GLOBAL CHANGE BIOLOGY
BIOENERGY

GCB Bioenergy (2016) **8**, 481–494, doi: 10.1111/gcbb.12289

# Influence of corn, switchgrass, and prairie cropping systems on soil microbial communities in the upper Midwest of the United States

EDERSON DA C. JESUS[1,2], CHAO LIANG[3,4], JOHN F. QUENSEN[1], ENDANG SUSILAWATI[1,5], RANDALL D. JACKSON[3], TERESA C. BALSER[3] and JAMES M. TIEDJE[1]

[1]*Center for Microbial Ecology and DOE Great Lakes Bioenergy Research Center, Michigan State University, 540 Plant and Soil Sciences Building, East Lansing, MI 48824-1325, USA,* [2]*Embrapa Agrobiologia, BR 465, km 7, Seropédica, Rio de Janeiro 23890-000, Brazil,* [3]*Department of Agronomy and DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, 1575 Linden Drive, Madison, WI 53706, USA,* [4]*State Key Laboratory of Forest and Soil Ecology, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110164, China,* [5]*Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON M5S 3E5, Canada*

## Abstract

Because soil microbes drive many of the processes underpinning ecosystem services provided by soils, understanding how cropping systems affect soil microbial communities is important for productive and sustainable management. We characterized and compared soil microbial communities under restored prairie and three potential cellulosic biomass crops (corn, switchgrass, and mixed prairie grasses) in two spatial experimental designs – side-by-side plots where plant communities were in their second year since establishment (i.e., *intensive* sites) and regionally distributed fields where plant communities had been in place for at least 10 years (i.e., *extensive* sites). We assessed microbial community structure and composition using lipid analysis, pyrosequencing of rRNA genes (targeting fungi, bacteria, archaea, and lower eukaryotes), and targeted metagenomics of *nifH* genes. For the more recently established intensive sites, soil type was more important than plant community in determining microbial community structure, while plant community was the more important driver of soil microbial communities for the older extensive sites where microbial communities under corn were clearly differentiated from those under switchgrass and restored prairie. Bacterial and fungal biomasses, especially biomass of arbuscular mycorrhizal fungi, were higher under perennial grasses and restored prairie, suggesting a more active carbon pool and greater microbial processing potential, which should be beneficial for plant acquisition and ecosystem retention of carbon, water, and nutrients.

*Keywords:* bacterial communities, biofuel crops, fungal communities, lipid analysis, *nifH*, pyrosequencing

*Received 24 January 2015; revised version received 23 April 2015 and accepted 5 June 2015*

## Introduction

To help reduce dependence on fossil fuels, there is great interest in using plant biomass for energy. Significant efforts are underway to understand what biomass crops should be grown, where they should be grown, and how they can be managed in sustainable ways (Kim *et al*., 2012; Gao *et al*., 2013; Werling *et al*., 2014). Annual crops such as corn continue to be the most readily abundant and available crops for biofuel production in the United States (US-DOE, 2011), while canola and soybeans have been used to produce biodiesel (IEA, 2007). However, these crops are monocultures that require high-energy inputs to maintain, lose soil and nutrients, and serve as key food crops for humans and/or livestock. Their use for biofuel may increase food costs and may encourage more land to be converted to agriculture, which has negative ramifications for ecosystem carbon balance, wildlife habitat, and a host of other ecosystem services (Fargione *et al*., 2008).

The disadvantages listed above have encouraged a focus on perennial grasses such as *Panicum virgatum* (switchgrass) for lignocellulosic ethanol production (IEA, 2007). Switchgrass produces high amounts of biomass and is a native species of North America that should require less intensive agricultural management than annual crops (Wright & Turhollow, 2010). Biologically diverse, low-input systems consisting of mixtures of native grasses are also candidates for cellulosic

Correspondence: James M. Tiedje, tel. 1 517 353 9021, fax 1 517 353 2917, e-mail: tiedjej@msu.edu

biomass crops because they can be highly productive, while conserving nutrients (Jach-Smith & Jackson, 2015), can positively affect the diversity of other groups of organisms (Werling *et al.*, 2014), and result in lower greenhouse gas emissions than annuals and perennial monocultures (Tilman *et al.*, 2006; Oates *et al.*, 2015). Theoretically, more diverse plant communities will improve sustainability by requiring fewer inputs than monocultures, although support for this hypothesis in agricultural production settings is scant (but see Tilman *et al.*, 2006; Webster *et al.*, 2010). In more productive soils, it may be more important to have a particular plant community, for example, one that includes the most productive taxa, to minimize inputs, while improving stability and resilience (Adler *et al.*, 2009).

All agricultural activity affects biodiversity, soil fertility, and water resources (Groom *et al.*, 2008), and these factors should be assessed when evaluating the sustainability of biofuel cropping systems. Considering the impacts on microbial soil communities is an important component of this assessment because most soil-based ecosystems services such as organic matter degradation, nitrogen fixation, nitrification, denitrification, soil aggregation, and water retention are driven by microbial activity (Swift *et al.*, 2004). As has been shown for other crops, the cultivation of biofuel crops can be expected to influence soil microbial communities, thus affecting the key ecosystem processes and the services they provide (Groom *et al.*, 2008; Liang *et al.*, 2012). Corn is presently the major crop used for ethanol production in the United States, and soil microbial communities under corn and prairie have often been contrasted, showing differences in microbial community composition and improved carbon storage and soil aggregation under prairie (Bailey *et al.*, 2002; Allison *et al.*, 2005; Bach *et al.*, 2010; Fierer *et al.*, 2013; Murphy & Foster, 2014). The growing interest in using switchgrass and other grasses for cellulosic biomass production has sparked similar investigations contrasting soil microbial communities under these grasses with those under corn (Jesus *et al.*, 2010; Mao *et al.*, 2011, 2013; Liang *et al.*, 2012). The ways and extent to which these crops have been found to influence soil microbial communities varied according to the methods used, spatial sampling schemes, soil and environmental variables, land management and land history, but a general finding has been that cultivation of perennial grasses stimulates communities more similar to those under prairies (Liang *et al.*, 2012). This is a desirable outcome, because systems with soil microbial communities similar to those under prairies should require fewer external inputs and, for this reason, be more sustainable. But for the most part, these studies have been performed in local settings and the need to carry out studies at larger geographic scales, including sites with different management types, different times since crop establishment, and for a range of soil conditions is necessary to better examine shifts in microbial communities.

For this reason, we used a more holistic approach for our study. Our initial hypothesis was that the cultivation of switchgrass and mixed grasses would lead to microbial communities more similar to those under prairie species, implying a more sustainable system. To test this hypothesis, we sampled over a larger geographic scale, including sites in two states with a range of soil conditions, with different times since crop conversion, and under two different sampling strategies. We compared microbial communities in soils cultivated with three potential biofuel crops (corn, switchgrass, and mixed grasses) and with prairie species, and in two spatial experimental designs. One design consisted of side-by-side plots where plant communities were in their second year since establishment (i.e., *intensive* sites), and the other consisted of regionally distributed fields where plant communities had been in place for at least 10 years (i.e., *extensive* sites). We assessed the microbial communities using three different methods: lipid analysis, pyrosequencing of ribosomal genes (that target fungi, bacteria, archaea, and lower eukaryotes), and targeted metagenomics of a gene important for a key ecological function, *nifH* coding for nitrogen reductase ($N_2$ fixation). Our main questions were as follows: (i) how do the different biofuel crops affect soil microbial communities, that is, are soil microbial communities under switchgrass and mixed grasses more similar to those under prairie, (ii) how are any effects modified by location and soil type, and (iii) how do alternative soil microbial assay methods compare in revealing community differences?

## Materials and methods

### *Site description and soil sampling*

Soil samples were collected from sites in southern Michigan and southern Wisconsin under two different designs that have been used for other studies by the Great Lakes Bioenergy Research Center (GLBRC) (Fig. S1). These two designs are *intensive* and *extensive*.

The intensive plots were located at the Kellogg Biological Station (KBS) in Michigan and at the Arlington Agricultural Research Station (AARS) in Wisconsin. The plots were arrayed in randomized complete blocks designs consisting of five replicated 30 × 40 m plots of each of four plant communities – corn, switchgrass, mixed grasses, and restored prairie – and were harvested annually for biomass. We sampled from three of the five blocks 2 years after their establishment.

The extensive sites were fields located on working farms or reserves in Michigan and Wisconsin and were selected from

among those studied by Werling *et al.* (2014) to cover the range of soil types and conditions of the southern regions of both states. Nine fields were sampled in each state – three in corn, three in switchgrass, and three in restored prairie – but only the corn fields were harvested. All sites had been under their respective vegetation for at least 10 years.

Three composite samples were taken at random from each of the sampling units. Each composite sample consisted of five soil cores taken to a depth of 10 cm. All samples were transported on ice to the laboratory and then stored at −20 °C until processing.

### *Soil analysis*

Soil samples were analyzed for elemental composition (Al, B, Ca, Cu, Fe, K, Mg, Mn, Na, P, S, and Zn), total C, total N, pH, and soil texture as previously described (Liang *et al.*, 2012).

### *Lipid analysis*

Microbial community composition was determined using a hybrid procedure of phospholipid fatty acid (PLFA) and fatty acid methyl ester (FAME) analysis as previously described (Liang *et al.*, 2012). The total biomass of bacteria (B), fungi (F), and protozoa was estimated. Bacteria were further subdivided into Gram-positive ($Gm^+$) and Gram-negative ($Gm^-$) categories, and the fungi, into arbuscular mycorrhizal fungi (AMF) and saprophytic fungi (SF) (*Liang et al.*, 2012).

### *DNA extraction*

DNA was extracted from each well-mixed 500 mg soil sample using MoBio's Power Soil DNA Isolation Kit (Mobio Laboratories Inc., Carlsbad, CA, USA) according to the manufacturer's instructions. The DNA was quantified with a Nanodrop ND-1000 spectrophotometer (Nanodrop Technology, Wilmington, DE, USA) and stored at −20 °C until use.

### *Preparation of 16S/18S (SSU) rRNA gene amplicon libraries for pyrosequencing*

The V6-V8 region of smal subunit (SSU) rRNA was amplified from the template DNA using primers 926F (5′- cct atc ccc tgt gtg cct tgg cag tct cag AAA CTY AAA KGA ATT GRC GG- 3′) and 1392R (5′ - cca tct cat ccc tgc gtg tct ccg act cag - <**XXXXX**> - ACG GGC GGT GTG TRC - 3′). Primer sequences were modified by the addition of 454 A or B adapter sequences (lower case). In addition, the reverse primer included a 5 bp bar code (designated by <XXXXX> above) for multiplexing of samples during sequencing. Twenty microliter PCRs were performed in duplicate and pooled to minimize PCR bias using 0.4 μl Advantage GC 2 Polymerase Mix (Advantage-2 GC PCR Kit, Clonetech, Mountain View, CA, USA), 4 μl 5× GC PCR buffer, 2 μl 5 M GC Melt Solution, 0.4 μl 10 mM dNTP mix (MBI Fermentas, Amherst, MA, USA), 1.0 μl of each 25 nM primer, and 10 ng sample DNA. The thermal cycler protocol was 95 °C for 3 min, 25 cycles of 95 °C for 30 s, 50 °C for 45 s, and 68 °C for 90 s, and a final 10-min extension at 68 °C. PCR amplicons were purified using SPRI Beads and quantified using a Qubit fluorometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). Samples were diluted to 10 ng $\mu l^{-1}$ and mixed in equal concentrations. Emulsion PCR and sequencing of the PCR amplicons were performed following the Roche 454 GS FLX Titanium technology manufacturer's instructions. Sequencing tags were analyzed using the software tool PYROTAGGER (Kunin & Hugenholtz, 2010) using a 180 bp sequence length threshold as described in Engelbrektson *et al.* (2010).

16S/18S rRNA gene nucleotide sequences were deposited in the European Nucleotide Archive (http://www.ebi.ac.uk/ena) as part of study PRJEB6704 under accession numbers ERR571396 through ERR571438.

### *Preparation of 28S (LSU) rRNA gene amplicon libraries for pyrosequencing*

PCR amplification of template DNA was also performed using the primers LR3 (5′-CCGTGTTTCAAGACGGG-3′) and LR0R (5′-ACCCGCTGAACTTAAGC-3′) (Liu *et al.*, 2012). These primers target a 625 bp fragment of the large subunit (LSU) rRNA gene in fungi. Detailed amplification and purification protocols are given in Penton *et al.* (2013). Adapters and bar codes were ligated to the amplicons prior to sequencing at Utah State University using Lib-L kits.

28S rRNA gene nucleotide sequences were deposited in the European Nucleotide Archive (http://www.ebi.ac.uk/ena) as part of study PRJEB6704 under accession numbers ERR571439 through ERR571456.

### *Preparation of nifH amplicon libraries for pyrosequencing*

The extracted DNA also served as template to prepare *nifH* gene libraries as described in Wang *et al.* (2013). The primers were based on those of Poly *et al.* (2001), which target an approximately 320 bp region of the *nifH* gene. *NifH* gene libraries were sequenced by the Research Technology Support Facility (RTSF) at Michigan State University (East Lansing).

*NifH* nucleotide sequences were deposited in the European Nucleotide Archive (http://www.ebi.ac.uk/ena) as part of study PRJEB6704 under accession numbers ERR571353 through ERR571395.

### *Data analysis*

Principal component analysis (PCA) was used to display distances between sites based on their soil attributes. For these analyses, the arc sin transformation was first applied to percentages of sand and silt, and then all soil variables were standardized to zero mean and unit variance prior to PCA on the correlation matrix using the R (R Core Team, 2012) package VEGAN's (Oksanen *et al.*, 2012) *rda* function. To aid interpretation, vectors for the soil variables were added to the PCA plots using vegan's *envfit* function.

Total carbon concentrations were compared by analysis of variance (ANOVA) using the *lm* function of the R package stats and a cutoff value of $\alpha = 0.05$.

Lipid data were used to estimate microbial biomass, and the following ratios were calculated: fungi/bacteria (F/B), arbuscular mycorrhizal fungi/saprophytic fungi (AMF/SF), and Gram-positive bacteria/Gram-negative bacteria ($Gm^+/Gm^-$). Data were displayed per crop, and mean and errors bars were calculated for each of the experiments types. Treatment differences were tested in the same manner as total carbon. When ANOVA was significant, treatment contrasts were made with the *TukeyHSD* function of the R package stats.

All 28S LSU rRNA sequences were first processed through RDP's pyrosequencing initial processing tool (http://pyro.cme.msu.edu/). Because the amplicons were ligated with the adapters and bar codes, both primer sequences were entered in the forward primer box. Filter parameters were 0 mismatches to the forward primer, 250 bp length filter, maximum number of N's = 0, and minimum quality score of 20. Because some sequences were read from each direction, it was not possible to align them. The sequences were therefore classified directly using the Ribosomal Database Project (RDP) classifier (Wang *et al.*, 2007) with a manually curated LSU gene training set v1 (Liu *et al.*, 2012) also used in Penton *et al.* (2013) which provides additional detail. Sequences were binned by genus if identified with confidence of 0.5 or greater, or otherwise to the lowest rank category for which confidence was at least 0.5, resulting in 639 categories.

All 16S/18S rRNA gene sequences that passed the quality controls of the GL FLX software were uploaded on the PYROTAGGER pipeline (Kunin & Hugenholtz, 2010). Raw sequences were sorted by bar code, trimmed, filtered to remove sequences of low quality (10% threshold), and aligned. The minimum sequence length allowed was 150 bp. Potential chimeras were identified and excluded from downstream analysis. Sequences were clustered at the level of 97% identity, and the best hit in Greengenes (for prokaryotes) and Silva (for eukaryotes) databases was determined for each cluster. The output $OTU_{0.03}$ tables were used for statistical analysis.

All *nifH* sequences also were initially processed using the pyrosequencing pipeline tools on RDP's Web site. Reads passing the initial filters were frame shift corrected and translated into NifH protein sequences using the RDP FrameBot tool (Wang *et al.*, 2013). About 15% of the sequences had frame shift errors detected and corrected by FrameBot, such that more than 99% of the sequences were retained for analysis. The protein sequences were aligned using the HMMER3 aligner, clustered at 95% identity, and the representative sequences for each cluster classified using the FunGene Pipeline (Fish *et al.*, 2013) to find the nearest match among 675 protein sequences in a curated reference set (Wang *et al.*, 2013).

Good's coverage (Good, 1953) was calculated as a percentage for each sample from the 16S/18S rRNA, 28S rRNA, and NifH data as 100 times the quantity one minus the number of singletons divided by the total number of sequences.

PCA was used to display distances between sites based on all four data types (lipid, 16S/18S rRNA, 28S rRNA, and NifH). For these analyses, the Hellinger transformation (Legendre & Gallagher, 2001) was applied to the OTU count data using vegan's *decostand* function prior to PCA on the variance–covariance matrix using vegan's *rda* function. Multivariate analysis of variance by permutation (PMANOVA) was used to test for significant differences in dispersion among groups and for differences between group centroids (Anderson, 2001, 2006) using vegan's *betadisper* and *adonis* functions. The factors considered in these analyses were state (location), crop, and the interaction between them. Here, 'state' is actually a proxy for several correlated soil attributes, differing between intensive and extensive experiments, as explained in the Results section.

The sequences contributing most significantly to the ordinations were identified using Biodiversity. R's *ordiequilibriumcircle* function (Kindt & Coe, 2005), and indicator group analysis (Dufrene & Legendre, 1997) was used to identify OTUs whose occurrences were linked to specific crop types using labdsv's *indval* function (Roberts, 2012) and the package QVALUE (Dabney *et al.*, 2012) to assign statistical significance. Procrustes analysis (Cox & Cox, 2001) was applied to determine whether there were significant correlations between ordinations based on the four types of data using vegan's *protest* function.

## Results

### *Soil chemical and physical analysis*

Michigan and Wisconsin intensive sites differed markedly in their physical and chemical soil attributes (Fig. S2, Tables S1 and S2). Samples from the two states were separated along the first PCA axis, which explained 75.1% of the variance and represent a sand/silt gradient (Fig. S2). Except for pH and Cu, the measured soil variables, which are all linked to nutrient concentrations in the soil, were positively correlated with higher percentages of silt. Of the variables measured, pH had the highest projection on the second PCA axis, which explained only 8.4% of the variance.

A sand/silt gradient also separated the extensive sites by state, but fewer of the soil variables were strongly correlated with this gradient (Fig. S3). Indeed, Mg, Ca, and pH were orthogonal to the gradient, while K, Na, S, and total C were nearly so. However, none of these soil variables separated sites by crop.

There were important differences between states and experiments in total soil carbon. For the intensive sites, total soil carbon was higher in the Wisconsin samples for each of the studied crops (ANOVA, $P < 0.001$, Fig. 1b). For the extensive sites, total carbon tended to be higher in the Wisconsin samples for each crop, but the differences were less pronounced and not significant (ANOVA, $P = 0.15$).

### *Lipid analysis*

In most cases, small sample size and high variance precluded detection of statistically significant among

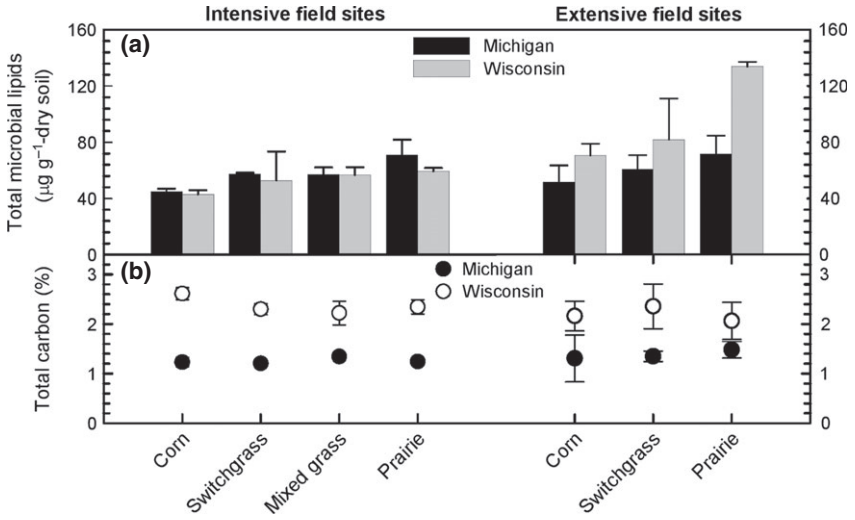


**Fig. 1** Total lipid biomass (a) and total carbon (b) under different biofuel cropping systems at intensive and extensive sites in Wisconsin and Michigan. Bars represent ± 1 standard error.

treatment differences in lipids characterizing various microbial groups. For this reason, we are limited to discussing trends in the data, the strengths of which may be judged from the relative error bars in Figs 1–3.

For the intensive sites, total lipid biomass per treatment did not differ between states (Fig. 1a). We observed a trend for higher total microbial biomass under perennial species than under corn for both states. For Michigan, this modest difference was accounted for by increases of both bacteria and fungi under the perennial grasses, with a slightly higher proportion of AMF in mixed grasses and prairie (Fig. 2a,b). For Wisconsin,

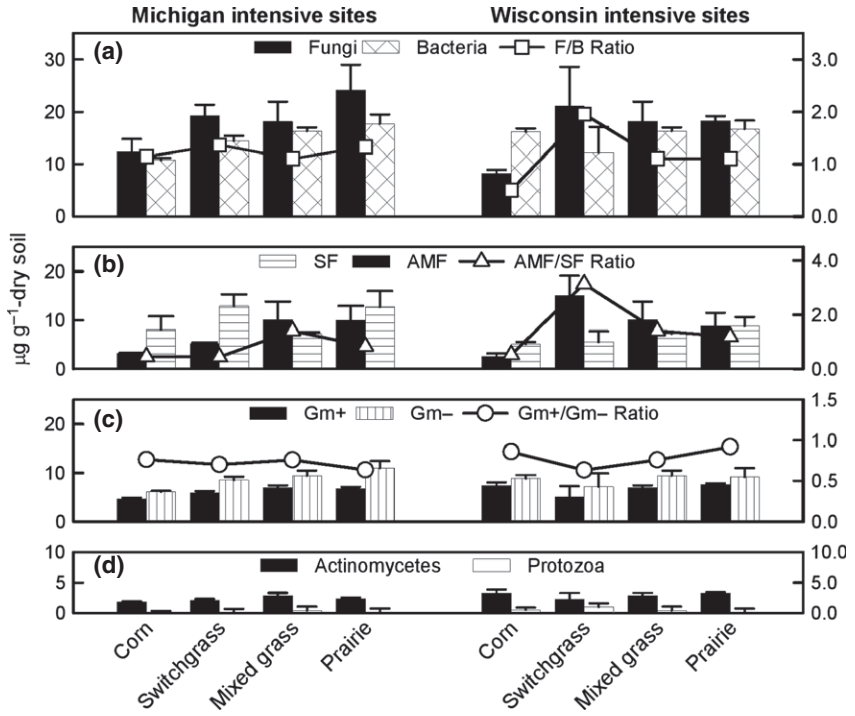


**Fig. 2** Abundance (left axis) of fungi and bacteria (a), arbuscular mycorrhizal (AMF) and saprophytic fungi (SF) (b), Gram-positive ($Gm^+$) and Gram-negative bacteria ($Gm^-$) (c), actinomycetes and protozoa (d), and the F/B, AMF/SF and $Gm^+/Gm^-$ ratios (right axis) of microbial communities under corn, switchgrass, mixed grasses, and prairie at intensive sites in Michigan and Wisconsin. Bars represent ± 1 standard error.

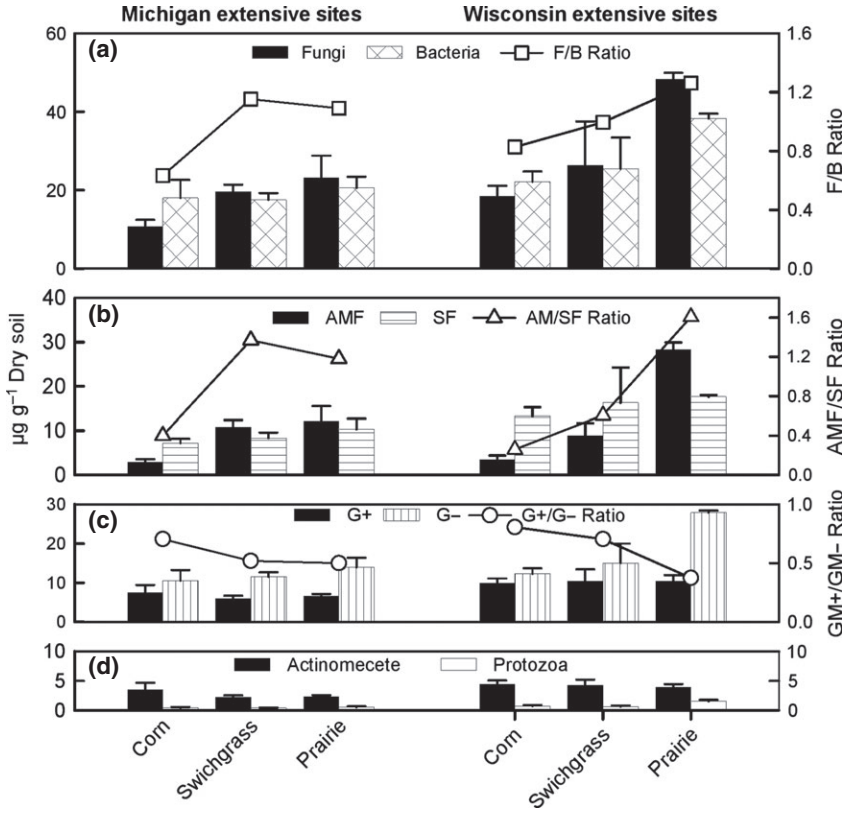


**Fig. 3** Abundance (left axis) of fungi and bacteria (a), arbuscular mycorrhizal (AMF) and saprophytic fungi (SF) (b), Gram-positive ($Gm^+$) and Gram-negative bacteria ($Gm^-$) (c), actinomycetes and protozoa (d), and the F/B, AMF/SF and $Gm^+/Gm^-$ ratios (right axis) of microbial communities under corn, switchgrass, and prairie at the extensive sites in Michigan and Wisconsin. Bars represent ± 1 standard error.

this increased total biomass under the perennials was accounted for mainly by the fungi, and of the fungi, mainly by the AMF, especially in switchgrass. There were no differences in actinomycetes or protozoa lipid markers among treatments for either state (Fig. 2d).

The extensive sites exhibited much greater differences in total biomass between the two states (Fig. 1a). Microbial biomass per treatment was higher for the extensive sites in Wisconsin than for those in Michigan, but markedly so only for the prairie sites. It also tended to increase among treatments from corn to switchgrass to prairie, especially in Wisconsin. Wisconsin extensive sites exhibited higher microbial biomass per treatment than corresponding intensive sites, especially for the corn and prairie sites.

For the extensive sites in Wisconsin, the higher biomass under perennial grasses was due to an increase in the biomass of both fungi and bacteria, and more so to fungi, but for those sites in Michigan, it was due to fungi only (Fig. 3a). For both states, the Gram-positive/Gram-negative ($Gm^+/Gm^-$) ratio was lower (Fig. 3c) and the AMF/SF ratio was greater (Fig. 3b) under the perennial grasses. AMF increased from 3.5 ± 0.9 $\mu g\ g^{-1}$ in Wisconsin and 2.9 ± 0.6 $\mu g\ g^{-1}$ in Michigan under corn to 28.3 ± 1.6 $\mu g\ g^{-1}$ and 12.1 ± 3.4 $\mu g\ g^{-1}$ in the prairie. Actinomycetes and protozoa markers were more abundant in Wisconsin; however, there were no apparent differences in these two groups among treatments for either state (Fig. 3d).

### *28S rRNA gene pyrosequencing*

We obtained 124 654 28S rRNA gene sequences for the 16 extensive sites samples retained in the study, with an average read length of 446 bp. Samples MIE.Co.16 and WIE.Co.2, extensive corn sites in MI and WI, respectively, were excluded for yielding too few sequences. Sequences not identified as fungi by the RDP classifier with the confidence filter set at 0.5 were removed, leaving a total of 119 793 sequences in 632 categories, 94 of which were universal singletons. Sequences per sample ranged from 1979 to 13 280. Good's coverage was high, ranging from 97.2 to 99.6 with a mean of 98.9%.

The three most abundant identifiable phyla were Ascomycota (67% of total sequences), Basidiomycota

(15%), and Chytridiomycota (4%) (Fig. S4). Fungi unclassified at the phylum level made up 13% of the sequences. Basidiomycota were most abundant at the Michigan prairie, and Chytridiomycota were most abundant at the Wisconsin prairie.

By IndVal analysis, the genera *Ascobolus, Podospora, Coprinellus*, *Ascodesmis*, and *Byssonectria* characterized the corn sites, with *Ascobolus* being the most abundant (Table S3). Of the many genera characteristic of the prairie, unclassified Helotiales, *Clavaria*, and *Tricladium* were the most abundant. *Beauveria* was characteristic of the switchgrass sites, but weakly so because it was not abundant and was also found at prairie sites.

As arbuscular mycorrhizal fungi (AMF) were identified by lipids analysis as an abundant group, we sought to identify fungal sequences belonging to this group in the 28S rRNA gene pyrosequencing data. We found that only 0.13% of the sequences were classified to Glomeromycota. Of these, *Paraglomus* sequences were by far the most abundant in the dataset. Most of the AMF sequences were recovered from prairie soils, which also presented the highest detected richness, with a combination of at least four genera per site. In contrast, just one or two genera could be found in soils cultivated with switchgrass, with a predominance of *Paraglomus*. The same applies to soils cultivated with corn in Wisconsin. No AMF sequences were recovered from soils cultivated with corn in Michigan.

### *16S/18S rRNA gene pyrosequencing*

We obtained 167 848 16S/18S rRNA gene sequences with an average of 7570 $\pm$ 894 sequences per sample and a minimum read length of 150 bp. These samples were aligned and clustered into 10 092 clusters (OTUs) at a distance of 3%; 6628 of these clusters were global singletons with Good's estimated sample coverage of 52 to 92% and averaging 85%.

Prokaryote sequences accounted for 86.3% of the recovered sequences and eukaryote sequences accounted for 9.2% (Fig. S5). Unassigned sequences accounted for 4.5%. Prokaryote sequences (86.3%) were mostly bacterial with only 0.02% belonging to Archaea. The most abundant bacterial phyla in the libraries were *Proteobacteria*, *Actinobacteria,* and *Acidobacteria*, at both intensive and extensive sites. Fungi, Metazoa, and Cercozoa were the more abundant eukaryotic phyla at both intensive and extensive sites.

No significant differences among locations or treatments were evident at the phylum level. Although such differences were observed for OTUs, interpretation was problematic due to the large number of clusters, most containing few sequences, and due to the poor identification of representative sequences, with many not being identified past the phylum level.

### *NifH gene pyrosequencing*

We obtained 195 385 NifH sequences for the 41 samples retained in the study, with a mean of 4765 sequences per sample and a standard deviation of 811. The average read length was 320 bp. After frame shift correction and translation into amino acids, they were clustered at a distance of 0.05 yielding 2799 OTUs. Of these, 773 were global singletons, with Good's estimated sample coverage varying among samples from 93.1 to 98.4 with a mean of 96.4%. FrameBot, included in the Fungene Pipeline (Fish *et al*., 2013), was used to match representative sequences from each cluster to 187 of 782 unique NifH reference sequences. These 187 matches fell into 100 genera.

More than 95% of the recovered NifH sequences were assigned to *Proteobacteria* (Fig. S6). Within this phylum, closest matches to the *Alphaproteobacteria* and *Betaproteobacteria* were generally more abundant and *Gammaproteobacteria* least abundant. Variances were large, but there was a tendency for *Alphaproteobacteria* affiliates to be higher in the extensive sites and *Betaproteobacteria* to be higher in the intensive sites, the latter especially for switchgrass. Matches to the *Deltaproteobacteria* were most abundant in the Michigan extensive switchgrass sites, dominating all three replicates. Unidentified environmental sequences and sequences known in *Actinobacteria*, *Bacteroidetes*, *Chlorobi*, *Cyanobacteria*, *Euryarchaeota*, *Firmicutes*, *Fusobacteria*, *Nitrospirae*, *Spirochaetes*, *Synergistetes*, and *Verrucomicrobia* were also detected.

For the NifH data, sequences contributing the most significantly to ordination of the intensive sites were closest matches to *Azospirillum*, *Bradyrhizobium*, *Rubrivivax*, *Leptothrix*, *Dechloromonas*, and *Geobacter* (Fig. S7). Of these, *Geobacter*-like sequences were present in all samples, but they were especially abundant in Michigan, representing more than 20% of the sequences in soils under prairie, mixed grasses, and switchgrass. *Bradyrhizobium* and *Rubrivivax*-related sequences were more characteristic of Wisconsin: *Rubrivivax*-like sequences accounted for more than 17% of the sequences from soils under corn, prairie, and switchgrass in that state. *Azospirillum*, *Dechloromonas,* and *Leptothrix*-like sequences did not distinguish samples by state, being more related to the pH gradient.

Genera contributing most significantly to the ordination of the extensive sites were *Geobacter*- and *Hyphomicrobium*-related (Fig. S8). *Geobacter*-like sequences represented 60% of the sequences from Michigan switchgrass samples, while *Hyphomicrobium*-like sequences were most abundant in prairie soil from both states.

*Microbial community structure*

Microbial data (lipid, 16S/18S rRNA, and NifH data) from the intensive sites were analyzed by PCA and PMANOVA (Table 1 and Fig. 4) and for differences in dispersion among factors. PMANOVA indicated that the 16S/18S rRNA and NifH data separated the samples by location only, and centroids are drawn for location only in Fig. 4b,c. The only significant difference in dispersion ($\alpha = 0.05$) was for the NifH data by location (Fig. 4c), indicating that the Wisconsin samples were also more variable. For the lipid data, however, the interaction term was significant, indicating a crop effect differing by state. Centroids drawn for all treatment combinations (Fig. 4a) depict no separation of Michigan samples by crop, but do reveal a separation of Wisconsin samples by crop. In particular, corn and to a lesser extent mixed grasses are separated from prairie and switchgrass.

In contrast to what was observed for the intensive sites, all four data types (lipid, 28S rRNA, 16S/18S rRNA, and NifH) separated communities by crop in the extensive sites (Table 2 and Fig. 5), and there were no significant differences in dispersion ($\alpha = 0.05$). The 28S rRNA gene data separated prairie and corn sites from each other, but samples from the switchgrass sites overlapped both (Fig. 5a). The lipid data separated corn from switchgrass and prairie (Fig. 5b). The 16S/18S rRNA gene data separated all three crops (Fig. 5d). For the NifH data, there was some overlap between the corn and Wisconsin switchgrass samples, but otherwise crops were separated (Fig. 5e). Additionally, the lipid and NifH data separated the samples by location (Table 2 and Fig. 5c,f).

*Procrustes analysis*

We performed Procrustes analyses to determine whether there were significant correlations between ordinations based on the four types of data (lipid, 28S rRNA, 16S/18S rRNA, and NifH). For the intensive sites, ordination by the lipid data was correlated with those by 16S/18S rRNA and by NifH, but ordinations by 16S/18S rRNA and NifH differed significantly. The difference was primarily due to two Michigan corn samples having a greater distance from their centroid by 16S/18S rRNA gene data than by NifH data (Fig. 4b, c), but this did not influence interpretation of results. For the extensive sites, ordinations were correlated with the exceptions of 16S/18S rRNA vs. NifH data and 16S/18S rRNA vs. 28S rRNA.

**Table 1** Results of multivariate analysis of variance by permutation (function *adonis* in VEGAN package) of sequence data for the intensive sites

| Source of variation | Lipids | | rRNA | | NifH | |
|---|---|---|---|---|---|---|
| | df† | *F* statistics | df | *F* statistics | df | *F* statistics |
| Location | 1 | $5.2^{***}$ | 1 | $2.9^{***}$ | 1 | $6.3^{***}$ |
| Crop | 3 | $2.2^{*}$ | 3 | $1.2^{NS}$ | 3 | $1.3^{NS}$ |
| Interaction | 3 | $2.7^{**}$ | 3 | $1.1^{NS}$ | 3 | $1.0^{NS}$ |
| Residuals | 13 | | 14 | | 15 | |
| Total | 20 | | 21 | | 22 | |

Significance codes: ***0.001; **0.01; *0.05; NS, non significant.
†Degrees of freedom.

## Discussion

The three methodological approaches we used to characterize microbial communities provided complimentary insights. Lipid analysis provided general taxonomic information coupled to biomass estimates that gave insight into ecosystem function (Kirk *et al.*, 2004). Moreover, the lipid data proved more sensitive to cropping system treatments showing the importance of management on ecosystem processes. Pyrosequencing the rRNA gene provided in-depth taxonomic information (Roesch *et al.*, 2007), and gene-targeted metagenomics provided information on a subset of the community responsible for a certain function (Iwai *et al.*, 2010, 2011). In our case, we targeted the *nifH* gene, which codes for dinitrogenase reductase, a component of nitrogenase, the enzyme responsible for $N_2$ fixation. $N_2$-fixing bacteria were chosen as a model to test the effect of cultivation on an important functional group as opposed to information provided by taxonomic markers. Indeed, there is evidence that perennial grasses with potential for biofuel production, such as *Miscanthus*, may be associated with $N_2$-fixing bacteria (Tjepkema & Burris, 1976; Davis *et al.*, 2010; Mao *et al.*, 2013; Keymer & Kent, 2014), which points to $N_2$-fixing microorganisms as an important target group. Additionally, we expected NifH to be less conserved than ribosomal genes, thus giving us a contrast to the highly conserved rRNA gene.

All three approaches revealed similar differences among cropping systems in community structure. This agreement between approaches indicates that similar factors are shaping the structure of bacteria, fungi, and $N_2$-fixing communities under our studied conditions and that disparate taxa are being affected similarly by cultivation, soil type, and land use. Differences in community structure could be linked to both treatment and environmental factors, but the relative importance of the linkages differed between intensive and extensive sites.

Our initial hypothesis that soil microbial communities under switchgrass and mixed grasses would be more

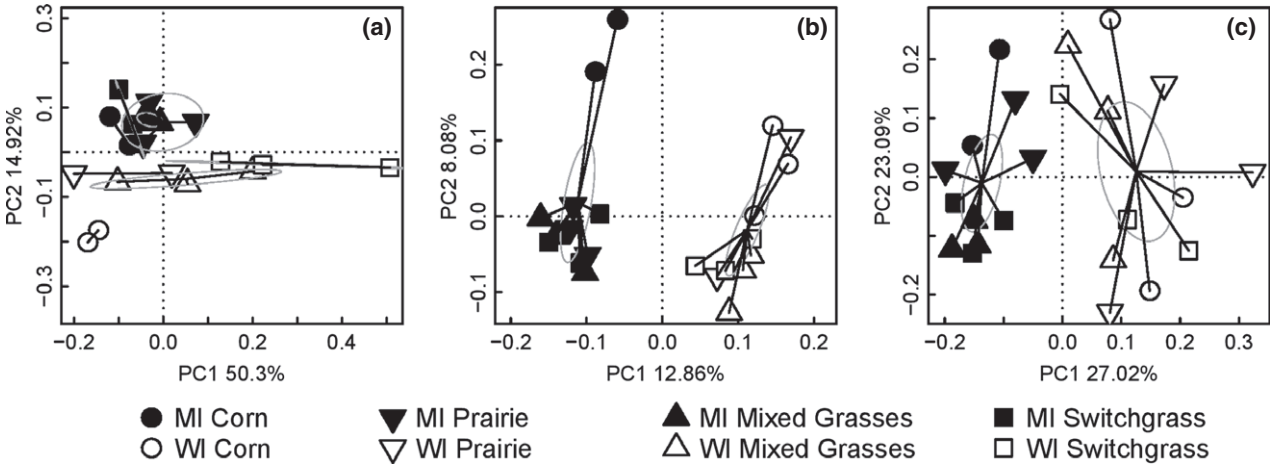


**Fig. 4** Principal components analysis of soil microbial communities from Michigan and Wisconsin intensive sites as evaluated by sequencing of SSU rRNA genes (a), lipid analysis, (b) and NifH sequences (c). Ellipses are 95% confidence intervals about centroid means and were drawn to indicate the main factors related to community structure.

**Table 2** Results of multivariate analysis of variance by permutation (function *adonis* in VEGAN package) of sequence data for the extensive sites

| | 28S rRNA | | Lipids | | 16S rRNA | | NifH | |
|---|---|---|---|---|---|---|---|---|
| Source of variation | df† | *F* statistics | df | *F* statistics | df | *F* statistics | df | *F* statistics |
| Location | 1 | $1.5^{NS}$ | 1 | $3.2^{*}$ | 1 | $1.0^{NS}$ | 1 | $2.5^{**}$ |
| Crop | 2 | $2.2^{***}$ | 2 | $8.0^{***}$ | 2 | $1.8^{**}$ | 2 | $2.6^{***}$ |
| Interaction | 2 | $1.0^{NS}$ | 2 | $1.7^{NS}$ | 2 | $1.0^{NS}$ | 2 | $1.3^{NS}$ |
| Residuals | 10 | | 11 | | 11 | | 12 | |
| Total | 15 | | 16 | | 16 | | 17 | |

Significance codes: ***0.001; **0.01; *0.05; NS, non significant.
†Degrees of freedom.

like those under prairies was confirmed in the older, extensive sites, but not in the young, intensive sites. In the second case, soil type was a stronger predictor of community structure and composition. Here, soil type is confounded with location (state), meaning that we cannot separate the effects of soil type from the effects of geographical distance. We assume, however, that soil type is the key factor because of the sharp difference in soil texture and fertility between our intensive sites in the two states (Fig. S2). The Michigan soils we studied are sandier and have lower fertility than the Wisconsin soils, which are loess-derived. Our results agree with those of Mao *et al.* (2013) who also compared microbial soil communities under biofuel crops by pyrosequencing 16S rRNA and *nifH* genes and found that site-to-site variation surpassed variation stemming from plant type.

In contrast, communities at the extensive sites tended to group more strongly by crop, indicating that plant species had a stronger influence on microbial communities as the plant communities effects on soil microbes accumulated over time. Our results were similar to those of Allison *et al.* (2005) and Mao *et al.* (2013) in that communities under corn were separated from those under perennial grasses by lipids and 16S rRNA gene analysis, respectively. Mao *et al.* (2013), however, were not able to detect differences in $N_2$-fixing communities between crops, while we did for the extensive sites.

The differential response to crops observed between intensive and extensive experiments is likely related to the length of time the crops had been grown at the sites. Previous experiments by Murphy & Foster (2014) and Buckley & Schmidt (2003) demonstrated that despite changes in plant cover and management, soil microbial communities remained similar even after 6 and 7 years, respectively. In another experiment, Jangid *et al.* (2011) found an even longer historical effects lasting through 17 years of succession in a previously cultivated field. At the time of our sampling, crops at the intensive sites had been cultivated for only 2 years, which likely was

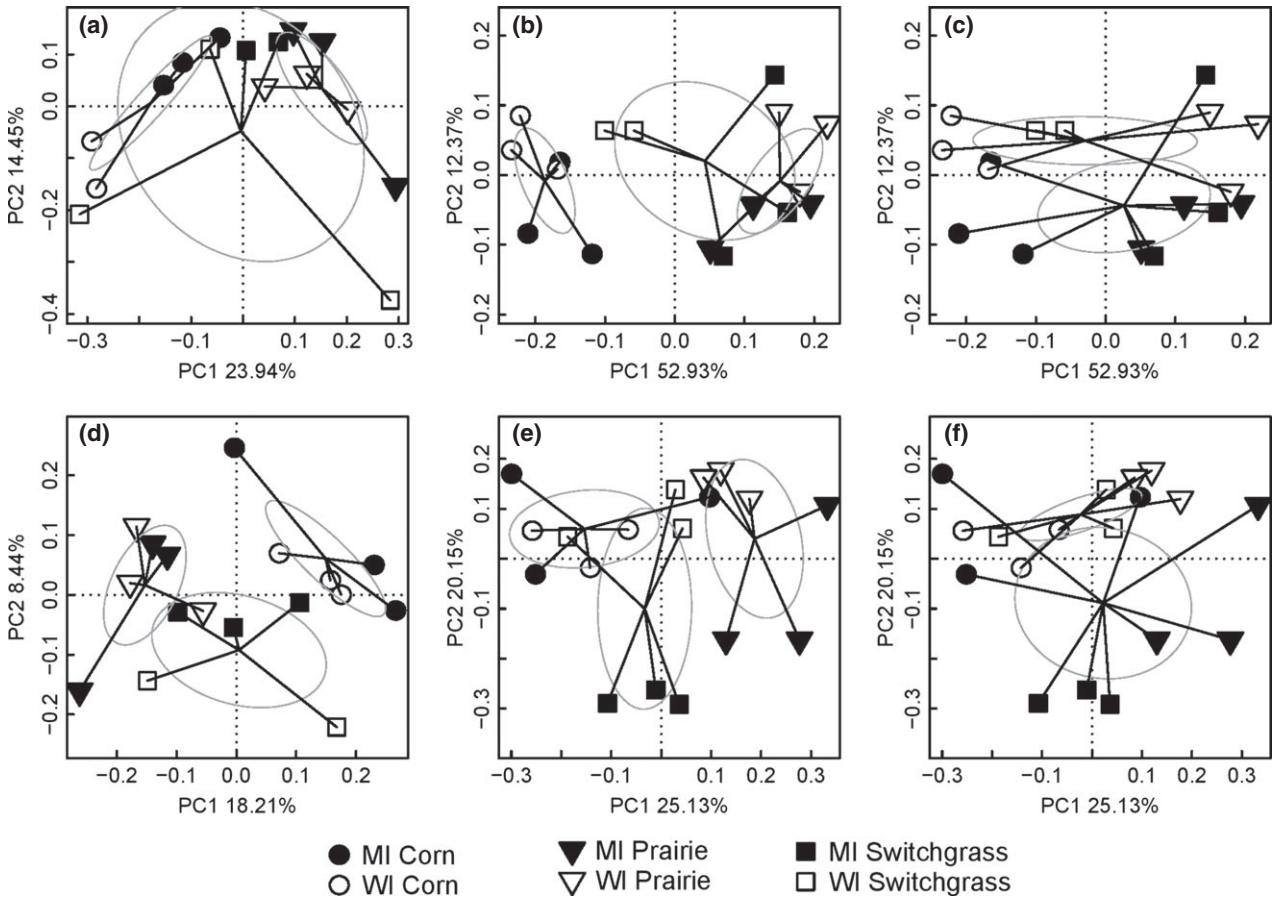


**Fig. 5** Principal components analysis of soil microbial communities from Michigan and Wisconsin extensive sites as evaluated by 28S rRNA gene sequences (a), lipid analysis (b and c), SSU rRNA gene sequences (d), and NifH protein sequences(e and f). Ellipses are 95% confidence intervals about centroid means and were drawn to indicate the main factors related to community structure.

not enough time to imprint significant differences on the structure of the communities. In contrast, the extensive sites were much older, >10 years, allowing time for a more pronounced differentiation of community structure according to crop type. Our findings reinforce the previous findings of these authors, but in the context of a larger geographical scale, and including two different cultivation settings and a diverse range of soil types, especially for the extensive sites.

The one exception to this generalization was that the lipid data did reveal a crop effect on microbial community structure for the intensive Wisconsin sites. This may have to do with the differences in relative proportion of AMF biomass between treatments for the two states. Bacterial and SF biomasses were similar among treatments for intensive sites in both states. For the richer Wisconsin soils, however, there was a greater relative difference in AMF biomass between corn and the other crops, especially switchgrass (Fig. 2b), which plots farthest from corn in Fig. 4a. Herzberger *et al.* (2014) also reported total biomass, and especially AMF biomass, was higher under restored prairie than corn two years after establishment at Wisconsin intensive sites. Community differences due to AMF would not be revealed by the rRNA or NifH data.

As previously observed (Liang *et al.*, 2012), when compared to corn, perennial grasses favored the accumulation of microbial biomass as well as an increase in F/B and AMF/SF ratios in both states, indicating that these grasses favor the accumulation of biomass and fungi, especially AMF. This was especially true among the extensive sites in Wisconsin, which had more time to become established than the intensive sites and had higher C (Fig. 1b) and clay (Table S1) contents than the Michigan sites. Soils with higher C and clay contents are often associated with higher microbial biomass because there is more C available for microbial growth (Bach *et al.*, 2010). On the other hand, the higher abundance of Gram-positive bacteria and actinomycetes under corn indicates a more stressful environment, because these organisms are known for thriving in

stressful environments and for producing spores (Yao *et al*., 2000; Fierer *et al*., 2003).

The perennial character of switchgrass and prairie assemblages likely contributes to the accumulation of fungal biomass, especially AMF biomass, while hyphae are disrupted by tillage at corn sites. Tillage breaks the hyphae of AMF and increases the decomposition of organic matter, which might decrease the abundance of AMF and favor saprophytic fungi. Mycorrhizal fungi are known to enhance P absorption and utilization, so an increase in AMF implies an improvement in the absorption and utilization efficiency of this nutrient. Additionally, recently reported results showed that mycorrhizae increased the amount of total N in shoots of switchgrass (Schroeder-Moreno *et al*., 2012), indicating that mycorrhizae can also contribute to the increased utilization of this nutrient. The higher fungal biomass under perennial grasses implies greater potential for C accumulation under these crops (Blanco-Canqui, 2010) because it is assumed that root production is higher (Zan *et al*., 2001), soil aggregates form more quickly and to a higher degree (Jastrow, 1987; Jarchow & Liebman, 2012) and that fungi produce C compounds more difficult to degrade (Allison *et al*., 2005). Tilman *et al*. (2006) also noted that C sequestration was higher under native grassland perennials than under corn, and Bailey *et al*. (2002) observed larger quantities of C and larger activity ratios in the soil of a restored prairie compared to a neighboring corn farmland.

Fungal species specific to each crop were all saprophytic fungi. Only a few sequences belonging to AMF were identified, but they did reveal some interesting patterns. Among these, there was a greater abundance and richness (number of OTUs) of AMF under prairies for both states, including groups known to have distinct functional traits, such as Paraglomerales and Diversisporales (van der Heijden & Scheublin, 2007). The higher number of OTUs observed under prairie may be a consequence of its higher plant diversity, because a greater diversity of plant hosts creates a greater diversity of niches for AMF. In addition, it is possible that the lower AMF richness under corn stemmed from higher soil disturbance in the cultivation of this crop, because there is evidence that soil disturbance was responsible for reducing the phylotype richness of AMF communities under seminatural grasslands (Schnoor *et al*., 2011). AMF was linked previously to higher plant productivity (Maherali & Klironomos, 2007), so this greater AMF richness under prairies may positively influence the productivity of the prairie vegetation. It is worth noting that most of the AMF sequences found at our studied sites, and especially those in the prairie soils, belonged to the genus *Paraglomus*. Previous researchers found Paraglomerales are difficult to detect in roots and soils and that commonly used primers for AMF fail to amplify Paraglomerales sequences (Lumini *et al*., 2010; Gosling *et al*., 2014). Thus, our data indicate that Midwestern prairies may be good places to study the diversity and ecology of this lesser known AMF genus.

There was an inconsistency in the AMF prevalence indicated by the lipid method where AMF were 41% of the fungal biomass and the pyrosequencing methods where only 0.13% of the fungal 28S rRNA sequences were assigned to Glomeromycota. A similarly low fraction of AMF was observed in the 18S rRNA data. Both SSU and LSU primers used were perfect matches to most known strains, so mismatch is not a likely explanation, although other biases have been observed in competitive rRNA gene amplification. One possible explanation is that the biomass (hyphae) measured by lipid may not be filled with protoplasm or with nuclei. Cytoplasmic streaming is known to occur in soil fungi, often resulting in evacuated hyphae because their protoplasm is concentrated at the growing tips (Klein & Paschke, 2004). The reason for this large discrepancy is important to resolve for proper accounting of this important group of soil fungi.

While we could specifically link SSU rRNA gene sequences to crops and soils, and given that these sequences provide better taxonomical resolution, the detection of large numbers of uncultured organisms and taxa with no clearly defined roles limited physiological and functional interpretation. Furthermore, the large number of sequences and OTUs made it difficult to detect relevant indicator organisms based on rRNA pyrosequencing data. This is a problem common to many SSU rRNA sequencing studies, especially for environmental samples. On the other hand, sequence assignment and the identification of indicator organisms were more informative with the NifH sequences, although some horizontal gene transfer may cloud precise taxonomic resolution.

Contrary to what we expected, the faster evolving protein coding gene *nifH* provided no better resolution than the other methods after conversion of nucleotide sequences to amino acid sequences, which was necessary to correct for sequencing errors as well as to reflect function. The low diversity we found for NifH may be the result of the primers we used not amplifying all *nifH* variants. The primers used are reported as being selective for *Proteobacteria* sequences (Diallo *et al*., 2008). However, it is worth noting that these authors used a previous PCR step, with different primers, which might have introduced extra bias into their PCR. In a more recent study, Gaby & Buckley (2012) reported that no *nifH* primers were comprehensive for the known *nifH* genes and that the primers developed by Poly *et al*. (2001) do exclude certain groups. A trade-off exists

between finding primers and conditions that give the best coverage for groups important in the habitat and reliable amplification. We decided the Poly primers were the current best choice because amplification with the broader coverage Zehr & McReynolds (1989) primers was troublesome, probably because of their high degeneracy.

We conclude that location, a proxy primarily for soil type but also including site history, landscape, and climate, was the major factor determining microbial communities in our 2-year-old intensive sites and that these study sites were not under cultivation long enough for the crop to impose a strong signature on the microbial communities. The only exception to this was that the lipid data revealed a crop effect in the richer Wisconsin soil. In contrast, when the same crop had been grown on a site for 10 years or longer, a crop effect was observed, with communities under corn clearly differentiated from those under perennial grasses. Both presence of perennial plants and higher plant diversity likely favored the accumulation of microbial biomass and fungi, especially AMF, under switchgrass, mixed grasses, and prairie, leading to a more stable environment and highlighting that these alternatives to corn for biofuels may improve soil functional stability and sustainability.

## Acknowledgements

This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494). We thank Gregg Sanford and Joe Simmons for managing the intensive cropping systems experiments in Wisconsin and Michigan, respectively. Thanks to Doug Landis and Ben Werling for help with sampling and providing information on the extensive sites in Michigan; Tim Meehan for help with sampling on the extensive sites in Wisconsin; Susanna Tringe, Stephanie Malfatti, Tijana Galvina del Rio at the Joint Genome Institute for SSU rRNA pyrotag sequencing; James Cole from the Ribosomal Database Project for support and comments with sequence data analysis; Harry Read for lipid analysis; and David Duncan for soil physicochemical analysis.

## References

Adler PR, Sanderson MA, Weimer PJ, Vogel KP (2009) Plant species composition and biofuel yields of conservation grasslands. *Ecological Applications*, **19**, 2202–2209.

Allison VJ, Miller RM, Jastrow JD, Matamala R, Zak DR (2005) Changes in soil microbial community structure in a tallgrass prairie chronosequence. *Soil Science Society of America Journal*, **69**, 1412–1421.

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, **26**, 32–46.

Anderson MJ (2006) Distance-based tests for homogeneity of multivariate dispersions. *Biometrics*, **62**, 245–253.

Bach EM, Baer SG, Meyer CK, Six J (2010) Soil texture affects soil microbial and structural recovery during grassland restoration. *Soil Biology and Biochemistry*, **42**, 2182–2191.

Bailey VL, Smith JL, Bolton H (2002) Fungal-to-bacterial ratios in soils investigated for enhanced C sequestration. *Soil Biology and Biochemistry*, **34**, 997–1007.

Blanco-Canqui H (2010) Energy crops and their implications on soil and environment. *Agronomy Journal*, **102**, 403–419.

Buckley DH, Schmidt TM (2003) Diversity and dynamics of microbial communities in soils from agro-ecosystems. *Environmental Microbiology*, **5**, 441–452.

Cox TF, Cox MAA (2001) *Monographs on Statistics and Applied Probability. Multidimensional Scaling*. Chapman & Hall/CRC, Boca Raton.

Dabney A, Storey JD, Warnes R (2012) *qvalue: Q-value estimation for false discovery rate control*. Available at: http://www.bioconductor.org/packages/release/bioc/html/qvalue.html (accessed 15 October 2012).

Davis SC, Parton WJ, Dohleman FG, Smith CM, Del Grosso S, Kent AD, DeLucia EH (2010) Comparative biogeochemical cycles of bioenergy crops reveal nitrogen-fixation and low greenhouse gas emissions in a *Miscanthus x giganteus* agro-ecosystem. *Ecosystems*, **13**, 144–156.

Diallo MD, Reinhold-Hurek B, Hurek T (2008) Evaluation of PCR primers for universal *nifH* gene targeting and for assessment of transcribed *nifH* pools in roots of *Oryza longistaminata* with and without low nitrogen input. *FEMS Microbiology Ecology*, **65**, 220–228.

Dufrene M, Legendre P (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs*, **67**, 345–366.

Engelbrektson A, Kunin V, Wrighton KC, Zvenigorodsky N, Chen F, Ochman H, Hugenholtz P (2010) Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME Journal*, **4**, 642–647.

Fargione J, Hill J, Tilman D, Polasky S, Hawthorne P (2008) Land clearing and the biofuel carbon debt. *Science*, **319**, 1235–1238.

Fierer N, Schimel JP, Holden PA (2003) Variations in microbial community composition through two soil depth profiles. *Soil Biology and Biochemistry*, **35**, 167–176.

Fierer N, Ladau J, Clemente JC *et al.* (2013) Reconstructing the microbial diversity and function of pre-agricultural tall grass prairie soils in the United States. *Science*, **342**, 621–624.

Fish JA, Chai B, Wang Q, Sun Y, Brown CT, Tiedje JM, Cole JR (2013) FunGene: the functional gene pipeline and repository. *Frontiers in Microbiology*, **4**, 291.

Gaby JC, Buckley DH (2012) A comprehensive evaluation of PCR primers to amplify the *nifH* gene of nitrogenase. *PLoS ONE*, **7**, e42149.

Gao J, Thelen KD, Hao XM (2013) Life cycle analysis of corn harvest strategies for bioethanol production. *Agronomy Journal*, **105**, 705–712.

Good IJ (1953) The population frequencies of species and the estimation of population parameters. *Biometrika*, **40**, 237–264.

Gosling P, Proctor M, Jones J, Bending GD (2014) Distribution and diversity of *Paraglomus* spp. in tilled agricultural soils. *Mycorrhiza*, **24**, 1–11.

Groom MJ, Gray EM, Townsend PA (2008) Biofuels and biodiversity: principles for creating better policies for biofuel production. *Conservation Biology*, **22**, 602–609.

van der Heijden MGA, Scheublin TR (2007) Functional traits in mycorrhizal ecology: their use for predicting the impact of arbuscular mycorrhizal fungal communities on plant growth and ecosystem functioning. *New Phytologist*, **174**, 244–250.

Herzberger AJ, Duncan DS, Jackson RD (2014) Bouncing back: plant-associated soil microbes respond rapidly to prairie establishment. *PLoS ONE*, **9**, e115775.

IEA (2007) Biofuel Production - ETE02. Energy Technology Essentials. Available at: http://www.iea.org/techno/essentials2.pdf (accessed 23 March 2012).

Iwai S, Chai BL, Sul WJ, Cole JR, Hashsham SA, Tiedje JM (2010) Gene-targeted-metagenomics reveals extensive diversity of aromatic dioxygenase genes in the environment. *ISME Journal*, **4**, 279–285.

Iwai S, Chai B, Jesus E, Penton CR, Lee TK, Cole JR, Tiedje JM (2011) Gene-targeted metagenomics (GT Metagenomics) to explore the extensive diversity of genes of interest in microbial communities. In: *Handbook of Molecular Microbial Ecololgy I - Metagenomics and Complementary Approaches* (ed. De Bruijn FJ), pp. 235–243. Wiley-Blackwell, Hoboken, NJ.

Jach-Smith LC, Jackson RD (2015) Switchgrass and prairie response to N fertilizer and harvest time on productive soils of upper Midwest. *Agriculture, Ecosystems and Environment*, **204**, 62–71.

Jangid K, Williams MA, Franzluebbers AJ, Schmidt TM, Coleman DC, Whitman WB (2011) Land-use history has a stronger impact on soil microbial community composition than aboveground vegetation and soil properties. *Soil Biology and Biochemistry*, **43**, 2184–2193.

Jarchow ME, Liebman M (2012) Tradeoffs in biomass and nutrient allocation in prairies and corn managed for bioenergy production. *Crop Science*, **52**, 1330–1342.

Jastrow JD (1987) Changes in soil aggregation associated with tallgrass prairie restoration. *American Journal of Botany*, **74**, 1656–1664.

Jesus ED, Susilawati E, Smith SL *et al.* (2010) Bacterial communities in the rhizosphere of biofuel crops grown on marginal lands as evaluated by 16S rRNA gene pyrosequences. *Bioenergy Research*, **3**, 20–27.

Keymer DP, Kent AD (2014) Contribution of nitrogen fixation to first year *Miscanthus x giganteus*. *Global Change Biology Bioenergy*, **6**, 577–586.

Kim S, Dale BE, Ong RG (2012) An alternative approach to indirect land use change: allocating greenhouse gas effects among different uses of land. *Biomass and Bioenergy*, **46**, 447–452.

Kindt R, Coe R (2005) *Tree Diversity Analysis. A Manual and Software for Common Statistical Methods for Ecological and Biodiversity Studies*. World Agroforestry Centre (ICRAF), Nairobi.

Kirk JL, Beaudette LA, Hart M, Moutoglis P, Khironomos JN, Lee H, Trevors JT (2004) Methods of studying soil microbial diversity. *Journal of Microbiological Methods*, **58**, 169–188.

Klein DA, Paschke MW (2004) Filamentous fungi: the indeterminate lifestyle and microbial ecology. *Microbial Ecology*, **47**, 224–235.

Kunin V, Hugenholtz P (2010) PyroTagger: a fast, accurate pipeline for analysis of rRNA amplicon pyrosequence data. The Open Journal, Article 1. Available at: http://www.theopenjournal.org/toj_articles/1#5 (accessed 15 January 2015).

Legendre P, Gallagher ED (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia*, **129**, 271–280.

Liang C, Jesus ED, Duncan DS, Jackson RD, Tiedje JM, Balser TC (2012) Soil microbial communities under model biofuel cropping systems in southern Wisconsin, USA: impact of crop species and soil properties. *Applied Soil Ecology*, **54**, 24–31.

Liu KL, Porras-Alfaro A, Kuske CR, Eichorst SA, Xie G (2012) Accurate, rapid taxonomic classification of fungal large-subunit rRNA genes. *Applied and Environmental Microbiology*, **78**, 1523–1533.

Lumini E, Orgiazzi A, Borriello R, Bonfante P, Bianciotto V (2010) Disclosing arbuscular mycorrhizal fungal biodiversity in soil through a land-use gradient using a pyrosequencing approach. *Environmental Microbiology*, **12**, 2165–2179.

Maherali H, Klironomos JN (2007) Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science*, **316**, 1746–1748.

Mao YJ, Yannarell AC, Mackie RI (2011) Changes in N-transforming Archaea and Bacteria in soil during the establishment of bioenergy crops. *PLoS ONE*, **6**, e24750.

Mao YJ, Yannarell AC, Davis SC, Mackie RI (2013) Impact of different bioenergy crops on N-cycling bacterial and archaeal communities in soil. *Environmental Microbiology*, **15**, 928–942.

Murphy CA, Foster BL (2014) Soil properties and spatial processes influence bacterial metacommunities within a grassland restoration experiment. *Restoration Ecology*, **22**, 685–691.

Oates LG, Duncan DS, Robertson GP, Gelfand I, Miller N, Jackson RD (2015) Nitrous oxide emissions during establishment of eight alternative cellulosic bioenergy crops in the North Central United States. *Global Change Biology Bioenergy*, doi: 10.1111/gcbb.12268.

Oksanen J, Blanchet FG, Kindt R *et al.* (2012) vegan Community Ecology Package. Available at: http://CRAN.R-project.org/package=vegan (accessed 12 October 2012).

Penton CR, St Louis D, Cole JR *et al.* (2013) Fungal diversity in permafrost and tallgrass prairie soils under experimental warming conditions. *Applied and Environmental Microbiology*, **79**, 7063–7072.

Poly F, Monrozier LJ, Bally R (2001) Improvement in the RFLP procedure for studying the diversity of *nifH* genes in communities of nitrogen fixers in soil. *Research in Microbiology*, **152**, 95–103.

R Core Team (2012) *R: A language and environment for statistical computing*, Vienna, Austria, R Foundation for Statistical Computing. Available at: http://www.R-project.org/ (accessed 15 October 2012).

Roberts DW (2012) *labdsv: Ordination and Multivariate Analysis for Ecology*. Available at: http://CRAN.R-project.org/package=labdsv (accessed 15 October 2012).

Roesch LF, Fulthorpe RR, Riva A *et al.* (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME Journal*, **1**, 283–290.

Schnoor TK, Lekberg Y, Rosendahl S, Olsson PA (2011) Mechanical soil disturbance as a determinant of arbuscular mycorrhizal fungal communities in semi-natural grassland. *Mycorrhiza*, **21**, 211–220.

Schroeder-Moreno MS, Greaver TL, Wang SX, Hu SJ, Rufty TW (2012) Mycorrhizal-mediated nitrogen acquisition in switchgrass under elevated temperatures and N enrichment. *Global Change Biology Bioenergy*, **4**, 266–276.

Swift MJ, Izac AMN, van Noordwijk M (2004) Biodiversity and ecosystem services in agricultural landscapes - are we asking the right questions? *Agriculture Ecosystems and Environment*, **104**, 113–134.

Tilman D, Hill J, Lehman C (2006) Carbon-negative biofuels from low-input high-diversity grassland biomass. *Science*, **314**, 1598–1600.

Tjepkema JD, Burris RH (1976) Nitrogenase activity associated with some Wisconsin prairie grasses. *Plant and Soil*, **45**, 81–94.

US-DOE (2011) *U.S. Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry*, Oak Ridge, TN, Oak Ridge National Laboratory. Available at: http://www1.eere.energy.gov/bioenergy/pdfs/billion_ton_update.pdf (accessed 15 January 2015).

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, **73**, 5261–5267.

Wang Q, Quensen JF, Fish JA, Lee TK, Sun YN, Tiedje JM, Cole JR (2013) Ecological patterns of nifH genes in four terrestrial climatic zones explored with targeted metagenomics using FrameBot, a new informatics tool. *mBio*, **4**, e00592–00513.

Webster CR, Flaspohler DJ, Jackson RD, Meehan TD, Gratton C (2010) Diversity, productivity and landscape-level effects in North American grasslands managed for biomass production. *Biofuels*, **1**, 451–461.

Werling BP, Dickson TL, Issacs R *et al.* (2014) Perennial grasslands enhance biodiversity and multiple ecosystem services in bioenergy landscapes. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 1652–1657.

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

Yao H, He Z, Wilson MJ, Campbell CD (2000) Microbial biomass and community structure in a sequence of soils with increasing fertility and changing land use. *Microbial Ecology*, **40**, 223–237.

Zan CS, Fyles JW, Girouard P, Samson RA (2001) Carbon sequestration in perennial bioenergy, annual corn and uncultivated systems in southern Quebec. *Agriculture Ecosystems and Environment*, **86**, 135–144.

Zehr JP, McReynolds LA (1989) Use of degenerate oligonucleotides for amplification of the *nifH* gene from the marine Cyanobacterium *Trichodesmium thiebautii*. *Applied and Environmental Microbiology*, **55**, 2522–2526.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Figure S1.** Locations of the sampling sites. Intensive sites were located at the Arlington Agricultural Research Station (AARS) in Wisconsin and at the Kellogg Biological Station (KBS) in Michigan (triangles). Other symbols show locations of extensive sites. The AARS symbol masks the symbol for one corn extensive site (WIE.Co.1) and one switchgrass extensive site (WIE.Sw.1).
**Figure S2.** PCA biplots for the intensive sites based on their soil attributes. Scaling 1 depicts distances between sites. Scaling 2 depicts higher correlations between variables as smaller angles between vectors.
**Figure S3.** PCA biplots for the extensive sites based on their soil attributes. Scaling 1 depicts distances between sites. Scaling 2 depicts higher correlations between variables as smaller angles between vectors. Environmental data were not available for sample WIE.Sw.2.
**Figure S4.** Relative abundances of fungal phyla at extensive sites determined from 28S rRNA gene sequences. Others includes Blastocladiomycota, Glomeromycota, and Neocallimastigomycota, each <1% of the total sequences in any sample. MIE = Michigan extensive sites; WIE = Wisconsin extensive sites; Co = corn; Pr = prairie; Sw = switchgrass. Bars represent ±1 standard error.
**Figure S5.** Relative abundances of prokaryotic and eukaryotic phyla determined from SSU rRNA gene sequences. MII = Michigan intensive sites; MIE = Michigan extensive sites; WII = Wisconsin intensive sites; WIE = Wisconsin extensive sites; Co = corn; Mp = mixed grasses; Pr = prairie; Sw = switchgrass. Bars represent ±1 standard error.
**Figure S6.** Percentages of total NifH gene sequences, binned by taxonomic class, among sample categories. Others includes sequences from *Actinobacteria*, *Bacteroidetes*, *Chlorobi*, *Cyanobacteria*, *Euryarchaeota*, *Firmicutes*, *Fusobacteria*, *Nitrospirae*, *Spirochaetes*, *Synergistetes*, and *Verrucomicrobia*. MII = Michigan intensive sites; MIE = Michigan extensive sites; WII = Wisconsin intensive sites; WIE = Wisconsin extensive sites; Co = corn; Mp = mixed grasses; Pr = prairie; Sw = switchgrass. Bars represent ±1 standard error.
**Figure S7.** PCA for intensive site NifH data aggregated by genera. IndVal results indicated *Geobacter* was significantly associated with Michigan sites and *Bradyrhizobium* and *Rubrivivax* were significantly associated with Wisconsin sites.
**Figure S8.** PCA for extensive site NifH data aggregated by genera. Genera contributing the most to the ordination were *Geobacter*, most abundant in the Michigan switchgrass samples, and *Hyphomicrobium*, characteristic of prairie sites in both states.
**Table S1.** Carbon, nitrogen, pH and texture of soils from Michigan (MI) and Wisconsin (WI) intensive and extensive sites. Crops are corn (CO), mixed grasses (MP), prairie (PR) and switchgrass (SW).
**Table S2.** Chemical attributes of soils from Michigan (MI) and Wisconsin (WI) intensive and extensive sites. Crops are corn (CO), mixed grasses (MP), prairie (PR) and switchgrass (SW).
**Table S3.** IndVal results based on 28S rRNA sequences for genera distinguishing extensive sites by crop. Indval analysis was limited to clusters with at least 10 sequences. Percentages were calculated based on all OTUs initially present. q cutoff was 0.05.

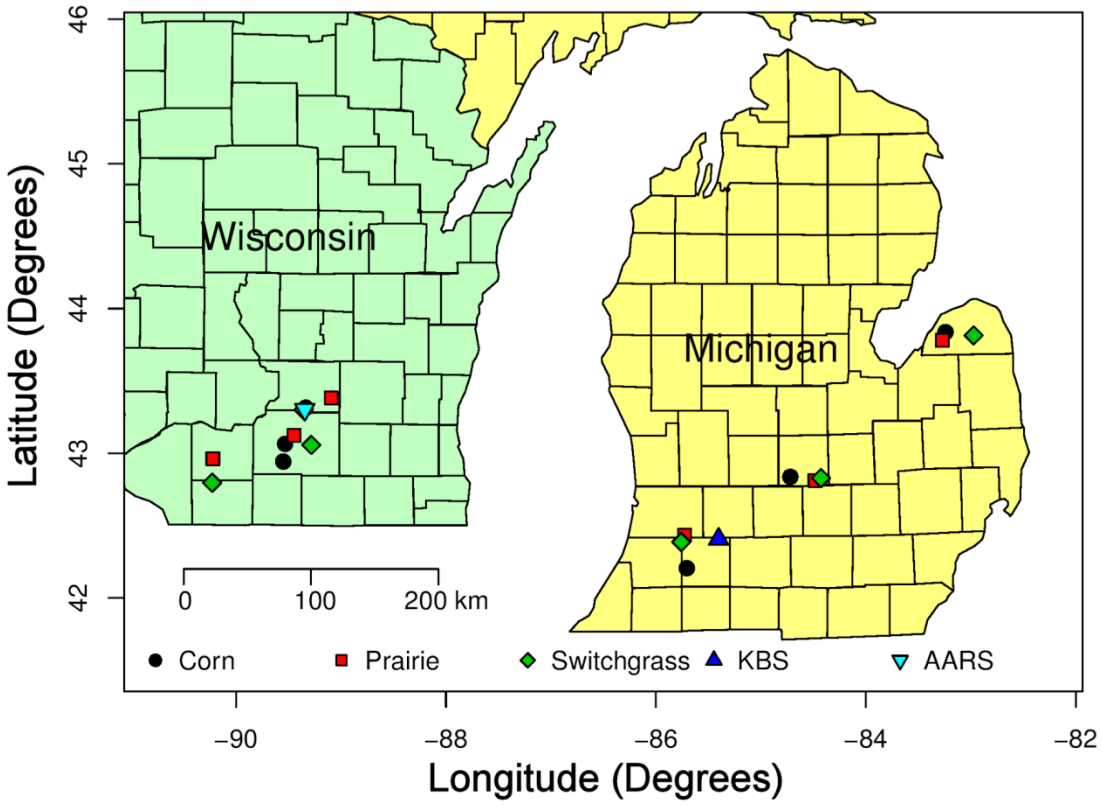


**Fig. S1: Locations of the sampling sites. Intensive sites were located at the Arlington Agricultural Research Station (AARS) in Wisconsin and at the Kellogg Biological Station (KBS) in Michigan (triangles). Other symbols show locations of extensive sites. The AARS symbol masks the symbol for one corn extensive site (WIE.Co.1) and one switchgrass extensive site (WIE.Sw.1).**

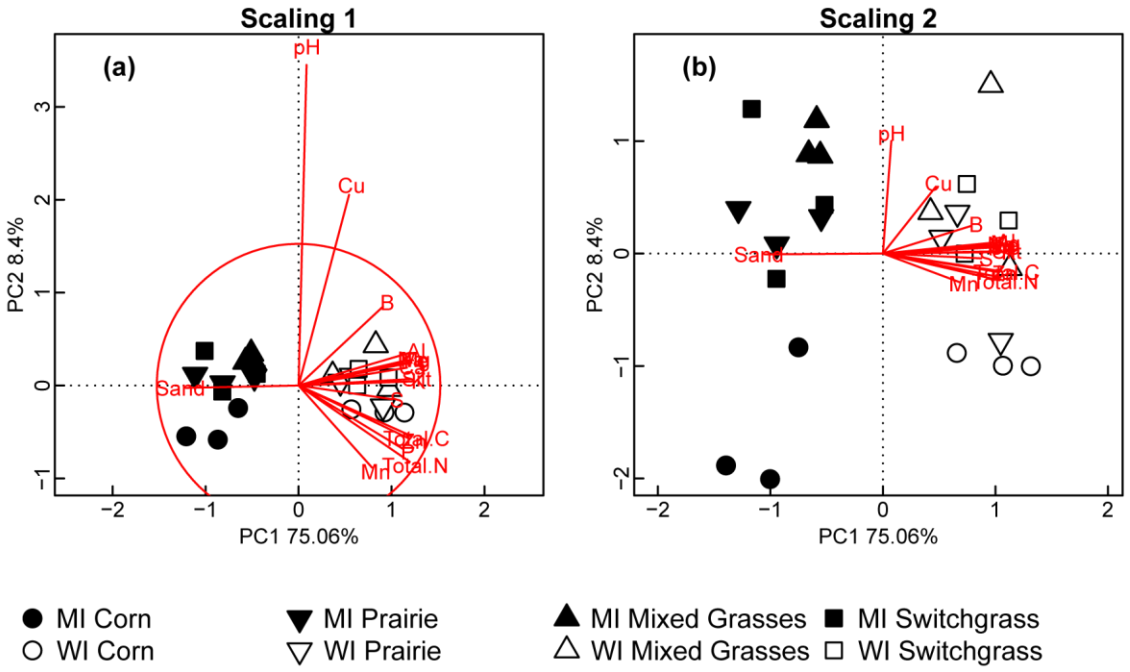


**Fig. S2: PCA biplots for the intensive sites based on their soil attributes. Scaling 1 depicts distances between sites. Scaling 2 depicts higher correlations between variables as smaller angles between vectors.**

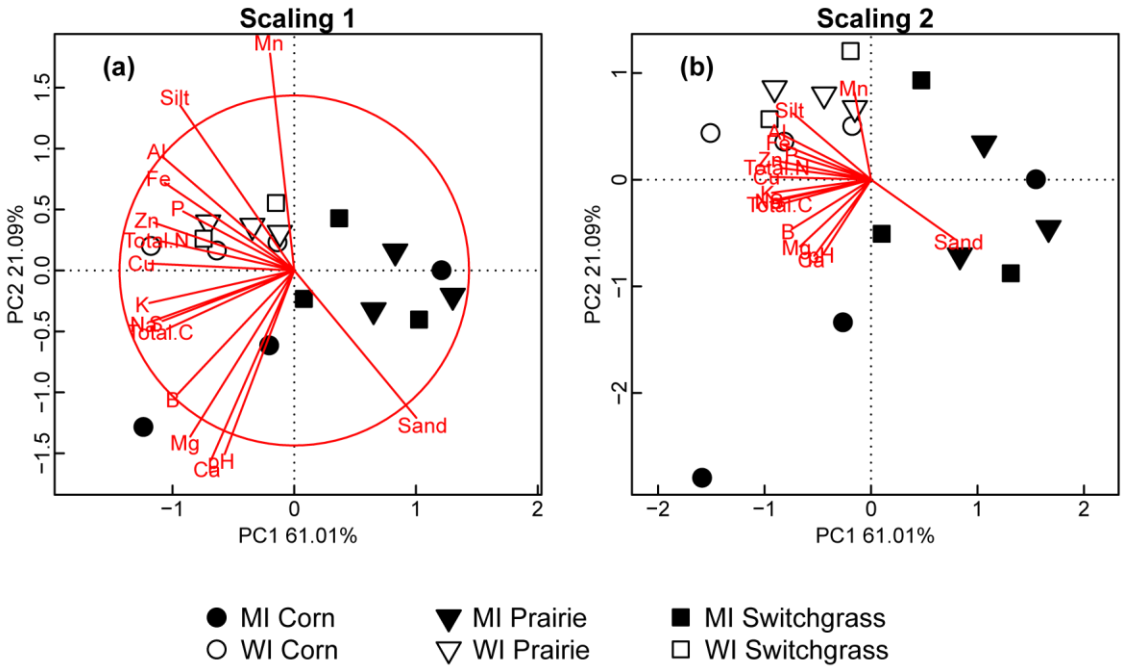


**Fig. S3: PCA biplots for the extensive sites based on their soil attributes. Scaling 1 depicts distances between sites. Scaling 2 depicts higher correlations between variables as smaller angles between vectors. Environmental data were not available for sample WIE.Sw.2.**

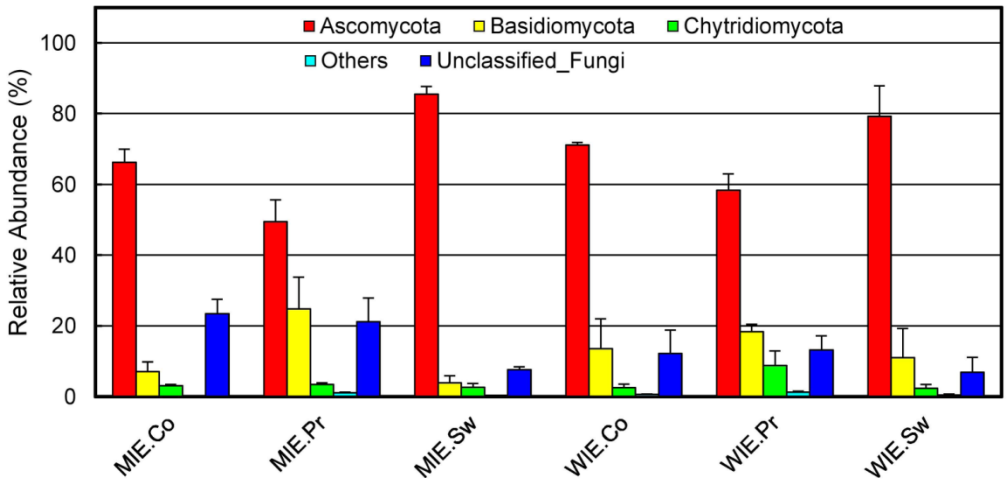


**Fig. S4: Relative abundances of fungal phyla at extensive sites determined from 28S rRNA gene sequences. Others includes Blastocladiomycota, Glomeromycota, and Neocallimastigomycota, each less than 1% of the total sequences in any sample. MIE = Michigan extensive sites; WIE = Wisconsin extensive sites; Co = corn; Pr = prairie; Sw = switchgrass. Bars represent ±1 standard error.**

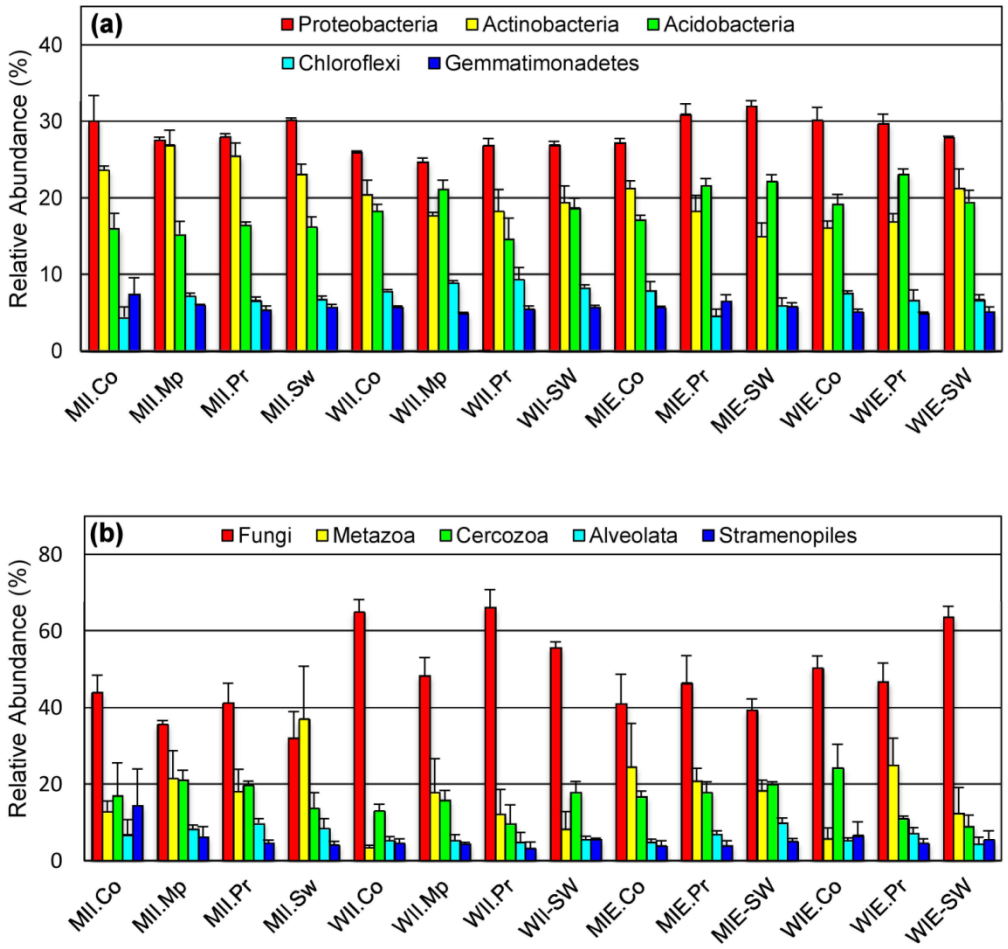


**Fig. S5: Relative abundances of prokaryotic and eukaryotic phyla determined from SSU rRNA sequences. MII = Michigan intensive sites; MIE = Michigan extensive sites; WII = Wisconsin intensive sites; WIE = Wisconsin extensive sites; Co = corn; Mp = mixed grasses; Pr = prairie; Sw = switchgrass. Bars represent ±1 standard error.**

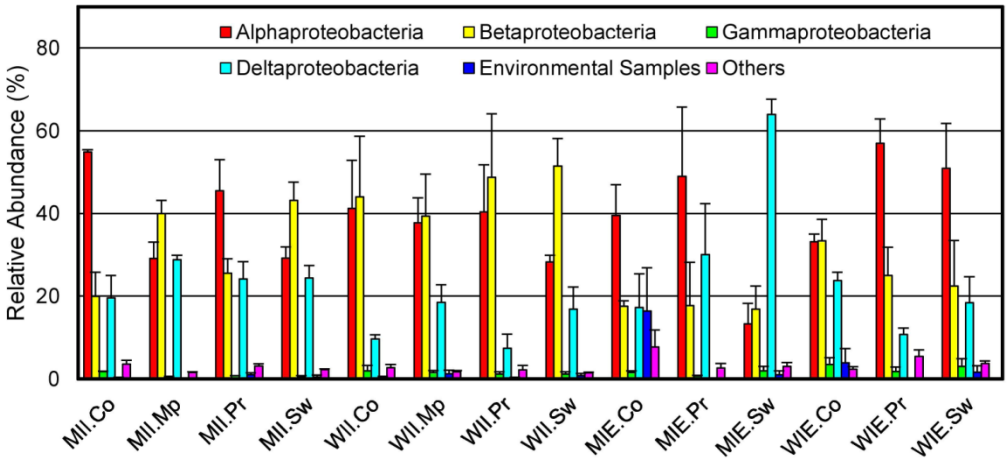


**Fig. S6: Percentages of total NifH sequences, binned by taxonomic class, among sample categories. Others includes sequences from *Actinobacteria*, *Bacteroidetes*, *Chlorobi*, *Cyanobacteria*, *Euryarchaeota*, *Firmicutes*, *Fusobacteria*, *Nitrospirae*, *Spirochaetes*, *Synergistetes*, and *Verrucomicrobia*. MII = Michigan intensive sites; MIE = Michigan extensive sites; WII = Wisconsin intensive sites; WIE = Wisconsin extensive sites; Co = corn; Mp = mixed grasses; Pr = prairie; Sw = switchgrass. Bars represent ±1 standard error.**

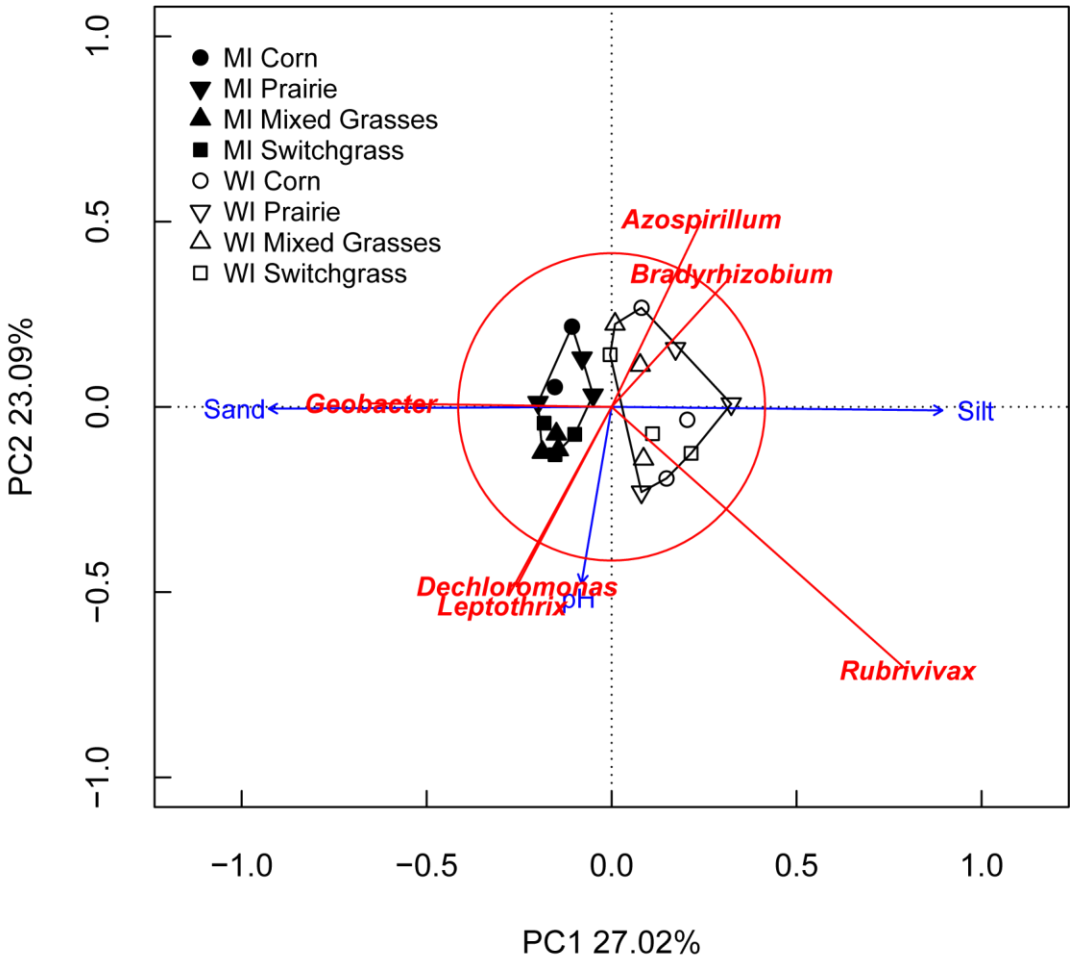


**Fig. S7: PCA for intensive site NifH data aggregated by genera. IndVal results indicated *Geobacter* was significantly associated with Michigan sites and *Bradyrhizobium* and *Rubrivivax* were significantly associated with Wisconsin sites.**

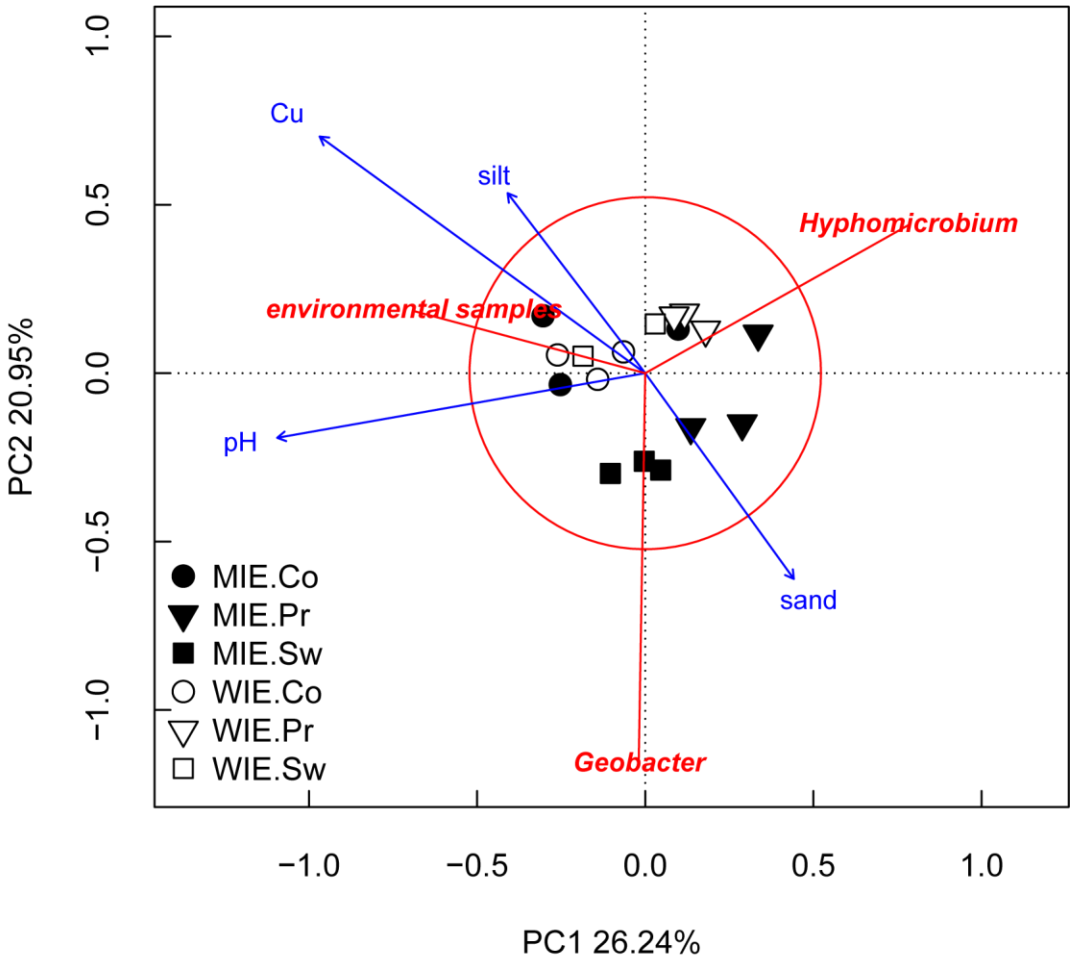


**Fig. S8: PCA for extensive site NifH data aggregated by genera. Genera contributing the most to the ordination were *Geobacter*, most abundant in the Michigan switchgrass samples, and *Hyphomicrobium*, characteristic of prairie sites in both states.**

**Table S1. Carbon, nitrogen, pH and texture of soils from Michigan (MI) and Wisconsin (WI) intensive and extensive sites. Crops are corn (CO), mixed grasses (MP), prairie (PR) and switchgrass (SW).**

| Intensive sites | Total C (%) | Total N (%) | pH | Sand (%) | Silt (%) | Clay (%) | Texture |
|---|---|---|---|---|---|---|---|
| MII-CO | 1.23 ± 0.09 | 0.11 ± 0.01 | 6.20 ± 0.21 | 56 ± 6 | 33 ± 6 | 11 ± 1 | Sandy Loam, Loam |
| MII-MP | 1.34 ± 0.04 | 0.11 ± 0.00 | 7.00 ± 0.06 | 47 ± 0 | 42 ± 1 | 11 ± 1 | Loam |
| MII-PR | 1.24 ± 0.06 | 0.10 ± 0.01 | 6.83 ± 0.03 | 54 ± 7 | 36 ± 6 | 10 ± 1 | Sandy Loam, Loam |
| MII-SW | 1.21 ± 0.07 | 0.09 ± 0.01 | 6.90 ± 0.06 | 53 ± 6 | 35 ± 5 | 12 ± 1 | Sandy Loam, Loam |
| WII-CO | 2.61 ± 0.12 | 0.27 ± 0.02 | 6.43 ± 0.03 | 9 ± 1 | 69 ± 1 | 22 ± 0 | Silt Loam |
| WII-MP | 2.22 ± 0.24 | 0.22 ± 0.03 | 6.97 ± 0.03 | 9 ± 1 | 68 ± 1 | 23 ± 2 | Silt Loam |
| WII-PR | 2.35 ± 0.14 | 0.24 ± 0.03 | 6.73 ± 0.12 | 9 ± 1 | 67 ± 1 | 24 ± 0 | Silt Loam |
| WII-SW | 2.30 ± 0.11 | 0.23 ± 0.02 | 6.87 ± 0.03 | 8 ± 0 | 67 ± 1 | 25 ± 1 | Silt Loam |
| Extensive sites | | | | | | | |
| MIE-CO | 2.07 ± 0.78 | 0.15 ± 0.05 | 7.10 ± 0.45 | 51 ± 8 | 32 ± 5 | 18 ± 4 | Sandy Loam, Loam |
| MIE-PR | 1.48 ± 0.17 | 0.11 ± 0.02 | 6.27 ± 0.28 | 60 ± 9 | 30 ± 7 | 10 ± 3 | Sandy Loam, Loamy Sand, Loam |
| MIE-SW | 1.44 ± 0.08 | 0.12 ± 0.01 | 6.67 ± 0.35 | 45 ± 10 | 40 ± 7 | 15 ± 4 | Sandy Loam, Loamy Sand, Loam |
| WIE-CO | 2.16 ± 0.30 | 0.22 ± 0.04 | 6.80 ± 0.21 | 12 ± 4 | 66 ± 5 | 21 ± 1 | Silt Loam |
| WIE-PR | 2.06 ± 0.37 | 0.20 ± 0.05 | 6.13 ± 0.13 | 13 ± 0 | 67 ± 3 | 20 ± 3 | Silt Loam |
| WIE-SW | 2.36 ± 0.46 | 0.24 ± 0.04 | 6.60 ± 0.20 | 10 ± 3 | 72 ± 1 | 18 ± 2 | Silt Loam |

**Table S2. Chemical attributes of soils from Michigan (MI) and Wisconsin (WI) intensive and extensive sites. Crops are corn (CO), mixed grasses (MP), prairie (PR) and switchgrass (SW).**

| | P | K | Ca | Mg | S | Zn | B | Mn | Fe | Cu | Al | Na |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Intensive sites | % | % | % | % | % | ppm | ppm | ppm | ppm | ppm | ppm | ppm |
| MII-CO | 0.04 ± 0.003 | 0.11 ± 0.01 | 0.20 ± 0.01 | 0.15 ± 0.01 | 0.01 ± 0.001 | 37.88 ± 1.20 | 6.31 ± 0.45 | 624 ± 94 | 9816 ± 746 | 5.17 ± 0.18 | 10775 ± 1052 | 49 ± 3 |
| MII-MP | 0.04 ± 0.001 | 0.15 ± 0.00 | 0.24 ± 0.01 | 0.19 ± 0.00 | 0.02 ± 0.000 | 43.73 ± 0.90 | 6.21 ± 0.54 | 634 ± 27 | 11118 ± 152 | 7.88 ± 0.37 | 13323 ± 134 | 52 ± 1 |
| MII-PR | 0.04 ± 0.003 | 0.13 ± 0.01 | 0.23 ± 0.02 | 0.17 ± 0.01 | 0.01 ± 0.001 | 39.67 ± 3.73 | 6.79 ± 0.25 | 515 ± 133 | 10279 ± 327 | 6.27 ± 0.34 | 11985 ± 551 | 50 ± 5 |
| MII-SW | 0.04 ± 0.001 | 0.13 ± 0.01 | 0.23 ± 0.01 | 0.17 ± 0.00 | 0.01 ± 0.001 | 39.31 ± 2.93 | 6.04 ± 0.19 | 691 ± 148 | 10086 ± 401 | 7.04 ± 0.65 | 11870 ± 616 | 53 ± 3 |
| WII-CO | 0.10 ± 0.014 | 0.29 ± 0.01 | 0.39 ± 0.02 | 0.29 ± 0.01 | 0.02 ± 0.002 | 80.79 ± 7.75 | 9.53 ± 0.23 | 858 ± 53 | 14761 ± 279 | 7.39 ± 0.62 | 18844 ± 309 | 77 ± 3 |
| WII-MP | 0.08 ± 0.018 | 0.27 ± 0.02 | 0.39 ± 0.04 | 0.31 ± 0.01 | 0.02 ± 0.002 | 66.94 ± 9.51 | 14.04 ± 2.82 | 701 ± 72 | 14876 ± 439 | 6.67 ± 0.62 | 18656 ± 749 | 86 ± 6 |
| WII-PR | 0.08 ± 0.007 | 0.28 ± 0.01 | 0.36 ± 0.01 | 0.29 ± 0.01 | 0.02 ± 0.002 | 68.91 ± 4.22 | 9.54 ± 0.65 | 759 ± 31 | 14931 ± 96 | 7.16 ± 0.25 | 18876 ± 358 | 74 ± 4 |
| WII-SW | 0.08 ± 0.010 | 0.29 ± 0.01 | 0.41 ± 0.06 | 0.33 ± 0.03 | 0.02 ± 0.000 | 72.33 ± 3.82 | 10.55 ± 0.40 | 768 ± 21 | 15338 ± 327 | 7.23 ± 0.45 | 19584 ± 517 | 76 ± 1 |
| Extensive sites | | | | | | | | | | | | |
| MIE-CO | 0.06 ± 0.007 | 0.26 ± 0.08 | 1.36 ± 0.91 | 0.58 ± 0.32 | 0.02 ± 0.011 | 57 ± 6.25 | 12.33 ± 3.97 | 429 ± 142 | 11881 ± 1850 | 9.10 ± 2.54 | 13556 ± 2738 | 84 ± 27 |
| MIE-PR | 0.04 ± 0.003 | 0.14 ± 0.05 | 0.19 ± 0.01 | 0.18 ± 0.03 | 0.01 ± 0.002 | 37 ± 1.10 | 7.36 ± 2.07 | 594 ± 150 | 9654 ± 919 | 3.70 ± 1.17 | 10012 ± 1358 | 54 ± 8 |
| MIE-SW | 0.05 ± 0.013 | 0.17 ± 0.04 | 0.24 ± 0.05 | 0.20 ± 0.04 | 0.01 ± 0.001 | 44 ± 6.39 | 7.61 ± 1.48 | 572 ± 238 | 10733 ± 1496 | 5.64 ± 2.12 | 11627 ± 2504 | 61 ± 12 |
| WIE-CO | 0.09 ± 0.028 | 0.30 ± 0.04 | 0.44 ± 0.08 | 0.36 ± 0.02 | 0.03 ± 0.004 | 64 ± 5.67 | 11.23 ± 0.72 | 806 ± 44 | 15711 ± 564 | 11.43 ± 1.57 | 22956 ± 213 | 98 ± 9 |
| WIE-PR | 0.08 ± 0.006 | 0.29 ± 0.02 | 0.32 ± 0.03 | 0.32 ± 0.02 | 0.02 ± 0.004 | 62 ± 5.52 | 10.92 ± 0.80 | 785 ± 94 | 16490 ± 419 | 10.03 ± 0.63 | 22525 ± 628 | 83 ± 1 |
| WIE-SW | 0.13 ± 0.065 | 0.23 ± 0.04 | 1.14 ± 0.77 | 0.30 ± 0.02 | 0.12 ± 0.096 | 57 ± 6.58 | 10.51 ± 1.63 | 875 ± 279 | 14394 ± 1213 | 9.83 ± 0.94 | 17132 ± 4172 | 80 ± 8 |

**Table S3. IndVal results based on 28S rRNA sequences for genera distinguishing extensive sites by crop. Indval analysis was limited to clusters with at least 10 sequences. Percentages were calculated based on all OTUs initially present. q cut-off was 0.05.**

| | | Indval Result | | | | Mean Percentage of Sequences per Sample | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| OTU | genus | group | indval | pvalue | freq | MIE.Co | MIE.Pr | MIE.Sw | WIE.Co | WIE.Pr | WIE.Sw |
| OTU0665 | Ascobolus | Corn | 0.9292 | 0.0052 | 14 | 2.5860 | 0.0064 | 0.0593 | 8.6197 | 0.1450 | 0.1907 |
| OTU1029 | Podospora | Corn | 0.9044 | 0.0086 | 13 | 0.3414 | 0.0147 | 0.0129 | 1.1564 | 0.0059 | 0.0309 |
| OTU1683 | Coprinellus | Corn | 0.8276 | 0.0012 | 12 | 0.2015 | 0.0024 | 0.0041 | 0.4254 | 0.0180 | 0.0083 |
| OTU0668 | Ascodesmis | Corn | 0.8257 | 0.0026 | 7 | 0.4348 | 0.0024 | 0.0000 | 0.0382 | 0.0614 | 0.0000 |
| OTU0706 | Byssonectria | Corn | 0.7889 | 0.0043 | 5 | 0.1062 | 0.0000 | 0.0044 | 0.2758 | 0.0000 | 0.0000 |
| OTU1803 | Camarophyllopsis | Prairie | 1.0000 | 0.0003 | 6 | 0.0000 | 0.0388 | 0.0000 | 0.0000 | 0.0350 | 0.0000 |
| OTU1337 | unclassified_Ceratobasidiaceae | Prairie | 0.9742 | 0.0016 | 8 | 0.0000 | 0.2608 | 0.0132 | 0.0000 | 0.1517 | 0.0055 |
| OTU1715 | Lagarobasidium | Prairie | 0.9606 | 0.0013 | 11 | 0.0035 | 0.2356 | 0.0148 | 0.0066 | 0.2889 | 0.0055 |
| OTU0601 | Tricladium | Prairie | 0.9563 | 0.0000 | 11 | 0.0439 | 2.0780 | 0.1914 | 0.0000 | 1.6480 | 0.1155 |
| OTU0035 | Scolecobasidiella | Prairie | 0.9518 | 0.0024 | 8 | 0.0212 | 0.0977 | 0.0066 | 0.0000 | 0.3181 | 0.0000 |
| OTU2275 | unclassified_Glomeromycetes | Prairie | 0.9412 | 0.0004 | 7 | 0.0000 | 0.0432 | 0.0000 | 0.0000 | 0.0314 | 0.0055 |
| OTU2115 | unclassified_Tremellaceae | Prairie | 0.9363 | 0.0016 | 13 | 0.0231 | 0.7304 | 0.0281 | 0.0266 | 0.1225 | 0.0496 |
| OTU1723 | Sebacina | Prairie | 0.9298 | 0.0013 | 7 | 0.0000 | 0.2071 | 0.0750 | 0.0000 | 0.3873 | 0.0000 |
| OTU1444 | unclassified_Entolomataceae | Prairie | 0.9243 | 0.0001 | 14 | 0.0123 | 0.8034 | 0.0880 | 0.0133 | 0.2614 | 0.0215 |
| OTU2195 | Kappamyces | Prairie | 0.8821 | 0.0024 | 14 | 0.0196 | 0.6017 | 0.0947 | 0.0199 | 0.0761 | 0.0268 |
| OTU2208 | Spizellomyces | Prairie | 0.8505 | 0.0028 | 8 | 0.0000 | 0.0717 | 0.0000 | 0.0266 | 0.0463 | 0.0053 |
| OTU2120 | unclassified_Tremellomycetes | Prairie | 0.8470 | 0.0055 | 12 | 0.0141 | 0.4020 | 0.0838 | 0.1562 | 0.2264 | 0.0168 |
| OTU1722 | Piriformospora | Prairie | 0.8442 | 0.0015 | 8 | 0.0266 | 0.1270 | 0.0000 | 0.0000 | 0.0220 | 0.0000 |
| OTU0185 | Neottiosporina | Prairie | 0.8333 | 0.0023 | 5 | 0.0000 | 0.0805 | 0.0000 | 0.0000 | 0.1079 | 0.0000 |
| OTU1900 | unclassified_Tricholomataceae | Prairie | 0.8318 | 0.0005 | 12 | 0.0141 | 0.1761 | 0.0417 | 0.0066 | 0.1887 | 0.0141 |
| OTU1339 | Uthatobasidium | Prairie | 0.8246 | 0.0095 | 7 | 0.0000 | 1.0559 | 0.0132 | 0.0000 | 0.0888 | 0.0107 |
| OTU0223 | Septoria | Prairie | 0.8213 | 0.0032 | 6 | 0.0000 | 0.0224 | 0.0066 | 0.0000 | 0.1499 | 0.0000 |
| OTU1788 | uncultured_Thelephoraceae | Prairie | 0.8207 | 0.0044 | 6 | 0.0000 | 0.0514 | 0.0000 | 0.0066 | 0.1631 | 0.0000 |
| OTU2273 | unclassified_Diversisporales | Prairie | 0.8197 | 0.0012 | 8 | 0.0000 | 0.0473 | 0.0000 | 0.0066 | 0.0112 | 0.0028 |
| OTU0905 | unclassified_Chaetosphaeriaceae | Prairie | 0.8182 | 0.0015 | 10 | 0.0000 | 0.0674 | 0.0287 | 0.0000 | 0.0930 | 0.0163 |
| OTU2109 | Asterotremella | Prairie | 0.8170 | 0.0081 | 10 | 0.0071 | 0.1021 | 0.0133 | 0.0266 | 0.0538 | 0.0056 |
| OTU1436 | Entoloma | Prairie | 0.8114 | 0.0027 | 6 | 0.0000 | 0.1046 | 0.0044 | 0.0000 | 0.0295 | 0.0000 |
| OTU0331 | Cladophialophora | Prairie | 0.8030 | 0.0073 | 12 | 0.0318 | 0.8913 | 0.1920 | 0.0000 | 0.3597 | 0.0270 |
| OTU0200 | Paraphaeosphaeria | Prairie | 0.7986 | 0.0056 | 6 | 0.0000 | 0.1006 | 0.0000 | 0.0000 | 0.0951 | 0.0107 |
| OTU0337 | Sarcinomyces | Prairie | 0.7957 | 0.0094 | 7 | 0.0141 | 0.0160 | 0.0265 | 0.0000 | 0.4729 | 0.0000 |
| OTU1913 | unclassified_Agaricomycetes | Prairie | 0.7933 | 0.0107 | 17 | 0.1753 | 0.7931 | 0.7391 | 0.3905 | 4.2167 | 1.2139 |
| OTU1346 | Clavaria | Prairie | 0.7903 | 0.0049 | 6 | 0.0000 | 2.3848 | 0.0000 | 0.0000 | 1.9438 | 0.4317 |
| OTU1511 | Myxarium | Prairie | 0.7802 | 0.0033 | 7 | 0.0088 | 0.1021 | 0.0000 | 0.0000 | 0.0715 | 0.0056 |
| OTU2246 | leaf | Prairie | 0.7799 | 0.0023 | 15 | 0.0370 | 0.5289 | 0.3685 | 0.0582 | 0.5531 | 0.0641 |
| OTU2270 | Paraglomus | Prairie | 0.7797 | 0.0043 | 11 | 0.0000 | 0.1367 | 0.0483 | 0.0000 | 0.1528 | 0.0993 |
| OTU1164 | unclassified_Xylariales | Prairie | 0.7654 | 0.0004 | 13 | 0.0833 | 0.2391 | 0.0482 | 0.0249 | 0.1240 | 0.0162 |
| OTU0624 | Lachnum | Prairie | 0.7652 | 0.0046 | 8 | 0.0000 | 0.2176 | 0.1990 | 0.0000 | 0.2107 | 0.0053 |
| OTU1131 | Hilberina | Prairie | 0.7520 | 0.0098 | 7 | 0.0000 | 0.0719 | 0.0949 | 0.0000 | 0.4228 | 0.0000 |
| OTU2119 | unclassified_Tremellales | Prairie | 0.7323 | 0.0017 | 12 | 0.0106 | 0.1453 | 0.0751 | 0.0631 | 0.0804 | 0.0053 |
| OTU0638 | Mitrula | Prairie | 0.7313 | 0.0058 | 9 | 0.0000 | 0.0761 | 0.1083 | 0.0000 | 0.0943 | 0.0055 |
| OTU0642 | Torrendiella | Prairie | 0.7246 | 0.0067 | 6 | 0.0071 | 0.0259 | 0.0000 | 0.0000 | 0.0303 | 0.0000 |
| OTU1107 | Papulosa | Prairie | 0.7161 | 0.0101 | 8 | 0.0035 | 0.0629 | 0.0000 | 0.0000 | 0.0803 | 0.0295 |
| OTU0647 | unclassified_Helotiales | Prairie | 0.6885 | 0.0053 | 17 | 1.1430 | 2.3123 | 1.0967 | 0.6464 | 2.2541 | 0.4398 |
| OTU0305 | unclassified_Dothideomycetes | Prairie | 0.6674 | 0.0082 | 17 | 0.2034 | 0.3317 | 0.1334 | 0.1329 | 0.5583 | 0.1431 |
| OTU0937 | Beauveria | Switchgrass | 0.8372 | 0.0076 | 9 | 0.0000 | 0.0123 | 0.1160 | 0.0000 | 0.0199 | 0.0334 |
| OTU1161 | unclassified_Sordariales | Switchgrass | 0.6238 | 0.0086 | 17 | 0.2174 | 0.2109 | 1.1582 | 0.1811 | 0.2754 | 0.3632 |