environmental microbiology

Environmental Microbiology (2021) **23**(11), 6828–6843 doi:10.1111/1462-2920.15785

The influence of alfalfa-switchgrass intercropping on microbial community structure and function

Gyuhyon Cha,[1] **Kelley A. Meinhardt**,[2†] **Luis H. Orellana**,[1‡] **Janet K. Hatt**,[1] **Manmeet W. Pannu**,[2§] **David A. Stahl**[2] **and Konstantinos T. Konstantinidis**[1*]

[1]*School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332.*

[2]*Civil and Environmental Engineering, University of Washington, Seattle, WA 98195.*

Summary

The use of nitrogen fertilizer on bioenergy crops such as switchgrass results in increased costs, nitrogen leaching and emissions of N_2O, a potent greenhouse gas. Intercropping with nitrogen-fixing alfalfa has been proposed as an environmentally sustainable alternative, but the effects of synthetic fertilizer versus intercropping on soil microbial community functionality remain uncharacterized. We analysed 24 metagenomes from the upper soil layer of agricultural fields from Prosser, WA over two growing seasons and representing three agricultural practices: unfertilized switchgrass (control), fertilized switchgrass and switchgrass intercropped with alfalfa. The synthetic fertilization and intercropping did not result in major shifts of microbial community taxonomic and functional composition compared with the control plots, but a few significant changes were noted. Most notably, mycorrhizal fungi, ammonia-oxidizing archaea and bacteria increased in abundance with intercropping and fertilization. However, only betaproteobacterial ammonia-oxidizing bacteria abundance in fertilized plots significantly correlated to N_2O emission and companion qPCR data. Collectively, a short period of intercropping elicits minor but significant changes in the soil microbial community toward nitrogen preservation and that intercropping may be a viable alternative to synthetic fertilization.

Received 20 October, 2020; accepted 17 September, 2021. *For correspondence. E-mail kostas@ce.gatech.edu; Tel. 404-385-3628; Fax 404-894-8266. Present addresses: †Los Angeles Pacific University, San Dimas, CA, 91773. ‡Max-Planck-Institut für Marine Mikrobiologie, Celsiusstrasse 1, D-28359, Bremen, Germany. §Orange County Water District, Anaheim, CA, 92807

Introduction

As part of the continuous efforts to shift energy dependency from petroleum hydrocarbons to renewable sources, switchgrass (*Panicum virgatum* L.) has been studied for its potential as a biomass feedstock (i.e. bioenergy crop) for biofuel production to replace corn-based ethanol. Comparable biomass yield to corn, ability to grow on marginal lands and low nutrient and water requirements are the main reasons for the increased interest in switchgrass (McLaughlin and Kszos, 2005; Collins *et al*., 2010; Kimura *et al*., 2015). Several studies have reported increased biomass yields of switchgrass upon nitrogen fertilization (Madakadze *et al*., 1999; Lemus *et al*., 2008). However, the application of synthetic fertilizer contributes to increased nitrate (NO_3^-) leaching to groundwater and nitrous oxide (N_2O) emissions from soils (Mosier *et al*., 1991; Venterea *et al*., 2012). Notably, N_2O is a major greenhouse gas with higher warming potential than CO_2 and causes stratospheric ozone depletion (Dalal *et al*., 2003; Ravishankara *et al*., 2009).

Reactive ammonium fertilizer not assimilated by crops or soil microbes can be a major source of N_2O emissions as a byproduct of ammonia oxidation carried out by ammonia-oxidizing archaea (AOA) and ammonia-oxidizing bacteria (AOB) among other abiotic and biotic processes (Sanford *et al*., 2012). Due to the vast spatial heterogeneity and diversity of soil microbiota, the abundance and activity of the AOA and AOB varies, depending on site location, soil type, moisture content, soil pH, nutrient levels, as well as biotic factors such as species–species interactions and viral predation (Erguder *et al*., 2009; Pannu *et al*., 2019). This complexity makes it challenging to track the relative contributions of each microbial group to ammonia oxidation and N_2O emissions in different environments. Generally, the AOA populations are reported to be more abundant than AOB in soils (Leininger *et al*., 2006; Chen *et al*., 2008) but recent studies have reported a larger contribution by AOB to N_2O emission, especially in cases of excessive nitrogen fertilization (Hink *et al*., 2017; Hink *et al*., 2018; Meinhardt *et al*., 2018).

Intercropping – simultaneously growing at least two crop species in proximity – not only affects the crop yield but can also alter the structure and activity of the soil microbiome (Duchene *et al*., 2017; Yu *et al*., 2019). Intercropping of switchgrass with legumes like alfalfa (*Medicago sativa*), aiming to serve as a nitrogen-supplier to the system, has been shown to be effective in reducing N_2O emissions while still supporting comparable crop biomass yield relative to the application of synthetic fertilizers (Wright and Turhollow, 2010; Barton *et al*., 2011). On the other hand, a recent study of the same site as that used in our present study (Prosser, WA, USA) reported an increase in the relative abundance of AOA in intercrop plots (i.e. switchgrass and alfalfa) compared with monocrop fertilized plots, but this increase did not correlate to lower N_2O emissions (Pannu *et al*., 2019), making it challenging to interpret the observed increase in nitrogen and, more specifically, ammonia with relatively low N_2O emissions from intercrop plots. At another site, intercropping wheat (*Triticum aestium* L.) and soybean (*Glycin max* L.) were reported to change the dominant species of AOB and increase the expression of nitrogen cycle related functional genes in the rhizosphere (Wallenstein *et al*., 2006). Most previous studies were based on PCR-generated gene amplicons (Meinhardt *et al*., 2018; Pannu *et al*., 2019). Thus, it remains unclear if novel diversity was missed due to primer specificity to only known sequences used for primer design (Meinhardt *et al*., 2015). The combined investigation of soil microbiomes and crop management strategies can provide new insights into the questions described above and expand current understanding of the responses of soil microbial communities to synthetic fertilization and intercropping treatments.

In the present study, using shotgun-metagenomic sequencing over two growing seasons in 2012 and 2014, we evaluated the responses of the soil microbial communities in switchgrass fields in Prosser, WA, USA to the nitrogen supply from either synthetic fertilizer or intercropping with alfalfa, focusing primarily on the ammonia-oxidizing populations and N_2O emissions. Our main goal was to assess whether intercropping switchgrass with nitrogen-fixing alfalfa can be a viable and environmentally friendly alternative to applying synthetic fertilizer. Specifically, and based on the previously reported lower N_2O emissions from the intercropped plots compared with the fertilized plots and lack of correlation of emissions to AOA abundance, we hypothesized that the shifts in microbial community structure in response to intercropping, especially in AOB abundance and as captured by shotgun metagenomic sequencing, would show significant correlation with N_2O emissions. Further, we expected to capture novel gene and genome diversity of ammonia oxidizers within our metagenomes that may have been missed previously due to issues related to PCR primer specificity. Finally, we hypothesized that the presence of alfalfa elicits significant changes in the functional gene content of the corresponding microbial communities that will be discoverable by our metagenomic effort, despite the relatively short period (one growing season) since the onset of intercropping.

Results and discussion

Treatment description and physiochemical measurements

Switchgrass was established in June 2010 at the Prosser field site (silt loam) in Eastern Washington State, in control (i.e. unfertilized) and fertilized plots. Fertilized plots received synthetic nitrogen–phosphorus–potassium (NPK) fertilizer (16–16–16) twice a year, in April and July (total agronomic rate of 224 kg ha^{-1} $year^{-1}$) starting in the year 2011. For intercropped plots, alfalfa (*Medicago sativa*) was planted in open rows between planted switchgrass rows in September 2010 with the ratio of switchgrass:alfalfa to be 70:30. The final experimental design included 3 $\times$ 6 m plots in a randomized complete block with four replicates per treatment. Our first soil sampling took place in April 2012, which was 1 and 1.5 years after the first synthetic fertilization and the establishment of intercropped plots respectively (Tables S1 and S2). For comparative purposes, previously determined metagenomes from various representative soils (0–15 cm depth) across the USA were included in our analyses (Table S1). These included the native tallgrass prairie soils (NTP) (Mackelprang *et al*., 2018), soils from the national Critical Zone Observatories (CZOs) (Brewer *et al*., 2019) and agricultural soils from Illinois (Orellana *et al*., 2018).

Differences in the soil chemical parameters were observed among the three treatments at the Prosser site based on 2012–2013 samples (Table S1). Ammonium and nitrate levels were highest in fertilized plots, as expected (5.963 $\pm$ 9.037 NH_4-N mg kg^{-1} and 7.355 $\pm$ 13.40 NO_3-N mg kg^{-1}), compared with controls (1.528 $\pm$ 1.020 NH_4-N mg kg^{-1} and 1.192 $\pm$ 1.682 NO_3-N mg kg^{-1}) and intercropped plots (1.010 $\pm$ 0.3608 NH_4-N mg kg^{-1} and 2.533 $\pm$ 0.9534 NO_3-N mg kg^{-1}) (Fig. 1).

It is noteworthy that both ammonium and nitrate concentrations in the fertilized plot samples have larger SD values compared with other treatment plots from the fertilized samples from July 2013. This difference could be explained by the soil heterogeneity and the lack of (additional) replicate measurements (please refer to the *Experimental procedures* section for further details on replicate sampling). To test whether the differences of ammonium and nitrate concentrations between three

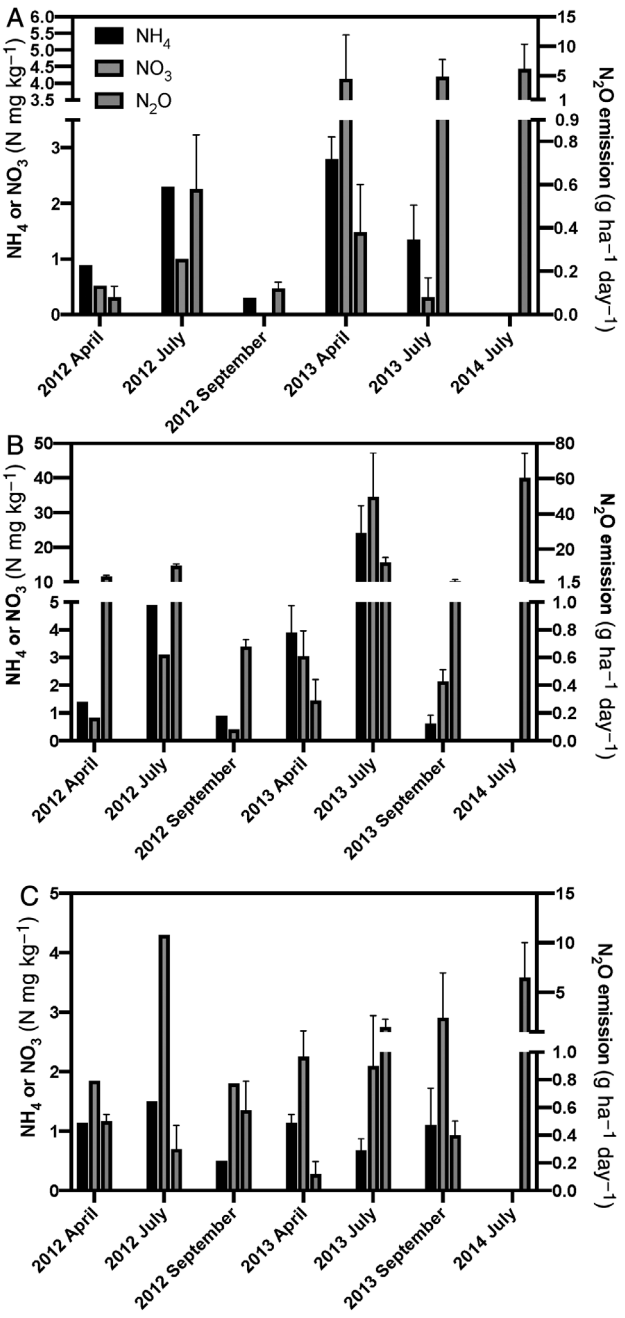

Fig. 1. Soil ammonium/nitrate concentrations and nitrous oxide emission from (A) unfertilized control, (B) fertilized and (C) intercropped plots. Note the absence of ammonium and nitrate concentrations from July 2014 samples.

treatment types were significant, one-way ANOVA was applied followed by multiple comparisons using Tukey's test. Both ammonium and nitrate concentrations between treatments were not statistically significant (ANOVA, $R^2 = 0.1742$, $p = 0.2619$/$R^2 = 0.1156$, $p = 0.4233$ respectively).

Fertilized plots also emitted the largest amounts of nitrous oxide (N_2O) gas in 2012–2014 with an average of 12.99 ± 8.107 g ha^{-1} day^{-1}, which was 6.4 and 9.2-fold higher than the control (2.025 ± 1.113 g ha^{-1} day^{-1}) and intercropped plots (1.414 ± 0.8653 g ha^{-1} day^{-1}) respectively. Nonetheless, the differences of N_2O emissions between three treatment plots were not statistically significant, presumably due to large variation among measurements (one-way ANOVA followed by Tukey's test: $R^2 = 0.1705$, $p = 0.2042$). Soil pH was alkaline in all Prosser plots, with an average around pH 8.3, which was significantly higher than the surface soils from NTP, CZOs and Illinois, which averaged 6.6–6.8 (Table S1). The pH differences between three treatment types were not statistically significant (one-way ANOVA, $R^2 = 0.05961$, $p = 0.7821$). Notably, agricultural soils in Illinois with a long history of fertilization had sevenfold and sixfold higher ammonium and nitrate levels, respectively, compared with the Prosser soils.

The yields of dry matter biomass from the control, fertilized and intercropped plots are shown in Table S3 as a proxy of switchgrass productivity. After machine-harvesting of the aboveground biomass twice per year in 2012 and 2013 from the centre of each plot, the dry biomass was determined using 0.5–1.0 kg subsamples (no such data were collected for 2014). In both 2012 and 2013, the average biomass yields from the fertilized and intercropped plots were higher than the control plots. Notably, the yield from the intercropped plots in 2013 did not significantly differ from that of the fertilized plots ($p = 0.06$). Considering that those differences were significant in the previous year 2012, the increase of dry mass yield from the intercropped plots could be a result of the gradual development of interconnections between the root systems of switchgrass and alfalfa. Therefore, intercropping has good potential as an environmentally friendly alternative to fertilizer addition considering the lower N_2O emissions, which we investigated more fully using shotgun metagenomics.

Shotgun metagenome statistics

In the Prosser soil dataset, 23 to 42 million reads per metagenome were obtained from a total of 24 samples, and the average length of the short-reads after trimming was 129 ± 3 bp (Table S2). The Chao-1 index of rarefied datasets to the lowest coverage observed revealed about 2000–2500 16S rRNA gene-based operational taxonomic units (or OTUs), defined at a 97% level for nucleotide sequence identity, in each of the Prosser datasets. The Prosser and Illinois datasets had higher species richness compared with the majority of the CZOs datasets (Fig. S1A) according to the Chao-1 index based on 16S rRNA gene OTUs (nonparametric Student's *t*-test using Monte Carlo permutations, $p < 0.05$); however, the differences among other datasets (e.g. Prosser vs. Illinois or

Prosser vs. NTP) were not statistically significant. In terms of intra-sample diversities, the Nonpareil diversity (N_d) (Rodriguez-R *et al*., 2018) and Shannon diversity (H') of all datasets were 22–24 and 8.0–9.0, respectively (Fig. S1B); these values are within the range typically observed for highly diverse soil samples. The NTP dataset had the highest sequencing effort which was in line with the highest average coverage of the sampled microbial community by sequencing (~50%) as estimated by Nonpareil. The other three datasets, including the Prosser metagenomes, resulted in lower Nonpareil-estimated coverage values of around 20%. As noted previously (Rodriguez-R and Konstantinidis, 2014), comparing the abundance of features (taxa, pathways or genes) between samples is robust as long as the Nonpareil coverage between the datasets does not differ more than twofold; higher than twofold difference in coverage could result in substantial false positive signal. The coverage values obtained did not violate these assumptions; hence, our comparison between datasets was robust.

Microbial community structure and diversity shifts

A clear separation of the four metagenomic dataset series from the different studies analysed here was observed using kmer-based, whole-community Mash pairwise distances (Ondov *et al*., 2016) (permutational multivariate ANOVA [Adonis] and analysis of similarities [Anosim] $p = 0.001$) (Fig. 2A). The Bray–Curtis dissimilarity plot constructed using the 16S rRNA and fungal 18S rRNA gene-based OTUs also showed a distinct separation of the Prosser metagenomes from those of other sites. Furthermore, the separation between three treatments within the Prosser dataset was statistically significant (Anosim static $R = 0.3735$, $p = 0.001$ in Table S4) based on 16S rRNA gene-based OTUs (Fig. 2B); however, the difference between treatments based on 18S rRNA gene-based OTUs (Fig. 2C) was not significant (Anosim static $R = -0.08163$, $p = 0.821$).

Focusing on only beta diversity within the Prosser datasets enabled us to elucidate the small but significant differences between them (Fig. 1, plots on the right side). The intercropped switchgrass and alfalfa samples were clearly separated from the control and fertilized samples when using total metagenomic sequence diversity (Mash) or distances based on 16S rRNA gene-based taxonomic profiles. In fact, the separation between control and fertilized samples by the 16S and 18S rRNA gene taxonomic profiles (Adonis and Anosim, $p = 0.001$) were clearer than by Mash distances (Table 1). Further, the highest Adonis R^2 and Anosim static R value 'treatment type' compared with 'sampled year' and 'month' revealed a stronger effect of the former in shaping the 16S rRNA taxonomic composition than the latter (Table S5). These findings indicated that a large portion of the microbial community was mostly unaffected by seasonal changes, consistent with other previously studied year-round soil community data (Lennon and Jones, 2011; Orellana *et al*., 2018), and that the treatment caused discoverable changes to the corresponding microbial communities, which we investigated further below.

Taxonomic composition shifts

Best-match analysis of protein-coding genes against the TrEMBL database showed that the Prosser datasets were dominated by bacteria (> 98%) (Fig. 3C). Differences in the relative abundance of several abundant bacterial phyla (i.e. ≥ 0.1% of the total, on average, across all 24 Prosser datasets) were observed between intercropped or fertilized metagenomes and the control (Fig. 3A). None of the bacterial phyla had significant difference between treatments (one-way ANOVA $p > 0.05$). In general, the relative abundances of the four major archaeal phyla were similar in all treatment types, and *Thaumarchaeota* and *Euryarchaeota* were relatively more abundant than other phyla. Compared with the control samples, both fertilization and intercropping resulted in the decrease of *Thaumarchaeota* relative abundance. The relative abundance of *Euryarchaeota* decreased by 19.4% with fertilization but increased 47.6% with intercropping ($p < 0.05$).

Intercropping switchgrass with alfalfa significantly shifted the composition of the fungal community compared with control switchgrass (Fig. 4 and Fig. S2). Between 1% and 5% of all recovered small subunit rRNA gene fragments corresponded to fungi depending on the specific sample considered. Intercropping with alfalfa was expected to increase the abundance of arbuscular mycorrhizal (AM) fungi, which are known to be dominant in grassy plants and can serve as nitrogen scavengers from alfalfa to switchgrass (Govindarajulu *et al*., 2005). However, only one among the top 13 relatively abundant fungal 18S rRNA gene OTUs was taxonomically assigned to AM fungi as *Glomeromycetes*. Notably, the ectomycorrhizal (ECM) fungi of the *Agaricomycotina* subphylum, typically associated with woody trees and not with grassy plants (Agerer, 2001; Smith and Read, 2010), were more abundant in intercropped samples compared with the two other treatments. Fertilization caused a decrease in the relative abundance of mycorrhizal fungi compared with control, but intercropping with alfalfa led to a 2.5-fold increase, mainly represented by *Agaricomycotina*. It is important to note that our soil samples originated from bulk soil as opposed to the rhizosphere; hence, the increase of ECM fungal abundance as a response to intercropping represented a somewhat unexpected finding that

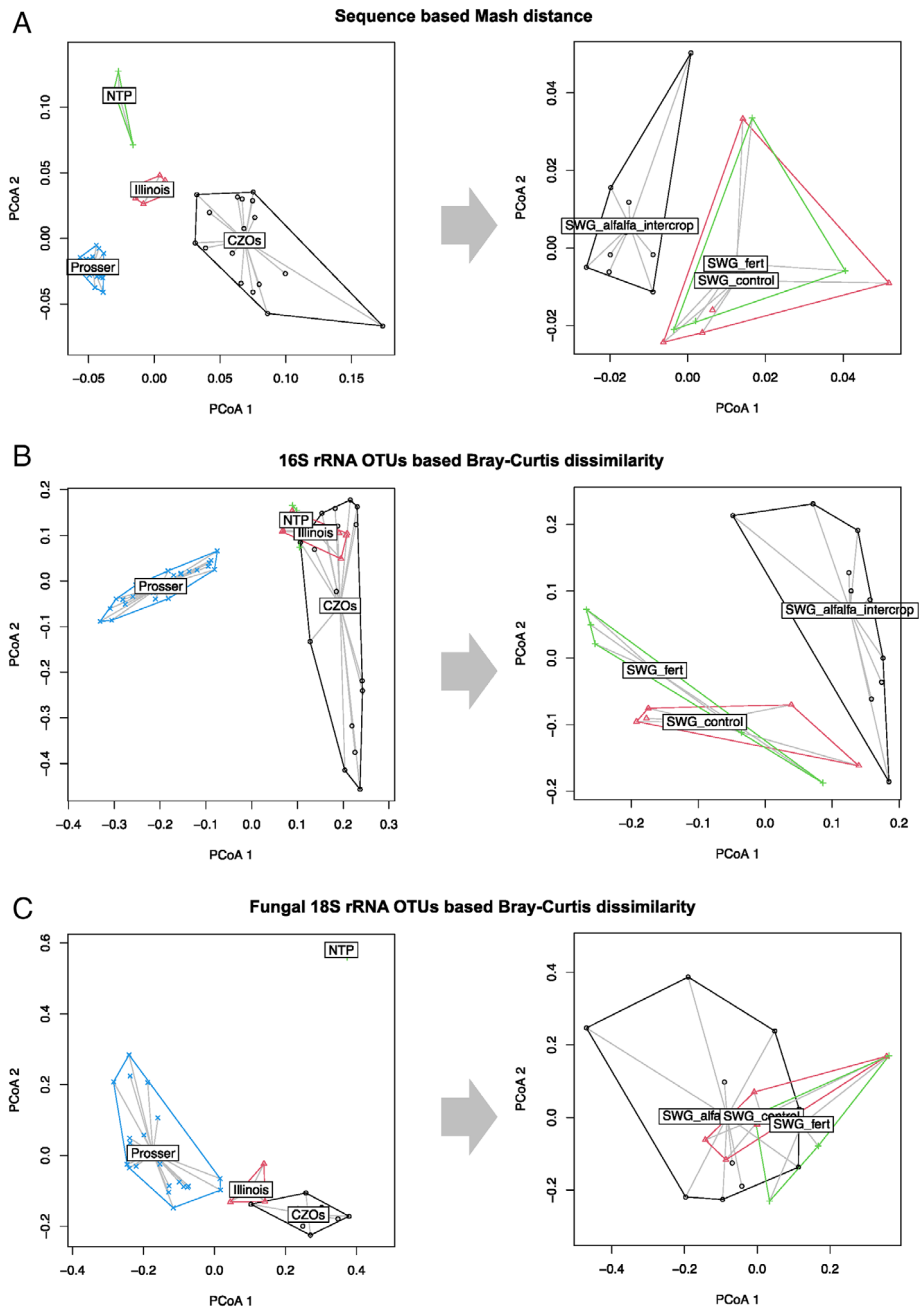

Fig. 2. Distinct clustering of surface-soil samples depending on ecosystem of origin and treatment types. Principal coordinate analysis (PCoA) plot of (A) a Mash distance matrix generated based on quality-trimmed metagenomes (only forward reads were used). A Mash sketch size of 100 000 was used. (B) 16S (bacterial) and (C) 18S (fungal) rRNA gene-based OTU community structure based on a Bray–Curtis distance matrix of rRNA gene-encoding fragments recovered with SortMeRNA and processed in the QIIME software package, as described in the *Experimental procedures*. One centroid per cluster is retained. Statistical significance of each clustering is given in Table S4.

should be further investigated in the future. Overall, *Plectosphaerellaceae*, a family encompassing both saprobes and plant pathogens, was the most abundant fungal family in all Prosser soil samples, with an average relative abundance of 18.4 $\pm$ 7.2% of total fungal 18S rRNA gene sequences, and did not shift in abundance with treatment.

Functional gene shifts

In addition to the shifts in microbial community taxonomic structure mentioned above, 14 Gene Ontology (GO) pathways were $\geq$ 10% more abundant in intercropped soil populations than those of other treatments, and 6 pathways were significantly different (Wald test, $p < 0.05$) (Fig. 5B). These pathways included catabolic processes for (i) adenosine, (ii) sulfur compounds and (iii) nucleotides; (iv) metabolic processes of phosphatidylinositol, purine ribonucleoside monophosphate, inosine, mycothiol and mannitol; (v) nitrogen compound transport and (vi) viral entry into its host cell. In particular, the nitrogen transport pathway involves movement of nitrogen-containing compounds in/out of the cell by a transporter or porin, and thus the increase in relative

Table 1. The effects of season and nitrogen availability on soil microbial community structure.

		Anosim			Adonis		
Underlying data	Distance metric	Treatment	Year	Month	NH_4	NO_3	N_2O
Entire dataset							
k-mer composition	Mash	**0.003**	**0.032**	**0.035**	**0.047**	**0.063**	0.191
16S-based OTUs	Bray–Curtis	**0.001**	**0.029**	**0.054**	0.550	0.969	**0.045**
18S-based OTUs	Bray–Curtis	**0.001**	0.882	0.153	0.504	0.993	0.114
Fungal 18S-based OTUs	Bray–Curtis	0.502	0.247	0.659	0.621	0.834	0.439
Control							
k-mer composition	Mash	-	0.400	0.467	0.100	**0.008**	0.767
16S-based OTUs	Bray–Curtis	-	0.200	0.267	0.525	**0.075**	0.458
18S-based OTUs	Bray–Curtis	-	0.800	0.267	0.450	**0.083**	0.717
Fungal 18S-based OTUs	Bray–Curtis	-	0.500	0.867	0.167	**0.025**	0.342
Fertilized							
k-mer composition	Mash	-	0.400	0.133	**0.022**	**0.009**	0.596
16S-based OTUs	Bray–Curtis	-	0.100	1.000	0.585	0.589	0.786
18S-based OTUs	Bray–Curtis	-	0.200	0.633	0.187	0.406	0.367
Fungal 18S-based OTUs	Bray–Curtis	-	1.000	**0.067**	0.176	0.206	0.475
Intercrop							
k-mer composition	Mash	-	0.112	**0.030**	0.494	0.731	0.635
16S-based OTUs	Bray–Curtis	-	**0.034**	**0.045**	0.195	0.666	0.519
18S-based OTUs	Bray–Curtis	-	**0.086**	0.219	0.971	0.927	0.732
Fungal 18S-based OTUs	Bray–Curtis	-	0.297	0.808	0.758	0.861	0.360

Analysis of similarity (Anosim) was used to test the seasonality effect, and permutational multivariate ANOVA (Adonis) was used to test the effect of nitrogen availability (e.g. NH_4, NO_3 and N_2O) on each community metric. Numerical values represent probability scores (i.e. *p*-values) resulting from each test. $p < 0.1$ are in bold to highlight significant and near-significant differences.

abundance could be related to higher nitrogen availability due to intercropping. Indeed, the higher abundance of this specific pathway appeared to be attributable to the higher abundance of nitrate/nitrite transporter (*narK*) and ammonium transporter (*amt*) genes in the intercrop soil populations compared with control. Meanwhile, the mycothiol (MSH) biosynthetic process pathway was highly abundant in both intercrop and switchgrass monocrop populations. MSH is known to be specific to *Actinomycetales* and functions as a reserve of cysteine for the detoxification of alkylating agents, reactive oxygen and nitrogen species and antibiotics.

No significant difference in the relative abundance of the nitrogen-cycle related GO biological pathways was noted overall among treatments (Wald test, *p*-value and Bonferroni adj. $p > 0.05$) (Fig. 4A). However, several pathways were more abundant in either fertilized or intercropped samples relative to the control. Namely, the 'nitrogen utilization' related pathways were more abundant in both fertilized and intercropped soil populations compared with control, which is likely due to the additional input of reactive nitrogen. 'Anammox' and 'urea metabolic processes' were detected and were slightly more abundant in intercropped soil populations than control (log2-fold difference ranged between 0.2 and 0.3). The relative abundance of 'Denitrification cycle', 'nitrogen fixation' and 'nitrate assimilation' pathways in fertilized and intercrop soils did not significantly differ relative to the control (log2-fold difference < 0.1). Therefore, it appears that the relative abundance of the N-cycle related functional potential was mostly similar among the three treatments, revealing a relatively limited response to additional nitrogen inputs. Overall, the taxonomic and functional composition of the Prosser soils did not show dramatic shifts at the DNA level as an effect of different nitrogen management strategies, presumably due to the relatively short history of the treatments and the slow growth kinetics of microbes in bulk soil. In fact, the first soil sampling took place 1 and 1.5 years after the first synthetic fertilizer application and intercrop establishment respectively. The relatively narrow shifts observed are in line with the low seasonal variation observed in the Illinois soils in metagenomes collected across a year (Orellana *et al*., 2018).

The effect of fertilization and intercropping on ammonia-oxidizing populations

We specifically examined the effect of fertilization and intercropping on ammonia-oxidizing communities by counting the metagenomic reads carrying the gene of interest. Archaeal *amoA* genes were found to be ~4-fold more abundant, on average, than their bacterial counterparts in all soil samples (Fig. 6A), which matches with results of a previous study of the same site (Pannu *et al*., 2019). To evaluate whether or not the AOA and AOB abundances increased with time for each treatment, linear regression

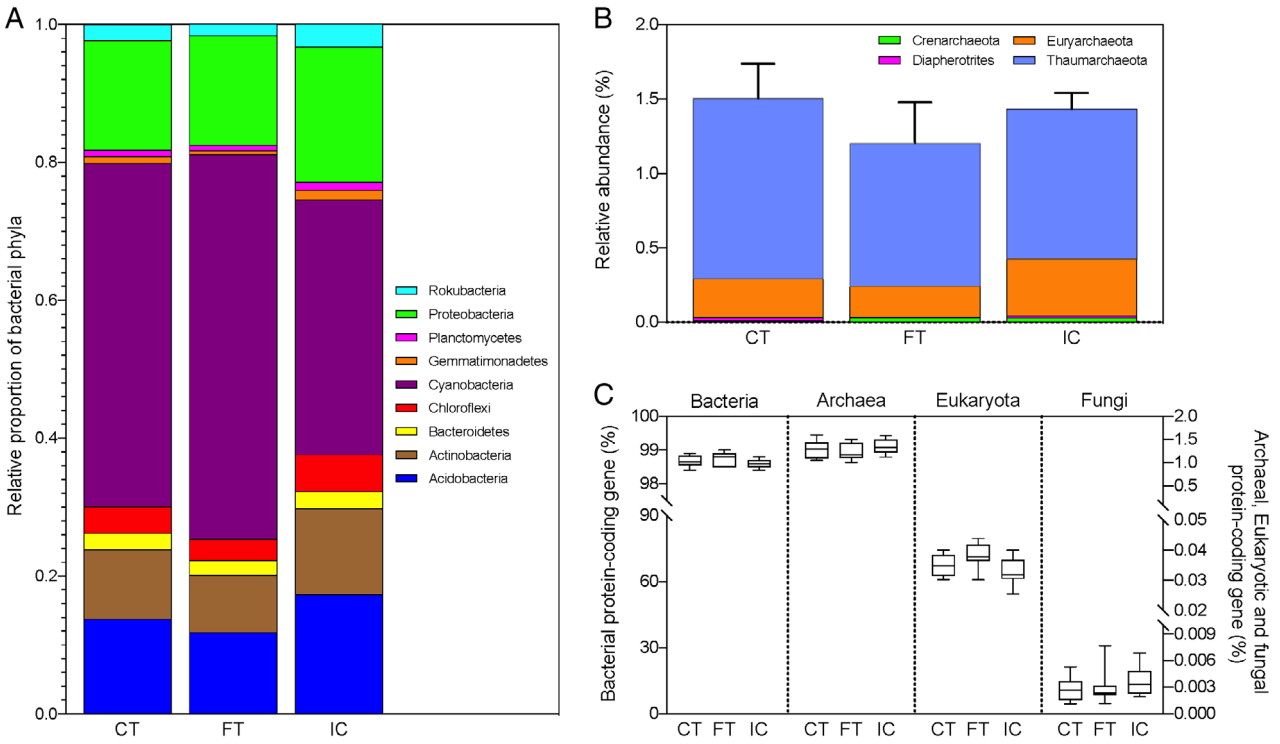

Fig. 3. Taxonomic composition of control, fertilized and intercrop sample groups.
A. The mean relative abundance of bacterial phyla for each treatment type based on 16S rRNA gene-carrying reads from the metagenomic datasets. Values represent the abundance of each bacterial phylum as a proportion of the total bacterial community. Only phyla with a mean relative abundance higher than 0.1% across all 24 datasets are displayed.
B. The mean relative abundance of archaeal phyla for each treatment type. Values represent the abundances (as a percentage) of each archaeal phylum relative to the total abundance of all recovered bacterial and archaeal 16S rRNA gene-carrying reads (i.e. the total prokaryotic community). Error bars represent the mean ± standard error of the mean (SEM) ($n = 6$, 7 and 11 of control, fertilized and intercrop, respectively) for cumulative (total) relative *Archaea* abundance.
C. Percent of the three domains and fungi in three treatment types. CT, FT and IC are control, fertilized and intercrop respectively. The best-matching taxonomy of protein-coding genes against TrEMBL database downloaded in February 2020 was used.

analysis was performed, followed by the analysis of covariance (ANCOVA) to assess the statistical significance of the differences between slopes of fitted lines (Fig. S7). In case of AOA, the slope of the control samples had a negative value of −7.32e−0.005; in contrast, fertilized and intercropped samples had positive slopes at 7.99e−0.005 and 9.91e−4 respectively. Especially, the intercropped samples showed a robust linear fit ($R^2 = 0.86$) with a significant non-zero slope ($p = 0.0081$). Further, the differences between slopes of control and intercropped were significantly different ($p = 0.0385$) based on ANCOVA, which was not the case between control and fertilized ($p = 0.7635$). Notably, the genome equivalent values (i.e. what fraction of the total cells sampled carry the gene of interest) of archaeal *amoA* genes showed a continuous increase through the entire sampling (27 months) in response to the intercropping treatment, which resulted in a ∼2.7-fold higher *amoA* abundance in July 2014 relative to April 2012. In contrast, the abundance of archaeal *amoA*-carrying reads of control and fertilized soil populations were rather stable across time. On the other hand, bacterial *amoA* genes in the fertilized plot showed a seasonal shift; increasing from April to July and decreasing from July to September. Notably, this abundance pattern matched with the fertilization timepoints (i.e. April and July). A similar trend was observed in intercropped soil samples having a peak in bacterial *amoA* abundance (genome equivalents) in July with slightly lower abundance than the populations of the fertilized plots ($R^2 = 0.4845$, $p = 0.0412$).

Contrary to the lack of statistically significant correlation between the archaeal *amoA* abundance and the level of reactive nitrogen (i.e. ammonium and nitrate) in all three treatment plots, a significant positive correlation between bacterial *amoA* and both ammonium ($R^2 = 0.78$, $p < 0.05$) and nitrate ($R^2 = 0.56$, $p < 0.05$) was revealed for fertilized soil populations (Fig. S8), consistent with our expectations and previous report on Prosser soils (Pannu *et al*., 2019). Fertilized plots emitted the highest average N_2O flux as 12.99 ± 8.107 g ha^{-1} day^{-1} among all three treatments,

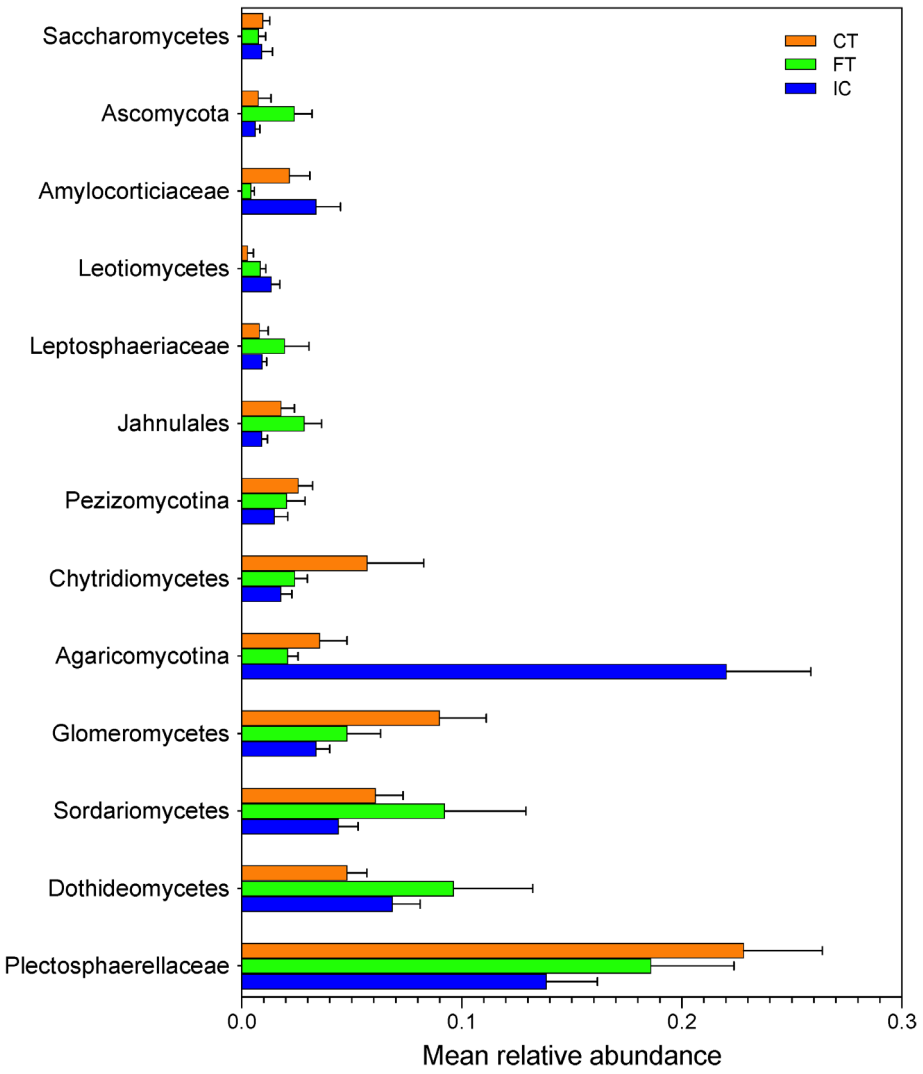

Fig. 4. The mean relative abundance of top 13 fungal taxa in each treatment type. Values represent the mean abundance of fungal taxa identified by taxonomic annotation of the total fungal 18S rRNA gene-carrying reads against the SILVA 18S database v.132 for each treatment. Error bars represent the mean ± SEM ($n = 6$, 7 and 11 of control, fertilized and intercrop, respectively) for cumulative (total) relative abundance.

which represented 6.4 and 9.2-fold higher emissions than control and intercropped plots respectively. N_2O fluxes measured from all Prosser soil samples had statistically significant correlations with bacterial *amoA* genome equivalents ($R^2 = 0.19$, $p < 0.05$) but not with archaeal *amoA* ($p = 0.3616$) or *nosZ* ($p = 0.4659$) (Fig. S5). Collectively, these results indicated that AOB could be a major contributor to N_2O emissions, consistent with previous reports (Meinhardt *et al*., 2018; Pannu *et al*., 2019).

Taxonomic affiliation of responding ammonia-oxidizing taxa

The phylogenetic placement of the recovered short-reads carrying archaeal or bacterial *amoA* gene on the phylogenetic trees of known reference AmoA protein sequences as terminal nodes allowed the taxonomic assessment of ammonia oxidizers in each treatment (Fig. 6B and C). In all three treatment plots, the majority of the archaeal *amoA* reads (77.6–85.3% of the total, depending on the sample considered) were placed within the I.1b clade. Compared with other treatments, the fertilized plot had the largest number of archaeal *amoA* reads placed in the I.1a clade (22.4%) and the smallest number placed in the I.1b clade (77.6%); control and intercropped plots showed a similar distribution of reads to clades I.1b and I.1a. The phylogenetic placement of the bacterial *amoA* reads further corroborated the abovementioned finding of more dramatic shifts compared with the archaeal counterparts. For instance, datasets from fertilized plots showed a decreased number of bacterial *amoA* reads placed within the soil *Nitrospira* (phylum *Nitrospirae*) clade (from 52.6% to 38.6%) and an increase in reads placed within the *Betaproteobacteria* class (from 36.6% to 50.4%) compared with the control plots.

Notably, the lower abundance of commamox *Nitrospira* amoA genes under fertilized versus unfertilized (control) conditions corroborated previous finding that commamox

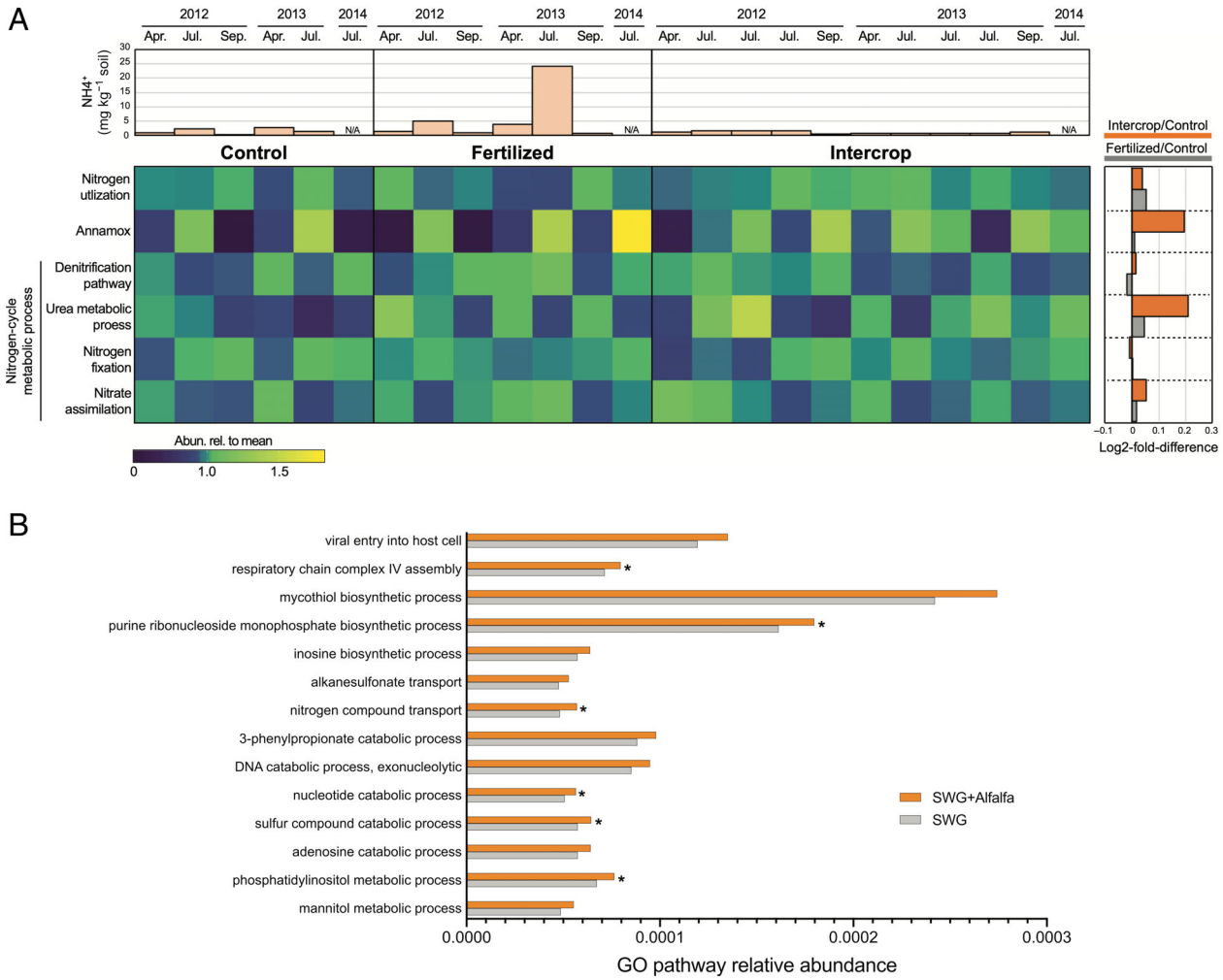

Fig. 5. The effect of fertilizer application and intercropping on functional gene diversity.
A. The relative abundance of nitrogen-cycle related Gene Ontology (GO) terms. For a detailed description of GO terms please see *Experimental procedures*. To emphasize differences in gene abundances between datasets, the abundance values for each gene were divided by the mean relative abundance of each gene in all 24 soil metagenomes (by rows). Barplots to the right of the heatmap represent the log2-fold-difference of the normalized gene abundance values between control and intercrop plots (in orange) or between control and fertilized plots (in grey) calculated for each gene. Barplots above the heatmap represents the NH_4^+ concentration for each sample.
B. A comparison of the mean GO pathway abundance in intercrop soils relative to the control soils. The relative abundance of different pathways was determined by matching Swiss-Prot references to GO Biological Process terms and dividing the number of annotations for each term by the total number of Swiss-Prot annotations found. Only pathways that have mean relative abundance > 0.005% and differed between control and intercrop plots by ≥ 10% are shown. Bars marked with *denote a significant difference (Wald test, $p < 0.05$).

Nitrospira are typically adapted to living in oligotrophic conditions (Kits *et al*., 2017; Sakoula *et al*., 2021), and implied that these organisms may largely contribute to nitrification in non-fertilized soils. Nevertheless, a few recent studies have reported positive correlations between high nitrogen inputs and the abundance of comammox *Nitrospira* in the soils (Li *et al*., 2019; Wang *et al*., 2019), which could indicate that terrestrial comammox *Nitrospira* may possess a broad range of ecological niches in terms of substrate availability. In our metagenomes, the Comammox *Nitrospira* represented 9.40% (genome equivalent) of total bacterial communities (Fig. S10), and were mostly (98.4%) assigned to the soil *Nitrospira* HD021 clade recovered from an Illinois agricultural soil previously (Orellana *et al*., 2018). MAG HD021 is a close relative of Ca. *Nitrospira inopinata* (66.3% average amino-acid identity [AAI]) (Orellana *et al*., 2018), known to have high affinity for ammonia and the low maximum rate of ammonia oxidation (Kits *et al*., 2017). Due to the lack of cultured representatives, the relative contributions and adaptive mechanisms of ammonia oxidizers, including the recently discovered commamox *Nitrospira* groups, to fertilization and intercropping are yet not well understood.

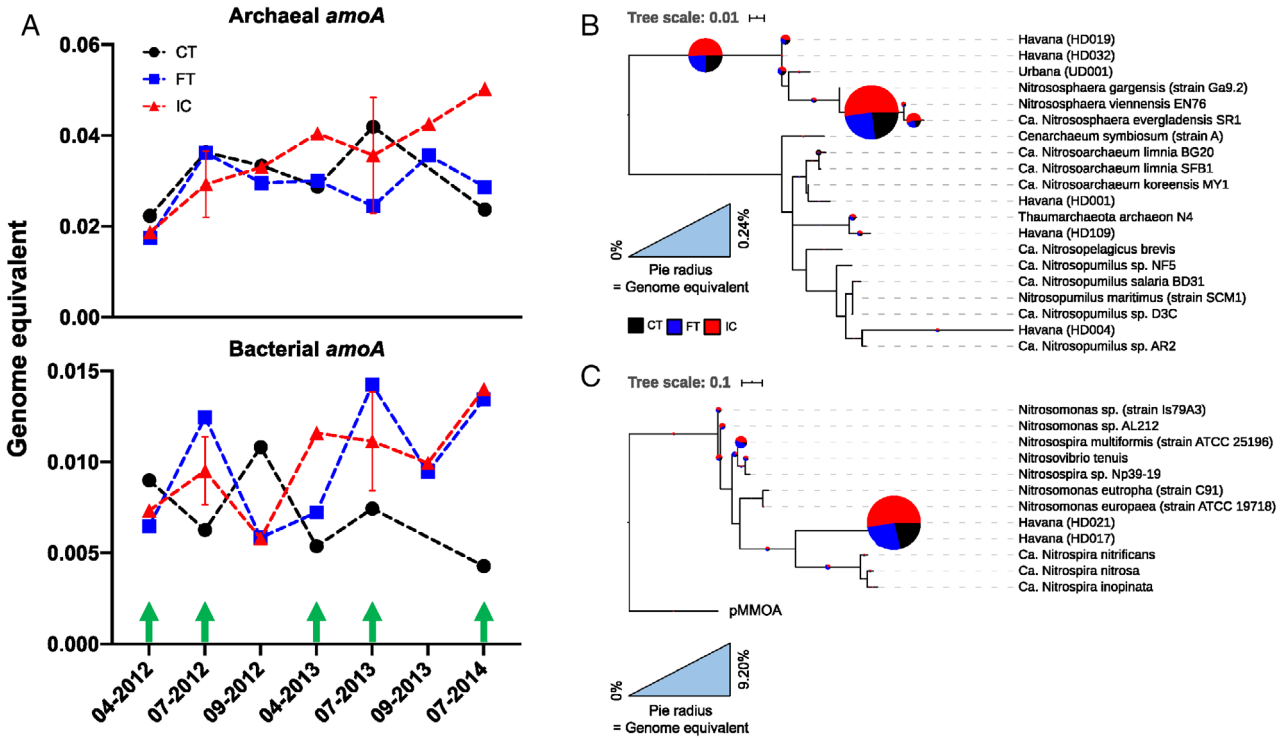

Fig. 6. Abundance and phylogenetic diversity of ammonia oxidizers. A. Seasonal abundance shifts, expressed as genome equivalents, of archaeal and bacterial *amoA* genes. Note that a synthetic fertilizer was applied to fertilized plots (FT) twice per year in April and July (green arrows). Phylogenetic reconstruction of archaeal (B) and bacterial (C) AmoA protein sequences recovered from contigs. The pie charts represent the number of recovered *amoA*-carrying reads (expressed as genome equivalents), and the colours of the slices represent the treatment types from which the metagenomic reads originated. Reads were placed on the tree using the RAxML v8.0.19 EPA algorithm (Stamatakis, 2014). AmoA protein sequences predicted from the MAGs recovered from Illinois soil metagenome previously (e.g. Havana and Urbana) (Orellana *et al*., 2018b) are included as references.

On the other hand, it is worth noting that the effect of viruses on the abundance patterns of prokaryotic ammonia oxidizers was not considered in this study. Despite no available studies about interactions between viruses and ammonia oxidizers in soil, a few studies have reported the discovery of putative AOA proviruses (e.g. *Caudovirales* and class nov., *Thaspviridae*) in soil and marine environment (Krupovic *et al*., 2011; Kim *et al*., 2019).

Abundance correlation to N_2O emissions

We next examined which taxa (species-level, defined at 95% nucleotide identity based on the alignment of the sequences used in the phylogenetic tree) correlated the most to N_2O emissions as a proxy for activity and responded to fertilization. The correlation between the abundance of *amoA* reads assigned to specific taxa and measured N_2O emission flux was not always the highest for the most abundant taxa. In control plots, *Nitrososphaera viennensis* EN76, to which 2.5% of the archaeal *amoA* reads from the control plots mapped, was the only taxon among all archaeal and bacterial ammonia oxidizers of the control plots with significant correlation to N_2O emission ($R^2 = 0.7113$, $p = 0.0174$). On the other hand, the N_2O emissions measured from the fertilized plots were greatly correlated with an *amoA* gene sequence recovered from a previously reported MAG (ID: HD004) ($R^2 = 0.8697$, $p = 0.0067$) (Orellana *et al*., 2018), recruiting 3.0% of the archaeal *amoA* reads from fertilized plots. Lastly, none of the AOA or AOB significantly correlated to N_2O emissions in intercropped plots, after removing the intercropped sample from July 2014 with the abnormally high N_2O emissions.

In summary, despite the 3.4 to 5.8-fold higher abundance of archaeal *amoA* reads placed within the I.1b as opposed to the I.1a clade, the relative abundance of I.1b representatives did not correlate to N_2O emission in control or fertilized plots, unlike their I.1a counterparts whose abundance moderately correlated to N_2O emissions. The bacterial ammonia-oxidizers appear to be much more responsive to fertilization with the abundance of *Nitrosovibrio tenuis*-like sequences correlating strongly to N_2O emissions from fertilized and intercropped plots.

This was also consistent with previous studies (Orellana *et al*., 2019; Pannu *et al*., 2019) reporting an increased abundance of betaproteobacterial ammonia oxidizers upon fertilization treatments. The recently discovered complete ammonia-oxidizer (Comammox) *Nitrospira* has been hypothesized to be a dominant genus of AOB in terrestrial environments; however, its ecological niche and the role in N_2O emission remains unclear (Daims *et al*., 2015; van Kessel *et al*., 2015; Aigle *et al*., 2019). The HD021 MAG-carrying *amoA* gene phylogenetically closest to a Comammox soil *Nitrospira* and discovered in fertilized soil from the Havana site, Illinois, was found to be the most dominant AOB in all three treatment plots at the Prosser site (30.5–38.2%). Notably, its abundance also correlated significantly with N_2O emission in intercropped plots ($R^2 = 0.7894$, $p = 0.0038$). Future functional work (e.g. transcriptome or enzyme assay) studies of these AOA and AOB taxa is needed to corroborate the correlation results reported here and further quantify their contributions to N_2O emissions.

Correlation to qPCR data

In addition to metagenome-based assessment of abundances, we also quantified the AOA/AOB abundances using qPCR and assessed the extent to which the two data series correlated through simple linear regressions (Fig. S6). It should be noted that only the betaproteobacterial *amoA* was considered for this analysis, to the exclusion of the Comammox *Nitrospira amoA*, since the primer set GenAOB (Meinhardt *et al*., 2015) used to target the AOB was not designed to amplify Comammox *amoA* genes. The genome equivalents of the entire betaproteobacterial AOB group showed a positive linear regression with the corresponding AOB qPCR counts when all samples were considered, but the archaeal *amoA* gene and AOA taxa counts did not (Fig. S6A), revealing that the two methods may provide qualitatively consistent results. We also examined in more detail the specific ammonia-oxidizing (nitrifying) taxa with statistically significant positive correlation to the N_2O emission under the assumption that these taxa may contribute more to ammonia oxidation and/or are more active than other taxa that did not correlate. Considering only the archaeal *amoA* gene-based OTUs (> 95% nucleotide identity level) with significant positive correlation with N_2O emission ($p < 0.05$), the R^2 value of the linear regression between metagenome- and qPCR-based abundances did not improve (Fig. S6C). Alignment of the recovered *amoA* short-reads against the sequences of the GenAOA/GenAOB PCR primer used (Meinhardt *et al*., 2015) revealed that among the reads that sample the priming site (i.e. BLASTn matches with alignment length equal to the primer length), 98% and 32% of the archaeal and bacterial *amoA* reads, respectively, had more than two mismatches. Further and considering that none of the *amoA* short-reads perfectly matched to the primer sets of both archaeal and bacterial *amoA* genes, the lack of improvement in the linear regression statistics could be due to inadequate specificity of the PCR primer sets to recover the responding ammonia oxidizers. The DNA sequences reported here could be used to further improve the PCR primer sequence to avoid mismatches, and thus more effectively target the key nitrifying taxa and their activities (RNA level) in future studies.

Conclusions

Neither synthetic fertilization nor intercropping with alfalfa led to major shifts of overall taxonomic and functional composition of soil microbial communities in switchgrass plots in Prosser, WA, but a few minor significant changes were noted. The minor changes are likely due to the general slow growth rates of soil microbes, estimated at 1–2 generations per year for the bulk soil sampled here (Gray and Williams, 1971). Thus, it likely takes several years of treatment before major shifts become apparent at the DNA level assessed here. Consistent with these results and interpretations, a previous study of similar agricultural treatments also reported minor compositional shifts of ammonia-oxidizing populations (Bertagnolli *et al*., 2016). Interestingly, the relative abundance of the mycorrhizal fungi increased 2.5-fold due to intercropping, mainly represented by *Agaricomycotina*. Further, the N_2O emissions from Prosser soils were significantly correlated to the abundance of bacterial *amoA* genes but generally, not to their archaeal counterparts. The specific ammonia-oxidizing taxa with potential contributions to N_2O emission were identified as members of the AOA I.1a archaeal clade, Comammox soil *Nitrospira*, *Nitrosovibrio tenuis* and *Nitrosospira* (both *Betaproteobacteria*), and should be targets of future investigation to better quantify their role in N_2O emissions. Collectively, these results indicated that intercropping could be a viable alternative to fertilization since it did not dramatically alter the microbial community structure and function, and the few shifts observed may be related to the preservation of nitrogen in the system and lower N_2O emissions. However, longer time-series studies are probably needed for more robust conclusions.

Experimental procedures

Site description and sampling protocol

The Prosser site (46°15′10.0″N 119°44′14.0″W; 203 m) was established and is maintained by the Washington State University Agriculture Research and Extension

Center in Prosser, WA, USA. The Blackwell switchgrass and Pioneer 54H11 alfalfa (*Medicago sativa*) were planted in June and September 2010 respectively. More detailed information about the site is provided elsewhere (Bertagnolli *et al*., 2015; Pannu *et al*., 2019). Three types of applied treatments were unfertilized switchgrass (control), fertilized switchgrass and switchgrass-alfalfa intercropping at a ratio of 70:30. Synthetic fertilization was applied at a total agronomic rate of 224 kg ha^{-1} $year^{-1}$ to fertilized plots as NPK commercial fertilizer (16–16–16) using a hand-held broadcast spreader. Fertilization happened twice per year, once in April before the switchgrass breaks dormancy, and again in July after the annual harvest.

In 2012, soil and greenhouse gases were sampled following automated irrigation. In 2013, rather than automatic irrigation, we took manual control of water application early in the season; thus, gas and soil were sampled immediately following irrigation. All soil cores were collected at 0–15 cm depth in triplicate using a 1.2 cm diameter steel probe, randomly from each plot. After pooling and homogenizing, the soil samples were stored at 4°C for 24 h before returning to the laboratory and stored at −80°C before DNA extraction. For chemical analysis, a portion (5 g) of the homogenized soil was used for the measurement of ammonium and nitrate concentrations. Detailed information about the field gas sampling procedure and chemical analysis (i.e. NH_4^+/NO_3^- analysis) is given elsewhere (Pannu *et al*., 2019).

DNA extraction and shotgun sequencing

Soil DNA was extracted from 0.5 g of soil using MO BIO Powersoil DNA Isolation Kit (MO BIO, Carlsbad, CA), as previously described (Pannu *et al*., 2019). Dual indexed libraries were prepared using the Nextera XT DNA Library Prep Kit (Illumina) following the manufacturer's protocol until isolation of cleaned double-stranded libraries. Quality of the libraries was assessed prior to sequencing by determining the concentration using the Qubit dsDNA HS Assay Kit (Thermofisher) and the insert size distribution determined using a High Sensitivity DNA Kit (Agilent) and 2100 Bioanalyzer (Agilent). Sequencing of the samples was performed at the Molecular Evolution Core facility at Georgia Tech on a HiSeq 2500 platform in a rapid run mode using the HiSeq Rapid PE Cluster Kit v2 (Illumina) and the HiSeq Rapid SBS Kit v2 (Illumina) with 2 × 150 bp paired-end format. Adapters were trimmed and libraries demultiplexed by the instrument.

Use of publicly available surface soil metagenomes

Three publicly available surface soil metagenomes that represent the 0–10 cm depth from nine CZOs located across the United States (CZOs) (Christina, Shale Hills, Reynolds, Calhoun, Luquillo, Boulder, IML, Catalina Jemez and Sierra) (Brewer *et al*., 2019), the 0 cm depth from paired cultivated-corn and native prairie soils in Wisconsin, Iowa and Kansas (NTP) (Mackelprang *et al*., 2018), and the 0–5 cm depth from corn and soybean soils in Havana and Urbana, Illinois (Orellana *et al*., 2018) were used. Because the sequencing platform varied among datasets and resulted in differing lengths of raw reads, the quality-trimmed forward reads from all datasets including Prosser were truncated to the minimum average length of the samples (76 bp) for consistency. Only forward reads of the raw Illumina reads were analysed for consistency and quality trimmed with SolexaQA v3.1.3 (Cox *et al*., 2010) using a Phred score cutoff of 20 and minimum length of 50 bp. The resulting reads were used for the alpha and beta diversity comparison between datasets. Other types of analysis, specific to the Prosser datasets, employed both sister reads of the Prosser dataset following the procedures described below. A summary of samples from each site is provided in Table S1.

Taxonomic and functional profiling

Raw sequencing reads for each of the 24 Prosser soil samples were merged using PEAR v0.9.0 (options: -p 0.001). Both merged and unmerged reads were quality trimmed as described above. These sets of trimmed reads were used for the subsequent 16S and 18S rRNA gene-based taxonomic annotations. Functional profiling and identification of nitrogen-cycle genes were conducted using quality-trimmed paired-end reads without merging. For taxonomic profiling of the soil community, 16S and 18S rRNA gene fragments were first extracted from quality-trimmed merged and unmerged forward reads using SortMeRNA v2.1 (Kopylova *et al*., 2012) with default options. For closed-reference OTU picking, a 97% nucleotide identity threshold was used against the SILVA SSU database v132 (Quast *et al*., 2012) dereplicated at 95% nucleotide identity based on the UCLUST (Edgar, 2010) algorithm as implemented in QIIME v1.9.1 (Caporaso *et al*., 2010). Bray–Curtis dissimilarities were calculated from the resulting OTU tables and visualized through principal coordinate analysis (PCoA) plots. Statistical significance of treatment types, seasonal effect and nitrogen substrate availability on the OTU differences were assessed using Analysis of similarity (Anosim) and permutational multivariate ANOVA (Adonis) as implemented in the R package Vegan v2.5-6 (Oksanen *et al*., 2019). The Kruskal–Wallis test was used on the rarified 16S and 18S rRNA OTU tables to determine which bacterial families and fungal classes were significantly different in abundance among the three types of

treatment plots. Prokaryotic phylum abundances were determined by aligning the recovered 16S rRNA gene fragments against the SILVA SSU database v132 using BLASTn v2.2.29 (Camacho *et al*., 2009) with default options. To further corroborate the fraction of each domain annotated based on 16S and 18S rRNA genes, the taxonomic assignment to the protein-coding genes (i.e. non-rRNA-carrying reads rejected by SortMeRNA) was also performed by aligning these sequences against their best match in the TrEMBL protein database (Consortium, 2019) (downloaded in February 2020) with DIAMOND BLASTx v0.9.22.123 (Buchfink *et al*., 2015) (option: -k 1 -e 1E-3 -sensitive). The DIAMOND BLASTx outputs were filtered for best match only and a minimum bit score of > 55. Retained protein-coding genes were taxonomically annotated as bacterial, archaeal, eukaryotic and fungal as described previously (Johnston *et al*., 2016). For gene function profiling, the quality-trimmed short-reads without merging were searched against the Swiss-Prot database (Consortium, 2019) using DIAMOND BLASTx (options: -k 1 -e 1E-5 -sensitive) and the search outputs were filtered for best match only and a minimum amino-acid identity of > 40% and bit score of > 55 to minimize the false-positives matches. Swiss-Prot annotations were consolidated into GO pathway categories as described elsewhere (Johnston *et al*., 2019). The differentially abundant pathways in fertilized and intercropped plots compared with the control were determined on the resulting raw counts table of genes and GO pathways with the DESeq2 software package (Love *et al*., 2014).

Quantification and annotation of nitrogen cycle genes

ROCker models (Orellana *et al*., 2016) of the nitrogen cycle genes, including archaeal and bacterial ammonia monooxygenase alpha subunit (AmoA), hydroxylamine dehydrogenase (HAO), nitrite oxidoreductase alpha subunit (NxrA), nitrate reductase (NarG), nitrite reductase (NirK and NirS), nitric oxide reductase beta subunit (NorB), nitrous oxide reductase (NosZ) and nitrite reductase (NrfA), with lengths of 125 bp (the average of all quality-trimmed Prosser metagenomic reads) were built using the manually curated protein reference databases organized as described elsewhere (Orellana *et al*., 2018). Models are available here: http://enve-omics.ce.gatech.edu/rocker/. Quality-trimmed short-reads without merging were aligned against the in-house protein databases of each nitrogen cycle gene using BLASTx with default options. BLASTx hits with an e-value lower than 0.001 were retained and subsequently filtered using previously constructed ROCker models (Orellana *et al*., 2018) to maximize the recovery of true positive and minimize false positive matches, which enables an accurate detection of metagenomic reads related to a single function of interest. The ROCker-filtered BLASTx outputs were further normalized by the median length of the reference sequence of each nitrogen cycle gene and by the average coverage of single-copy universal genes in each metagenome dataset using MicrobeCensus v1.1.0 (Nayfach and Pollard, 2015) (options: -n 100000000) to provide the 'genome equivalent' values. Therefore, the final normalized gene abundance had a unit of 'fraction of total cells sampled that carries the gene of interest assuming a single-copy of the gene per genome'. The detailed procedure of the construction of the phylogenetic reference trees and the placement of the recovered short-reads harbouring nitrogen cycle genes on those reference trees were described elsewhere (Orellana *et al*., 2018).

Accession number

Raw metagenomic soil datasets have been submitted to the European Nucleotide Archive under study number PRJNA669067.

Acknowledgements

This research was partially supported by the Genomic Science and Technology for Energy and the Environment (Grant DE-SC0006869) from the Department of Energy (to DAS) and Award #1831582 from NSF's Dimensions program (to KTK). GC was supported, in part, by a doctoral Fulbright fellowship.

References

Agerer, R. (2001) Exploration types of ectomycorrhizae. *Mycorrhiza* **11**: 107–114.

Aigle, A., Prosser, J.I., and Gubry-Rangin, C. (2019) The application of high-throughput sequencing technology to analysis of amoA phylogeny and environmental niche specialisation of terrestrial bacterial ammonia-oxidisers. *Environ Microbiome* **14**: 3.

Barton, L., Butterbach-Bahl, K., Kiese, R., and Murphy, D.V. (2011) Nitrous oxide fluxes from a grain–legume crop (narrow-leafed lupin) grown in a semiarid climate. *Glob Chang Biol* **17**: 1153–1166.

Bertagnolli, A.D., Meinhardt, K.A., Pannu, M., Brown, S., Strand, S., Fransen, S.C., and Stahl, D.A. (2015) Influence of edaphic and management factors on the diversity and abundance of ammonia-oxidizing thaumarchaeota and bacteria in soils of bioenergy crop cultivars. *Environ Microbiol Rep* **7**: 312–320.

Bertagnolli, A.D., McCalmont, D., Meinhardt, K.A., Fransen, S.C., Strand, S., Brown, S., and Stahl, D.A. (2016) Agricultural land usage transforms nitrifier population ecology. *Environ Microbiol* **18**: 1918–1929.

Brewer, T.E., Aronson, E.L., Arogyaswamy, K., Billings, S.A., Botthoff, J.K., Campbell, A.N., *et al*. (2019) Ecological

and genomic attributes of novel bacterial taxa that thrive in subsurface soil horizons. *MBio* **10**: e01318–e01319.

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

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

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., *et al*. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Chen, X.P., Zhu, Y.G., Xia, Y., Shen, J.P., and He, J.Z. (2008) Ammonia-oxidizing archaea: important players in paddy rhizosphere soil? *Environ Microbiol* **10**: 1978–1987.

Collins, H.P., Smith, J., Fransen, S., Alva, A., Kruger, C., and Granatstein, D. (2010) Carbon sequestration under irrigated switchgrass (*Panicum virgatum* L.) production. *Soil Sci Soc Am J* **74**: 2049–2058.

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

Cox, M.P., Peterson, D.A., and Biggs, P.J. (2010) SolexaQA: at-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics* **11**: 485.

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

Dalal, R.C., Wang, W., Robertson, G.P., and Parton, W.J. (2003) Nitrous oxide emission from Australian agricultural lands and mitigation options: a review. *Soil Res* **41**: 165–195.

Duchene, O., Vian, J.-F., and Celette, F. (2017) Intercropping with legume for agroecological cropping systems: complementarity and facilitation processes and the importance of soil microorganisms. A review. *Agr Ecosyst Environ* **240**: 148–161.

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

Erguder, T.H., Boon, N., Wittebolle, L., Marzorati, M., and Verstraete, W. (2009) Environmental factors shaping the ecological niches of ammonia-oxidizing archaea. *FEMS Microbiol Rev* **33**: 855–869.

Govindarajulu, M., Pfeffer, P.E., Jin, H., Abubaker, J., Douds, D.D., Allen, J.W., *et al*. (2005) Nitrogen transfer in the arbuscular mycorrhizal symbiosis. *Nature* **435**: 819–823.

Gray, T., and Williams, S. (1971) Microbial productivity in soil. *Soc Gen Microbiol 21st Symp Microbes Biol Prod* **21**: 255–286.

Hink, L., Nicol, G.W., and Prosser, J.I. (2017) Archaea produce lower yields of N_2O than bacteria during aerobic ammonia oxidation in soil. *Environ Microbiol* **19**: 4829–4837.

Hink, L., Gubry-Rangin, C., Nicol, G.W., and Prosser, J.I. (2018) The consequences of niche and physiological differentiation of archaeal and bacterial ammonia oxidisers for nitrous oxide emissions. *ISME J* **12**: 1084–1093.

Johnston, E.R., Rodriguez-R, L.M., Luo, C., Yuan, M.M., Wu, L., He, Z., *et al*. (2016) Metagenomics reveals pervasive bacterial populations and reduced community diversity across the Alaska tundra ecosystem. *Front Microbiol* **7**: 579.

Johnston, E.R., Hatt, J.K., He, Z., Wu, L., Guo, X., Luo, Y., *et al*. (2019) Responses of tundra soil microbial communities to half a decade of experimental warming at two critical depths. *Proc Natl Acad Sci* **116**: 15096–15105.

van Kessel, M.A., Speth, D.R., Albertsen, M., Nielsen, P.H., den Camp, H.J.O., Kartal, B., *et al*. (2015) Complete nitrification by a single microorganism. *Nature* **528**: 555–559.

Kim, J.-G., Kim, S.-J., Cvirkaite-Krupovic, V., Yu, W.-J., Gwak, J.-H., López-Pérez, M., *et al*. (2019) Spindle-shaped viruses infect marine ammonia-oxidizing thaumarchaea. *Proc Natl Acad Sci* **116**: 15645–15650.

Kimura, E., Collins, H., and Fransen, S. (2015) Biomass production and nutrient removal by switchgrass under irrigation. *Agron J* **107**: 204–210.

Kits, K.D., Sedlacek, C.J., Lebedeva, E.V., Han, P., Bulaev, A., Pjevac, P., *et al*. (2017) Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle. *Nature* **549**: 269–272.

Kopylova, E., Noé, L., and Touzet, H. (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**: 3211–3217.

Krupovic, M., Spang, A., Gribaldo, S., Forterre, P., and Schleper, C. (2011) A thaumarchaeal provirus testifies for an ancient association of tailed viruses with archaea. *Biochem Soc Trans* **39**: 82–88.

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

Lemus, R., Parrish, D.J., and Abaye, O. (2008) Nitrogen-use dynamics in switchgrass grown for biomass. *Bioenergy Res* **1**: 153–162.

Lennon, J.T., and Jones, S.E. (2011) Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Microbiol* **9**: 119–130.

Li, C., Hu, H.-W., Chen, Q.-L., Chen, D., and He, J.-Z. (2019) Comammox *Nitrospira* play an active role in nitrification of agricultural soils amended with nitrogen fertilizers. *Soil Biol Biochem* **138**: 107609.

Love, M.I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Mackelprang, R., Grube, A.M., Lamendella, R., Jesus, E. D.C., Copeland, A., Liang, C., *et al*. (2018) Microbial community structure and functional potential in cultivated and native tallgrass prairie soils of the Midwestern United States. *Front Microbiol* **9**: 1775.

McLaughlin, S.B., and Kszos, L.A. (2005) Development of switchgrass (*Panicum virgatum*) as a bioenergy feedstock in the United States. *Biomass Bioenergy* **28**: 515–535.

Madakadze, I., Stewart, K., Peterson, P., Coulman, B., and Smith, D. (1999) Cutting frequency and nitrogen fertilization effects on yield and nitrogen concentration of switchgrass in a short season area. *Crop Sci* **39**: 552–557.

Meinhardt, K.A., Bertagnolli, A., Pannu, M.W., Strand, S.E., Brown, S.L., and Stahl, D.A. (2015) Evaluation of revised polymerase chain reaction primers for more inclusive quantification of ammonia-oxidizing archaea and bacteria. *Environ Microbiol Rep* **7**: 354–363.

Meinhardt, K.A., Stopnisek, N., Pannu, M.W., Strand, S.E., Fransen, S.C., Casciotti, K.L., and Stahl, D.A. (2018) Ammonia-oxidizing bacteria are the primary N_2O producers in an ammonia-oxidizing archaea dominated alkaline agricultural soil. *Environ Microbiol* **20**: 2195–2206.

Mosier, A., Schimel, D., Valentine, D., Bronson, K., and Parton, W. (1991) Methane and nitrous oxide fluxes in native, fertilized and cultivated grasslands. *Nature* **350**: 330–332.

Nayfach, S., and Pollard, K.S. (2015) Average genome size estimation improves comparative metagenomics and sheds light on the functional ecology of the human microbiome. *Genome Biol* **16**: 51.

Oksanen, J., Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., *et al*. (2019) Vegan: Community Ecology Package (Version 2.5-6); R. In: CRAN.

Ondov, B.D., Treangen, T.J., Melsted, P., Mallonee, A.B., Bergman, N.H., Koren, S., and Phillippy, A.M. (2016) Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol* **17**: 132.

Orellana, L.H., Rodriguez-R, L.M., and Konstantinidis, K.T. (2016) ROCker: accurate detection and quantification of target genes in short-read metagenomic data sets by modeling sliding-window bitscores. *Nucleic Acids Res* **45**: e14.

Orellana, L.H., Chee-Sanford, J.C., Sanford, R.A., Löffler, F. E., and Konstantinidis, K.T. (2018) Year-round shotgun metagenomes reveal stable microbial communities in agricultural soils and novel ammonia oxidizers responding to fertilization. *Appl Environ Microbiol* **84**: e01646-01617.

Orellana, L.H., Hatt, J.K., Iyer, R., Chourey, K., Hettich, R.L., Spain, J.C., *et al*. (2019) Comparing DNA, RNA and protein levels for measuring microbial dynamics in soil microcosms amended with nitrogen fertilizer. *Sci Rep* **9**: 17630.

Pannu, M.W., Meinhardt, K.A., Bertagnolli, A., Fransen, S. C., Stahl, D.A., and Strand, S.E. (2019) Nitrous oxide emissions associated with ammonia-oxidizing bacteria abundance in fields of switchgrass with and without intercropped alfalfa. *Environ Microbiol Rep* **11**: 727–735.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., *et al*. (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590–D596.

Ravishankara, A.R., Daniel, J.S., and Portmann, R.W. (2009) Nitrous oxide (N_2O): the dominant ozone-depleting substance emitted in the 21st century. *Science* **326**: 123–125.

Rodriguez-R, L.M., and Konstantinidis, K.T. (2014) Estimating coverage in metagenomic data sets and why it matters. *ISME J* **8**: 2349–2351.

Rodriguez-R, L.M., Gunturu, S., Tiedje, J.M., Cole, J.R., and Konstantinidis, K.T. (2018) Nonpareil 3: fast estimation of metagenomic coverage and sequence diversity. *mSystems* **3**: e00039-00018.

Sakoula, D., Koch, H., Frank, J., Jetten, M.S., Van Kessel, M.A., and Lücker, S. (2021) Enrichment and physiological characterization of a novel comammox *Nitrospira* indicates ammonium inhibition of complete nitrification. *ISME J* **15**: 1010–1024.

Sanford, R.A., Wagner, D.D., Wu, Q., Chee-Sanford, J.C., Thomas, S.H., Cruz-García, C., *et al*. (2012) Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils. *Proc Natl Acad Sci* **109**: 19709–19714.

Smith, S.E., and Read, D.J. (2010) *Mycorrhizal Symbiosis*: London, UK: Academic Press.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Venterea, R.T., Halvorson, A.D., Kitchen, N., Liebig, M.A., Cavigelli, M.A., Grosso, S.J.D., *et al*. (2012) Challenges and opportunities for mitigating nitrous oxide emissions from fertilized cropping systems. *Front Ecol Environ* **10**: 562–570.

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

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

Wright, L., and Turhollow, A. (2010) Switchgrass selection as a ''model'' bioenergy crop: a history of the process. *Biomass Bioenergy* **34**: 851–868.

Yu, L., Tang, Y., Wang, Z., Gou, Y., and Wang, J. (2019) Nitrogen-cycling genes and rhizosphere microbial community with reduced nitrogen application in maize/-soybean strip intercropping. *Nutr Cycl Agroecosyst* **113**: 35–49.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. Alpha diversity and sequence coverage of the microbial communities studied. (A) Species richness presented as the number of observed OTUs (left panel) and Chao-1 index (right panel), (B) intra-sample diversity measured by Shannon's diversity (H') of 16S rRNA gene-based OTUs (left panel) and Nonpareil diversity (N_d) of shotgun metagenomic reads (right panel), and (C) 16S rRNA gene-based OTUs coverage (Observed OTUs/Chao-1 index) and Nonpareil-estimated coverage (left panel), and sequencing effort applied (right panel). The underlying data for the Observed OTUs, Chao-1 index, and Shannon's diversity (H') is the 16S rRNA gene-based OTU table rarefied to the lowest level of coverage as 600. Nonpareil diversity (N_d) and coverage are estimated from the quality trimmed metagenome datasets without rarefaction. The box plots represent the 25th and 75th percentile and whiskers the min and max values in (A) and (B). The line within each box is the median and the '+', the mean. Error bars in (C) represent the standard deviation from the mean.

Fig. S2. Relative abundance of fungal taxa in each Prosser dataset. Values represent the relative abundance of each fungal taxa that were taxonomically annotated using SILVA's 18S rRNA gene database v.132 by UCHIME embedded in QIIME1 closed OTU picking pipeline. Abundance was estimated as proportion of the total fungal 18S rRNA gene-carrying read sequences found in each dataset.

Fig. S3. Comparison of the average genome size estimate with the coverage of the rpoB and 16S rRNA genes between treatment groups. (A) Average genome size estimated from MicrobeCensus (in Gbp) and (B) Coverage of the rpoB and 16S rRNA genes in $\times 1000$ short reads carrying each gene. Correlation between the three metrics used to express relative gene abundance as genome equivalents, i.e. % of total cells encoding the gene of interest, was significant (paired *t*-test, $p < 0.05$).

Fig. S4. Differential abundance of Gene Ontology (GO) pathways between treatment groups. Significant differences in abundance of GO pathways between intercrop (switchgrass and alfalfa) and fertilized (switchgrass alone) soils were identified using a negative binomial test as implemented in DESeq2. Selected pathways showing log2-fold change ≥ 2 or ≤ -2 and *p*-values < 0.05 are shown.

Fig. S5. Correlation of N_2O emission with (A) archaeal *amoA*, (B) comammox *Nitrospira amoA*, (C) betaproteobacterial *amoA* and (D) *nosZ* abundances for all Prosser soil samples (left four plots) or without the fertilized plot samples from July 2014 (right four plots). R^2 and *p* values of the correlation analysis are shown on each panel. The N_2O emission flux value represents the mean of 12 values (four replicate plots and three random samples per plot using static chambers). The right four plots were drawn to evaluate the effect of the fertilized sample from July 2014 with the highest N_2O emission of 60.39 g ha^{-1} day^{-1}. Notably, the correlation analysis N_2O fluxes remained statistically significant ($p < 0.05$) correlation with betaproteobacterial *amoA* abundance (measured as genome equivalents).

Fig. S6. Linear regression of qPCR AOA/AOB counts (cells g^{-1} soil) with metagenome-based abundance of (A) archaeal (B) betaproteobacterial *amoA* OTUs or (C) only archaeal OTUs with positive correlation ($p < 0.05$) to N_2O emission. Archaeal *amoA* OTUs with significant correlation to N_2O emissions included *Nitrososphaera viennensis* EN76 and Havana (HD004). None of the betaproteobacterial *amoA* OTUs significantly correlated to the N_2O emission. Best-fit lines, 95% confidence bands of the best-fit line (in dotted lines), and R^2 values are marked in each plot. OTU abundance is measured as genome equivalents. Note that for the qPCR data, the number of genes has been divided by 2.4 to get to cells for AOB, as that is the average number of copies of the *amoA* gene found in AOB genomes. Besides, typical AOA genomes carry a single copy of *amoA* gene, so AOA cells per gram soil is equivalent to *amoA* gene copies per gram soil.

Fig. S7. Linear regression of (A) archael and (B) bacterial amoA genome equivalent versus time. Linear regression analysis was followed by the analysis of covariance (ANCOVA) test to assess the statistical significance of the differences between slopes of fitted lines. In case of AOA, the control samples had a negative slope of $-7.32e{-}0.005$; in contrast, fertilized and intercropped samples had positive slopes as $7.99e{-}0.005$ and $9.91e{-}4$, respectively. Especially, the intercropped samples had a good linear fit ($R^2 = 0.8568$) with a significant non-zero slope ($p = 0.0081$). Also, the differences between slopes of control and intercropped was significant ($p = 0.0385$) based on the ANCOVA result, while the difference between control and fertilized was not ($p = 0.7635$). Note that the fertilized and intercropped samples from September 2013 were excluded from this analysis since we did not have the control sample from that year.

Fig. S8. Linear regression of bacterial *amoA* genome equivalent in (A) all fertilized samples or (B) excluding the July 2013 sample, against either ammonium or nitrate concentrations. The purpose of the right figure was to visually inspect whether a few (or a single) outlier datapoint drove a positive correlation since the July 2013 sample had the highest bacterial *amoA* abundance, ammonium, and nitrate concentrations among the fertilized plot samples. The tables attached to each figure summarize the statistics of the linear regression and also the Pearson's correlation analysis. Excluding the July 2013 fertilized sample decreased the statistical significance; however, this did not change the positive correlations.

Fig. S9. Correlation between N_2O emission and abundance of the archaeal amoA gene. Gene abundance was measured as genome equivalents. The amoA gene sequence used as reference was recovered from a previously reported MAG HD004 Orellana *et al*., 20182018. Note that, the fertilized sample from July 2014 with the highest N_2O emission was removed prior to the analysis.

Fig. S10. Abundance and phylogenetic diversity of comammox *Nitrospira* AmoA proteins. Phylogenetic reconstruction of comammox *Nitrospira* AmoA protein sequences recovered from contigs. The approximately-maximum-likelihood phylogenetic reference tree was constructed using RAxML v8.0.19 (option: -f E) Stamatakis, 20182014. Reference sequences were collected from the GenBank database; AmoA protein sequences predicted from MAG sequences recovered from an Illinois soil metagenome previously (i.e. Havana and Urbana) Orellana *et al*., 2018 were also included. *Nitrosomonas europaea* was used as an outgroup (marked with an arrow). The pie chart represents the number of recovered comammox *amoA*-carrying reads (expressed as genome equivalents), and the colours of the slices represent the treatment types from which the metagenomic reads originated (CT, control; FT, fertilized; IC, intercropped). Reads were placed on the tree using the RAxML v8.0.19 EPA algorithm.

Table S1. Metadata of soil samples determined by this study.

Table S2. Statistics of the metagenomic datasets determined by this study.

Table S3. Annual dry biomass yields (Mg ha^{-1}) for Prosser soils in years 2012 and 2013. The significant differences are denoted by different letters as superscripts. Significance was assessed based on the non-parametric multiple comparison Wilcoxon test.

Table S4. Differences in diversity among datasets used in this study.

Table S5. Statistical significance of each variable in shaping the 16S rRNA gene-based taxonomic composition of the Prosser soil metagenomes.

Table S6. Soil nitrate concentrations with depth in the Prosser plots in October 2012. Data were collected from four replicate plots.