Edaphic correlates of feedstock-associated diazotroph communities

CHINMAY SOMAN1, DANIEL P. KEYMER2,* and ANGELA D. KENT1,2

1Department of Natural Resources and Environmental Sciences, University of Illinois at Urbana-Champaign, W-503 Turner Hall, 1102 South Goodwin Ave, Urbana, IL 61801, USA, 2Energy Biosciences Institute, University of Illinois at Urbana-Champaign, W-503 Turner Hall, 1102 South Goodwin Ave, Urbana, IL 61801, USA

Abstract

Miscanthus × giganteus and Panicum virgatum are potential promising bioenergy feedstock crops suitable for the temperate zone. The energy efficiency and sustainability of bioenergy production could be improved by reducing their fertilizer inputs—particularly energy intensive nitrogen fertilizers. Miscanthus is known to benefit from nitrogen fixation by associative diazotrophs. However, because the effects of edaphic-, management-, and plant-related factors on feedstock-associated diazotroph communities have not yet been characterized, it is not currently possible to optimize the nitrogen contribution to feedstock crops from associated diazotroph communities. To address this critical knowledge gap, we characterized the bacterial and diazotroph communities in the rhizosphere and endophytic compartments of both species at eight research sites across Illinois. We also quantified the nifH gene abundance in the rhizosphere soil as well as a range of soil chemistry parameters at these sites. Multivariate statistical analyses revealed that diazotroph and bacterial communities in the rhizosphere varied primarily among sites, with very small differences between host species. Conversely, diazotroph and bacterial communities in the endophytic compartments differed significantly between plant species, but did not vary substantially among sites. Finally, nifH gene abundance in the rhizospheres of both species varied substantially from site to site and was positively correlated with soil iron concentration as well as soil ammonium concentration, and negatively correlated with abundance of other soil nutrients including calcium, total nitrogen, and nitrates. These results indicate the potential edaphic drivers of associative diazotroph communities in feedstock rhizospheres and suggest that manipulating bioavailable iron content in the soil is a potential direction for investigating the optimization of these communities to improve their nitrogen contribution to crops.

Keywords: diazotrophs, endophytic bacteria, miscanthus, nifH gene, rhizosphere microbiome, soil chemistry, switchgrass

Received 9 June 2017; accepted 6 November 2017

Introduction

Diazotrophs are bacteria and archaea that convert atmospheric nitrogen to ammonia, thus making it bioavailable to other life forms. Communities of diazotrophs in soils and in plant tissues could play an integral role in enhancing the sustainability of biofuels production by reducing the nitrogen fertilizer needed for producing feedstock crops. Biofuels are expected to constitute a growing portion of the global energy infrastructure due to government mandates and incentives (Dimaranan & Laborde, 2012; Timilsina, 2014). Cellulosic biofuels, such as those from Miscanthus × giganteus (miscanthus) and Panicum virgatum (switchgrass), have the potential to have significantly lower—perhaps even negative—carbon intensity compared with conventional biofuels (Qin et al., 2006; Adler et al., 2007; Davis et al., 2009; Dimaranan & Laborde, 2012; Sanscartier et al., 2014; Timilsina, 2014). However, nitrogen fertilizers, which can improve feedstock yields (Arundale et al., 2014a), constitute significant energy expenditure in feedstock production. Reducing the need for nitrogen fertilizer can therefore substantially decrease the net carbon intensity of biofuels (Melillo et al., 2009; Davis et al., 2011). Improving the nitrogen contribution from associative diazotroph communities to feedstock crops is also desirable for enabling cellulosic feedstock production on low-nutrient soils (Chakraborty, 2014; Lowman et al., 2015) like nonprime farmland and marginal land, thereby reducing the competition between biofuels and food crops for prime farmland (Hardy et al., 2007; Saraf & Hastings, 2010).

Diazotroph communities are associated with the rhizosphere and endophytic compartments of miscanthus and switchgrass, similar to many other grasses (Bagwell
et al., 1998; Erica et al., 2014; Gupta et al., 2014; Sarathambal et al., 2015), sugarcane (Thaweenut et al., 2010; Taulé et al., 2011), and cereals (Prakashang et al., 2009; Venieraki et al., 2010; Rangjaroen et al., 2013; Rodríguez-Blanco et al., 2015). The composition of diazotroph communities in the rhizosphere and the endophytic compartment varies with edaphic and management variables, as well as between hosts. We have demonstrated that for genetically identical M. × giganteus cultivated in the United States across large geographic distances, and diazotroph community structure in the rhizosphere was affected by edaphic variables, while communities in the endophytic compartment remained largely similar across sites (Li et al., 2015). Among diazotroph communities associated with native and uncultivated relatives of M. × giganteus – namely M. floridulus and M. sinensis – rhizosphere communities have been shown to be affected largely by edaphic factors in the rhizosphere but endophytic communities to vary primarily between host plant species in the endophytic compartment (Li, 2015). We have also shown that 16% of the nitrogen in the first-year growth of M. × giganteus was derived from biological nitrogen fixation (BNF) (Keymer & Kent, 2014). Optimizing feedstock-associated diazotroph communities to increase BNF contribution to the 80% common in forage legumes (Carlsson & Huss-Danell, 2003) may be unrealistic. However, the 60% maximum BNF from associative diazotrophs observed in sugarcane (Taulé et al., 2011; Urquiaga et al., 2011), or 70% in elephant grass (Morais et al., 2011) suggest that there could be substantial room for improving the BNF contribution from feedstock-associated diazotroph communities. While BNF is conventionally expected from soil-dwelling diazotrophs, it has also been observed in the endophytic compartment of many plant species including wild rice (Elbeltagy et al., 2001), sugarcane (Boddy et al., 2003), grasses (Dalton et al., 2004), and switchgrass (Lowman et al., 2015). Together, these previous findings suggest that in order to enhance the nitrogen efficiency of cellulosic biofuels production, and it may be possible to optimize diazotroph communities in the feedstock rhizosphere by manipulating edaphic variables such as micronutrient concentration (Fe, Ca) organic matter content, and pH as well as conceivably those in the endophytic compartment through targeted plant breeding and engineering.

The essential first step toward optimizing feedstock-associated diazotroph communities is a thorough understanding of the effects of edaphic, biotic, and management variables on these communities. The structure and function of soil microbial communities are known to be affected strongly by edaphic variables including pH, soil organic matter, and soil type (Girvan et al., 2003; Lauber et al., 2009; Fierer et al., 2012; Kuramae et al., 2012; Schreiter et al., 2014). Nitrogen application (Wakelin et al., 2007; Shu et al., 2011; Stepiet et al., 2014), plant species (Mao et al., 2011, 2013; Li, 2015), and management practices like soil amendments, residue retention, and pasture grazing (Wakelin et al., 2007, 2009) are also known to affect diazotroph abundance and diversity. Therefore, it is reasonable to expect that edaphic and management variables can affect rhizosphere feedstock-associated diazotroph communities.

To identify potential edaphic drivers of variation in diazotroph communities associated with M. × giganteus and P. virgatum, we characterized the structure of the associated diazotroph and bacterial communities at eight research sites across Illinois with variation in soil chemical properties. We quantified the soil chemical parameters at each site and the abundance of the nifH gene in the rhizosphere. Additionally, variability in these communities was characterized across host species and habitats (i.e., rhizosphere vs. endophytic compartment). Our findings reveal significant variation in diazotroph community structure correlated with variation in soil chemical parameters. Furthermore, diazotroph abundance in the rhizosphere, as indicated by nifH gene abundance, was found to be strongly correlated with soil iron and ammonium concentrations. These findings indicate potential directions for further investigating the parameters related to optimal feedstock-associated diazotroph communities and could, over the long-term, inform site selection and optimization to enhance beneficial diazotroph communities in the biofuels feedstock rhizosphere.

Materials and methods

Site description

The research sites for this study have been described in detail elsewhere (Heaton et al., 2009; Arundale et al., 2014a). In brief, field trials were established in 2002 (Dekalb, Urbana, Dixon Springs), 2004 (Havana, Orr, Brownstown, Fairfield), and 2005 (Dudley Smith Farm) with the locations chosen to represent the variety of soil chemical parameters and climate across Illinois. Each field trial consisted of randomized complete block design with four 100 m² plots each of M. × giganteus (var. Illinois) and P. virgatum (var. Cave-in-Rock) (except for Dudley Smith Farm, which consisted of four 0.2 ha plots of the two plant species), and four nitrogen application rates from 0 to 200 kg hectare⁻¹. Soil characteristics, agronomic management, and yield from these plots have been described previously (Heaton et al., 2008, 2009; Arundale et al., 2014a). A small number of miscanthus samples were collected from subplots that received fertilizer application at 200 kg hectare⁻¹ and switchgrass subplots received fertilizer applications ranging from zero to 13 kg hectare⁻¹. Key soil chemistry and site parameters are tabulated in Table S1.
Sample collection

Soil and root samples were collected in August 2009. Three rhizome samples with adhering rhizosphere soil were collected for switchgrass and miscanthus from each of the four unfertilized subplots at each site for a total of 96 rhizosphere samples from unfertilized switchgrass and miscanthus plots and an additional 15 samples from fertilized miscanthus plots. Six bulk soil cores for quantifying soil chemical parameters were collected to 6” depth from each subplot. All soil cores and plant tissue samples were placed in sealed plastic bags on ice while in the field and transported back to the laboratory.

Sample processing

Representative subsamples from each sampled plant rhizome were aseptically transferred to sterile 1 l Nalgene bottles with sterile water. Adhering rhizosphere soil was eluted from the root surface by vigorously shaking the Nalgene bottles. The eluent soil slurry was collected and stored at −20 °C until it could be lyophilized. Extraction of endophytic community DNA was carried out as described by Keymer & Kent (2014), a modification of the process described by Brulc et al. (2009). Rhizomes were washed vigorously with distilled water, placed in 1 l Nalgene bottles with sufficient 5.25% sodium hypochlorite solution and shaken for 15 min, and subsequently washed twice with sterile distilled water. The cleaned tissues were homogenized in a blender with 40 ml phosphate-buffered saline + 0.15% with Tween 80. The homogenized slurry was poured into 50 ml tubes with glass beads and shaken on ice for 1 h. The extract was passed through a coarse sieve (No. 25 US standard test sieve, pore size 710 μm; Newark Wire Cloth Company, Clifton, NJ, USA) to remove large pieces of plant tissue, and finer plant debris was removed by centrifugation for 5 min at 453 × g. The resulting supernatant was centrifuged at 3900 × g for 10 min in a Sorvall RC-5B refrigerated superspeed centrifuge with SS-34 rotor (DuPont) to recover endophytic bacterial cells in the reserved pellet.

DNA extraction

DNA from the endophytic microbial communities was extracted from the pellet resuspended in sterile water using the FastDNA kit from MP Biomedicals (Solon, OH). DNA was extracted from lyophilized rhizosphere soil samples using the FastDNA for Soil kit from MP Biomedicals (Solon, OH). The extracted DNA samples were purified using CTAB (cetyltrimethylammonium bromide) and 1:24 chloroform:isoamyl alcohol followed by ethanol precipitation to remove humic contaminants. Extracted and purified DNA was quantified using NanoDrop 1000 (Nanodrop Technologies, LLC, Wilmington, DE). The DNA was stored at −20 °C until used for ARISA, T-RFLP, and qPCR analyses as described below.

Nitrogenase gene abundance

As a marker of nitrogen fixation potentially available to the feedstock crops, quantitative polymerase chain reaction was used to measure abundance of the nifH gene in the rhizosphere. nifH was amplified using the primers PolF (5′-TCCGAYCCSAARGCBGACTC-3′) and PolR (5′-ATSGCC ATCATYTCRCCGGA-3′) (Poly et al., 2001). The 25 μl reaction volumes were prepared using 1× Sybr Green PCR MasterMix (Applied Biosystems, Foster City, CA, USA), 400 nm of each nifH primer, 20 ng of template DNA, and 500 μg ml−1 bovine serum albumin. The thermal cycling program consisted of an initial denaturation at 94 °C for 5 min, 40 cycles of 94 °C for 30 s, annealing temperature for 30 s, 72 °C for 30 s. Annealing temperatures were stepped down during the initial cycles (1 cycle at 64 °C, 2 cycles at 62 °C, 3 cycles at 60 °C, and 4 cycles at 58 °C, and final 30 cycles at 56 °C) to improve specificity. Negative controls and duplicate standard curves generated by serial dilution of purified PCR products were included in each plate. Copies of nifH were normalized on per nanogram of DNA determined by template DNA quantification using the Quant-iT™ dsDNA HS Assay Kit with a Qubit® fluorometer (Invitrogen/Molecular Probes, Eugene, OR, USA).

Diazotroph community structure

The effects of site variables, plant species, and plant habitat on the community of nitrogen fixing bacteria were assessed with terminal-restriction fragment length polymorphism (T-RFLP) analysis of the nitrogenase gene. The gene was amplified from the extracted and purified microbial DNA using primers PolF (5′-FAM-TCCGAYCCSAARGCBGACTC-3′) and PolR (5′-HEX ATSGCCATCATYTCRCCGGA-3′), which generate amplicons of approximately 360 base pairs (sequence position 115–476 relative to A. vinelandii nifH coding sequence (GenBank accession number M20568)) (Poly et al., 2001). The PolF/PolR primers were chosen for effective performance for fragment analysis as well as qPCR for the communities of interest in agricultural soils (Gaby & Buckley, 2012, 2017). PCRs were conducted as described in (Keymer & Kent, 2014). The resulting amplicons were purified with MinElute 96 UF PCR Purification Kit (Qiagen, Valencia, CA). Purified amplicons were digested for 16 h at 37 °C using the restriction enzyme MboI and MnlI following the procedures recommended by New England Biolabs (Beverly, MA) (Keymer & Kent, 2014). The lengths of digested gene fragments were quantified by denaturing gel capillary electrophoresis at the W. M. Keck Center for Comparative and Functional Genomics, University of Illinois, Urbana-Champaign, with the ABI 3730xl Genetic Analyzer (PE Biosystems, Foster City, CA, USA). ROX 1000 size standard (Applied Biosystems, Foster City, CA, USA) was used during capillary electrophoresis. Fragment size was determined for the assay as described below.

Microbial community structure

To understand the effects of site characteristics, plant species, and habitat on the bacterial community, we analyzed the community structure in each sample with automated ribosomal intergenic spacer analysis (ARISA) (Fisher & Tripplett, 1999). The bacterial rRNA operon intergenic transcribed spacer (ITS) region was amplified from extracted and purified microbial
DNA with PCR using 5' 6-FAM labeled forward primer 1406f (5'-TGYACACACCGCCGT-3' – universal, 16S rRNA gene), and reverse primer 23Sr (5'-GGGTBCCCATTCGGC-3' – bacteria-specific, 23S rRNA gene). Denaturing capillary electrophoresis of the amplicons was performed at the W. M. Keck Center for Comparative and Functional Genomics, University of Illinois, Urbana-Champaign, IL. The ROX 1000 size standard (Applied Biosystems, Foster City, CA, USA) was used during capillary electrophoresis.

Fragment size analysis

Sizes of the ITS and T-RFLP amplicons were determined in GENEMARKER software (Version 2.4, SoftGenetics, LLC, PA). Minor run-to-run variations in observed vs. actual fragment length that results from capillary electrophoresis were resolved using the allele-calling features in GENEMARKER before analysis. To include the maximum number of peaks while excluding background fluorescence, a threshold of 200 fluorescence units was used. The area of each peak was normalized to account for run-to-run variations in signal detection by dividing the area of individual peaks by the total fluorescence (area) detected in each sample, expressing each peak as a proportion of the observed community (Yannarell & Triplett, 2005).

Soil chemical analysis

Soil chemistry parameters, including pH, total nitrogen, NO₃⁻-nitrogen, NH₄⁺-nitrogen, total carbon percent, total organic matter, and Mehlich-3 phosphorus, potassium, calcium, and iron (Mehlich, 1984), were quantified from bulk soil cores that were obtained at the same time as the rhizosphere and plant tissue samples. Analysis was carried out by the Iowa State Soil and Plant Analysis Laboratory (Ames, IA). Soil moisture was quantified in-house with gravimetric soil moisture analysis, where samples were weighed before and after oven drying at 60 °C for 24 h.

Statistical analyses

ARISA, T-RFLP, qPCR, and edaphic data of each of the samples were analyzed statistically using multivariate correlational and ordination methods in R (R Core Team, 2017) and with R packages vegan version 2.4-2 (Oksanen et al., 2017) and labdsv version 1.8-0 (Roberts, 2016). Multiple samples obtained from each 100 m² plot were treated as independent samples rather than technical replicates, to capture within-plot variation in the communities. Significant local variation in the communities is expected within these large experimental plots due to within-plot differences in edaphic conditions (Tripathi et al., 2014; Livermore & Jones, 2015; Nunan, 2017). The correlation of nifH abundance in the rhizosphere with environmental and soil chemical parameters was assessed with Pearson’s product-moment correlation coefficient (Puth et al., 2014). The effect size and statistical significance of site, feedstock species, and habitat on the structures of the microbial and diazotrophic communities were determined with permutational multivariate analysis of variance (PERMANOVA) (Anderson, 2005). Community structure (ARISA and nifH T-RFLP) was analyzed with partial correspondence analysis (CA) (Legendre & Legendre, 2012) and visualized as ordinations to facilitate the assessment of community similarities. The correlation of soil chemistry parameters with community structure variation across sites was quantified with the function envfit and plotted on the ordinations, as well as tabulated (Table 2).

Table 1 The effect of cultivation site as well as feedstock species on the structure of diazotrophic communities and on total bacterial communities

| Factor      | Rhizosphere | Endophytic compartment |
|-------------|-------------|------------------------|
|             | R²          | P ≤                    | R²          | P ≤                    |
| nifH T-RFLP |             |                        |             |                        |
| Species     | 0.03        | 0.001                  | 0.14        | 0.001                  |
| Site        | 0.42        | 0.001                  | 0.14        | 0.001                  |
| Species : Site | 0.07        | 0.001                  | 0.07        | 0.001                  |
| ARISA       |             |                        |             |                        |
| Species     | 0.02        | 0.001                  | 0.20        | 0.001                  |
| Site        | 0.31        | 0.001                  | 0.09        | 0.001                  |
| Species : Site | 0.07        | 0.001                  | 0.07        | 0.001                  |

n = 207 rhizosphere samples, n = 202 for endophytic compartment samples for the two feedstock species and eight sites combined.
Finally, \textit{nifH} abundance was significantly correlated with variation in community structure. The strong statistical correlation of phosphorus concentration with diazotroph community structure may be due to the strong gradient in phosphorus concentration – from 48.75 to 5 ppm, across our study sites – a greater percent difference than any other soil chemical parameter. The graphical representation of the edaphic factors (Fig. 1) is useful to depict their relative trends with respect to each other across the geographic area studied (e.g., \textit{nifH} abundance and Fe concentration are positively correlated with each other, but negatively correlated with total N and most other soil chemical parameters), the tabulation is provided for the quantitative understanding of strength and significance of these correlations with the structure of the communities.

### Bacterial communities in the rhizosphere

The variation in the structure of bacterial communities in the feedstock rhizosphere was influenced by site, and differed by a very small, though statistically significant, extent between the host plant species (Table 1). The interaction of site and host species had a small effect on the variation between rhizosphere diazotroph communities.

Among edaphic and environmental factors, site latitude was strongly correlated with the variation in bacterial community structure (Table 2; Fig. 1b), reflecting the effect of the combination of edaphic factors on the site-to-site variation of bacterial communities. Among specific soil chemical properties, the ones most correlated with the variation in bacterial community structure were pH, phosphorus, and calcium. In contrast with diazotroph communities, total nitrogen content as well as nitrogen species was correlated with variation in bacterial community structure. \textit{nifH} abundance was also significantly correlated with variation in bacterial community structure.

### Table 2 Correlation of edaphic and endogenous factors with microbial community structure variation across study sites, assessed with function envfit

|                           | Rhizosphere diazotroph communities | Rhizosphere bacterial communities |
|---------------------------|-----------------------------------|----------------------------------|
|                           | \( R^2 \)  | \( P \leq \) | \( R^2 \)  | \( P \leq \) |
| Latitude                  | 0.29    | 0.001     | 0.71    | 0.001     |
| P (ppm)                   | 0.35    | 0.001     | 0.38    | 0.001     |
| Fe (ppm)                  | 0.19    | 0.001     | 0.27    | 0.001     |
| K (ppm)                   | 0.18    | 0.001     | 0.25    | 0.001     |
| pH                        | 0.16    | 0.001     | 0.32    | 0.001     |
| C (%)                     | 0.08    | 0.001     | 0.22    | 0.001     |
| \(\text{NH}_4\) (ppm)    | 0.08    | 0.001     | 0.13    | 0.001     |
| \(\text{NO}_3\) (ppm)    | 0.07    | 0.003     | 0.22    | 0.001     |
| Ca (ppm)                  | 0.06    | 0.005     | 0.13    | 0.001     |
| N (%)                     | 0.03    | 0.065     | 0.13    | 0.001     |
| \textit{nifH} abundance   | 0.39    | 0.001     | 0.36    | 0.001     |

\( n = 207 \) rhizosphere samples for the two feedstock species across the eight research sites.

\( n = 207 \) rhizosphere samples for the two feedstock species across the eight research sites.

\( n = 207 \) rhizosphere samples for the two feedstock species across the eight research sites.

Fig. 1  Variation in the structure of communities in the rhizosphere of bioenergy feedstocks across Illinois, visualized with partial correspondence analysis (controlling for the small variation related to host species). Communities of diazotrophs (a) as well as all bacteria (b) are distinct between sites and correlated with soil chemical parameters. Ellipses display standard deviation of the sample set for each site around the centroid. Increasing soil concentration of iron, ammonium, and \textit{nifH} gene abundance is positively correlated with each other and negatively correlated with other edaphic factors. (\( n = 207 \)).
Communities in the endophytic compartment

Diazotroph communities in the endophytic compartment (Table 1; Fig. 2) varied between the host species as well as among sites. Endophytic bacterial communities were dissimilar between the two host species and were affected to a smaller extent by site. As expected, habitat (rhizosphere vs. endophytic compartment) was a significant factor in the variation in community structure for diazotrophs as well as bacterial communities.

Nitrogenase gene abundance in the rhizosphere

Abundance of nifH gene in the feedstock rhizosphere revealed interesting patterns of correlation with edaphic factors (Table 3; Figures S1 and S2). Notably, nifH abundance was greater in sites with greater iron content and was also correlated with greater concentration of ammonia in the soil. Conversely, sites with greater nitrate concentration exhibited lower nifH abundance, though the correlation was less than statistically significant for communities in the miscanthus rhizosphere. Similarly, nifH abundance was generally lower in the miscanthus plots that received nitrogen fertilizer treatment. Greater concentration of conventional markers of soil fertility – such as soil organic matter, total nitrogen, potassium, and phosphorus, as well as calcium – was also correlated with lower nifH abundance. The marked north-south variation in soil chemical properties is reflected in the correlation of nifH abundance with site latitude.

Comparing across host species, nifH gene abundance was not significantly different in the rhizosphere of the two hosts (P = 0.24).

Discussion

Informed site selection and soil amendments appear to be promising tools for optimizing rhizosphere diazotrophic communities and thereby improving the sustainability of cellulosic biofuels production. The systematic analysis of the differences in the correlations of edaphic factors with the variations in overall bacterial communities vs. diazotroph communities is a novel and notable finding of this work. For future research, it will be important to take into account the distinct responses of diazotroph communities and total bacterial communities to edaphic factors.

The quantification of absolute and relative differences in the correlations of edaphic factors with community structure variation was possible due to the strong north-south gradient in soil chemical parameters in Illinois which stems from its distinctive paleogeography. The greater age and weathering of the southern and western sites in this study – Dixon Springs, Brownstown, Fairfield, and Orr – corresponds to their greater iron content and lower calcium, carbon, nitrogen, phosphorous, and potassium content (Martinson et al., 1987; Curry & Baker, 2000). Long-term weathering of soils generally increases their extractable iron content (Arduino et al., 1986), while calcium carbonate content of soils

Fig. 2  Partial correspondence analysis ordinations, of the community structure of diazotrophs (a) and bacteria (b) as revealed by nifH T-RFLP and ARISA, respectively. The ordination controls for site-to-site variation to highlight the effect of habitat and host species (M. × giganteus – MxG, Panicum virgatum – PV). The difference in community structure between rhizosphere communities of the two host species is small, though statistically significant. (Table 1 lists effect sizes.) (n = 207 rhizosphere samples, n = 202 for endophytic compartment samples).
hours. Given the strength of the correlation of iron with diazotroph community abundance and the iron requirements of soil diazotrophs, it is likely that iron availability was a driving factor behind the variation in abundance at these sites. However, the quantified correlation of individual soil chemical parameters with diazotroph community properties presented here offers novel directions for future research into optimization of feedstock production.

It is particularly notable that soil iron concentration was relatively more strongly correlated than most other edaphic factors with the site-to-site variation in the structure of diazotroph communities, with pH being the factor best correlated with variation in bacterial community structure. The overriding effect of site variables on the feedstock rhizosphere microbiome – relative to the effect of plant species – is in agreement with previous findings by our group (Li, 2015; Li et al., 2015) and others (Girvan et al., 2003; Kuramae et al., 2012; Schreiter et al., 2014). However, the quantified correlation of individual soil chemical parameters with diazotroph community properties presented here offers novel directions for future research into optimization of feedstock production.

It is particularly notable that soil iron concentration was relatively more strongly correlated than most other edaphic factors with the site-to-site variation in the structure of diazotroph communities, with pH being the factor best correlated with variation in bacterial community structure. The overriding effect of site variables on the feedstock rhizosphere microbiome – relative to the effect of plant species – is in agreement with previous findings by our group (Li, 2015; Li et al., 2015) and others (Girvan et al., 2003; Kuramae et al., 2012; Schreiter et al., 2014). However, the quantified correlation of individual soil chemical parameters with diazotroph community properties presented here offers novel directions for future research into optimization of feedstock production.

It is particularly notable that soil iron concentration was relatively more strongly correlated than most other edaphic factors with the site-to-site variation in the structure of diazotroph communities, with pH being the factor best correlated with variation in bacterial community structure. The overriding effect of site variables on the feedstock rhizosphere microbiome – relative to the effect of plant species – is in agreement with previous findings by our group (Li, 2015; Li et al., 2015) and others (Girvan et al., 2003; Kuramae et al., 2012; Schreiter et al., 2014). However, the quantified correlation of individual soil chemical parameters with diazotroph community properties presented here offers novel directions for future research into optimization of feedstock production.

It is particularly notable that soil iron concentration was relatively more strongly correlated than most other edaphic factors with the site-to-site variation in the structure of diazotroph communities, with pH being the factor best correlated with variation in bacterial community structure. The overriding effect of site variables on the feedstock rhizosphere microbiome – relative to the effect of plant species – is in agreement with previous findings by our group (Li, 2015; Li et al., 2015) and others (Girvan et al., 2003; Kuramae et al., 2012; Schreiter et al., 2014). However, the quantified correlation of individual soil chemical parameters with diazotroph community properties presented here offers novel directions for future research into optimization of feedstock production.

It is particularly notable that soil iron concentration was relatively more strongly correlated than most other edaphic factors with the site-to-site variation in the structure of diazotroph communities, with pH being the factor best correlated with variation in bacterial community structure. The overriding effect of site variables on the feedstock rhizosphere microbiome – relative to the effect of plant species – is in agreement with previous findings by our group (Li, 2015; Li et al., 2015) and others (Girvan et al., 2003; Kuramae et al., 2012; Schreiter et al., 2014). However, the quantified correlation of individual soil chemical parameters with diazotroph community properties presented here offers novel directions for future research into optimization of feedstock production.
including those of total nitrogen, ammonium, and nitrites at these sites. We hypothesize that greater nifH abundance compensates, in large part, for the site-to-site variation in nitrogen and mitigates the effect of low soil fertility on feedstock yield. Nitrogen fertilization has been found to have conflicting effects on miscanthus productivity. Across the research sites covered in this study, with stand age greater than 5 years, nitrogen fertilization at 202 kg ha\(^{-1}\) was found to be correlated with 25% increase in \(M. \times giganteus\) yield relative to unfertilized plots (Arundale et al., 2014b). However, a long-term trial of nitrogen inputs between zero and 120 kg ha\(^{-1}\) was found not to affect miscanthus yields for 14 successive harvests (Christian et al., 2008). Efficiency analysis has revealed that nitrogen use efficiency and energy use efficiency are simultaneously highest for miscanthus when no fertilizer was applied (Lewandowski & Schmidt, 2006). The high energy and nitrogen efficiency of miscanthus appear to be due to high levels of BNF associated with miscanthus as revealed by modeling (Davis et al., 2009) as well as experimental (Keymer & Kent, 2014) studies. On the other hand, in previous studies involving switchgrass, nutrient balance modeling (Davis et al., 2009) did not suggest a significant contribution of BNF to nitrogen balance, and optimum biomass yields were achieved at the 120 kg ha\(^{-1}\) nitrogen fertilization (Vogel et al., 2002). It is plausible that the differences between the contribution of BNF to miscanthus and switchgrass can be attributable in part to interspecies differences in associating productively with rhizosphere and endophytic diazotrophs.

In the endophytic compartments of \(M. \times giganteus\) and \(P. virginatum\), the community structure of bacteria and diazotrophs is distinct between the two species, with a substantially smaller site-to-site variation – similar to previous reports by us and others (Rangjaroen et al., 2013; Li, 2015; Li et al., 2015; Rodriguez-Blanco et al., 2015), suggesting that interspecies physiological differences play a role in shaping these distinct endophytic communities.

In the context of prior findings, the results presented here indicate that there may be an advantage to prioritizing feedstock production at sites that are low-fertility for conventional agriculture, but have a high potential for supporting beneficial associative diazotroph communities. Reducing the need for fertilizers would simultaneously reduce potential nitrate leaching (Behnke et al., 2012) as well as improve the carbon intensity of biofuels production (Lewandowski & Schmidt, 2006). Improving the productivity of biofuels feedstock production on low-fertility land would also minimize competition with production of food crops (Hardy et al., 2007; Saraf & Hastings, 2010). In addition to site selection, the findings here also suggest a potential for shaping the diazotroph communities through specific soil amendments that are correlated with potential desirable outcomes such as greater nifH abundance. Sites that are otherwise suitable for feedstock production but have low bioavailable iron could be amended to increase iron content and reduce soil pH in order to promote an increase in the population of diazotrophic microbes, potentially increasing the productivity of bioenergy feedstock crops.

Conclusions
This study demonstrates that soil chemical properties, particularly iron and calcium concentration, are significantly correlated with nitrogenase gene abundance as well as rhizosphere microbial community structure across a large geographic scale. These findings imply that informed site selection as well as soil amendments that improve associative diazotroph communities could improve the sustainability of biofuels feedstock production. The potential enhancements in cellulosic biofuel feedstock production from such efforts may improve the economic feasibility and productivity of feedstock farming on nonprime land, reduce potential competition for prime farmland with food and commodity crops and thus improve the energy infrastructure without adversely affecting food prices.

Acknowledgements
The authors acknowledge Dongfang Li and Ariane Peralta for assistance with fieldwork and sample processing. This study was funded by the Energy Biosciences Institute at the University of Illinois. Chinmay Soman was funded by a National Science Foundation Postdoctoral Fellowship award (#1 314 064).

References
Adler PR, Grosso SJ, Parton WJ (2007) Life-cycle assessment of net greenhouse-gas flux for bioenergy cropping systems. Ecological Applications, 17, 675–691.
Anderson MJ (2005) Permutational multivariate analysis of variance. Department of Statistics, University of Auckland, Auckland.
Arduino E, Barberis E, Ajmone Marsan F, Zanini E, Franchini M (1986) Iron oxides and clay minerals within profiles as indicators of soil age in northern Italy. Geoderma, 37, 45–55.
Arundale R (2012) The higher productivity of the bioenergy feedstock Miscanthus \(\times giganteus\) relative to Panicum virgatum is seen both into the long term and beyond Illinois.
Arundale RA, Dohleman FG, Heaton EA, McGrath JM, Voigt TB, Long SP (2014a) Yields of Miscanthus \(\times giganteus\) and Panicum virgatum decline with stand age in the Midwestern USA. GCB Bioenergy, 6, 1–13.
Arundale RA, Dohleman FG, Voigt TB, Long SP (2014b) Nitrogen fertilization does significantly increase yields of stands of Miscanthus \(\times giganteus\) and Panicum virgatum in multiyear trials in Illinois. BioEnergy Research, 7, 408–416.
Bagwell CE, Picono YM, Ashburne-Lucas A, Lovell CR (1998) Physiological diversity of the rhizosphere diazotroph assemblages of selected salt marsh grasses. Applied and Environmental Microbiology, 64, 4276–4282.
Behnke GD, David MB, Voigt TB (2012) Greenhouse gas emissions, nitrate leaching, and biomass yields from production of Miscanthus \(\times giganteus\) in Illinois, USA. BioEnergy Research, 5, 801–813.
across rice landraces grown in the highlands of northern Thailand. Archives of Microbiology, 196, 35–49.

Roberts DW (2016) Labeled: Ordination and multivariate analysis for ecology. Retrieved from https://CRAN.R-project.org/package=labeled

Rodriguez-Blanco A, Sicardi M, Proni L (2015) Plant genotype and nitrogen fertilization effects on abundance and diversity of diazotrophic bacteria associated with maize (Zea mays L.). Biology and Fertility of Soils, 51, 391–402.

Rosconi F, Souza EM, Pedrosa FO et al. (2006) Iron depletion affects nitrogenase activity and expression of nifH and nifA genes in Herbaspirillum seropedicae. FEMS Microbiology Letters, 258, 214–219.

Sanscarter D, Deen B, Dias G, MacLean HL, Dadfar H, McDonald I, Kludze H (2014) Implications of land class and environmental factors on life cycle GHG emissions of Miscanthus as a bioenergy feedstock. GCB Bioenergy, 6, 401–413.

Saraf M, Hastings A (2010) Biofuels, the role of biotechnology to improve their sustainability and profitability. In: Biodiversity, Biofuels, Agroforestry and Conservation Agriculture (ed Lichtfouse E), pp. 123–148. Springer, Netherlands.

Sarathambal C, Ilamurugu K, Balachandar D, Chinnadurai C, Gharde Y (2015) Characterization and crop production efficiency of diazotrophic isolates from the rhizosphere of semi-arid tropical grasses of India. Applied Soil Ecology, 87, 1–10.

Saraf M, Hasting A (2010) Biofuels, the role of biotechnology to improve their sustainability and profitability. In: Biodiversity, Biofuels, Agroforestry and Conservation Agriculture (ed Lichtfouse E), pp. 123–148. Springer, Netherlands.

Schreiter S, Ding G-C, Hexer H et al. (2014) Effect of the soil type on the microbiome in the rhizosphere of field-grown lettuce. Frontiers in Microbiology, 5, 144. https://doi.org/10.3389/fmicb.2014.00144.

Shah VK, Brill WJ (1977) Isolation of an iron-molybdenum cofactor from nitrogenase. Proceedings of the National Academy of Sciences, 74, 3249–3253.

Shu W, Pablo GP, Juen Y, Danfeng H (2011) Abundance and diversity of nitrogen-fixing bacteria in rhizosphere and bulk paddy soil under different duration of organic management. World Journal of Microbiology and Biotechnology, 28, 493–503.

Stepien W, Górska EB, Pietkiewicz S, Kalaji MH (2014) Long-term mineral fertilization impact on chemical and microbiological properties of soil and Miscanthus × giganteus yield. Plant Soil and Environment, 60, 117–122.

Taule C, Mareque C, Barlocco C, Hackemburch F, Reis VM, Sicardi M, Battistoni F (2011) The contribution of nitrogen fixation to sugarcane (Saccharum officinarum L.), and the identification and characterization of part of the associated diazotrophic bacterial community. Plant and Soil, 356, 35–49.

Thaweunruit N, Hachisuika Y, Ando S, Yanagiwasa S, Yoneyama T (2010) Two season’s study on nifH gene expression and nitrogen fixation by diazotrophic endophytes in sugarcane (Saccharum spp. hybrids): expression of nifH genes similar to those of rhizobia. Plant and Soil, 338, 435–449.

Timilsina GR (2014) Biofuels in the long-run global energy supply mix for transportation. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 372, 20120323.

Tripathi BM, Lee-Cruz I, Kim M et al. (2014) Spatial scaling effects on soil bacterial communities in Malaysian tropical forests. Microbial Ecology, 68, 247–258.

Urquiza S, Xavier KP, Morais RFD et al. (2011) Evidence from field nitrogen balance and 15N natural abundance data for the contribution of biological N2 fixation to Brazilian sugarcane varieties. Plant and Soil, 356, 5–21.

Venieraki A, Dimou M, Pergalis P, Kafalogianni I, Chatzivarup lidis I, Katinakis P (2010) The genetic diversity of cultivable nitrogen-fixing bacteria in the rhizosphere of wheat. Microbial Ecology, 61, 277–285.

Vogel KP, Brejda JJ, Walters DT, Buxton DR (2002) Switchgrass biomass production in the Midwest USA. Agronomy Journal, 94, 413–420.

Wakelin SA, Colloff MJ, Harvey PR, Marschner P, Gregg AL, Rogers SL (2007) The effects of stubble retention and nitrogen application on soil microbial community structure and functional gene abundance under irrigated maize. FEMS Microbiology Ecology, 59, 661–670.

Wakelin SA, Gregg AL, Simpson RJ, Li GD, Riley IT, McKay AC (2009) Pasture management clearly affects soil microbial community structure and N-cycling bacteria. Pedobiologia, 52, 237–251.

Wakelin SA, Gupta VSVR, Forrester ST (2010) Regional and local factors affecting diversity, abundance and activity of free-living, N2-fixing bacteria in Australian agricultural soils. Pedobiologia, 53, 391–399.

Yamarell AC, Triplet EW (2005) Geographic and environmental sources of variation in lake bacterial community composition. Applied and Environmental Microbiology, 71, 227–239.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1. Correlation of nifH abundance (Y axis) in the Miscanthus rhizosphere communities with various edaphic factors (X axes) at the eight bioenergy research sites in Illinois. The Pearson correlation is denoted by “r” and correlations significant at the P < 0.001 level are denoted by ***. Samples from each site are denoted by the colored dots as in the legend on the left. BST: Brownstown, DKB: Dekalb, DSP: Dixon Springs, DUF: Dudley Field, FRF: Fairfield, HVN: Havana, ORR: Urb, URB: Urbana

Figure S2. Correlation of nifH abundance (Y axis) in the switchgrass rhizosphere communities with various edaphic factors (X axes) at the eight bioenergy research sites in Illinois. The Pearson correlation is denoted by “r” and correlations significant at the P < 0.001 level are denoted by ***. Samples from each site are denoted by the colored dots as in the legend on the left. BST: Brownstown, DKB: Dekalb, DSP: Dixon Springs, DUF: Dudley Field, FRF: Fairfield, HVN: Havana, ORR: Urb, URB: Urbana

Table S1. Average values of soil chemistry parameters at each of the research sites.

© 2018 The Authors. Global Change Biology Bioenergy Published by John Wiley & Sons Ltd., 10, 343–352