found in the Nitrospira, is estimated to be more energy-efficient than the cytoplasmic version, enabling growth at lower nitrite concentrations<sup>32</sup>. Whereas the proteobacterial *Nitrobacter* fix CO<sub>2</sub> via the Calvin cycle, at least some Nitrospira can use organic C<sup>33</sup>.

The inherent difficulty in growing these slow-growing microorganisms in lab culture has previously impeded research. A better understanding of the various nitrifier groups in agricultural soil, essential for effective management of the N cycle, is now being developed using metagenomics and metatranscriptomics. On Rothamsted farm, we have shown that AOB numbers increase in response to added N fertilizer<sup>34</sup> and that AOA were more abundant in tilled soils than undisturbed areas<sup>16</sup>. In arable soil receiving regular N fertilizer applications, the number of *amoA* copies from AOA and AOB was shown to be similar, at 5 x 10<sup>6</sup> copies g<sup>-1</sup> dry soil, using qPCR<sup>35</sup>. However, qPCR is limited by the specificity of primers, may not capture all groups that are present and can be skewed by primer bias. In contrast, sequencing the soil metagenome and metatranscriptome provides a representative sample of the genomes present and the genes that are being expressed although there may be an initial bias in nucleic acid extraction efficiency from different organisms, a problem for all molecular methods and subsequent analyses<sup>36</sup>. Current next-generation sequencing methods present different challenges, to identify genes of interest due to the short reads and large data sets. Here, we investigate metagenomic and metatranscriptomic datasets from Rothamsted soils with

contrasting long-term management including N-fertilizer treatments, together with assays for nitrification potential, to compare the abundance and activity of bacteria and archaea involved in different nitrification steps.

## Results

In the three years prior to the experiment mean DNA yields for bare fallow, arable and grassland were 21, 95 and 298  $\mu\text{g g}^{-1}$  dry soil, respectively with no significant differences between replicate plots within each treatment, highly significant differences between treatments ( $F_{2,18} = 646$ ,  $P < .001$ , all means significantly different at  $\alpha=0.05$ ) and no interaction between treatments and years, indicating that the effects of treatments and time were independent. The corresponding 16S rRNA gene copy numbers were  $3.4 \times 10^8$ ,  $14.1 \times 10^8$  and  $33.3 \times 10^8$  copies  $\text{g}^{-1}$  dry soil, equivalent to approximately  $85 \times 10^6$ ,  $353 \times 10^6$  and  $831 \times 10^6$  prokaryotic cells  $\text{g}^{-1}$  dry soil assuming an average of four 16S rRNA gene copies per cell<sup>37</sup>, highly significantly different between treatments ( $F_{2,18}=235$ ,  $P<0.001$ ; all means significantly different at  $\alpha=0.05$ ) with no effect of the sampling year and no treatment by year interaction.

In 2011 the mean DNA yields for bare fallow, arable and grassland were 22, 104 and 275  $\mu\text{g g}^{-1}$  dry soil, respectively, all significantly different from each other but not from the same plots for the previous three years with no interaction between treatment and years, which justified using the three-year mean estimate of prokaryotic cells.

Metagenomic DNA sequencing generating a 149 nt average read length provided  $\sim 42$  Gb from bare fallow and arable plots and 74 Gb from the grassland soil, corresponding to  $3 \times 10^8$  and  $5 \times 10^8$  sequences, respectively. The RNA extracted at the same time (average read length 87 nt) gave  $1 \times 10^8$ ,  $2 \times 10^8$  and  $3 \times 10^8$  sequences corresponding to 5.4, 2.9 and 8.6 Gb from bare fallow, arable and grassland soil, respectively. MEGAN6 identified over 5,500 prokaryotic taxa in each of the sequenced samples (bare fallow – 5 617; arable – 5 563; grass – 5 516); not significantly different

between treatments although the grassland sequences required the smallest mass of soil and bare fallow the largest to achieve the same amount of DNA for sequencing.

Functional nitrification assays showed that after 7 days, 70% of the added ammonium-N had been converted to nitrate in the arable soil compared to only 20% in bare fallow and grassland soils. The daily rate of nitrate production was approximately threefold higher in arable soil; DCD, a specific inhibitor of bacterial AMO, halved the rate of nitrate production in all soils (Figure 1).

Metagenomes and metatranscriptomes were interrogated using DIAMOND/MEGAN6 for all sequences belonging to known guilds AOA, AOB, NOB, comammox and anammox. NOB sequences were significantly most abundant overall in the metagenome with similar numbers for AOA and AOB, and significantly fewer for comammox and anammox bacteria ( $F_{4,30} = 139$ ,  $P < .001$ ); NOB and AOA sequences had similar abundance in the transcriptome, significantly more than AOB which in turn was significantly more abundant than comammox and anammox ( $F_{4,30} = 230$ ,  $P < .001$ ; (Figure 2 A, B). Overall, DNA sequences assigned to nitrifying groups were significantly more abundant in the arable and grassland soils than in bare fallow ( $F_{2,30} = 119$ ,  $P < .001$ ; Figure 2 A) whereas there were most mRNA sequences in the arable soil with fewer in grass and fewest in the bare fallow soil ( $F_{2,30} = 111$ ,  $P < .001$ ; Figure 2 B). When the results were expressed as ‰, NOB and AOA DNA formed a significantly greater proportion of the total prokaryotic sequences in arable and bare fallow compared to the grassland soil ( $F_{2,30} = 139$ ,  $P < .001$ ; Supplementary Figure 1 A), and mRNA ‰ was significantly highest in arable and lowest in bare fallow soil ( $F_{2,30} = 230$ ,  $P < .001$ ; Supplementary Figure 1 B).

Comparison of DNA and mRNA reads of known nitrifier guilds identified using DIAMOND/MEGAN6 revealed that the NOB were dominated by *Nitrospira*, with *Nitrobacter* comprising 2 – 20% and the other NOB groups less than 2% (Figure 3 A, B). For AOA, the NS groups *Ca. Nitrosocosmicus* and *Nitrososphaera* were most numerous, with *Ca. Nitrosocosmicus* more dominant in the mRNA and fewer than 5% of sequences assigned to other groups (Figure 3 C, D). The most abundant AOB in soil



metagenomes and metatranscriptomes was *Nitrosospira* (50 - 80%) with fewer assigned to *Nitrosomonas* (13 – 41%) and less than 10% identified as gammaproteobacteria (Figure 3 E, F).

Anammox bacterial DNA and mRNA was dominated by *Ca. Brocadia* but all known genera were detected (Figure 3 G, H). However, the overall proportion of comammox and anammox bacteria was low compared to the AOA, AOB and NOB (Figure 2 A, B). Of NOB identified as comammox, *N. inopinata* and *Ca. N. nitrificans* were present with similar abundance in DNA reads with *Ca. N. nitrificans* present at 20 – 28% but in mRNA, *Ca. N. nitrificans* was most abundant at 55 – 83 % of the total (Figure 3 I, J). The named comammox species were present at around 5% of total *Nitrosospira* DNA and 3% of mRNA sequences, but they were closer in abundance to the AOB in bare fallow soils (~40% of DNA and mRNA sequences, 20% DNA 10 % RNA in arable and grassland). A full list of mean DNA and mRNA values for each group within each guild, and the guild totals, expressed both as reads g<sup>-1</sup> dw soil and as % total prokaryotic DNA or mRNA sequences in each sample, together with ANOVA results and Tukey's post-hoc test for significantly different means, is given in Supplementary Tables 1 – 4.

Further analyses identified individual sequence reads with homology to *amoA* genes from AOA, AOB and comammox clades A and B in the metagenomes and metatranscriptomes. The proportion of reads assigned to the component groups of each guild, expressed as % total prokaryotic sequences, are shown in Figure 4 A, B and Supplementary Table 5. Results expressed as reads g<sup>-1</sup> dw soil are shown in Supplementary Figure 2 A, B and Supplementary Table 6. The number of genomic sequences assigned to AOA and AOB *amoA* was < 0.1% of the overall total number of DNA sequences assigned to the genome of each group. However, number of mRNA reads was higher, with AOB *amoA* reads at 6% of the total for AOB in arable soil, 2% in bare fallow and 1.5% in grassland soil; for AOA *amoA* there were < 1% total AOA mRNA reads in all soils. AOA *amoA* could not be divided into sequences belonging to either *Nitrososphaera* or *Ca. Nitrosocosmicus*, but the NS group was the most abundant around 80% of DNA and 90% of mRNA reads, with NP comprising 8 – 16% of DNA and 5 – 13 % mRNA reads. AOB *amoA* DNA sequences included *Nitrosospira* and

*Nitrosomonas* comprising 60% of DNA reads with comammox clades A and B contributing 12 – 16%. The Nitrosomonadaceae also contributed 60% of mRNA reads with commamox clade A contributing 20 – 27% and gammaproteobacterial 10 – 14%. No reads were detected from the gammaproteobacterial AOB *Ca. Nitrosoglobulus terrae* and *Nitrosococcus*.

## Discussion

Our previous finding, based on 16S rRNA amplicon sequencing, indicated that AOA constitute a higher proportion of the microbial community in tilled soils than undisturbed areas<sup>16</sup> and that the *Nitrospira* were more abundant than the AOA, with highest numbers in the arable plots<sup>35</sup>. The AOA and AOB were more abundant in the Highfield arable soils compared to grassland and bare fallow, using qPCR, where  $\sim 5 \times 10^6$  copies of AOA-*amoA* and AOB-*amoA* g<sup>-1</sup> dry arable soil with fewer in the other plots was reported<sup>35</sup>. The relative abundance of AOA and AOB reads in the metagenomic and metatranscriptome gave similar results, with an estimated of  $\sim 3 \times 10^6$  AOA and AOB in arable soil. It also indicates that the AOA are more transcriptionally active than the AOB, possibly because AOA remain active when N levels in soil are low and AOB are inactive, likely to be the case when soil was sampled prior to fertilizer addition in October. The ability of AOA to survive heterotrophically, as well as their high ammonia affinity enabling utilization of very low concentrations, may contribute to this. However, the relative abundance of *amoA* DNA and mRNA reads from AOA and AOB was similar, with mRNA comprising a much larger proportion of the metatranscriptome (2‰) than DNA reads of the metagenome (0.003‰), indicating its importance for these nitrifiers. The AOB-*amoA* mRNA included a relatively high proportion of comammox (40 - 50%), broadly consistent with the relative abundance of the total number of reads assigned to the Nitrosomonadaceae and clade A comammox *Nitrospira* overall (10 – 60%). This suggests that comammox previously assumed to inhabit aquatic and marine environments are present and active in relatively high numbers in soil. Results may be influenced by the assignment of reads from Nitrosomonadaceae to comammox

Nitrospira, or *vice-versa*, due to similarities between their *amoA* genes <sup>20</sup>, and potential biases in databases used for the assignment.

The functional assay for nitrification potential, where ammonium was added to soil at a relatively high concentration equivalent to 250 kg N ha<sup>-1</sup>, showed much higher enzyme activity attributed to AOB (i.e. inhibited by DCD) in the arable soil. Nitrification activity in grassland and bare fallow soil was similar although AOB were more abundant in grassland. The concentration of DCD was calculated to be equivalent to a level inhibitory to the AOB *Nitrosospira amoA* but not to affect the AOA *Nitrososphaera amoA*. The susceptibility of comammox Nitrospira to DCD is not known but it is reported that their abundance is not reduced by addition of DCD to arable soil, in contrast to the Nitrosomonadaceae <sup>29</sup>. Nitrification commences rapidly after substrate addition and the assay measures AMO already present in cells, rather than *de-novo* transcription of *amoA* and synthesis of new enzyme, or cell division<sup>38</sup>.

The study shows that AOA are both more numerous and active than AOB, especially in N-fertilized arable soils compared to bare fallow with no N inputs and grassland containing large populations of competitive heterotrophic bacteria. The AOB are likely to be responsible for most of the rapid nitrification observed when N fertilizer is added to soil but appear to be less active when N-availability is low; AOA in contrast remain active at low soil N concentrations and may provide a background level of nitrification. It is unclear if the comammox AOB are contributing to nitrification although they are present and active at around 50% of the number of Nitrosomonadaceae in the arable soil. The Nitrosomonadaceae are obligate autotrophs and obtain C only from CO<sub>2</sub> fixation via the Calvin cycle; the AOA and comammox Nitrospira use different and more efficient C fixation pathways and at least some can utilize simple organic substrates <sup>27</sup>. Therefore, the AOA and comammox AOB have a survival advantage in soil conditions where organic C is available and the paucity of N limits the Nitrosomonadaceae. However, when N fertilizer is applied, they are less able to take advantage of the increased ammonia concentration than the Nitrosomonadaceae, which increase in numbers.

The NOB are present and active in relatively high numbers compared to the AOA and AOB, which may explain why nitrite oxidation is rapid and rarely limiting in most soil. Nitrospira are the most abundant group, with *Nitrobacter*, previously considered to be the archetypal nitrifier, at < 20%. This suggests that further investigation of the role of Nitrospira in soil is important. In contrast to the cytoplasmic NXR of *Nitrobacter*, the periplasmic Nitrospira NXR maintains oxidation activity at very low nitrite concentrations <sup>32</sup>. The Nitrospira fix CO<sub>2</sub> more efficiently than *Nitrobacter* (which, like the proteobacterial AOB, use the Calvin cycle) and, as mentioned above, can utilize some organic C <sup>27</sup> potentially providing an alternative energy source <sup>33</sup>. This would enable Nitrospira to compete in low nutrient soils and benefit from nitrite generated by AOA also active in those conditions. Meanwhile, it is likely that the proteobacterial nitrifiers Nitrosomonadaceae and *Nitrobacter* take advantage of the relatively high ammonia levels after N fertilizer is applied to arable soil with a growth boost but then die down as plants remove the N and become less abundant than the AOA and Nitrospira at the end of the season, as in October when soil was sampled. Our previous work had shown such episodic increases in AOB in response to fertilizer applications <sup>34</sup>. Thus, AOA and AOB, Nitrospira and *Nitrobacter* may not compete directly for resources in soil, but instead occupy complementary niches.

The anammox bacteria were not thought to play a role in aerobic environments such as soil, and their numbers are relatively low, but they appear to be present and active and it is possible that they inhabit anaerobic, saturated soil pores and have an as-yet undetermined role in the terrestrial N cycle. The study also revealed groups previously reported in marine environments, including the NP AOA Nitrosopumilis and the NOB Nitrospina, were present at low abundance in terrestrial systems.

The analysis of soil metagenomes and metatranscriptomes offers a new approach to understanding soil microbial ecology and this study indicates that the results are consistent with those obtained using other methods for previously-identified groups, e.g. qPCR. However, it also reveals the presence and activity of less well-known groups which have not been grown in lab culture and which

do not have extensive genomic sequence information, leading to a more comprehensive understanding of the functioning of the complex soil microbiome.

## **Methods**

### *Soil sampling*

The Highfield Ley-Arable experiment at Rothamsted Research contrasts different long-term land management: permanent grassland; continuous arable (wheat) that receives regular fertilizer applications; and bare fallow, where all plant growth has been removed by regular tillage. Located at 00:21:48W, 51:48:18N, the soil is a silty loam over clay and is classified as a Chromic Luvisol according to FAO criteria. After >50 y, the plots have developed distinct microbial communities and divergent soil organic C and structure<sup>39</sup>. Soil was collected from the Highfield permanent plots in October 2011 to 10 cm depth using a 3 cm diameter corer; the top 2 cm containing root mats and other plant detritus was discarded. Ten cores per plot were pooled, thoroughly mixed whilst sieving through 2 mm mesh and placed in liquid N<sub>2</sub> within 5 minutes of collection; then samples were frozen at -80 °C prior to nucleic acid extraction; this provided three true replicate soil samples per plot treatment. All implements were cleaned with 70% ethanol between sampling/sieving soil from each plot.

### *Nucleic acid extraction and analysis*

To ensure that extracted DNA and RNA came from the same soil sample, community DNA and RNA was extracted from a minimum of 2 g soil using the MoBio RNA PowerSoil® Total RNA isolation kit followed by the RNA PowerSoil®DNA Elution Accessory kit, with three replicates for each soil treatment. All RNA samples were DNAase treated with Ambion Turbo DNA-free™. When necessary, extracts were pooled to provide sufficient material for sequencing. For the three years prior to this study, soil 16S rRNA gene copy numbers were estimated by qPCR with universal primers to allow normalisation of results<sup>40</sup>.

Full metagenomic sequencing of >10 µg DNA from each replicate soil treatment was provided by Illumina®, Cambridge, UK using a HiSeq 2000, generating 150bp paired end reads. RNA was subjected to ribodepletion and sequenced by The Genome Analysis Centre (TGAC), Norwich, UK using a HiSeq 2000, generating 100 bp paired end reads. Sequences were quality checked using the FASTX-Toolkit (version 0.0.13.2, [http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html)) with a quality threshold of 25, minimum length 100 bp for DNA and 80 bp for RNA. Reads were assigned to taxa using DIAMOND <sup>41</sup> and analysed in MEGAN6 <sup>42</sup>; Megan assigned KO (KEGG Orthology) molecular functional identifiers were used to extract individual sequences for each KEGG function (k04561 : norB, k10944 : amoA). Alignment and phylogenic assignation of retrieved sequences from metagenomic libraries was performed using MAFFT v7.450<sup>43</sup> translation alignment and Geneious v10.2.3 Tree Builder (Jukes-Cantor distance model, Neighbor-Joining tree build, random seed 1000 bootstraps) against archetypal reference sequences across the diversity of the targeted gene. Sequences retrieved from the soil metagenomes have been deposited in the ENA database (<http://www.ebi.ac.uk/ena/data/view/xxxxxxxxxx>).

The relative presence of each group or taxon assigned to taxon by DIAMOND/MEGAN6 was expressed as a proportion per 1000 (‰) of all prokaryotic reads in the relevant sample; the relative abundance of taxa in soil was estimated as the product of (number of prokaryotic cells g<sup>-1</sup> dry soil) x (proportion of all reads assigned to taxon by DIAMOND/MEGAN6 as a proportion of prokaryotic reads). Taxa assumed to represent nitrifying groups are indicated in the introduction.

#### *Potential Nitrification activity*

Soil was collected from each plot in October 2014, mixed and sieved as described above then 25 g aliquots were placed in Ziploc bags. These were stored in the dark at 4 °C for 7 days and 20 °C for a further 7 days. Meanwhile samples were taken to estimate field water holding capacity and dry weight. A solution was added at levels estimated to provide 100 mg NH<sub>4</sub><sup>+</sup>-N g<sup>-1</sup> dry soil with water to standardise moisture at 60% of field capacity; for controls, water alone was added. The nitrification

inhibitor DCD was added to half of the bags to provide  $16.8 \mu\text{g mL}^{-1}$  in the pore water ( $\approx 200 \mu\text{M}$ ).

Bags were incubated in the dark at  $20^\circ\text{C}$  and 5 g samples taken at the start, after 1, 4 and 7 days and frozen at  $-20^\circ\text{C}$ . Soils were extracted in 2 M KCl ( $5 \text{ mL g}^{-1}$  dw soil) by vigorous shaking (300 rpm) for 2 h then left to stand for 45 min before filtering through Whatman no 1 paper. Nitrate and ammonium in the filtrate were analysed simultaneously using a Skalar SAN<sup>PLUS</sup> System continuous flow analyser; nitrite was measured in a separate run. There were three plot treatment replicates and at least two experimental replicates per treatment.

### *Statistical Analyses*

Statistical analysis was performed using one-factor and general ANOVA in GenStat 19<sup>th</sup> Edition (VSN International Ltd., Hemel Hempstead, UK). To check that each set of measured values met the assumptions of ANOVA and were normally distributed, residuals were plotted. If they did not show normal distribution, data was log-transformed and again checked for normal distribution of residuals. Treatment comparisons with *F* statistics with  $P < 0.05$  were considered significant,  $P < 0.001$  highly significant. Means were compared using Tukey's *post-hoc* method in the GenStat multiple comparison menu with 95% confidence; means are considered significantly different at  $\alpha = 0.05$  and where appropriate, are represented by different letters.

## **Acknowledgements**

We thank Rodger White for help with statistical analyses and we are grateful to the DeepSoil consortium that included Illumina, Rothamsted Research and Janet Jansson at the Lawrence Berkeley National Laboratory/Pacific Northwest National Laboratory U.S.A for providing the metagenomic data and to TGAC for the metatranscriptome (CCC-108). Rothamsted Research receives strategic funding from the Biotechnology and Biological Research Council of the UK, this work was supported by BBS/E/C/00005196, BBS/E/C/00010310, and the CINAg collaborative grant with China BB/NO13468/1.

## **Author Contributions**

IMC and PRH conducted the initial experiment; QF and PRH measured soil nitrification, QF, MA, and IMC extracted soil DNA and RNA and performed qPCR; IMC, PRH, MA and DH analysed the data and produced the first draft, all contributed to the final manuscript.

## **Additional Information**

The authors declare no competing interests



## Figure legends

**Figure 1** Production of nitrate in response to 100 mg ammonium-N g<sup>-1</sup> dry soil added to microcosms incubated for 7 d with no inhibitors (solid lines) or 200 µM dicyandiamide (DCD) (dashed lines); standard error of difference of means ( $P < 0.001$ ) shown as bar.

**Figure 2** Mean relative abundance: A - DNA and B - mRNA reads g<sup>-1</sup> dry soil assigned to nitrifying guilds in each soil treatment. Standard errors of differences of means (s.e.d.) shown for all groups; different letters indicate significantly different means.

**Figure 3.** Mean proportion of DNA reads (left) and mRNA (right) reads assigned to the main nitrifying guilds shown in Figure 2, expressed as per 1000 of total. The proportions formed by the component groups of each guild are shown (ANOVA statistics of component groups are shown in supplementary tables 1 and 2). Overall standard errors of differences of means (s.e.d.) are shown for each guild; different letters indicate significantly different means for each guild in each treatment. A, B – NOB; B, C – AOA; E, F – AOB; G, H – anammox; I, J – comammox.

**Figure 4.** Mean proportion of A - DNA reads and B - mRNA *amoA* reads assigned to the AOA and AOB (including comammox). The proportion of the total formed by each sub-group identified is shown (the NS AOA could not be separated into *Nitrososphaera* and *Ca. Nitrosocosmicus*). ANOVA statistics of the component groups are shown in supplementary table 5.

## References

- 1 Spiertz, J. H. J. Nitrogen, sustainable agriculture and food security. A review. *Agron. Sustain. Dev.* **30**, 43-55, doi:10.1051/agro:2008064 (2010).
- 2 Kowalchuk, G. A. & Stephen, J. R. Ammonia-oxidizing bacteria: a model for molecular microbial ecology. *Annu. Rev. Microbiol.* **55**, 485-529, doi:10.1146/annurev.micro.55.1.485 (2001).
- 3 Zumft, W. G. Cell biology and molecular basis of denitrification. *Microbiol. Mol. Biol. Rev.* **61**, 533-616 (1997).
- 4 Gelfand, I. & Yakir, D. Influence of nitrite accumulation in association with seasonal patterns and mineralization of soil nitrogen in a semi-arid pine forest. *Soil Biol. Biochem.* **40**, 415-424, doi:10.1016/j.soilbio.2007.09.005 (2008).
- 5 Subbarao, G. V. *et al.* Scope and strategies for regulation of nitrification in agricultural systems-challenges and opportunities. *Crit. Rev. Plant Sci.* **25**, 303-335, doi:10.1080/07352680600794232 (2006).
- 6 Shen, T., Stieglmeier, M., Dai, J., Ulrich, T. & Schleper, C. Responses of the terrestrial ammonia-oxidizing archaeon *Ca. Nitrososphaera viennensis* and the ammonia-oxidizing bacterium *Nitrosospira multififormis* to nitrification inhibitors. *FEMS Microbiol. Lett.* **344**, 121-129, doi:10.1111/1574-6968.12164 (2013).
- 7 Prosser, J. I., Head, I. M. & Stein, L. Y. in *The Prokaryotes - Alphaproteobacteria and Betaproteobacteria* (ed DeLong Rosenberg E., E.F., Lory, S., Stackebrandt, E., Thompson, F.) 901-918 (Springer-Verlag, 2014).
- 8 Hayatsu, M. *et al.* An acid-tolerant ammonia-oxidizing gamma-proteobacterium from soil. *ISME J.* **11**, 1130-1141, doi:10.1038/ismej.2016.191 (2017).
- 9 Prosser, J. I. & Nicol, G. W. Archaeal and bacterial ammonia-oxidisers in soil: the quest for niche specialisation and differentiation. *Trends Microbiol.* **20**, 523-531, doi:10.1016/j.tim.2012.08.001 (2012).

- 10 Pester, M. *et al.* *amoA*-based consensus phylogeny of ammonia-oxidizing archaea and deep sequencing of *amoA* genes from soils of four different geographic regions. *Environ. Microbiol.* **14**, 525-539, doi:10.1111/j.1462-2920.2011.02666.x (2012).
- 11 Spang, A. *et al.* The genome of the ammonia-oxidizing *Candidatus Nitrososphaera gargensis*: insights into metabolic versatility and environmental adaptations. *Environ. Microbiol.* **14**, 3122-3145, doi:10.1111/j.1462-2920.2012.02893.x (2012).
- 12 Daims, H. *et al.* Complete nitrification by *Nitrospira* bacteria. *Nature* **528**, 504-509, doi:10.1038/nature16461 (2015).
- 13 van Kessel, M. A. *et al.* Complete nitrification by a single microorganism. *Nature* **528**, 555-559, doi:10.1038/nature16459 (2015).
- 14 Junier, P. *et al.* Phylogenetic and functional marker genes to study ammonia-oxidizing microorganisms (AOM) in the environment. *Appl. Microbiol. Biotechnol.* **85**, 425-440, doi:10.1007/s00253-009-2228-9 (2010).
- 15 Leininger, S. *et al.* Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* **442**, 806-809, doi:10.1038/nature04983 (2006).
- 16 Zhalnina, K. *et al.* *Ca. Nitrososphaera* and *Bradyrhizobium* are inversely correlated and related to agricultural practices in long-term field experiments. *Front. Microbiol.* **4**, 104, doi:10.3389/fmicb.2013.00104 (2013).
- 17 Pjevac, P. *et al.* *AmoA*-Targeted Polymerase Chain Reaction Primers for the Specific Detection and Quantification of Comammox *Nitrospira* in the Environment. *Front. Microbiol.* **8**, 1508, doi:10.3389/fmicb.2017.01508 (2017).
- 18 Palomo, A., Dechesne, A. & Smets, B. F. Genomic profiling of *Nitrospira* species reveals ecological success of comammox *Nitrospira*. *bioRxiv*, 612226, doi:10.1101/612226 (2019).
- 19 Poghosyan, L. *et al.* Metagenomic recovery of two distinct comammox *Nitrospira* from the terrestrial subsurface. *Environ. Microbiol.* **21**, 3627-3637, doi:10.1111/1462-2920.14691 (2019).

- 20 Palomo, A. *et al.* Comparative genomics sheds light on niche differentiation and the evolutionary history of comammox *Nitrospira*. *ISME J.* **12**, 1779-1793, doi:10.1038/s41396-018-0083-3 (2018).
- 21 Strous, M. *et al.* Deciphering the evolution and metabolism of an anammox bacterium from a community genome. *Nature* **440**, 790-794, doi:10.1038/nature04647 (2006).
- 22 De Boer, W. & Kowalchuk, G. A. Nitrification in acid soils: micro-organisms and mechanisms. *Soil Biol. Biochem.* **33**, 853-866, doi:10.1016/s0038-0717(00)00247-9 (2001).
- 23 Tourna, M. *et al.* *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 8420-8425, doi:10.1073/pnas.1013488108 (2011).
- 24 Arp, D. J., Chain, P. S. G. & Klotz, M. G. The impact of genome analyses on our understanding of ammonia-oxidizing bacteria. *Annu. Rev. Microbiol.* **61**, 503-528 (2007).
- 25 Simon, J. & Klotz, M. G. Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations. *Biochim. Biophys. Acta* **1827**, 114-135, doi:10.1016/j.bbabbio.2012.07.005 (2013).
- 26 Walker, C. B. *et al.* *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 8818-8823, doi:10.1073/pnas.0913533107 (2010).
- 27 Koch, H., van Kessel, M. A. H. J. & Lückner, S. Complete nitrification: insights into the ecophysiology of comammox *Nitrospira*. *Appl. Microbiol. Biotechnol.* **103**, 177-189, doi:10.1007/s00253-018-9486-3 (2019).
- 28 Cassman, N. A. *et al.* Nitrification inhibitors effectively target N<sub>2</sub>O-producing *Nitrososphaera* spp. in tropical soil. *Environm. Microbiol.* **21**, 1241-1254, doi:10.1111/1462-2920.14557 (2019).
- 29 Li, C. Y., Hu, H. W., Chen, Q. L., Chen, D. L. & He, J. Z. Comammox *Nitrospira* play an active role in nitrification of agricultural soils amended with nitrogen fertilizers. *Soil Biol. Biochem.* **138**, doi:10.1016/j.soilbio.2019.107609 (2019).

- 30 Daims, H., Lucker, S. & Wagner, M. A new perspective on microbes formerly known as nitrite-oxidizing bacteria. *Trends Microbiol.* **24**, 699-712, doi:10.1016/j.tim.2016.05.004 (2016).
- 31 Castelle, C. J. *et al.* Extraordinary phylogenetic diversity and metabolic versatility in aquifer sediment. *Nat. Commun.* **4**, 2120, doi:10.1038/ncomms3120 (2013).
- 32 Sorokin, D. Y. *et al.* Nitrification expanded: discovery, physiology and genomics of a nitrite-oxidizing bacterium from the phylum *Chloroflexi*. *ISME J.* **6**, 2245-2256, doi:10.1038/ismej.2012.70 (2012).
- 33 Lucker, S. *et al.* A *Nitrospira* metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13479-13484, doi:10.1073/pnas.1003860107 (2010).
- 34 Mendum, T. A., Sockett, R. E. & Hirsch, P. R. Use of molecular and isotopic techniques to monitor the response of autotrophic ammonia-oxidizing populations of the beta subdivision of the class Proteobacteria in arable soils to nitrogen fertilizer. *Appl. Environ. Microbiol.* **65**, 4155-4162 (1999).
- 35 Hirsch, P. R. *et al.* Soil resilience and recovery: rapid community responses to management changes. *Plant Soil* **412**, 283-297, doi:10.1007/s11104-016-3068-x (2017).
- 36 Hirsch, P. R., Mauchline, T. H. & Clark, I. M. Culture-independent molecular techniques for soil microbial ecology. *Soil Biol. Biochem.* **42**, 878-887, doi:10.1016/j.soilbio.2010.02.019 (2010).
- 37 Vetrovsky, T. & Baldrian, P. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One* **8**, e57923, doi:10.1371/journal.pone.0057923 (2013).
- 38 Bollmann, A., Schmidt, I., Saunders, A. M. & Nicolaisen, M. H. Influence of starvation on potential ammonia-oxidizing activity and *amoA* mRNA levels of *Nitrosospira briensis*. *Appl. Environ. Microbiol.* **71**, 1276-1282, doi:10.1128/aem.71.3.1276-1282.2005 (2005).

- 39 Hirsch, P. R. *et al.* Starving the soil of plant inputs for 50 years reduces abundance but not diversity of soil bacterial communities. *Soil Biology and Biochemistry* **41**, 2021-2024, doi:10.1016/j.soilbio.2009.07.011 (2009).
- 40 Clark, I. M., Buchkina, N., Jhurrea, D., Goulding, K. W. & Hirsch, P. R. Impacts of nitrogen application rates on the activity and diversity of denitrifying bacteria in the Broadbalk Wheat Experiment. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 1235-1244, doi:10.1098/rstb.2011.0314 (2012).
- 41 Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Meth.* **12**, 59-60 (2015).
- 42 Huson, D. H., Auch, A. F., Qi, J. & Schuster, S. C. MEGAN analysis of metagenomic data. *Genome Res.* **17**, 377-386, doi:10.1101/gr.5969107 (2007).
- 43 Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772-780, doi:10.1093/molbev/mst010 (2013).

Figure 1

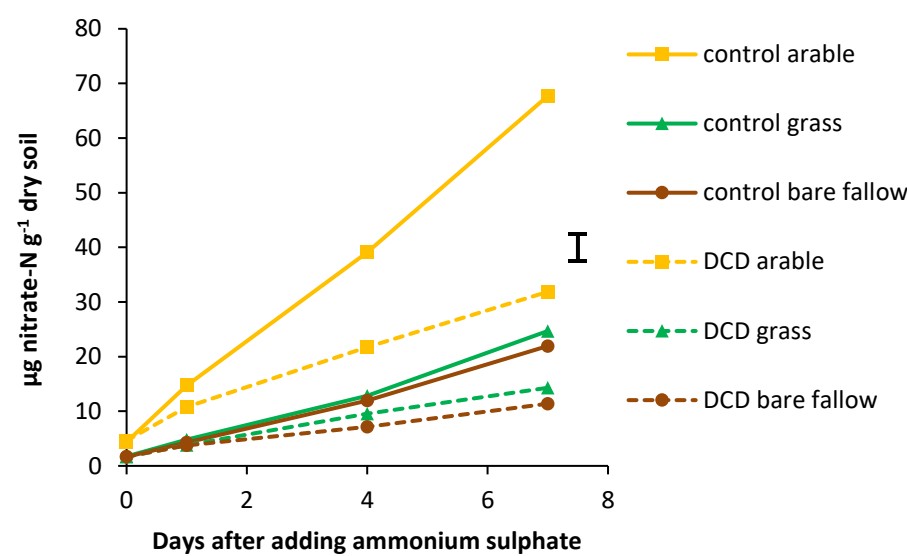


Figure 2

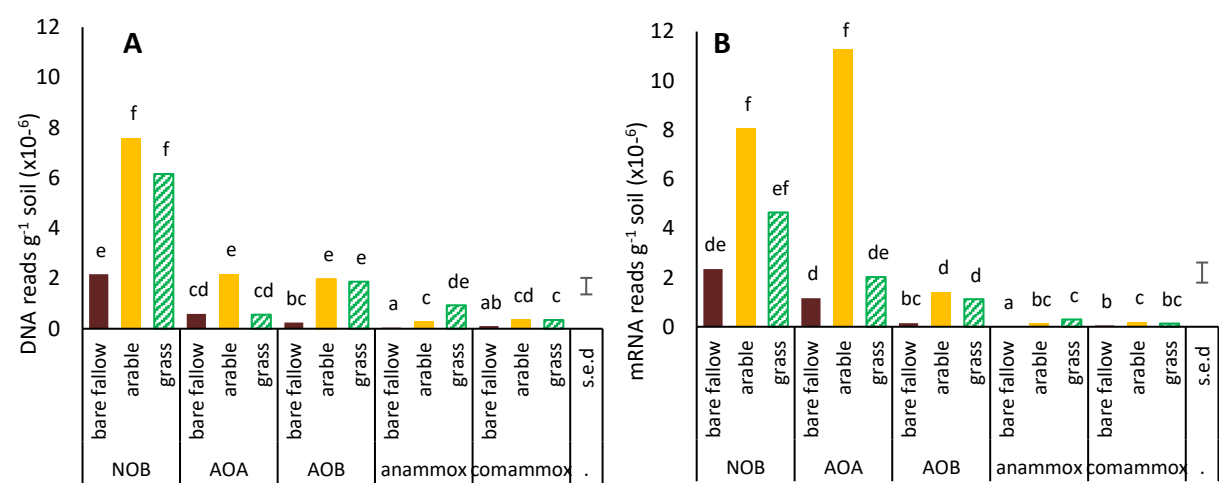


Figure 3

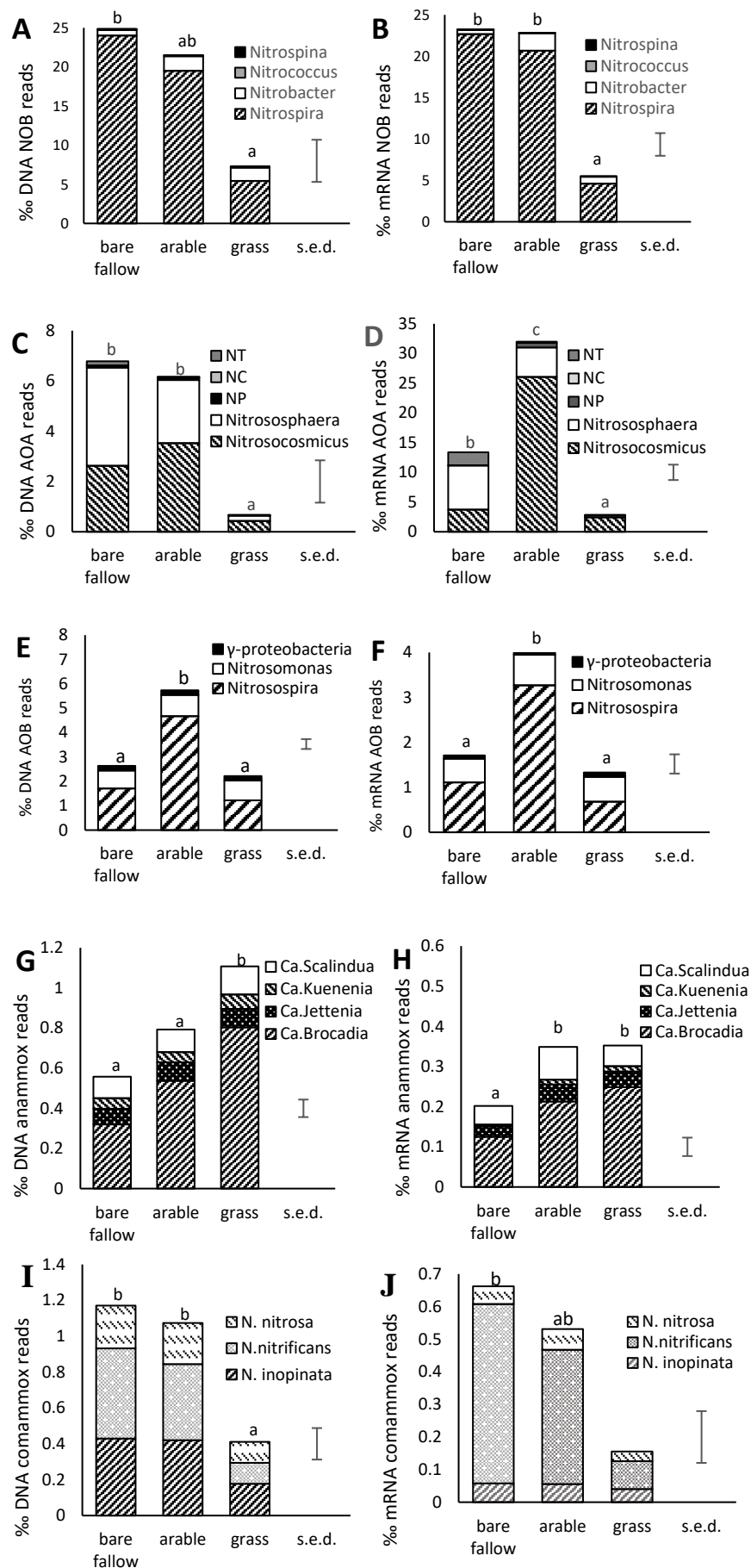




Figure 4

