Received: 7 July 2021 | Accepted: 21 August 2021

DOI: 10.1111/gcbb.12893

GCB-BIOENERGY BIOPRODUCTS FOR A SUSTAINABLE BIOECONOMY Open Access

WILEY

ORIGINAL RESEARCH

Temporal dynamics of free-living nitrogen fixation in the switchgrass rhizosphere

Darian N. Smercina[1,2] | **Sarah E. Evans**[3] | **Maren L. Friesen**[4,5] | **Lisa K. Tiemann**[2]

[1]Biological Sciences Division, Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, Washington, USA

[2]Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, Michigan, USA

[3]W.K. Kellogg Biological Station, Department of Integrative Biology, Michigan State University, Hickory Corners, Michigan, USA

[4]Department of Plant Pathology, Washington State University, Pullman, Washington, USA

[5]Department of Crop and Soil Sciences, Washington State University, Pullman, Washington, USA

Correspondence
Lisa K. Tiemann, 1066 Bogue St. A286 East Lansing, MI 48824, USA.
Email: ltiemann@msu.edu

Funding information
U.S. Department of Energy, Grant/Award Number: DE-FC02-07ER64494, DE-SC0014108 and DE-SC0018409; National Science Foundation Long-term Ecological Research Program, Grant/Award Number: DEB 1832042; Michigan State University AgBioResearch

Abstract

Free-living nitrogen fixation (FLNF) represents an important terrestrial N source and is gaining interest for its potential to contribute plant available N to bioenergy cropping systems. Switchgrass, a cellulosic bioenergy crop, may be particularly reliant on FLNF when grown on low N systems, like marginal lands. However, the potential contributions of FLNF to switchgrass as well as the controls on this process are not well understood. In this study, we evaluated drivers of FLNF rates and N-fixing microbial (diazotrophic) community composition in field-grown switchgrass systems over two growing seasons with high temporal sampling. We found that climate variables are strong drivers of FLNF rates in switchgrass systems, compared to other environmental and biological factors including soil nutrients and diazotrophic community composition. Increased soil moisture availability generally promoted FLNF rates, but extreme rainfall events were detrimental. These climate-related responses suggest a potential for loss of FLNF-derived N contributions under projected climate shifts. We found a significant, but weak correlation between diazotrophic community composition and FLNF rates. We also observed a significant shift in the diazotrophic community composition between 2017 and 2018 and similarly measured a significant difference in FLNF rates between growing seasons. Lastly, we found that seasonal FLNF N contributions, based on measurement with high temporal resolution, has the potential to meet up to 80% of switchgrass N demands suggesting that FLNF measurements extrapolated from fewer time points or locations may underestimate these potential N contributions.

KEYWORDS

diazotroph, free-living nitrogen fixation, marginal land, plant nitrogen demands, sustainable bioenergy, switchgrass

1 | INTRODUCTION

Perennial grasses are a key facet of sustainable biofuel production, envisioned to provide net-zero carbon energy with potential for climate mitigation via belowground carbon sequestration (Robertson et al., 2017). Switchgrass (*Panicum virgatum*), a North American native C4 perennial grass, is a promising bioenergy crop with high biomass yields even when grown on marginal lands with minimal inputs of fertilizer nitrogen (N; Gelfand et al., 2013; Mehmood et al., 2017; Robertson et al., 2017) potentially due to N contributions from free-living nitrogen fixation (FLNF).

Free-living nitrogen fixation, here defined as N fixation occurring in soils and the rhizosphere (in and around roots) without direct plant symbiosis, is increasingly recognized for its importance as an N input to both natural and managed systems (Bloch et al., 2020; Davies-Barnard & Friedlingstein, 2020; Reed et al., 2011). Recent evidence suggests that FLNF may be a significant N source for cellulosic bioenergy cropping systems, like miscanthus (*Miscanthus* × *giganteus*; Davis et al., 2010) and switchgrass (*Panicum virgatum*; Roley et al., 2018; Smercina et al., 2019b) providing an alternative to fertilizer N additions and potentially increasingly the sustainability of such cropping systems. The apparent reliance of switchgrass on FLNF to support plant N demands has garnered much interest in recent years; however, little is still known about the controls on FLNF and the conditions which promote FLNF in switchgrass systems.

Free-living nitrogen fixation is an energy-intensive process, transforming dinitrogen (N_2) gas into biologically available ammonia, that occurs readily in the rhizospheres of many grasses where roots exude easily accessible carbon (C; Chalk, 2016; Roley et al., 2018; Smercina et al., 2019b). FLNF in the rhizosphere is carried out by a diverse community of N-fixing organisms (diazotrophs) under complex and dynamic conditions (Smercina et al., 2019b). These complex conditions make understanding and predicting FLNF difficult, and, to date, this process has remained poorly understood.

Past work has identified several broad controls on FLNF including carbon (C) availability, macro- and micronutrient availability (e.g., N, phosphorus, and metals), and climate (Roley et al., 2018; Smercina et al., 2019b). Given the high-energy and therefore high C demands of FLNF, it is generally accepted that FLNF is constrained to regions of the soil where C is readily accessible, such as the rhizosphere (Cleveland et al., 1999; Smercina et al., 2019b). Because of these high-energy demands, FLNF is not likely to be a competitive N-acquisition strategy when external N is available and thus is downregulated as diazotrophs access external N in favor of fixed N (Norman & Friesen, 2017; Reed et al., 2011). Phosphorus and metal availability also influences FLNF with increased availability of these nutrients generally supporting more FLNF, yet these controls are not well studied in the rhizosphere (Smercina et al., 2019b). Lastly, only weak connections have been established between climate variables and FLNF, with limited evidence of any strong interactions between climate variables and FLNF rates to date (Cleveland et al., 1999; Davies-Barnard & Friedlingstein, 2020; Reed et al., 2011). While these past studies may speak to ecosystem-scale patterns, little mechanistic work at fine spatial and temporal scales has been carried out to untangle the controls on FLNF, particularly in the rhizosphere.

In this study, we used biogeochemical, plant, climate, and molecular data, collected with high temporal frequency over 2 years to explore biological and environmental controls on FLNF. We examined the impact of various soil characteristics, climate conditions, plant metrics, and diazotroph community structure on FLNF process rates. We hypothesized that over the growing season, climate variables and plant metrics would be major drivers of FLNF. Specifically, increased soil moisture and increased plant productivity would result in increased FLNF rates. We also hypothesized that soil N availability and diazotroph community composition would be major drivers of FLNF whereby increased soil N availability would reduce FLNF rates and there would be diazotroph community members directly associated with greater FLNF.

2 | METHODS

2.1 | Field site and weather data

Samples were collected from Great Lakes Bioenergy Research Center (GLBRC; https://www.glbrc.org/) Marginal Land Experiment (MLE) at Lux Arbor (LUX; 42.476365, −85.451887) in southern Michigan. Soils at the LUX site are classified as Typic Hapludalfs (Alfisol) with a Loam texture (Kasmerchak & Schaetzl, 2018). Soils have a pH of 5.8, 0.77% total C, 0.06% total N, and 12 ppm inorganic phosphorus (Kasmerchak & Schaetzl, 2018). The site contains four multitreatment experimental blocks that include plots of monoculture switchgrass (*Panicum virgatum* L.; cv. Cave-in-Rock), which are split into annually fertilized (+56 kg urea-N ha^{-1} $year^{-1}$) and unfertilized (no added N) halves. Experimental plots were established in 2013 with switchgrass planted from seed. Continuous weather data are collected via a site-based weather station as part of the GLBRC MLE. Average annual precipitation at the site is 842 mm and average annual temperature is 9.0°C based on 30-year averages (Kasmerchak & Schaetzl, 2018).

2.2 | Field sampling

Soil and plant samples were collected every 2 weeks over two growing seasons (2017 and 2018) from pre-emergence to post-harvest. Samples were collected from March 27 to November 13 in 2017 and from March 19 to November 15 in 2018. Fertilizer N as urea (SUPERU, Koch Industries, Inc.) was applied on May 8 in 2017 and 2018. At each sampling date, three pseudoreplicate soil cores (5 cm wide, 10 cm deep) were collected from each split-plot in all four experimental blocks for a total of 24 soils per sampling date. Sampling locations within each spilt-plot were randomly selected and analyses performed on the separate pseudoreplicate samples. Soils were placed in a cooler, returned to the lab, and processed within 3 days of collection by passing through a 4-mm mesh sieve. During processing and until analyses were done, soils were stored at 4°C. A subset of each sieved soil was frozen at −80°C for extracellular enzyme activities and microbial community analyses (described below). Soil moisture content was determined by drying 5 g of processed, field moist soil at 60°C for 48 h.

On each sampling date, we also identified three switchgrass plants that corresponded to the pseudoreplicate locations for soil sampling. At each sampling, we measured plant height, metrics including rates of photosynthesis and C assimilation, leaf and root C:N ratios, specific stem density, and specific and top leaf area. Plant height was measured in the field to the nearest millimeter using a standard meterstick as the distance from the ground to the tallest point on the plant. Photosynthesis parameters including Phi2 (e.g., quantum yield of photosystem II) and relative chlorophyll were measured in field using a MultiSpeQ (Photosnyq, Inc.) at three locations on the plant (low, mid, and upper canopy). Reported values for photosynthesis parameters are averages across canopy samplings for each plant.

Leaf and root C and N were measured on an ECS 4010 elemental analyzer (Costech Analytical). Washed roots collected during soil sieving and plant leaves were first dried and ground before elemental analysis. Specific stem density was measured on the tallest stem for each sampled plant. The tallest stem was first clipped from each measured plant. A 10-cm subset of stem, with leaves removed, was then weighed, dried at 50°C for 4–7 days, and then weighed again. Specific stem density is calculated as g dry stem g^{-1} wet stem. Specific leaf area was determined by scanning leaves from the clipped stem at 2400 and 1200 dpi using a flatbed scanner. Leaf area was calculated using ImageJ software. Specific leaf area was determined by dividing leaf area (cm^2) by leaf dry mass (g).

2.3 | Free-living nitrogen fixation

Free-living nitrogen fixation rates were measured on intact soil cores collected from each pseudoreplicate location per split-plot as described by Smercina et al. (2019a). Briefly, we added a C cocktail solution containing equal parts glucose, sucrose, citrate, and malate at a rate of 4 mg C g^{-1} dry soil. Volume and concentration of the C solution added varied by sample date such that all samples were adjusted to 60% water holding capacity while achieving the target 4 mg C g^{-1} dry soil addition rate. After C addition, vials were sealed and evacuated. Vial headspace was replaced with 1 ml of $^{15}N_2$ (Sigma-Aldrich, Inc.), 10% Ultra High Purity (UHP) oxygen, and balanced with UHP Helium. Reference samples received UHP-N_2 in place of $^{15}N_2$. Vials were incubated for 3 days at room temperature. Vials were then uncapped and samples were dried at 60°C for 48 h before grinding and weighing for ^{15}N analysis. Tinned samples were analyzed following standard procedures at Washington State University's Stable Isotope Core Laboratory. FLNF rates were calculated in μg N fixed g^{-1} dry soil day^{-1} as:

$$\frac{AE_i \times TN_i}{AE_{atm} \times t},$$

where AE_i represents atom percent access of sample against an unenriched reference sample, TN_i represents total nitrogen content in sample, AE_{atm} represents atom percent excess in the vial atmosphere (98 atom% in our case), and t is incubation time in days (Roley et al., 2018; Warembourg, 1993).

2.4 | Inorganic N, dissolved organic C and N, and N mineralization

We measured soil inorganic N pools, dissolved organic C and N pools, and net rates of N mineralization. First, we generated soil extracts by shaking 6 g of field moist soil in 30 ml of 0.5 M potassium sulfate (K_2SO_4) at 200 rpm for 1 h. Extracts were filtered through Whatman grade 202 filter paper to remove soil. The resulting filtrate was used to measure soil inorganic N (NH_4 and NO_3), dissolved organic C (DOC), and dissolved total N (TN). Soil NH_4 and NO_3 concentrations were determined via 96-well high-throughput colorimetric analyses following methods described by Rhine et al. (1998) and Campbell et al. (2006), respectively. Soil extracts were analyzed on a Vario TOC cube (Elementar) to determine DOC and TN concentration. Concentrations of inorganic N, DOC, and TN were corrected for extract volume and soil moisture content.

DON was determined by subtracting inorganic N from TN.

Nitrogen mineralization rates (NH_4 and NO_3) were measured via field deployed ion exchange membranes following standard protocols of the Kellogg Biological Station Long-term Ecological Research station (KBS-LTER; lter.kbs.msu.edu). The exchange capacity and ion affinity of these membrane mimics that of plant roots, allowing an approximation of total plant available inorganic N over the period of deployment, as well as calculation of net rates of mineralization. Briefly, 10 cm × 2.5 cm strips of cation and anion exchange membranes (Membrane International, Inc.) were cut and activated with washes of 0.5 M hydrochloric acid and 0.5 M sodium bicarbonate. Membrane strips were rinsed and stored with DI water until field deployment. Cation and anion membranes' strip pairs were deployed at each split-plot pseudoreplicate location and collected at 2-week intervals during the 2017 and 2018 growing seasons. Harvested membranes were collected, rinsed of any adhering soil in field with DI water, and then stored in DI water until extraction. Ammonium and nitrate collected on membranes were extracted by shaking cation and anion membranes' pairs in 1 M potassium chloride for 24 h. Ammonium and nitrate concentrations were then analyzed using high-throughput colorimetric methods as above (Campbell et al., 2006; Rhine et al., 1998).

2.5 | Soil extracellular enzyme activities

Activities of 10 extracellular enzymes including alanine aminopeptidase (ALA), arginine aminopeptidase (ARG), β-1,4-glucosidase (BG), β-D-1,4-cellobiosidase (CBH), glutamine aminopeptidase (GLU), N-acetyl-β-D-glucosaminidase (NAG), leucine aminopeptidase (LAP), acid phosphatase (PHOS), tyrosine aminopeptidase (TYR), and urease (UREA) were measured via high-throughput microplate assays (Saiya-Cork et al., 2002; Weintraub et al., 2007). Assays were carried out under optimal conditions with excess substrate to ensure activities were not limited by substrate availability; therefore, activity rates are a measure of enzyme potential and not absolute activity. Soil slurries for each sample were made by homogenizing 1 g of soil in 125 ml of distilled water with a hand blender (Cuisinart®) for 30 s. Slurries were found to match soil sample pH and did not require buffer. Slurries were stirred constantly while 200 μl was pipetted into 24 replicate wells of a 96-well microplate. For the fluorescent assays (ALA, ARG, BG, CBH, GLU, NAG, LAP, PHOS, TYR), 16 replicate wells of each sample received 50 μl of fluorogenic substrate associated with the target enzyme. The remaining eight replicate wells were used to determine quench coefficients by adding 50 μl of fluorogenic standard corresponding to the fluorescent molecule attached to the substrate, either 4-methylumbelliferone (MUB) for BG, CBH, NAG, and PHOS or 7-amino-4-methylcoumarin (MC) for ALA, ARG, GLU, LAP, and TYR. For colorimetric assays (UREA), 16 replicate wells of each samples received 10 μl of urea. The remaining eight replicates were used to assess background absorbance and received 10 μl of DI water. Both fluorescent and colorimetric assays included eight replicates each of blanks (200 μl slurry plus 50 μl DI water) and negative controls (200 μl substrate plus 50 μl DI water). All plates were incubated for 18–24 h and then read on a Syngery H1 plate reader (BioTek Instruments, Inc.). Fluorometric assays were read at excitation of 370 nm and emission of 455 nm for MUB-labeled substrates and excitation of 350 nm and emission of 430 nm for MC-labeled substrates. Urea plates were analyzed for ammonium production via the Berthelot method (Rhine et al., 1998) and read at absorbance of 610 nm. 100 μl from all wells of the urea plates was transferred to fresh 96-well clear plates before reading to reduce interference of soil particles. Enzyme activities were corrected for slurry volume and soil moisture content and reported as nmol activity g^{-1} dry soil h^{-1}.

2.6 | Soil diazotroph community composition

Soil DNA was extracted from 0.25 g soil via recommended procedures of the Qiagen DNeasy PowerSoil kit (QIAGEN). Soil diazotroph communities were characterized by sequencing the *nifH* functional gene that was amplified using the IGK3 (5′-GCIWTHTAYGGIAARGGIGGIATHGGIAA-3′) forward primer and DVV (5′-ATIGCRAAICCICCRCAIACIACRTC-3′) reverse primer. The IGK3/DVV primer pair has been identified as an optimal primer set for capturing the widest diversity of diazotrophs (Gaby & Buckley, 2012).

PCR reactions to amplify *nifH* were carried out in two stages to first amplify *nifH* genes (Stage 1) and then to attach linker sequences (Stage 2). Reaction mixtures and PCR programs were as follows. Stage 1 was a 15-μl reaction with 0.9 μl of DNA extract, 1X AmpliTaq Gold 360 Master Mix (Applied Biosciences), and 1 μM concentrations of both the forward and reverse primers. Stage 1 PCR reactions were carried out as follows: 95°C start for 10 min, 34 cycles of denaturation at 95°C for 30 s, annealing at 54°C for 45 s, and extension at 72°C for 40 s, final extension at 72°C for 7 min. Amplification of target gene was confirmed via gel electrophoresis (1.5% gel agar, 90 V,

45 min) before proceeding to Stage 2 reactions. Stage 2 was a 20-μl reaction with 1.2 μl of Stage 1 PCR product, 1X AmpliTaq Gold 360 Master Mix (Applied Biosciences), and 1 μM concentrations of both the forward and reverse primers. An IGK3/DVV primer pair with linker sequences was used for Stage 2 reactions. Stage 2 was carried out as follows: 95°C start for 10 min, four cycles of denaturation at 95°C for 30 s, annealing at 56°C for 45 s, and extension at 72°C for 40 s, final extension at 72°C for 7 min. Successful amplification and the absence of non-target products after Stage 2 were confirmed via gel electrophoresis (1.5% gel agar, 90 V, 45 min).

Stage 2 PCR products were submitted for sequencing after first quantifying product concentrations via Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific). Products were normalized to 180–200 ng μl^{-1} DNA before library prep. Samples were submitted to the Michigan State University RTSF Genomics Core Facility for library prep and sequencing on the Illumina MiSeq platform (Illumina) with MiSeq standard reagent kit v.2 and 2 × 250 bp paired-end reads.

Sequence processing was performed following a modified version of the NifMAP pipeline (Angel et al., 2018). Forward and reverse reads were merged via USEARCH v. 10.0.240 fastq_mergepairs resulting in 12,384,888 sequences. Sequences were quality and length filtered to maximum expected errors of 1 and minimum length of 380 bp via USEARCH v. 10.0.240 fastq_filter, filtered for non-*nifH* reads using four Hidden-Markov Models (HMM), and then frameshift corrected and translated to protein sequences using Framebot (Wang et al., 2013). Frameshift-corrected sequences were then filtered again for homologs using HMM as described in Angel et al. (2018). Filtered sequences (6,120,550) were then dereplicated using USEARCH v. 10.0.240 fastx_uniques and clustered using USEARCH v. 10.0.240 cluster_otus which also filters chimeras. In total, 6534 reference OTUs were identified. Sequences were then mapped back to reference OTUs using USEARCH v. 10.0.240 usearch_global at 97% similarity. 95.7% of filtered and dereplicated sequences successfully mapped to reference OTUs. Taxonomy was assigned to reference OTUs using Blast+v. 2.7.1 blastn command, queried against the Gaby and Buckley (2014) *nifH* sequence database. Taxonomy was assigned according to percent similarity using empirically derived cutoffs of 75% similarity for family, 88.1% for genus, and 91.9% for species (Gaby et al., 2018). All other taxonomic assignments matching at <75% similarity were only assigned at the order level. Finally, a phylogenetic tree was constructed by first aligning sequences to an amino acid reference alignment (Angel et al., 2018) using MAFFT v. 7.305 and then generating a tree via FastTree v. 2.1.9.

2.7 | Statistical analysis

Prior to statistical analysis, pseudoreplicate data from within each split-plot, in each block were averaged, so that the 24 total individual soil samples became eight samples (four replicates per fertilizer treatment). A total of 240 samples were collected over the two growing seasons (120 samples per growing season). FLNF rates were analyzed using a repeated measures ANOVA with Tukey's adjustment for multiple comparisons using the *lmerTest* R package with year, sampling date, fertilizer N treatment, and their interactions as fixed effects and fertilizer N treatment nested within field replicate as a random effect were considered significant at $p \leq 0.05$.

nifH OTU counts were first rarefied to an even sampling depth of 500 using *rarefy_even_depth* in the R *phyloseq* package. This was chosen based on the rarefaction curves which indicated that a reasonable amount of diversity was captured within this sample size, while limiting loss of samples due to low OTU counts. However, 30 samples were still removed from downstream analysis due to too few OTU counts. Interestingly, these tended to be from early season sampling dates (March–May) surrounding fertilizer addition. For example, in 2017, all four fertilized replicates sampled on May 15, the first sampling data following fertilizer application, were removed because of low OTU counts.

We used Bray–Curtis dissimilarity to generate distance matrices for evaluating beta diversity using the distance function in R *phyloseq* and then ordinated via principal coordinate analysis (PCoA) using the *ordinate* function in R phyloseq. We used *adonis* in the R *vegan* package to conduct PERMANOVA of the Bray–Curtis distance matrices by year and month of sample collection, fertilizer treatment, and field replicate with field replicate as a repeated measure. Differences between years, months, fertilizer treatment, and field replicates were considered significant at $p \leq 0.05$. We examined differences in alpha diversity metrics and relative abundance of diazotroph community classes by year using repeated measures ANOVAs followed by Tukey's post hoc test with year, fertilizer, and their interaction as fixed effects and fertilizer nested within field replicate as a random effect. Additionally, we examined differences in alpha diversity metrics and relative abundance of diazotroph community classes by sampling date using a repeated measures ANOVAs followed by Tukey's post hoc test with sampling date, fertilizer, and their interaction as fixed effects and fertilizer nested within field replicate as a random effect. Differences in relative abundance of diazotroph community classes by sampling date were examined within each year, respectively. For all repeated measures ANOVAs, results were considered significant at $p \leq .05$.

Lastly, we explored diazotroph community data for OTUs whose abundance is potentially associated with high FLNF rates using the *indicspecies* package in R. This function identifies indicator species based on both abundance (specificity) and presence/absence (fidelity) in a category (Dufrêne & Legendre, 1997). Specificity, fidelity, and indicator values calculated as specificity*fidelity*100 are reported (Dufrêne & Legendre, 1997). FLNF rates were first binned into three categories: <0.5, 0.5–1, and >1 μg N fixed g^{-1} dry soil day^{-1}. We then looked for OTUs associated with samples in the >1 FLNF rate bin of which only three samples were placed.

Environmental and biological controls on FLNF were evaluated using a random forest regression algorithm via the *randomForest* package in R using the following parameters: ntree = 5000, mtry = 11 (default), NA values replaced with the median (numeric predictors) or the mode (categorical predictors). Regressions were performed using all available data for each sampling date. Two regression analyses were carried out. Regression #1 was carried out using the full available dataset including all environmental and biological data as well as temporal and site-specific categorical metrics (e.g., sampling date, field replicate, and fertilizer treatment). Regression #2 examined the full dataset without inclusion of temporal and site data. Regression #1 was intended to identify spatial and temporal controls on FLNF in our study system, while regression #2 was intended to explore generalizable environmental controls on FLNF.

3 | RESULTS

3.1 | Free-living nitrogen fixation rates

FLNF rates were significantly greater in 2018 than in 2017 ($F = 159.6$, $p < .001$). Fertilizer treatment was overall not a significant predictor of FLNF ($F = 2.15$, $p = 0.14$; Figure 1a). Sampling date was a significant predictor of FLNF rates ($F = 33.35$; $p < 0.001$; Figure 1b) and there was a significant interaction between sampling date and fertilizer treatment ($F = 1.91$, $p = 0.007$) where on a few sampling dates early in the 2017 season, a significant fertilizer treatment was observed. In particular, FLNF rates were much more variable across the growing season during 2017, but surprisingly consistent during the 2018 growing season.

3.2 | Diazotroph community composition

Alpha diversity of diazotroph communities measured as Chao1 and Inverse Simpson differed significantly by year, but fertilizer treatment was not significant overall (Table S1). Shannon diversity was not significantly different by year or fertilizer treatment (Table S1). All three indices indicated that alpha diversity of samples was generally greater in 2018 than in 2017 as was observed for FLNF rates. We also found that alpha diversity, for all three indices, differed significantly by sampling date, but found no significant interaction with fertilizer treatment (Table S2).

We assessed beta diversity of soil diazotroph communities via ordination of Bray–Cutis dissimilarity and PERMANOVA (Table S3). Diazotroph communities differed significantly by field replicate, reflecting spatial distribution of replicates plots in the field (Figure 2a). Therefore, field replicate was included as a random effect in PERMANOVA analysis of diazotroph community beta diversity. Although diazotroph community composition was found to differ significantly by both sampling date and fertilizer, these effects only explained a small amount of the spatial distribution observed in the principal coordinates analysis (PCoA; 0.95% and 1.5% variance

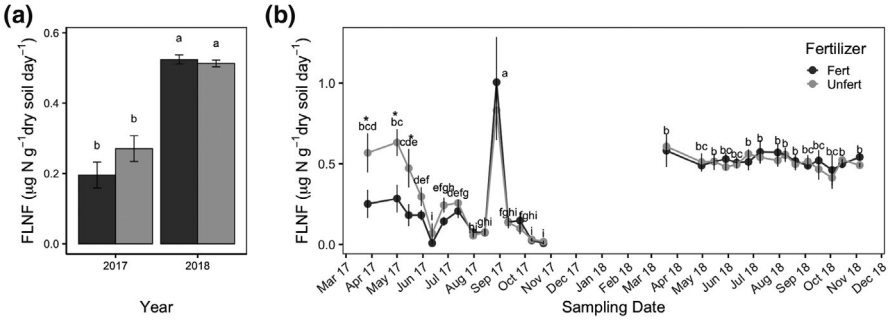

FIGURE 1 Free-living nitrogen fixation (FLNF) rates reported in μg N g^{-1} dry soil day^{-1} during the 2017 and 2018 growing seasons at the Lux Arbor field site. Samples were collected every 2 weeks from pre-emergence to post-harvest. This corresponded to sample collection from March 27 to October 23, 2017 and March 19 to November 15, 2018. (a) Average FLNF by year and fertilizer treatment ± standard error. Lowercase letters indicate significant difference at $p < 0.05$. (b) FLNF rates for each sampling data and fertilizer treatment. Each point represents average FLNF across four replicate field blocks ± standard error. Lowercase letters indicate significant differences by sampling date. Asterisks indicate significant difference by fertilizer treatment on associated sampling date

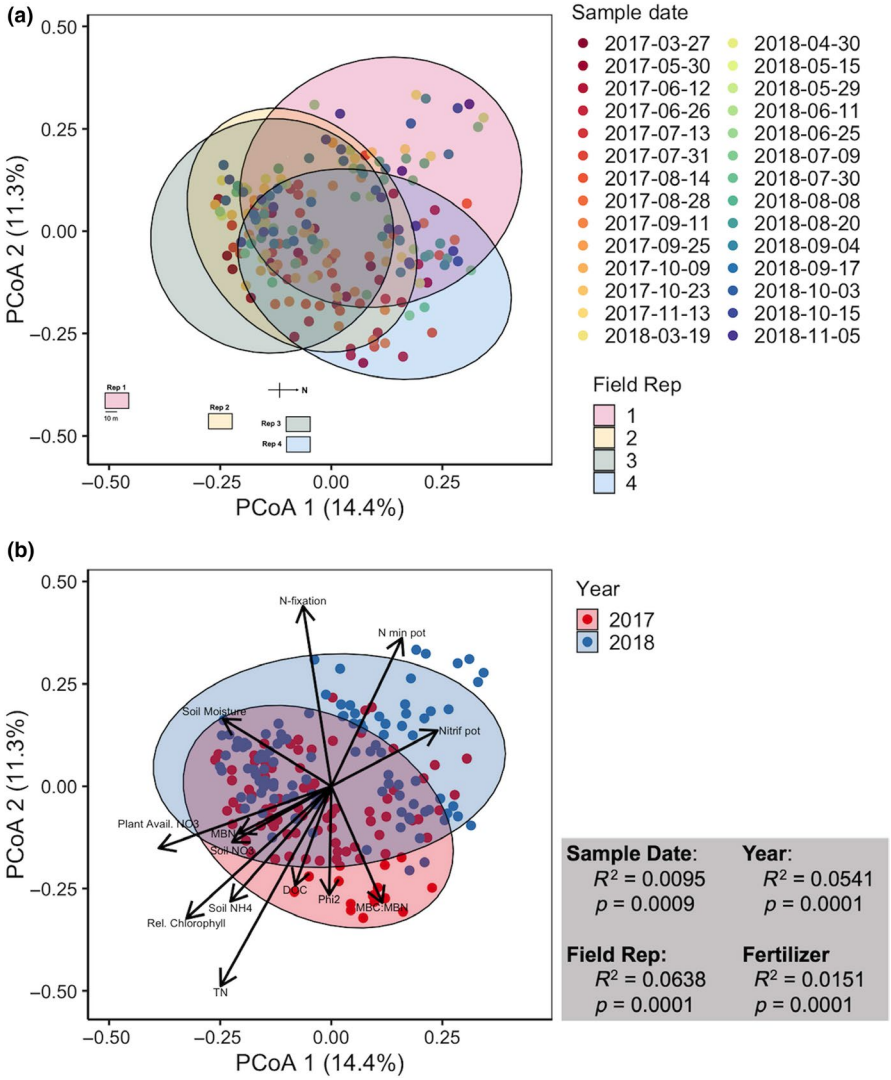

FIGURE 2 Principal coordinate analysis (PCoA) of soil diazotroph communities based on Bray–Curtis dissimilarly of relative abundances for each sample date. (a) PCoA ordination with 95% confidence ellipses representing each field replicate. Points are colored by sample date. Inset image shows distribution of field replicates at the field site. (b) PCoA ordination with 95% confidence ellipses representing growing season year and overlaid vectors of significantly correlated environmental factors ($p < 0.05$). Points are colored by growing season year. Diazotroph community composition is based on *nifH* amplicon sequencing.

explained for sampling date and field N treatment, respectively). Diazotroph community composition also differed significantly by year, accounting for ~6.3% of the variance (Figure 2b). To better understand the community members driving these temporal shifts in community composition, we also examined relative abundances of diazotroph classes across sampling dates (Figure S1A) and by year (Figure S1B). Community shifts from 2017 to 2018 showed increased abundance of Actinobacteria ($F = 39.65$, $p < 0.001$), Betaproteobacteria ($F = 62.02$, $p < 0.001$), Gammaproteobacteria ($F = 47.87$, $p < 0.001$), and Nostocales ($F = 21.49$, $p < 0.001$) and decreased abundance of Clostridia ($F = 7.35$, $p = 0.006$) and Delta/Epsilon proteobacteria ($F = 7.05$, $p = 0.007$). Given the large shifts in community composition between years, we only examined differences between collection date within their respective growing seasons. Differences in Alphaproteobacteria ($F = 3.37$, $p < 0.001$), Betaproteobacteria ($F = 3.75$, $p < 0.001$), Clostridia ($F = 2.69$, $p = 0.001$), and Nostocales ($F = 5.133$, $p < 0.001$) abundances were observed across the 2017 season. Differences in Gammaproteobacteria ($F = 2.42$, $p = 0.002$), Alphaproteobacteria ($F = 2.23$, $p = 0.005$), Bacilli ($F = 1.85$, $p = 0.027$), Betaproteobacteria ($F = 6.89$, $p < 0.001$), Nostocales ($F = 5.54$, $p < 0.001$), and Oscillatoriophycideae ($F = 2.69$, $p < 0.001$) abundances were observed across the 2018 season. There were no clear seasonal trends in these shifts, and in general, differences were the result of just a few individual dates (Figure S1A). Neither sampling date*fertilizer treatment nor year*fertilizer treatment interactions were found to have a significant impact on diazotroph community composition.

Bray–Curtis dissimilarity correlated significantly with several environmental variables including FLNF and several metrics of soil N availability (Figure 2b; Table S4). Dissimilarity also significantly correlated with metrics of plant productivity, including Phi2 and relative chlorophyll. Lastly, we identified 21 taxa potentially indicative of FLNF at the highest measured rates (>1 μg N fixed g^{-1} dry soil day; Table 1). All but two of these indicator taxa were identified as Proteobacteria and nearly half of these were members of the *Bradyrhizobium* genus.

3.3 | Random forest regression

We used random forest analysis to investigate potential environmental and biological controls on FLNF. Two analyses were conducted, the first to assess controls on FLNF specific to our study system by including temporal and site-specific variables (e.g., field replicate, fertilizer, year, and month of sampling) and the second to assess generalized controls on FLNF in which we did not include temporal or site-specific variables. Biological controls were represented by relative abundance of *nifH* phyla and classes. Environmental controls included measured soil and plant metrics and climate variables.

Random forest regression #1 revealed temporal variables as top predictors with year and month of sample identified as significant predictors of FLNF in our system (Table 2). Climate was also identified as a dominant control on FLNF rates with metrics of water availability (e.g., rainfall, soil moisture) and temperature representing the remaining significant predictors (Figure 3a). Overall, seven significant predictor variables ($p < 0.05$) were identified. Together these variables explained over 68% of the variance in FLNF and were able to very closely predict measured FLNF rates (Figure 3b).

Random forest regression #2 similarly identified several climate variables as significant, generalizable predictors of FLNF including soil moisture, rainfall, and air temperature (Figure 3c; Table 3). We also find that, generalized across temporal dynamics, metrics of plant productivity and microbial C and nutrient availability become top predictors (Figure 3c). Interestingly, the only significant biological control on FLNF identified was Gammaproteobacteria abundance, but only when generalized across temporal dynamics. Together, the top ten predictors from regression #2 explained ~53% of the variance in FLNF (Figure 3d). However, regression #2 did a relatively poor job at accurately predicting measured FLNF rates compared to regression #1 suggesting that temporal dynamics represent key controls on FLNF.

Random forest does not provide directionality nor does it account for covariance of predictors. Therefore, we explored correlations between FLNF rates and the top 10 predictors identified from regression #2 (Figure S2). We found FLNF correlated significantly and positively with soil moisture and TYR activity, but significantly and negatively with PHOS activity, Phi2, and microbial C:N. While not significant, there was a negative trend of FLNF with increasing air and soil temperature. Interestingly, though soil moisture and FLNF were positively correlated, we observed a negative trend between FLNF and rainfall. This appeared to be driven by outlying values suggesting extreme rainfall events are detrimental to FLNF activity. Lastly, we see a significant and positive correlation between Gammaproteobacteria and FLNF. However, this also appeared to be strongly influenced by outliers and is likely not a strong relationship.

4 | DISCUSSION

In this 2-year field study, we explored the relationships between switchgrass-associated FLNF, and various temporal, environmental, climatic, and biological factors. We identified several key factors that were strongly associated with FLNF rates and may improve our ability to predict and harness this important N source. Overall, we find that temporal variation, between and across growing seasons, water availability, and temperature were the strongest predictors of FLNF. Although many studies find no links between microbial community structure or functional gene abundance and function (Jansson & Hofmockel, 2018; Rocca et al., 2015), we find weak evidence for a link between diazotroph community structure (*nifH*) and FLNF rates.

As hypothesized, climate metrics, and specifically those relating to water availability, were key predictors of FLNF across both growing seasons. We see this evidenced in some of the seasonal patterns of FLNF rates. For example, FLNF measured on August 28, 2017 was significantly higher than all other sampling dates. This date also corresponded to the first significant rainfall event in several days, suggesting a sudden influx of moisture stimulated FLNF activity. Soil water availability is well known to influence microbial activity, with greater water availability generally increasing process rates (Harris, 1981; Tiemann & Billings, 2011; Wilson & Griffin, 1975; Zhang et al., 2020). FLNF has also been shown to be influenced by water availability as well as temperature, whereby FLNF is typically greater in warmer and wetter environments (Cleveland et al., 1999; Reed et al., 2011). In similar work on FLNF, observed episodic fluxes in FLNF process rates were thought to be associated with variation in soil moisture (Roley et al., 2019). Our results add to this body of work and highlight the need to understand water controls on FLNF, particularly in the face of altered precipitation regimes resulting from climate change.

Temperature, including air temperature and soil temperature of the top 10 cm of soil, was also a key predictor of FLNF identified by random forest regression (Figure 3). We found only weak correlations between FLNF and either measure of temperature, suggesting that temperature alone does not drive FLNF and is likely only significant when consider with other environmental and biological variables (Figure S2). Previous studies at larger scales disagree about the importance of climate variables on N-fixation rates. A recent global meta-analysis observed

TABLE 1 OTUs associated with the greatest FLNF rates measured (>1 μg N g^{-1} dry soil day^{-1}) identified by *indicspecies*

OTU	Specificity	Fidelity	Indicator value	*p*-value	Phylum	Class	Order	Family	Genus	Species
OTU6771	0.9618	1	0.981	.001	Proteobacteria	Alpha-proteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium	*Bradyrhizobium japonicum*
OTU7192	0.7933	1	0.891	.002	Proteobacteria	Alpha-proteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium	*Bradyrhizobium japonicum*
OTU323	0.8501	1	0.922	.003	Proteobacteria	Alpha-proteobacteria	Rhizobiales	Unknown	Unknown	Unknown
OTU6651	0.7333	1	0.856	.004	Proteobacteria	Alpha-proteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium	*Bradyrhizobium japonicum*
OTU4551	0.938	0.5	0.685	.005	Proteobacteria	Delta/epsilon subdivisions	Deltaproteobacteria	Desulfovibrionales	Unknown	Unknown
OTU439	0.9554	0.5	0.691	.006	Proteobacteria	Gamma-proteobacteria	Methylococcales	Methylococcaceae	Unknown	Unknown
OTU420	0.9788	0.5	0.7	.007	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Unknown	Unknown
OTU1860	0.9918	0.5	0.704	.01	Proteobacteria	Alpha-proteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium	*Bradyrhizobium japonicum*
OTU4501	0.6422	1	0.801	.011	Proteobacteria	Alpha-proteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium	*Bradyrhizobium japonicum*
OTU2244	0.6612	1	0.813	.011	Proteobacteria	Alpha-proteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium	Unknown
OTU1666	0.7813	1	0.884	.011	Proteobacteria	Alpha-proteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium	Unknown
OTU1346	0.9837	0.5	0.701	.014	Proteobacteria	Gamma-proteobacteria	Chromatiales	Ectothiorhodospiraceae	Unknown	Unknown
OTU6449	0.9697	0.5	0.696	.016	Proteobacteria	Alpha-proteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium	Bradyrhizobium sp. ORS 358
OTU3648	0.6318	1	0.795	.019	Proteobacteria	Alpha-proteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium	*Bradyrhizobium japonicum*
OTU20	0.6494	1	0.806	.021	Proteobacteria	Alpha-proteobacteria	Rhodobacterales	Rhodobacteraceae	Unknown	Unknown
OTU368	0.938	0.5	0.685	.022	Proteobacteria	Beta-proteobacteria	Burkholderiales	Burkholderiaceae	Unknown	Unknown
OTU216	0.8017	0.5	0.633	.026	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Paenibacillus	*Paenibacillus polymyxa*
OTU454	0.8331	0.5	0.645	.026	Proteobacteria	delta/epsilon subdivisions	Deltaproteobacteria	Desulfovibrionales	Unknown	Unknown
OTU2246	0.5584	1	0.747	.039	Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium	*Bradyrhizobium japonicum*
OTU1443	0.8194	0.5	0.64	.041	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylocystaceae	Unknown	Unknown
OTU836	0.7029	0.5	0.593	.05	Proteobacteria	Betaproteobacteria	Burkholderiales	Unknown	Unknown	Unknown

TABLE 2 Results of randomForest regression #1 including temporal and spatial variables. Predictors are sorted by level of importance based on percent increase in mean-squared error (%IncMSE)

Predictor	%IncMSE	*p*-value (%IncMSE)	INP	*p*-value (INP)
Year[*,†,§]	108.57	0.01	2.94	0.01
Month[*,†,§]	47.78	0.01	1.29	0.01
Soil moisture[*,†,§]	24.52	0.01	0.65	0.04
Air temperature[*,†,§]	23.57	0.01	0.40	0.01
Rainfall[*,†,§]	23.54	0.01	0.46	0.01
Soil temperature[*,†,§]	21.44	0.01	0.30	0.05
Days since last rainfall[*,†,§]	18.11	0.01	0.22	0.04
Phi2[*]	15.77	0.01	0.18	0.81
Gammaproteobacteria[*]	15.57	0.02	0.45	0.15
NAG activity[*]	15.21	0.02	0.23	0.97
PHOS activity[*]	14.68	0.01	0.34	0.34
GLU activity[*]	11.89	0.03	0.25	0.74
Field replicate[*,†]	11.57	0.01	0.19	0.09
Plant available ammonium[*]	10.61	0.04	0.15	1.00
Plant height	10.51	0.09	0.18	0.98
TYR activity	9.84	0.09	0.21	0.95
Microbial C:N	9.41	0.07	0.36	0.26
Specific leaf area	8.62	0.08	0.17	1.00
CBH activity	8.54	0.12	0.14	1.00
Shoot dry weight	8.31	0.20	0.28	0.37
Dissolved organic C	8.03	0.15	0.27	0.85
Soil nitrate	7.40	0.18	0.17	1.00
Soil ammonium	7.35	0.30	0.16	1.00
Alphaproteobacteria	7.29	0.12	0.12	1.00
Relative chlorophyll	7.21	0.08	0.29	0.39
Total N	7.17	0.21	0.33	0.41
Fertilizer treatment[*,†]	7.03	0.01	0.07	0.05
N mineralization potential	6.69	0.35	0.10	1.00
Microbial biomass C	6.62	0.14	0.20	0.99
LAP activity	6.38	0.27	0.11	1.00
Urease activity	5.98	0.13	0.22	0.99
ALA activity	5.74	0.40	0.10	1.00
Microbial biomass N	5.74	0.23	0.25	0.96
Actinobacteria_Abundance[*]	5.47	0.05	0.12	0.35
Nitrification potential	5.13	0.36	0.13	1.00
BG activity	5.11	0.41	0.19	1.00
Proteobacteria	5.09	0.24	0.07	1.00
Clostridia	4.58	0.11	0.09	0.68
Soil pH	4.43	0.18	0.08	0.08
pot_inorg	4.07	0.68	0.10	1.00
ARG activity	3.85	0.47	0.07	1.00
Nostocales	3.65	0.43	0.06	1.00

TABLE 2 (Continued)

Predictor	%IncMSE	*p*-value (%IncMSE)	INP	*p*-value (INP)
Actinobacteria (class)	3.60	0.10	0.11	0.44
Cyanobacteria	2.98	0.50	0.06	1.00
Betaproteobacteria	2.54	0.36	0.13	1.00
Firmicutes	2.20	0.24	0.16	0.44
Oscillatoriophycideae	2.16	0.26	0.01	1.00
Bacilli	2.09	0.24	0.23	0.13
Chlorobi	0.40	0.46	0.01	1.00
Spirochaetia	−0.02	0.38	0.00	0.98
Chlorobia	−0.06	0.51	0.01	1.00
Spirochaetes	−1.54	0.77	0.00	1.00
Delta/Epsilon	−7.44	0.98	0.33	0.17

*Denotes significant predictors at $p < 0.05$ based on percent increase in mean-squared error (%IncMSE).

†Denotes significant predictors at $p < 0.05$ based on increase in node purity (INP).

§Denotes significant predictors based on %IncMSE and INP significance.

only weak correlations between N fixation and climate variables, including temperature (Davies-Barnard & Friedlingstein, 2020). However, other globally focused work identified climate-related variables like evapotranspiration as drivers of N fixation suggesting global patterns of N-fixation and climate variables do not necessarily predict fine-scale (regional and smaller) measures of N fixation (Cleveland et al., 1999). Our findings suggest that climate is an important driver of FLNF when measured at fine spatial and temporal scales. Our results highlight the need for studies of FLNF that target high spatial and temporal resolution as FLNF may predominately be a hot spot, hot moment process sensitive to fine-scale patterns in environmental conditions (Roley et al., 2019; Smercina et al., 2020).

We expected that increasing N availability (fertilizer addition) would reduce FLNF rates by making FLNF less energetically favorable. It is generally accepted and has been demonstrated that increased external N availability decreases rates of FLNF (Hobbs & Schimel, 1984; Kox et al., 2016; Patra et al., 2007; Reed et al., 2011). Because N fixation is energy intensive and diazotrophs are not solely reliant on N fixation, it is likely that FLNF is downregulated as diazotrophs access fertilizer N or other high- and low-molecular weight N sources, including organic N forms (Norman & Friesen, 2017). However, we find little evidence of N availability controls on FLNF. For example, microbial C:N, a metric that highlights microbial C and N status, was negatively correlated with FLNF. This contrasts other recent studies that suggest FLNF is significantly and positively driven by C:N ratios (Dai et al., 2021; Zheng et al., 2019). In our study, N limitation alone did not seem to stimulate FLNF. This is similarly evidenced in the positive relationship between FLNF and TYR and a very weakly negative relationship between FLNF and NAG. TYR, an aminopeptidase, and NAG, a chitinase are enzymes that acquire both C and N and we hypothesized that such enzyme activities would be negatively correlated with FLNF as they increase microbial N availability. However, our observed relationships suggest that these enzymes may have led to predominately C rather than N acquisition. Like N availability, C availability alone does not seem to drive FLNF. Surprisingly, we found a significant, negative correlation between FLNF and Phi2, a measure of plant photosynthetic activity and potential for belowground C allocation. We hypothesized that increased plant productivity would promote greater FLNF through increased C availability. Although increasing Phi2 does seem to indicate increased microbial C availability as evidenced by a significant and positive relationship between these variables (Figure S3), this increased microbial C does not translate to greater FLNF. Overall, our results suggest that at a fine spatial and temporal scale, C or N availability alone are not good indicators of FLNF as suggested by larger scale studies (Cleveland et al., 1999; Reed et al., 2011; Vitousek et al., 2013; Zheng et al., 2019).

Lastly, we examined the relationships between the diazotrophic microbial community and FLNF rates. Overall diazotroph community composition correlated with FLNF as well as a variety of metrics relating to C and N availability. Random forest analysis indicates that this relationship may be driven by shifts in Gammaproteobacteria abundance; however, there is only a weak correlation between FLNF and Gammaproteobacteria. As with temperature above, this suggests that Gammaproteobacteria abundance alone does not drive FLNF but may be important when

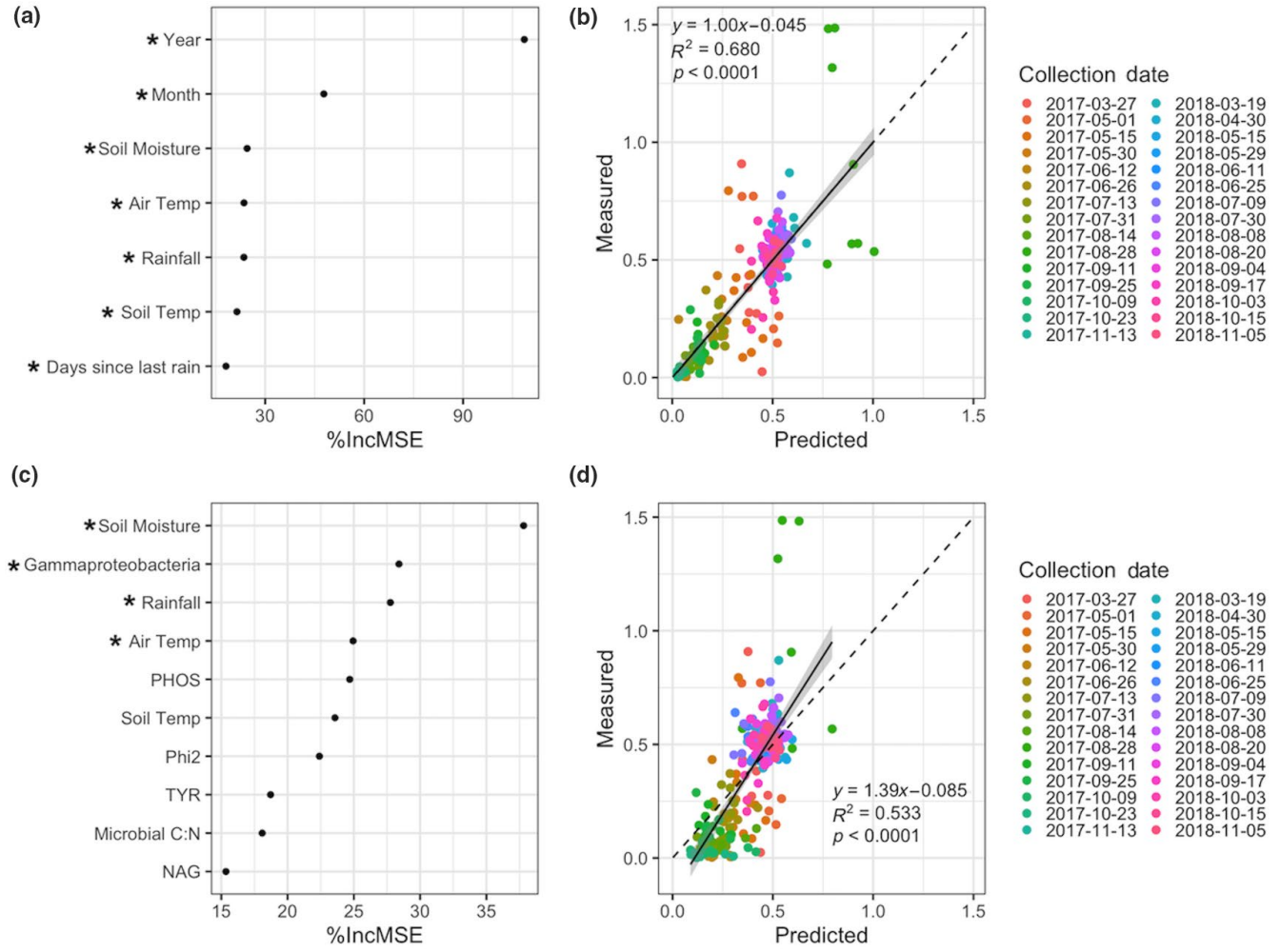

FIGURE 3 Results of random forest regression analysis for regression #1 (temporally and spatially explicit) and regression #2 (general). (a) Important predictors of FLNF identified by regression #1. Asterisks show predictors found to be significant by two metrics, percent increase in mean square error (%IncMSE; shown on figure) and increase in node purity. (b) Regression #1 model predicted versus measured FLNF rates. Points are colored by collection date. Dashed line represents 1:1 fit line. Solid line shows linear regression of predicted versus measured line with 95% confidence bands. (c) Top 10 environmental and biological predictors of FLNF identified by regression #2. Asterisks show predictors found to be significant by two metrics, percent increase in mean square error (%IncMSE; shown on figure) and increase in node purity. (d) Regression #2 model predicted versus measured FLNF rates. Points are colored by collection date. Dashed line represents 1:1 fit line. Solid line shows linear regression of predicted versus measured line with 95% confidence bands. Soil moisture, measured in the top 10 cm of soil, is reported in percent moisture by mass. Air and soil temperature, measured in the top 10 cm of soil, is reported in °C. Rainfall is report in mm. Gammaproteobacteria were reported as relative abundances. NAG, PHOS, and TYR activity is measured in nmol activity g^{-1} dry soil h^{-1}

considered with other influential factors. Additionally, there were no obvious patterns in diazotroph community composition along the FLNF correlation vector and no obvious shifts in community composition corresponding to patterns in FLNF rates.

There is a clear shift in diazotroph community composition between growing seasons like the significant difference in FLNF between years. Examining relative abundance of diazotroph classes between years (Figure S1B) suggests these shifts were driven by increased abundance of Actinobacteria, Betaproteobacteria, Gammaproteobacteria, and Nostocales in 2018 relative to 2017. Members of these classes may also play a role in the increased FLNF observed in 2018. Indeed, as noted above, Gammaproteobacteria abundance was identified as a key predictor of FLNF. These community differences provide evidence that community membership can impact rates of function and in the case of FLNF may suggest a lack of functional redundancy for this process. We also identify several potential indicator taxa associated with greater FLNF rates. These taxa were predominately Alphaproteobacteria of the order *Rhizobiales*, including several *Bradyrhizobium*. Interestingly, Alphaproteobacteria abundance did not differ significantly by year, suggesting that individual taxa

TABLE 3 Results of randomForest regression #2. Predictors are sorted by level of importance based on percent increase in mean-squared error (%IncMSE)

Predictor	%IncMSE	*p*-value (%IncMSE)	INP	*p*-value (INP)
Soil moisture*,†,§	37.80	0.01	1.09	0.01
Gammaproteobacteria*,†,§	28.40	0.01	0.83	0.02
Rainfall*,†,§	27.75	0.01	0.52	0.01
Air temperature*,†,§	24.94	0.01	0.36	0.05
PHOS activity*	24.69	0.01	0.65	0.06
Soil temperature*	23.57	0.01	0.33	0.08
Phi2*	22.40	0.01	0.34	0.32
TYR activity*	18.72	0.01	0.46	0.14
Microbial C:N*	18.08	0.02	0.60	0.07
NAG activity*	15.35	0.04	0.29	0.79
Total N*	14.59	0.01	0.49	0.12
Shoot dry weight*	14.04	0.03	0.34	0.31
Soil nitrate*	13.86	0.04	0.32	0.23
Alphaproteobacteria*	12.43	0.05	0.23	0.75
Days since last rainfall*	12.33	0.01	0.15	0.21
Clostridia*	12.28	0.02	0.20	0.12
BG activity*	12.17	0.02	0.30	0.60
Dissolved organic C	11.81	0.06	0.38	0.51
GLU activity	10.87	0.06	0.28	0.61
Actinobacteria (phylum)*	10.25	0.04	0.22	0.09
Plant height	10.10	0.09	0.23	0.73
Nitrification potential	9.50	0.11	0.23	0.95
Mineralization potential	8.63	0.26	0.15	1.00
Proteobacteria	8.12	0.10	0.13	1.00
Actinobacteria (class)*	7.85	0.04	0.22	0.08
Plant available nitrate	7.85	0.11	0.16	1.00
Soil ammonium	7.69	0.33	0.25	0.70
Specific leaf area	7.45	0.15	0.25	0.86
Microbial biomass C	7.34	0.14	0.29	0.62
Microbial biomass N	7.05	0.14	0.31	0.71
Betaproteobacteria	6.47	0.09	0.24	0.75
LAP activity	6.27	0.31	0.18	1.00
ALA activity	6.21	0.35	0.15	1.00
Relative chlorophyll	5.75	0.20	0.30	0.50
Soil pH	5.10	0.15	0.10	0.17
ARG activity	5.10	0.25	0.11	1.00
Urease activity	4.52	0.29	0.35	0.67
Nostocales	4.26	0.37	0.08	1.00
Plant available ammonium	4.08	0.54	0.14	1.00
Firmicutes	4.08	0.10	0.20	0.20
CBH activity	3.10	0.62	0.17	1.00
Cyanobacteria	2.66	0.57	0.08	1.00

(Continues)

TABLE 3 (Continued)

Predictor	%IncMSE	*p*-value (%IncMSE)	INP	*p*-value (INP)
Chlorobia	2.07	0.16	0.02	1.00
Chlorobi	1.25	0.23	0.02	1.00
Oscillatoriophycideae	0.84	0.42	0.02	1.00
Spirochaetes	−0.44	0.43	0.00	0.88
Spirochaetia	−1.11	0.62	0.00	0.82
Bacilli	−1.21	0.69	0.27	0.06
Delta/Epsilon a	−6.41	0.99	0.46	0.02

*Denotes significant predictors at $p < 0.05$ based on percent increase in mean-squared error (%IncMSE).

†Denotes significant predictors at $p < 0.05$ based on increase in node purity (INP).

§Denotes significant predictors based on %IncMSE and INP significance.

may play a large role in driving FLNF than higher level groupings. Previous work has identified *Bradyrhizobium* in the switchgrass rhizosphere, including as indicators of greater FLNF rates (Bahulikar et al., 2014; Dai et al., 2021; Roley et al., 2019; Smercina et al., 2020) and may suggest an important, yet overlooked role for these well-studied symbionts in grassland systems. In general, we find weak, but significant links between the diazotroph community and FLNF rates. Difficulties establishing mechanistic links between microbial community structure and process rates, even based on functional gene, are a well documented in soil microbial ecology (Jansson & Hofmockel, 2018; Rocca et al., 2015) and for FLNF in particular (Fürnkranz et al., 2008; Knief et al., 2012; Smercina et al., 2020). This suggests that FLNF may be controlled by many different taxa with different environmental controls, leading to seemingly spurious patterns in the whole community, or that rates are governed by short-term changes in activity rather than abundance of community members. Our work here further highlights the need to better understand the contributions of individual diazotroph community members, including those typically considered symbiotic N fixers, to switchgrass systems.

4.1 | Implications for sustainable bioenergy production

The results of this work have strong implications for switchgrass bioenergy crop management particularly in the context of climate change and marginal lands. First, using 2-week average FLNF rates from our high-resolution sampling, we determine that N inputs from FLNF have the potential to meet and exceed N deficits in switchgrass systems. By extrapolating our results from μg N fixed g^{-1} dry soil day^{-1} to kg N ha^{-1} $year^{-1}$, we estimate that FLNF contributed ~36 and ~89 kg N ha^{-1} in 2017 and 2018, respectively. Based on our extrapolated rates, FLNF has the potential to meet up to 80% of the N demands for optimal yields of fully established switchgrass, estimated at 109 kg N ha^{-1} $year^{-1}$ (Roley et al., 2018). Furthermore, these rates can meet ~87% to over 200% of the average N removed from unfertilized switchgrass plots at harvest (Roley et al., 2018). However, just because N is fixed does not mean this N is available or transferred to the plant. Previous work on FLNF in association with other grasses suggests that FLNF may meet upward of 50% of plant N demands (Chalk, 2016; Kuan et al., 2016; Ladha et al., 2016). A similar estimate for switchgrass was confirmed in a recent study which showed potential for switchgrass to acquire one-third of its N from FLNF (Wewalwela et al., 2020). In the context of previous studies, our work further highlights FLNF as an important N source to switchgrass cropping systems and underscores the potential for successful bioenergy production without fertilizer N addition.

Lastly, our results highlight the potential impacts of climate change on N inputs from FLNF and, indirectly, on the productivity of low-management switchgrass systems under changing climate conditions. Climate variables relating to temperature and water availability were key predictors of FLNF in our study. Our results suggest that FLNF is generally promoted by greater soil moisture but may be suppressed under extreme rainfall events. This suggests that FLNF within switchgrass systems is particularly sensitive to oxygen availability and thus may be dominated by aerobic and micro-aerobic organisms over strict anaerobes. We also observe a slightly negative trend in FLNF with increasing air and soil temperature, though this was not significant when considered independent of other variables (Figure S2). Taken together with predicted climate changes of a warmer, more drought-prone world punctuated with extreme precipitation events (IPCC et al., 2018), we can anticipate that climate change will reduce N inputs from FLNF. We can see anecdotal evidence of this by comparing our growing season years. We measured greater FLNF on average during the 2018 growing season

which also tended to be cooler and wetter than 2017 (Figure S4). Shifts in the diazotroph community between these growing seasons also suggest that climate change will reduce overall diversity and alter who contributes to FLNF. For example, lower abundances of cyanobacteria, like Nostocales members, in 2017 suggest that contributions from autotrophic diazotrophs may diminish with a changing climate. While we are still learning about the potential to harness FLNF for more sustainable crop production, contributions from this "free" N source may be dwindling under a changing climate. It is therefore vital that we investigate FLNF under the projected climate conditions of a warmer world in order to anticipate how we may benefit from this important biological process.

5 | CONCLUSIONS

The results of this study find that climate conditions are the most dominant predictors of FLNF activity compared to other environmental and biological variables explored here. We find precipitation and precipitation events to be strong predictors of FLNF rates with potential strong explanatory power for observed increases in FLNF process rates. In particular, these associations between climate metrics and FLNF suggest that N contributions from FLNF are vulnerable to climate change. Our findings also indicate that the relationship between soil N and FLNF at fine spatial and temporal scales may be more complex than previously thought when compared to ecosystem-scale patterns observed in global and meta-analysis studies. Thus, altering N fertilization regimes alone may not promote greater FLNF. We also find only weak evidence that changes in diazotroph community structure are directly associated with FLNF rates, though identify some specific members as potentially important to high rates, warranting further investigation. Overall, our work reveals the need for studies of FLNF at high spatial and temporal resolution. We find that FLNF contributions to switchgrass systems may be underestimated and the generalized controls on FLNF, like climate and soil nitrogen availability, are likely site and/or scale dependent.

ACKNOWLEDGMENTS

This work was supported by the MMPRNT project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research award DE-SC0014108. This work was also supported in part by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-SC0018409 and DE-FC02-07ER64494), by the National Science Foundation Long-Term Ecological Research Program (DEB 1832042) at the Kellogg Biological Station, and by Michigan State University AgBioResearch. We would like to thank all members of the MMPRNT team who helped to collect and process samples over the two growing seasons. We also thank Benli Chai and Jim Cole for assistance with bioinformatics. We would also like to thank the RTSF Genomics Core at Michigan State University for providing sequencing services, timely sample analysis, and expertise. Lastly, D.N.S. is also grateful for the support of the Linus Pauling Distinguished Postdoctoral Fellowship program.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

Data presented within this manuscript can be provided upon request to the corresponding author and will additionally be archived in the Dryad data repository (https://datadryad.org/stash/) upon project completion.

ORCID

Darian N. Smercina https://orcid.org/0000-0002-8484-3827
Sarah E. Evans https://orcid.org/0000-0001-6728-4499
Maren L. Friesen https://orcid.org/0000-0002-4274-8928
Lisa K. Tiemann https://orcid.org/0000-0003-0514-6503

REFERENCES

Angel, R., Nepel, M., Panhölzl, C., Schmidt, H., Herbold, C. W., Eichorst, S. A., & Woebken, D. (2018). Evaluation of primers targeting the diazotroph functional gene and development of NifMAP—A bioinformatics pipeline for analyzing nifH amplicon data. *Frontiers in Microbiology*, *9*, 703. https://doi.org/10.3389/fmicb.2018.00703

Bahulikar, R. A., Torres-Jerez, I., Worley, E., Craven, K., & Udvardi, M. K. (2014). Diversity of nitrogen-fixing bacteria associated with switchgrass in the native tallgrass prairie of Northern Oklahoma. *Applied and Environmental Microbiology*, *80*(18), 5636–5643. https://doi.org/10.1128/AEM.02091-14

Bloch, S. E., Ryu, M.-H., Ozaydin, B., & Broglie, R. (2020). Harnessing atmospheric nitrogen for cereal crop production. *Current Opinion in Biotechnology*, *62*, 181–188. https://doi.org/10.1016/j.copbio.2019.09.024

Campbell, W. H., Song, P., & Barbier, G. G. (2006). Nitrate reductase for nitrate analysis in water. *Environmental Chemistry Letters*, *4*(2), 69–73. https://doi.org/10.1007/s10311-006-0035-4

Chalk, P. M. (2016). The strategic role of ^{15}N in quantifying the contribution of endophytic N_2 fixation to the N nutrition of non-legumes. *Symbiosis*, *69*(2), 63–80. https://doi.org/10.1007/s13199-016-0397-8

Cleveland, C. C., Townsend, A. R., Schimel, D. S., Fisher, H., Howarth, R. W., Hedin, L. O., & Elseroad, A. (1999). Global patterns of terrestrial biological nitrogen (N_2) fixation in natural ecosystems. *Global Biogeochemical Cycles*, *13*(2), 623–645.

Dai, X., Song, D., Guo, Q., Zhou, W., Liu, G., Ma, R., Liang, G., He, P., Sun, G., Yuan, F., & Liu, Z. (2021). Predicting the influence of fertilization regimes on potential N fixation through their effect on free-living diazotrophic community structure in double rice cropping systems. *Soil Biology and Biochemistry*, *156*, 108220. https://doi.org/10.1016/j.soilbio.2021.108220

Davies-Barnard, T., & Friedlingstein, P. (2020). The global distribution of biological nitrogen fixation in terrestrial natural ecosystems. *Global Biogeochemical Cycles*, *34*(3), e2019GB006387. https://doi.org/10.1029/2019GB006387

Davis, S. C., Parton, W. J., Dohleman, F. G., Smith, C. M., Del Grosso, S., Kent, A. D., & DeLucia, E. H. (2010). Comparative biogeochemical cycles of bioenergy crops reveal nitrogen-fixation and low greenhouse gas emissions in a Miscanthus× giganteus agro-ecosystem. *Ecosystems*, *13*(1), 144–156.

Dufrêne, M., & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, *67*(3), 345–366. https://doi.org/10.2307/2963459

Fürnkranz, M., Wanek, W., Richter, A., Abell, G., Rasche, F., & Sessitsch, A. (2008). Nitrogen fixation by phyllosphere bacteria associated with higher plants and their colonizing epiphytes of a tropical lowland rainforest of Costa Rica. *The ISME Journal*, *2*(5), 561–570. https://doi.org/10.1038/ismej.2008.14

Gaby, J. C., & Buckley, D. H. (2012). A comprehensive evaluation of PCR primers to amplify the nifH gene of nitrogenase. *PLoS ONE*, *7*(7), e42149. https://doi.org/10.1371/journal.pone.0042149

Gaby, J. C., & Buckley, D. H. (2014). A comprehensive aligned nifH gene database: A multipurpose tool for studies of nitrogen-fixing bacteria. *Database*, *2014*. https://doi.org/10.1093/database/bau001

Gaby, J. C., Rishishwar, L., Valderrama-Aguirre, L. C., Green, S. J., Valderrama-Aguirre, A., Jordan, I. K., & Kostka, J. E. (2018). Diazotroph community characterization via a high-throughput nifH amplicon sequencing and analysis pipeline. *Applied and Environmental Microbiology*, *84*(4).

Gelfand, I., Sahajpal, R., Zhang, X., Izaurralde, R. C., Gross, K. L., & Robertson, G. P. (2013). Sustainable bioenergy production from marginal lands in the US Midwest. *Nature*, *493*(7433), 514–517.

Harris, R. (1981). Effect of water potential on microbial growth and activity. *Water Potential Relations in Soil Microbiology*, *9*, 23–95.

Hobbs, N. T., & Schimel, D. S. (1984). Fire effects on nitrogen mineralization and fixation in mountain shrub and grassland communities. *Rangeland Ecology & Management/Journal of Range Management Archives*, *37*(5), 402–405. https://doi.org/10.2307/3899624

IPCC. (2018). Summary for policymakers. In: V. Masson-Delmotte, P. Zhai, H. -O. Pörtner, D. Roberts, J. Skea, P. R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J. B. R. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor & T. Waterfield (Eds.), *Global warming of 1.5°C. An IPCC special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty*. World Meteorological Organization, 32 pp.

Jansson, J. K., & Hofmockel, K. S. (2018). The soil microbiome—From metagenomics to metaphenomics. *Current Opinion in Microbiology*, *43*, 162–168. https://doi.org/10.1016/j.mib.2018.01.013

Kasmerchak, C., & Schaetzl, R. (2018). *Soils of the GLBRC marginal land experiment (MLE) sites*. Kellogg Biological Station Long-Term Ecological Research Special Publication.

Knief, C., Delmotte, N., Chaffron, S., Stark, M., Innerebner, G., Wassmann, R., von Mering, C., & Vorholt, J. A. (2012). Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *The ISME Journal*, *6*(7), 1378–1390. https://doi.org/10.1038/ismej.2011.192

Kox, M. A. R., Lüke, C., Fritz, C., van den Elzen, E., van Alen, T., Op den Camp, H. J. M., Lamers, L. P. M., Jetten, M. S. M., & Ettwig, K. F. (2016). Effects of nitrogen fertilization on diazotrophic activity of microorganisms associated with *Sphagnum magellanicum*. *Plant and Soil*, *406*(1), 83–100. https://doi.org/10.1007/s11104-016-2851-z

Kuan, K. B., Othman, R., Abdul Rahim, K., & Shamsuddin, Z. H. (2016). Plant growth-promoting rhizobacteria inoculation to enhance vegetative growth, nitrogen fixation and nitrogen remobilisation of maize under greenhouse conditions. *PLoS ONE*, *11*(3), e0152478. https://doi.org/10.1371/journal.pone.0152478

Ladha, J. K., Tirol-Padre, A., Reddy, C. K., Cassman, K. G., Verma, S., Powlson, D. S., van Kessel, C., de B. Richter, D., Chakraborty, D., & Pathak, H. (2016). Global nitrogen budgets in cereals: A 50-year assessment for maize, rice and wheat production systems. *Scientific Reports*, *6*(1), 1–9. https://doi.org/10.1038/srep19355

Mehmood, M. A., Ibrahim, M., Rashid, U., Nawaz, M., Ali, S., Hussain, A., & Gull, M. (2017). Biomass production for bioenergy using marginal lands. *Sustainable Production and Consumption*, *9*, 3–21. https://doi.org/10.1016/j.spc.2016.08.003

Norman, J. S., & Friesen, M. L. (2017). Complex N acquisition by soil diazotrophs: How the ability to release exoenzymes affects N fixation by terrestrial free-living diazotrophs. *The ISME Journal*, *11*(2), 315–326. https://doi.org/10.1038/ismej.2016.127

Patra, A., Le Roux, X., Abbadie, L., Clays-Josserand, A., Poly, F., Loiseau, P., & Louault, F. (2007). Effect of microbial activity and nitrogen mineralization on free-living nitrogen fixation in permanent grassland soils. *Journal of Agronomy and Crop Science*, *193*(2), 153–156. https://doi.org/10.1111/j.1439-037X.2006.00247.x

Reed, S. C., Cleveland, C. C., & Townsend, A. R. (2011). Functional ecology of free-living nitrogen fixation: A contemporary perspective. *Annual Review of Ecology, Evolution, and Systematics*, *42*, 489–512. https://doi.org/10.1146/annurev-ecolsys-102710-145034

Rhine, E., Mulvaney, R., Pratt, E., & Sims, G. (1998). Improving the Berthelot reaction for determining ammonium in soil extracts and water. *Soil Science Society of America Journal*, *62*(2), 473–480. https://doi.org/10.2136/sssaj1998.03615995006200020026x

Robertson, G. P., Hamilton, S. K., Barham, B. L., Dale, B. E., Izaurralde, R. C., Jackson, R. D., Landis, D. A., Swinton, S. M., Thelen, K. D., & Tiedje, J. M. (2017). Cellulosic biofuel contributions to a sustainable energy future: Choices and outcomes. *Science*, *356*(6345). https://doi.org/10.1126/science.aal2324

Rocca, J. D., Hall, E. K., Lennon, J. T., Evans, S. E., Waldrop, M. P., Cotner, J. B., Nemergut, D. R., Graham, E. B., & Wallenstein, M. D. (2015). Relationships between protein-encoding gene abundance and corresponding process are commonly assumed

yet rarely observed. *The ISME Journal*, *9*(8), 1693–1699. https://doi.org/10.1038/ismej.2014.252

Roley, S. S., Duncan, D. S., Liang, D., Garoutte, A., Jackson, R. D., Tiedje, J. M., & Robertson, G. P. (2018). Associative nitrogen fixation (ANF) in switchgrass (*Panicum virgatum*) across a nitrogen input gradient. *PLoS ONE*, *13*(6), e0197320. https://doi.org/10.1371/journal.pone.0197320

Roley, S. S., Xue, C., Hamilton, S. K., Tiedje, J. M., & Robertson, G. P. (2019). Isotopic evidence for episodic nitrogen fixation in switchgrass (*Panicum virgatum* L.). *Soil Biology and Biochemistry*, *129*, 90–98. https://doi.org/10.1016/j.soilbio.2018.11.006

Saiya-Cork, K., Sinsabaugh, R., & Zak, D. (2002). The effects of long term nitrogen deposition on extracellular enzyme activity in an *Acer saccharum* forest soil. *Soil Biology and Biochemistry*, *34*(9), 1309–1315. https://doi.org/10.1016/S0038-0717(02)00074-3

Smercina, D. N., Evans, S. E., Friesen, M. L., & Tiemann, L. K. (2019a). Optimization of the $^{15}N_2$ incorporation and acetylene reduction methods for free-living nitrogen fixation. *Plant and Soil*, *445*(1), 595–611. https://doi.org/10.1007/s11104-019-04307-3

Smercina, D. N., Evans, S. E., Friesen, M. L., & Tiemann, L. K. (2019b). To fix or not to fix: Controls on free-living nitrogen fixation in the rhizosphere. *Applied and Environmental Microbiology*, *85*(6). https://doi.org/10.1128/AEM.02546-18

Smercina, D. N., Evans, S. E., Friesen, M. L., & Tiemann, L. K. (2020). Impacts of nitrogen addition on switchgrass root-associated diazotrophic community structure and function. *FEMS Microbiology Ecology*, *96*(12), fiaa208. https://doi.org/10.1093/femsec/fiaa208

Tiemann, L. K., & Billings, S. A. (2011). Changes in variability of soil moisture alter microbial community C and N resource use. *Soil Biology and Biochemistry*, *43*(9), 1837–1847. https://doi.org/10.1016/j.soilbio.2011.04.020

Vitousek, P. M., Menge, D. N., Reed, S. C., & Cleveland, C. C. (2013). Biological nitrogen fixation: Rates, patterns and ecological controls in terrestrial ecosystems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1621), 20130119.

Wang, Q., Quensen, J. F., Fish, J. A., Lee, T. K., Sun, Y., Tiedje, J. M., & Cole, J. R. (2013). Ecological patterns of nifH genes in four terrestrial climatic zones explored with targeted metagenomics using FrameBot, a new informatics tool. *mBio*, *4*(5).

Warembourg, F. R. (1993). Nitrogen fixation in soil and plant systems. *Nitrogen Isotope Techniques*, 127–156.

Weintraub, M. N., Scott-Denton, L. E., Schmidt, S. K., & Monson, R. K. (2007). The effects of tree rhizodeposition on soil exoenzyme activity, dissolved organic carbon, and nutrient availability in a subalpine forest ecosystem. *Oecologia*, *154*(2), 327–338. https://doi.org/10.1007/s00442-007-0804-1

Wewalwela, J. J., Tian, Y., Donaldson, J. R., Baldwin, B. S., Varco, J. J., Rushing, B., Lu, H., & Williams, M. A. (2020). Associative nitrogen fixation linked with three perennial bioenergy grasses in field and greenhouse experiments. *Global Change Biology Bioenergy*, *12*(12), 1104–1117. https://doi.org/10.1111/gcbb.12744

Wilson, J. M., & Griffin, D. (1975). Water potential and the respiration of microorganisms in the soil. *Soil Biology and Biochemistry*, *7*(3), 199–204. https://doi.org/10.1016/0038-0717(75)90038-3

Zhang, S., Yu, Z., Lin, J., & Zhu, B. (2020). Responses of soil carbon decomposition to drying-rewetting cycles: A meta-analysis. *Geoderma*, *361*, 114069. https://doi.org/10.1016/j.geoderma.2019.114069

Zheng, M., Zhou, Z., Luo, Y., Zhao, P., & Mo, J. (2019). Global pattern and controls of biological nitrogen fixation under nutrient enrichment: A meta-analysis. *Global Change Biology*, *25*(9), 3018–3030. https://doi.org/10.1111/gcb.14705

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Smercina, D. N., Evans, S. E., Friesen, M. L., & Tiemann, L. K. (2021). Temporal dynamics of free-living nitrogen fixation in the switchgrass rhizosphere. *Global Change Biology Bioenergy*, 13, 1814–1830. https://doi.org/10.1111/gcbb.12893