Drivers of Diazotroph Community Structure and Co-occurrence in Agricultural Pulse Crop Rotations: A Soil Microbiome and Ecosystem Health Study

Dr. Sofia Almeida[1]*, Dr. Carlos Navarro[1]

[1]Department of Environmental Health and Soil Microbiology, Hospital Clínico San Carlos, Madrid, Spain

Abstract

The Northern Great Plains of the United States and Canada have seen an increase in pulse crops like pea, lentil, and chickpea acreage in dryland and irrigated systems. Rotating pulse crops into cereal crop production systems could reduce the demand for nitrogen fertilizer due to the symbiotic relationship with nitrogen fixing bacteria called diazotrophs. Legume-diazotroph relationships are well studied, but little is known about the free-living bacteria of the bulk soil in the legume-cereal rotation systems. Here we determine the community structure of the overall (16S) and diazotroph (*nifH*) soil microbiome in pulse-cereal rotations across the state of Montana. Many variables influenced the microbial community structure, but and among them, irrigation caused a significant shift in the composition of the communities and co-occurrence network topology. Interestingly, the diazotrophic communities in dryland soil are more robust and contain more generalist compared to irrigated soils. Co-occurrence networks support the identification of different diazotrophic keystone taxa such as Firmicutes in irrigated and Proteobacteria in dryland soils. Linking farm management practices and crop productivity with microbial community structure dynamics is an initial

step in establishing a knowledge base for farmers to exploit the soil microbiota to maintain healthy soil for crop production.

1. Introduction

Farmers are increasingly choosing to rotate pulse crops, such as peas (*Pisum sativum L.*), with cereals in their production system due to the economic and environmental benefits. Because pulse crops are annual grain legumes which establish a symbiotic relationship with nitrogen-fixing bacteria (diazotrophs) in the soil, they can reduce the need for nitrogenous fertilizers and help to manage weeds and diseases when pulse crops are rotated with wheat (*Triticum aestivum*) (Gan et al., 2015). Rotating pulse crops with cereals also benefits the environment by reducing greenhouse gas emissions in the form of CO_2 and N_2O, while increasing soil organic carbon (Chen et al., 2012; Stagnari et al., 2017). With these benefits as well as adaptation to the Northern Great Plain environment, Montana has been the country's leader in pea and lentil production since 2011 (Lee, 2011).

Healthier soil and better crop yield in low input agricultural systems correlate to higher soil microbial diversity (Hartmann et al., 2015). Conventional wheat dominant agricultural systems are more homogenous and comparatively lack microbial diversity due to monoculture and agrochemical additions (Chaudhry et al., 2012; Lupatini et al., 2017). Diversified crop rotation systems offer heterogeneity to the soil with the addition of different crops (Nannipieri et al., 1999). The specific factors of rotational crops' influence on the soil microbiome differ by the quantity and composition of root exudates as well as the organic matter left by decaying biomass (Sul et al., 2013). A healthy soil microbiome benefits crop production by providing necessary mineral acquirements and protecting crops from diseases (Berg and Koskella, 2018; Finkel et al., 2017; Van Wees et al., 2008).

Pulse crop is often grown after wheat in a crop rotation sequence, and farmers regularly do not apply any nitrogen fertilizer to pulse crops. About 60-90% of the nitrogen in pulse crop biomass is taken up through symbiotic N-fixation from atmospheric N_2, and the rest of the 10-40% N came from soil (Peoples et al., 2009; St. Luce et al., 2015). The bulk soil after pulse harvest is an environment that is limited by reduced nitrogen leading to a selection of robust diazotrophs (Tan et al., 2003). An increase in diazotrophs in the bulk soil could potentially be benefiting to the cereals in the next season by increasing overall soil nitrogen and organic matter. Cereal crop grown after pulse crop have performed better than following a cereal crop (Chen et al., 2012). The rate of selection of robust diazotrophs and benefit of the free-living population to crops is still unknown, but a basic understanding of the community will allow farmers to establish best practices for the pulse crop rotation system.

Diazotrophs are phylogenetically diverse, including both symbiotic rhizobia bacteria and free-living soil bacteria (Reed et al., 2011). To determine the phylogeny of diazotrophs in the soil, the highly conserved gene *nifH*, a subunit of nitrogenase, can be used to distinguish different genera (Hsu and Buckley, 2009). In the bulk soil, free-living diazotrophs have access to multiple sources of fixed nitrogen and most likely only fix atmospheric nitrogen under N-limiting conditions (Norman and Friesen, 2017). Nitrogen fixation is an energy-intensive reaction requiring environmental conditions with high available carbon and a low nitrogen environment (Orr et al., 2012). The heterogeneity of soil and the intense requirements of nitrogen fixation requires species to evolve numerous different physiologies to occupy niches within the soil, and diazotrophs have diversified to fix nitrogen optimally in a variety of conditions.

The potential input or removal of nitrogen into the system (nitrogen increment) by pulse crops after harvest is variable and dependent on pulse crop species and geographical location (Walley et al., 2007). The effects of nitrogen increment on the microbial ecology of the bulk soil in pulse-wheat systems is understudied with little literature about drivers of microbial diversity past the rhizosphere. A better understanding of the microbial communities in the bulk soil and the factors that influence them will help determine how the microbial soil communities at large are impacted by pulse rotation.

Microbial communities are complex and dynamic, where each member is always competing for space and resources. Niche differentiation in the rhizosphere and bulk soil is continuously changing with influence from root exudates in the rhizosphere, mineral availability, and carbon recycling in the bulk soil (Kuzyakov and Blagodatskaya, 2015; Shi et al., 2016). How free-living diazotrophs interact and compete in a nitrogen-depleted niche is only recently being discovered (Norman and Friesen, 2017). Many abiotic factors have been shown to influence the diazotrophic microbiome from soil pH (Fan et al., 2018), and organic matter (Gupta et al., 2014), to soil moisture in both flooded riparian zones (Chen et al., 2019) and permafrost thawing regions (Penton et al., 2016). Here we analyzed the general and diazotrophic bacterial communities in the bulk soil taken after growing pea crop with different farm management practices such as irrigation at different geographical locations in Montana. We hypothesized that the edaphic, geographic, and farm management practice would have a larger influence on the bulk general and diazotrophic communities than pea variety. The results provide insight into the relationship between community structure and different farming

practices, soil properties, and environmental factors. The impact and contribution of free-living diazotrophs to crop productivity in agricultural soils are still unknown. Our study lays a foundation for further discovery in the complex system of diazotroph community and farm management practices.

2. Materials and Methods

2.1 Field locations

Fields used in this study are named for their location at Montana State University Agricultural Research Centers (ARCs) (Conrad, Huntley, Moccasin, Sidney, Corvallis, and Kalispell) as wells as a farmer's land near Richland (Fig. 1). Location were chosen as they are diverse in geographical and farm management practices. These reflect the variation in microclimates statewide, representing average rainfall across a growing season (April to August) ranging from 65.28 mm (Corvallis) to 264.67 mm (Sidney).

2.2 Soil sampling

Soil samples were taken using a soil hand probe after harvesting pea crop from a statewide pea variety trial at seven locations across Montana in late August to early September (Fig. 1). At each location (either dryland or irrigation site), a pea variety trial nursery had 6 varieties and 4 replications with total 24 plots. The plots were arranged as randomized complete block design. The plot size is 1.5 m x 6 m. Peas were planted at 86 pure live seeds per square meter and managed following recommended agronomic practices at each location, no fertilizer was added to the peas (Table S1). Soil cores were taken 0-30 cm depth within crop row from each plot, so that there were 4 soil samples (replications) per variety and total 24 samples for 6 varieties. Soil cores were taken from the top 0-30 cm depth within crop row, where the majority of crop roots and

microorganisms live and is most influenced by agronomic practice. Each soil core sample was placed in a Ziploc bag and labeled with the pea variety name, replication, location, and sampling date. The soil samples were then put into a cooler containing ice packs. After sampling all plots, the samples were transported to the laboratory and stored in a freezer at -20 ºC.

2.2 Soil analysis

Each sample collected representing a field replicate was partitioned into two containers for DNA extraction and chemical analysis. First, a volume of 50 mL was collected from each replicate separately and stored at -20°C for future DNA extraction. The rest of the soil was used for chemical analysis, due to a shortage of sample amount, 100 g soil from each of the four replicates of each pea variety was mixed thoroughly to a total amount of 400g. The mixed replicates sample was then sent to the University of Idaho Analytical Science Laboratory (uidaho.edu/asl) for chemical analysis. Two soil tests were performed on aliquots from the mixed sample: 1) The Extended Fertility Test was performed to measure moisture content, pH, available phosphorus, available potassium, nitrates and nitrites, ammonia, sulfates, organic matter, and boron using standard soil chemical analysis methods. 2) The Dissolved Metal Screen Test was used to measure elements: Ca, Cd, Co, Cr, Cu, Fe, K, Mg, Mn, Mo, Na, Ni, V, and Zn with inductively coupled plasma mass spectroscopy (ICP-MS).

2.3 DNA Extraction

The DNA extraction was performed with the DNeasy PowerSoil kit (MoBio Laboratories, Inc. Carlsbad, CA, USA). The Powersoil kit was used with the Mini-beadbeater-24 (Bio-Spec Products, Bartlesville, OK, USA) DNA was quantified with

absorbance 260/280 nm on a NanoDrop 2000 (ThermoFisher Scientific, Waltham, MA,USA). An aliquot of DNA was diluted to a concentration of 20 ng/μL. DNA was stored and shipped at -20°C. The DNA was tested for amplification against the 16S rRNA 515/806 primers, and the nifH gene was tested by the IGK and DVV primers using the following protocols.

2.4 PCR gene amplification

The 16S rRNA gene V4 variable region PCR primers 515/806, forward 515-GTGYCAGCMGCCGCGGTAA and reverse 806-GGACTACNVGGGTWTCTAAT, with a sample barcode on the forward primer, were used in a 30 cycle PCR using the HotStarTaq Plus Master Mix Kit (Qiagen, Germantown, MD, USA). Amplifications were performed using the following cycling parameters: an initial single step at 94°C for 3 min (denaturation) was followed by 30 cycles of the following: (a) 94°C for 30 s (denaturation), (b) 53°C for 40 s (primer annealing) and (c) 72°C for 1 min (elongation). A final single step at 72°C for 5 min followed these 30 cycles.

The *nifH* amplicon sequencing used primers adapted from the paper by Gaby and Buckley (Gaby and Buckley, 2012), forward IGK- AARGGNGGNATHGGNAA (Ohkuma and Kudo, 1996) and reverse DVV- ATIGCRAAICCICCRCAIACIACRTC (Ando et al., 2005). A Touchdown PCR was performed with the following conditions: 94ºC for 5 minutes, followed by 10 cycles starting at 94ºC for 90 seconds, 61ºC for 60 seconds, 72ºC for 90 seconds, the first ten cycles step down -1ºC every cycle till 53ºC, followed by 20 cycles of 94ºC for 90 sec, 53ºC for 60 sec, and 72ºC for 90 sec. A final elongation step at 72ºC for 10 minutes and storage at 4ºC. After amplification, PCR products were checked in 2% agarose gel to determine the success of amplification and

the relative intensity of bands. Multiple barcoded samples were pooled together in equal proportions based on their molecular weight and DNA concentrations. Pooled samples were purified using calibrated Ampure XP beads (Beckman Coulter, Indianapolis, IN, USA).

2.5 Sequencing

The pooled and purified PCR product is used to prepare a DNA library following the Illumina TruSeq DNA library preparation protocol (Illumina, San Diego, CA, USA). Sequencing was performed at MR DNA (www.mrdnalab.com, Shallowater, TX, USA) on a MiSeq following the manufacturer's guidelines. Sequence data were processed using MR DNA analysis pipeline (MR DNA, Shallowater, TX, USA). In summary, sequences were joined, depleted of barcodes then sequences <150bp were removed, and sequences with ambiguous base calls removed. Sequences were denoised, operational taxonomic units (OTUs) generated, and chimeras removed. OTUs were defined by clustering at a 3% divergence (97% similarity). Final OTUs were taxonomically classified using BLASTn against a curated database derived from GreenGenes (DeSantis et al., 2006), RDPII (Cole et al., 2014), and NCBI (www.ncbi.nlm.nih.gov).

2.6 Statistics and data analysis

Multivariate data analysis was performed in R.3.6.1 (https://www.r-project.org/) using different major packages (vegan, phyloseq, ggplot2). Data for primary analysis was rarefied and trimmed to remove OTUs that have less than three counts. Alpha and beta diversity metrics were performed in the phlyoseq package (McMurdie and Holmes, 2013). Statistics and model building were performed with the vegan package (Oksanen

et al., 2018). A more detailed explanation of the statistical methods can be found in the supplemental methods section.

2.7 Network construction

Networks were constructed using the R package SPIEC-EASI (Sparse InversE Covariance Estimation of ecological association inference) (Kurtz et al., 2015). Data was subsetted before network creation to a prevalence of 37% and an abundance great than five. The neighborhood (MB) selection framework was used as a covariance selection method with the minimal lambda ratio set to 0.001 and nlambda set to 30 for an adequate sample of the lambda path. All network parameters and network figures were produces using the R package iGraph (Csárdi and Nepusz, 2006).

3. Results

3.1 Soil microbial ecology across Montana

The most abundant phylum in the general bacteria community (based on the sequences of 16S ribosomal RNA amplicons) are *Actinobacteria, Proteobacteria, Acidobacteria,* and *Choloflexi* (Fig. 2A). While the diazotrophic community (based on sequences of *nifH* amplicons) is dominated by two main phyla, *Proteobacteria,* and *Firmicutes* (Fig. 2B). Within the diazotrophic community, the dominating genus is *Paenibacillus,* with *Geobacter* and *Rhizobium* also being dominant but not as ubiquitous (Fig. 2C). Richland site has the most extended pea production history and shows more abundance of Rhizobium than other locations. The Shannon diversity is significantly different across the state of Montana for both general bacteria (Fig. 3A, Table S4, Kruskal-Wallis, p = 8.8e-07) and diazotrophic (Fig. 3B, Table S5, Kruskal-Wallis, p = 4.2e-06) communities. Shannon diversity of the general community and Shannon

diversity of the diazotrophic community correlated across the state ($r = 0.44$, $p < 0.001$) (Fig. S1). Irrigated sites in Corvallis, Kalispell, and Sidney tend to have high Shannon diversity of general community and diazotrophic community, but some other locations do not follow this overall trend like Huntley irrigated site which has a high diversity of diazotrophs (average observed OTUs = 368) but normal average diversity of general bacteria (average observed OTUs = 2462) and Conrad dryland which has the opposite trend of high diversity of general bacteria and average diazotrophic community diversity (16S Observed OTUs = 2643, and *nifH* observed OTUs = 246) (Table S4,S5).

The influence of farm management and abiotic factors were investigated to determine their influence on alpha diversity among the bacterial communities. Pea varieties where shown not to have any correlation with alpha diversity in general ($p = 0.96$) and diazotrophic ($p = 0.95$) communities across the state. Soil chemical elements correlate poorly to the general bacterial Shannon index and the diazotrophic Shannon index, with most of the chemistry being nested into site location. Diazotrophic diversity was found correlated with moisture content statewide ($r = 0.31$ $p = 0.027$).

To compare the community structures at each site, a non-metric multidimensional scaling (NMDS) ordination plot was constructed for the general and diazotrophic communities (Fig. 3C and D). The plots were created from the Bray-Curtis dissimilarity matrix, and a stress test was accepted (16S stress = 0.07, *nifH* stress = 0.18). There is a distinct clustering of site location in both general (16S) and diazotroph (*nifH*) communities (Fig. 3C and D, colors). There is a differentiation in the irrigated and dryland plots along a bisecting axis in both general bacteria and diazotrophs communities (Fig. 3C and D, shapes). To determine the effects of farm management

and soil chemistry on beta diversity, the *envfit* function in vegan R package was used to correlate the environmental variables within the dissimilarity matrix. The function *envfit* fits the measured variables and factors to the NMDS space, which reflects the linear relationship between the factors and the Bray-Curtis distance. With this analysis, multiple factors significantly correlate to the ordination of the bacterial communities in the NMDS (Table S6). Overall location effect highly influences the community structure with the majority of variance explained by geographical location (16S r^2 = 0.7556 p = <0.001, *nifH* r^2=0.8274 p=<0.001). While the location is dominating overall, other farm management and chemical variables also influence the community structure. Farm management practices of tillage, previous crop, and irrigation all influence both general and diazotrophic communities significantly (16S: Tillage, r^2=0.25, p = <0.001, Previous Crop, r^2=0.26, p = <0.001, *nifH*: Tillage, r^2=0.21, p = <0.001, Previous Crop, r^2=0.35, p = <0.001,) but irrigation is responsible for 58% of the variability in 16S (p = <0.001) while it accounts for 44% of the variation in *nifH* (p=<0.001). Soil moisture content, which was measured during chemical analysis, influences the general and diazotrophic communities differently (16S r^2=0.14 p= 0.012, *nifH* r^2=0.048 p= 0.287). The general community is also correlated with grain yield (a proxy for pea crop health) while diazotrophs are not correlated (16S r^2=0.18 p= 0.008, *nifH* r^2=0.02 p= 0.562). A noticeable difference is in the large correlation of organic matter with the diazotrophs community (r^2=0.4791 p=<0.001) but not in the general bacterial (r^2=0.0463 p=0.303).

3.2 Irrigation and soil microbiology

In order to be more accurate in determining the influence of irrigation on the soil microbial ecology, we analyzed a subset of the ARC sites, Sidney and Huntley, which

contained two plots that were subject to irrigation or kept as dryland. By subsetting, the data overall geographical effect seen in the statewide survey was removed.

There is a slight influence of irrigation on the alpha diversity of the general bacterial community (Fig. 4A and S4), but there is a significant difference between the irrigation and dryland for the diazotrophic community (Huntley $p < 0.0001$, and Sidney $p < 0.001$) (Fig. 4B and S5). To determine the effect of irrigation on beta diversity, an ordination plot was generated using Detrended Correspondence Analysis (DCA), in order to better understand the irrigation gradient of the samples. In this analysis, the differentiation between the irrigation type appears diverging along the DCA1 axis with corresponding to 84.5% of the variance for the general bacterial community (Fig 4C) and 42.6% of the variation for the diazotrophic community (Fig 4D). The location influence is present with a smaller effect of 7.5% and 34.4% along the DCA2 axis for the general and diazotrophic community, respectively. For the general bacteria community, the samples have less dispersion, and the irrigated samples are highly clustered, representing less location effect for irrigated compared to dryland. The diazotrophic community is highly dispersed, which was seen before in the statewide analysis, most likely due to the higher abundance of low count unique species.

3.3 Constrained ordination of bacterial communities

To more precisely discover the drivers of the microbial communities, a model selection approach has been chosen to select the combination of chemical variables that drive the community structure. *Capscale* function in the vegan package was used to constrain the variables in a non-Euclidean Bray-Curtis dissimilarity space. Then the

function *ordiR2step* was used to build a correlation model by a forward selection approach.

The models explain 51% and 38% of the variance on the general bacteria and diazotrophic community structure statewide, respectively (Table S7). Using these models made in capsacle, an ***a priori*** vector fitting can be done on a CAP ordination (Fig. 4A (16S), 4B (*nifH*)). In this CAP ordination constrained by the model, specific differences between the general and diazotrophic bacterial communities can be established. First, organic matter, which was a significant factor in the envfit NMDS space, has a more significant driving force (larger arrow) on differentiating the Kalispell site and from others in the diazotrophic ordination. This reflects a more significant effect by organic matter on the diazotrophic bacterial community compared to the general bacterial community (Fig. 5A and B).

The same model building procedure was followed for the subsetted irrigation data. This shows a similar model which contains the same compilation of chemical factors that determined the statewide model (Table S8). However, subsetting the data provide a better insight into (1) what changes in chemical composition are associated with geographical sites, and (2) what effects are due to irrigation treatments. The general bacterial CAP plot (Figure 4C) shows farm amendments like potassium, nitrate, phosphorus, and organic matter all correlate to the irrigated fields following the CAP1 axis. While elemental differentiation like cobalt and calcium are strictly geographically in character following the CAP2 axis, which separated the ARCs. For the diazotrophic CAP plot (Figure 4D), a different model was produced using the same procedure. Ammonia, the product of nitrogen fixation, is correlated to the species composition of

the diazotrophs as well as the pH, which was not correlated with the 16S communities. Interestingly the two metals nickel and copper are differentiating along the CAP1 axis showing that copper is correlated with community structure in irrigated soils and nickel is correlated with community structure in dryland soils. The unique effectors of diazotrophs, ammonia, and pH are more correlated with community structure in irrigated plots than dryland.

By fitting highly correlated (r^2> 0.9) species onto the constrained CAP ordination, we can see which species in the general and the diazotrophic bacterial communities are affected most by the chemical gradients (Figure S3). For the general population, a diverse set of bacteria from 4 different phyla (Actinobacteria, Firmicutes, Verrucomicrobia, and Cyanobacteria) are all influenced by the chemical gradients while represented diazotrophs included bacteria from 2 different classes (Alphaproteobacteria, and Bacillus).

3.4 Irrigation management influences soil microbial ecology

There is substantial evidence to support that the irrigation management of the ARC affects the composition of the microbiome, its chemical drivers, and species (Fig. 4, 5, and S5). To more deeply interpret these effects of irrigation on the community structure of the diazotrophs, a co-occurrence network was created to determine its keystone species and the effect for irrigation on the robustness of the community.

The network topological structure of the dryland and irrigated diazotrophs show a similar trend, as was seen with the diversity analysis where irrigated communities are more similar while dryland communities are more dissimilar, showing unique co-occurrence. Overall the dryland network contains more edges (274), higher average

degree (3.58), shorter path length (5.4), and smaller betweenness centrality (223) (Table S9). The modularity and subgraphs of the networks were studied by removing all nodes that were not in a major community. These results show that the dryland network has a much higher density and three times more clusters than the irrigated network (Fig. 6 and Table S10). The modularity of the network is higher in dryland, and eigenvalue centrality shows that the dryland not only has more clusters but more interaction between clusters. The dryland network is also more robust than the irrigated network shown in node removal analysis (Fig. S4).

The diazotrophic keystone species were evaluated in both the dryland and irrigated networks (Fig. 7). Keystone species can be inferred from a network topology of high degree, high closeness, and low betweenness (Banerjee et al., 2018; Berry and Widder, 2014). In the dryland network, one OTU is a prominent highly interconnected node; this OTU is annotated to species *Paenibacillus graminis*. Overall, Bacilli is more enriched as high node degree in the dryland network (Fig. 7). The Irrigated network does not have such a precise keystone species with four OTUs with the highest degree; two of them *Rhizobium legumimpsarum,* and *Azotobacter vinelandii* have high closeness, but only *Rhizobium leguminosarum* and *Rhodopseudomonas palustris* have low betweenness. In contrast to dryland, high degree OTUs are dominated by Proteobacteria (alpha and gamma).

To determine if irrigation was selecting for a narrower range of diazotrophs as it is shown in the co-occurrence, a mean abundance and prevalence plot was made on the rarified data (Figure S5 and S6). The plot of all diazotrophs across the state show some specialist (low prevalence, high abundance) but more generalist (high prevalence,

high abundance). If the samples are subsetted to the irrigated plots and dryland plots across the state, almost all the specialists from across the state are in irrigated soils while dryland continues to maintain a high level of generalists.

4. Discussion

4.1 Geography and farm management shape microbial diversity

The geographic differences are a significant portion of the variance in-between bacterial communities. This is expected in a study across such a large area with such diverse environments. Even with this diversity of samples, some broad trends can be seen, with the predominant trend to be the influence of irrigation on soil microbial composition. Recently, a long-term study showed irrigation significantly reduced the native population of Pseudomonads in the field as well as impact the community composition due to changes in the water potential and pH caused by irrigation (Mavrodi et al., 2018). We have shown that the bulk soil is also influenced by irrigation, with a notable change in total bacterial community structure.

Irrigation overlaps with other farm management practices such as tillage, and previous crop. While confounding, we can suggest that both of these additional farm management practices influence the general and diazotrophic communities. The diazotrophic community composition has previously been shown to be impacted by tillage do to influence of biomass retention homogenization of soil, these findings are supported by our survey (Hsu and Buckley, 2009). Interestingly the previous crop (fallow, wheat, or barley) had more influence on the diazotrophic community composition than on the general bacteria community (Table S6). Though the presented survey has design constraints to claim any significance, previous work supports that

barley causes an increase in Shannon diversity of free-living diazotrophs in a potato-barley rotation (Orr et al., 2011). Most crop rotation studies only look at rhizosphere bacterial communities and fail to comprehend the farm-wide implications of pulse crop rotations (Larkin and Honeycutt, 2007; Niu et al., 2018).

4.2 Impacts of irrigation on microbial community structure

It is well known that the primary limiting factor of agricultural yield in a semiarid region is precipitation. Also, it has been well documented that changes in precipitation and irrigation regime effect microbial community structure and microbial activity (Castro et al., 2019; Chen et al., 2016; Lacerda-Júnior et al., 2019, Lüneberg et al., 2018; Mavrodi et al., 2018). We have shown that irrigation significantly impacts the general microbial community and the diazotrophic community diversity. While irrigated samples tend to have higher Shannon diversity then dryland samples, they are also more homogenous. This effect is amplified in the diazotrophic samples where all irrigated samples have higher Shannon diversity, but dryland samples are highly dissimilar (Fig. 4C and D).

Shifts in phylum-level relative abundances like those seen in a longer-term study of (Hartmann et al., 2017) were also shown in our survey. An increase of Proteobacteria and Firmicutes, and a decrease of Gemmatimonadetes and Actinobacteria were observed across the state and in controlled irrigation plots (Fig. 2A and B). This is following the copiotroph-oligotroph hypothesis presented by (Fierer et al., 2007) in which copiotrophic organisms which thrive in higher moisture and higher carbon environment versus oligotrophs which are slower growing and maintain in viability under limited substrates (Ho et al., 2017). Conversely, diazotrophs share some ecological

trends as the general population, such as the increase of *Geobacter* in high moisture and high carbon soils like Kalispell (Fig. 2). *Geobacter,* an anaerobic respiratory delta-proteobacteria, uses humic acid as an electron acceptor, allowing it to fix nitrogen in a copiotrophic environment (Lovley et al., 2011). Diazotrophs require a large amount of energy to maintain homeostasis while surviving on atmospheric nitrogen. Nitrogenase regulation across diazotrophs is stimulated by a large C:N ratio. Though diazotrophs need a high carbon level to maintain nitrogen fixation, they are still present in oligotrophic soils. While this topic is understudied, it has been intensely looked at in ocean ecosystems. Oceans are notoriously known as low carbon environments, while the diversity and the abundance of marine heterotrophic diazotrophs are considered to play a pivotal role in the nitrogen cycle (Gradoville et al., 2017; Zehr and Deniz, 2015). In the soil, we can postulate that strategies of mutualism with fungus and other heterotrophs will provide the needed carbon or diazotrophs utilizing nitrogenase as an electron sink could be possible in low carbon and low nitrogen soil niches.

4.3 Irrigation influences network topology

The co-occurrence of diazotrophs shows the interactions within the nitrogen fixing community or the concurrent influence of species to abiotic factors. Here we showed that specific farm management systems impact the network topology of the diazotrophic co-occurrence networks. The irrigated network has lower modularity, low density, and poor robustness, therefore, selecting for a more definite group of diazotrophs that are more susceptible to environmental variables (Röttjers and Faust, 2018). When plots are left as dryland, the community structure is more dissimilar, shown as a highly interconnected network. The network shows high modularity with

high network robustness as compared to the irrigated soils. The density of the dryland networks is also high, with a short path length showing an integrated network. Networks with these properties have been proposed to have a quick response to environmental perturbations (Thébault and Fontaine, 2010).

Habitat filtering affects the interpretation of co-occurrence networks, and environmental changes such as irrigation need to be made clear. As has been shown, irrigation selects for specialists in the co-occurrence networks, but it is difficult with this data to determine if this is due to microbe-microbe interaction or shared habitat preference. The specialists that are more abundant in the irrigated data and are highly connected in the network (high keystone properties) might be appearing in the samples due to fast growth in specific conditions and have little interaction with other bacteria in the network. Co-occurrence networks are a great start to deduce interactions, but real-world experiments must be used to confirm the interactions.

4.4 Irrigation effects diazotrophic community robustness

In this study, we have demonstrated that the diazotrophic community is sensitive to irrigation and irrigated soils are presumably is less resilient to changes of other edaphic, geographic, or farm management practices. It is difficult to determine the effects of the overall nitrogen fixing ability of the irrigated soil as compared to dryland soil due to the type of data collected. However, we can see that a shift in phylogenetic taxa compositions between the two soils leads to irrigation selecting for more specialized OTUs, such as *Geobacter*. The specialization of the irrigated soils can be understood as less robust and less responsive to environmental change at the taxa or phylogenetic level. The dryland soil communities are more dissimilar even though they

are less diverse than the irrigated soil communities. The dissimilarity offers the dryland an advantage with the “portfolio effect” as suggested by (Allison and Martiny, 2008), in which increasing the ability to fix nitrogen in one condition averages the overall decrease of nitrogen fixation. The portfolio effect and functional redundancy supported by dormancy of the diazotrophs allow an environment like dryland to adapt better to changes in growth conditions. The homogenous environment of agricultural soils could lead to more homogenous microbial communities that are prone to the inability to overcome disturbance such as the introduction of exogenous N or planting of leguminous plants such as peas (Zeng et al., 2016). Our hypothesis of irrigation sub selecting more specific diazotrophs that are less robust will need to be further tested. A time-series experiment of farm management practices in pulse rotation system where plots are alternated from irrigated and dryland should be conducted to understand the selection over time for specific diazotrophs.

5. Conclusion

The large-scale survey of diazotrophs present in this work lays the groundwork for future understanding and development of the nitrogen-fixing microbiome. We have demonstrated that geographic and edaphic characteristics influence the community structure and diversity of free-living diazotrophs in pulse crop rotations. Also, we have identified that irrigation has a significant role in shaping the diazotrophic communities. We see that the irrigated soil contains a more sparse co-occurrence network and select for more specialized species that in the class alpha-proteobacteria. While the dryland has a more interconnected and robustness co-occurrence network and has Bacilli as their keystone species. This shows that farm management factors can shape the free-

living diazotrophic community in the topsoil layer and opens the way for more intense study of the influences of this agriculturally and environmentally crucial microbial niche.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability

Sequencing data available at NCBI SRA BioProject PRJNA580458

R scripts and data available at https://github.com/alexander-alleman/MontanaPulseCrops

Acknowledgments

This study was supported by the funding from the Montana Research and Economic Development Initiative and Montana Agricultural Experiment Station. We would like to thank Assistant Professor Jed Eberly of Montana State Central Agricultural Research Center for editorial feedback.

References

Allison, S.D., Martiny, J.B.H., 2008. Resistance, resilience, and redundancy in microbial communities. Proceedings of the National Academy of Sciences 105, 11512–11519.

Banerjee, S., Schlaeppi, K., van der Heijden, M.G.A., 2018. Keystone taxa as drivers of microbiome structure and functioning. Nature Reviews Microbiology 16, 567–576.

Berg, M., Koskella, B., 2018. Nutrient- and dose-Dependent Microbiome-Mediated Protection against a Plant Pathogen. Current Biology 28, 2487-2492.

Berry, D., Widder, S., 2014. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. Frontiers in Microbiology 5, 1–14.

Castro, S.P., Cleland, E.E., Robert, W., Sawad, A.R., Lipson, D.A., 2019. Soil microbial responses to drought and exotic plants shift carbon metabolism. ISME Journal 13, 1776–1787.

Chaudhry, V., Rehman, A., Mishra, A., Chauhan, P.S., Nautiyal, C.S., 2012. Changes in Bacterial Community Structure of Agricultural Land Due to Long-Term Organic and Chemical Amendments. Microbial Ecology 64, 450–460.

Chen, C., Neill, K., Burgess, M., Bekkerman, A., 2012. Agronomic benefit and economic potential of introducing fall-seeded pea and lentil into conventional wheat-based crop rotations. Agronomy Journal 104, 215–224.

Chen, J., Wang, P., Wang, C., Wang, X., Miao, L., Liu, S., Yuan, Q., 2019. Dam construction alters function and community composition of diazotrophs in riparian soils across an environmental gradient. Soil Biology and Biochemistry 132, 14–23.

Chen, L.J., Feng, Q., Li, C.S., Song, Y.X., Liu, W., Si, J.H., Zhang, B.G., 2016. Spatial Variations of Soil Microbial Activities in Saline Groundwater-Irrigated Soil Ecosystem. Environmental Management 57, 1054–1061.

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A., Kuske, C.R., Tiedje, J.M., 2014. Ribosomal Database Project: Data and tools for high throughput rRNA analysis. Nucleic Acids Research 42, 633–642.

Csárdi, G., Nepusz, T., 2006. The igraph software package for complex network research. InterJournal 1695–1695.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., Andersen, G.L., 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Applied and environmental microbiology 72, 5069–72.

Fan, K., Weisenhorn, P., Gilbert, J.A., Shi, Y., Bai, Y., Chu, H., 2018. Soil pH correlates with the co-occurrence and assemblage process of diazotrophic communities in rhizosphere and bulk soils of wheat fields. Soil Biology and Biochemistry 121, 185–192.

Fierer, N., Bradford, M. A., Jackson, R.B., 2007. Toward an Ecological Classification of Soil Bacteria. Ecology 88, 1354–1364.

Finkel, O.M., Castrillo, G., Herrera Paredes, S., Salas González, I., Dangl, J.L., 2017. Understanding and exploiting plant beneficial microbes. Current Opinion in Plant Biology 38, 155–163.

Gan, Y., Hamel, C., O'Donovan, J.T., Cutforth, H., Zentner, R.P., Campbell, C.A., Niu, Y., Poppy, L., 2015. Diversifying crop rotations with pulses enhances system productivity. Scientific Reports 5, 1–14.

Gradoville, M.R., Bombar, D., Crump, B.C., Letelier, R.M., Zehr, J.P., White, A.E., 2017. Diversity and activity of nitrogen-fixing communities across ocean basins. Limnology and Oceanography 62, 1895–1909.

Gupta, V.V.S.R., Kroker, S.J., Hicks, M., Davoren, C.W., Descheemaeker, K., Llewellyn, R., 2014. Nitrogen cycling in summer active perennial grass systems in South Australia: non-symbiotic nitrogen fixation. Crop and Pasture Science 65, 1044–1044.

Hartmann, M., Brunner, I., Hagedorn, F., Bardgett, R.D., Stierli, B., Herzog, C., Chen, X., Zingg, A., Graf-Pannatier, E., Rigling, A., Frey, B., 2017. A decade of irrigation transforms the soil microbiome of a semi-arid pine forest. Molecular Ecology 26, 1190–1206.

Hartmann, M., Frey, B., Mayer, J., Mäder, P., Widmer, F., 2015. Distinct soil microbial diversity under long-term organic and conventional farming. ISME Journal 9, 1177–1194.

Ho, A., Di Lonardo, D.P., Bodelier, P.L.E., 2017. Revisiting life strategy concepts in environmental microbial ecology. FEMS microbiology ecology 93, 1–14.

Hsu, S.F., Buckley, D.H., 2009. Evidence for the functional significance of diazotroph community structure in soil. ISME Journal 3, 124–136.

Kurtz, Z.D., Müller, C.L., Miraldi, E.R., Littman, D.R., Blaser, M.J., Bonneau, R.A., 2015. Sparse and Compositionally Robust Inference of Microbial Ecological Networks. PLoS Computational Biology 11, 1–25.

Kuzyakov, Y., Blagodatskaya, E., 2015. Microbial hotspots and hot moments in soil: Concept & review. Soil Biology and Biochemistry 83, 184–199.

Lacerda-Júnior, G.V., Noronha, M.F., Cabral, L., Delforno, T.P., de Sousa, S.T.P., Fernandes-Júnior, P.I., Melo, I.S., Oliveira, V.M., 2019. Land Use and Seasonal Effects on the Soil Microbiome of a Brazilian Dry Forest. Frontiers in Microbiology 10, 1–14.

Larkin, R.P., Honeycutt, C.W., 2007. Effects of Different 3-Year Cropping Systems on Soil Microbial Communities and Rhizoctonia Diseases of Potato. Phytopathology 96, 68–79.

Lee, C., 2011. Montana's Pulse Industry: How it has Developed, Economic Impact & Potential for Growth.

Lovley, D.R., Ueki, T., Zhang, T., Malvankar, N.S., Shrestha, P.M., Flanagan, K.A., Aklujkar, M., Butler, J.E., Giloteaux, L., Rotaru, A.E., Holmes, D.E., Franks, A.E., Orellana, R., Risso, C., Nevin, K.P., 2011.

Geobacter. The Microbe Electric's Physiology, Ecology, and Practical Applications, 1st ed, Advances in Microbial Physiology. Elsevier Ltd.

Lüneberg, K., Schneider, D., Siebe, C., Daniel, R., 2018. Drylands soil bacterial community is affected by land use change and different irrigation practices in the Mezquital Valley, Mexico. Scientific Reports 8, 1–15.

Lupatini, M., Korthals, G.W., de Hollander, M., Janssens, T.K.S., Kuramae, E.E., 2017. Soil microbiome is more heterogeneous in organic than in conventional farming system. Frontiers in Microbiology 7, 1–13.

Mavrodi, D.V., Mavrodi, O.V., Elbourne, L.D.H., Tetu, S., Bonsall, R.F., Parejko, J., Yang, M., Paulsen, I.T., Weller, D.M., Thomashow, L.S., 2018. Long-Term Irrigation Affects the Dynamics and Activity of the Wheat Rhizosphere Microbiome. Frontiers in Plant Science 9.

McMurdie, P.J., Holmes, S., 2013. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE 8(4):e61217

Nannipieri, J., Ascher, J., Ceccherini M, T., Landi, L., Pietramellara, G., Renella, G., 1999. Microbial diversity and soil functions. European Journl of Soil Science 655–670.

Niu, Y., Bainard, L.D., May, W.E., Hossain, Z., Hamel, C., Gan, Y., 2018. Intensified pulse rotations buildup pea rhizosphere pathogens in cereal and pulse based cropping systems. Frontiers in Microbiology 9, 1–14.

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. ISME Journal 11, 315–326.

Oksanen, J., Blanchet, F.G., Michael Friendly, R.K., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2018. vegan: Community Ecology Package. 10, 631-637.

Orr, C.H., James, A., Leifert, C., Cooper, J.M., Cummings, S.P., 2011. Diversity and Activity of Free-Living Nitrogen-Fixing Bacteria and Total Bacteria in Organic and Conventionally Managed Soils. Applied and Environmental Microbiology 77, 911–919.

Orr, C.H., Leifert, C., Cummings, S.P., Cooper, J.M., 2012. Impacts of Organic and Conventional Crop Management on Diversity and Activity of Free-Living Nitrogen Fixing Bacteria and Total Bacteria Are Subsidiary to Temporal Effects. PLoS ONE 7.

Penton, C.R., Yang, C., Wu, L., Wang, Q., Zhang, J., Liu, F., Qin, Y., Deng, Y., Hemme, C.L., Zheng, T., Schuur, E.A.G., Tiedje, J., Zhou, J., 2016. NifH-harboring bacterial community composition across an alaskan permafrost thaw gradient. Frontiers in Microbiology 7, 1–11.

Peoples, M.B., Brockwell, J., Herridge, D.F., Rochester, I.J., Alves, B.J.R., Urquiaga, S., Boddey, R.M., Dakora, F.D., Bhattarai, S., Maskey, S.L., Sampet, C., Rerkasem, B., Khans, D.F., Hauggaard-Nielsen, H., Jensen, E.S., 2009. The contributions of nitrogen-fixing crop legumes to the productivity of agricultural systems. Symbiosis 48, 1–17.

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.

Röttjers, L., Faust, K., 2018. From hairballs to hypotheses–biological insights from microbial networks. FEMS Microbiology Reviews 42, 761–780.

Shi, S., Nuccio, E.E., Shi, Z.J., He, Z., Zhou, J., Firestone, M.K., 2016. The interconnected rhizosphere: High network complexity dominates rhizosphere assemblages. Ecology Letters 19, 926–936.

St. Luce, M., Grant, C.A., Zebarth, B.J., Ziadi, N., O'Donovan, J.T., Blackshaw, R.E., Harker, K.N., Johnson, E.N., Gan, Y., Lafond, G.P., May, W.E., Khakbazan, M., Smith, E.G., 2015. Legumes can reduce economic optimum nitrogen rates and increase yields in a wheat-canola cropping sequence in western canada. Field Crops Research 179, 12–25.

Stagnari, F., Maggio, A., Galieni, A., Pisante, M., 2017. Multiple benefits of legumes for agriculture sustainability: an overview. Chemical and Biological Technologies in Agriculture 4, 1–13.

Sul, W.J., Asuming-Brempong, S., Wang, Q., Tourlousse, D.M., Penton, C.R., Deng, Y., Rodrigues, J.L.M., Adiku, S.G.K., Jones, J.W., Zhou, J., Cole, J.R., Tiedje, J.M., 2013. Tropical agricultural land management influences on soil microbial communities through its effect on soil organic carbon. Soil Biology and Biochemistry 65, 33–38.

Tan, Z., Hurek, T., Reinhold-Hurek, B., 2003. Effect of N-fertilization, plant genotype and environmental conditions on *nifH* gene pools in roots of rice. Environmental Microbiology 5, 1009–1015.

Thébault, E., Fontaine, C., 2010. Stability of ecological communities and the architecture of mutualistic and trophic networks. Science 329, 1–25.

Van Wees, S.C., Van der Ent, S., Pieterse, C.M., 2008. Plant immune responses triggered by beneficial microbes. Current Opinion in Plant Biology 11, 443–448.

Zehr, J.P., Deniz, B., 2015. Marine Nitrogen Fixation: Organisms, Significance, Enigmas, and Future Directions. Biological Nitrogen Fixation 855–872.

Zeng, J., Liu, X., Song, L., Lin, X., Zhang, H., Shen, C., Chu, H., 2016. Nitrogen fertilization directly affects soil bacterial diversity and indirectly affects bacterial community composition. Soil Biology and Biochemistry 92, 41–49.

Figure Legends

Figure 1) Soil sampling locations across the state of Montana. Each site is a Montana State University Agricultural Research Center except the Richland site on a farmer's field.

Figure 2) The relative abundance of OTUS at each site across the state of Montana. The phylum-level relative abundance of the 16S sequences (a). The Diazotrophic sequences (*nifH*) at the phylum level (b) and the genus level (c).

Figure 3) Diversity of general and diazotrophic communities. The Shannon diversity index across the sites in Montana for the general bacterial community (a). The Shannon diversity index for the diazotrophic community (b). The NMDS orientation plot based on the Bray-Curtis dissimilarity for the general bacterial community comparing irrigated (triangle) and dryland (circle) and sites (color) (c). The NMDS orientation based on the Bray-Curtis dissimilarity of the diazotrophic community (d).

Figure 4) The Shannon diversity index is comparing irrigated and dryland sites for the general bacteria community (a) and the diazotrophic community (b). The DCA ordination plot based on the Bray-Curtis dissimilarity of the general (16S rRNA) (c) and diazotrophic (*nifH*) community (d). With irrigation and sites labeled as colors.

Figure 5) Chemical CAP Plots. CAP plots were constrained by an a priori selected capscale model. Constraining all the sites with capscale model in general (a) and diazotrophic (b) bacterial communities still shows separation between irrigated and dryland sites. Specific sites correlate best with chemical vectors in the plot as Kalispell, and organic matter, and Moccasin and Barium are two examples. The irrigation sites were then looked at by themselves with the same CAP/capscale model selection procedure for the general (c) and diazotrophic (d) bacterial communities. Av (Available)

Figure 6) Diazotrophic co-occurrence networks for the dryland and irrigated plots. The modularity of the irrigated and dryland networks was determined by the cluster walktrap algorithm and presented by colors. Edges that are interacting between modules are colored in red. Co-occurrence networks were built on species that occur in at least 37% of the samples and modularity was assigned to cluster with at least ten members.

Figure 7) Composition of diazotrophic OTUs (nodes) in dryland and irrigated plots distributed along degree against closeness centrality or betweenness. Nodes with a large degree, large closeness, and small betweenness can be identified as keystone species. The Diazotrophic class is colored on the plots with the Bacilli: *Paenibacillus graminis* having all characteristics of a keystone species in the dryland plots while the alphaproteobacteria: *Rhizobium leguminosarum* and *Rhodopseudomonas palustris* have the characteristics of a keystone species in the irrigated plot.

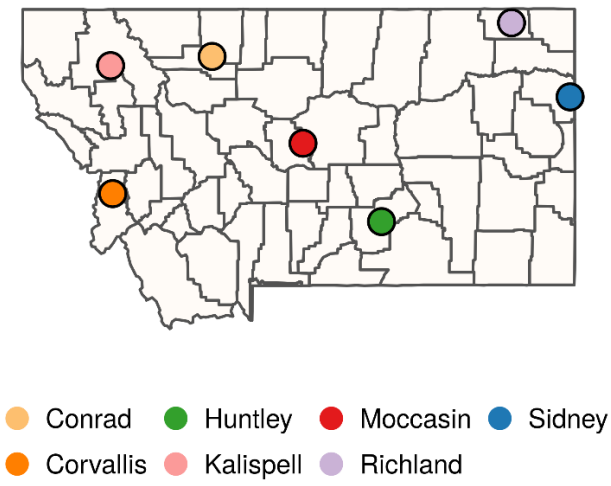
Conrad
Huntley
Moccasin
Sidney
Corvallis
Kalispell
Richland

Figure 1)

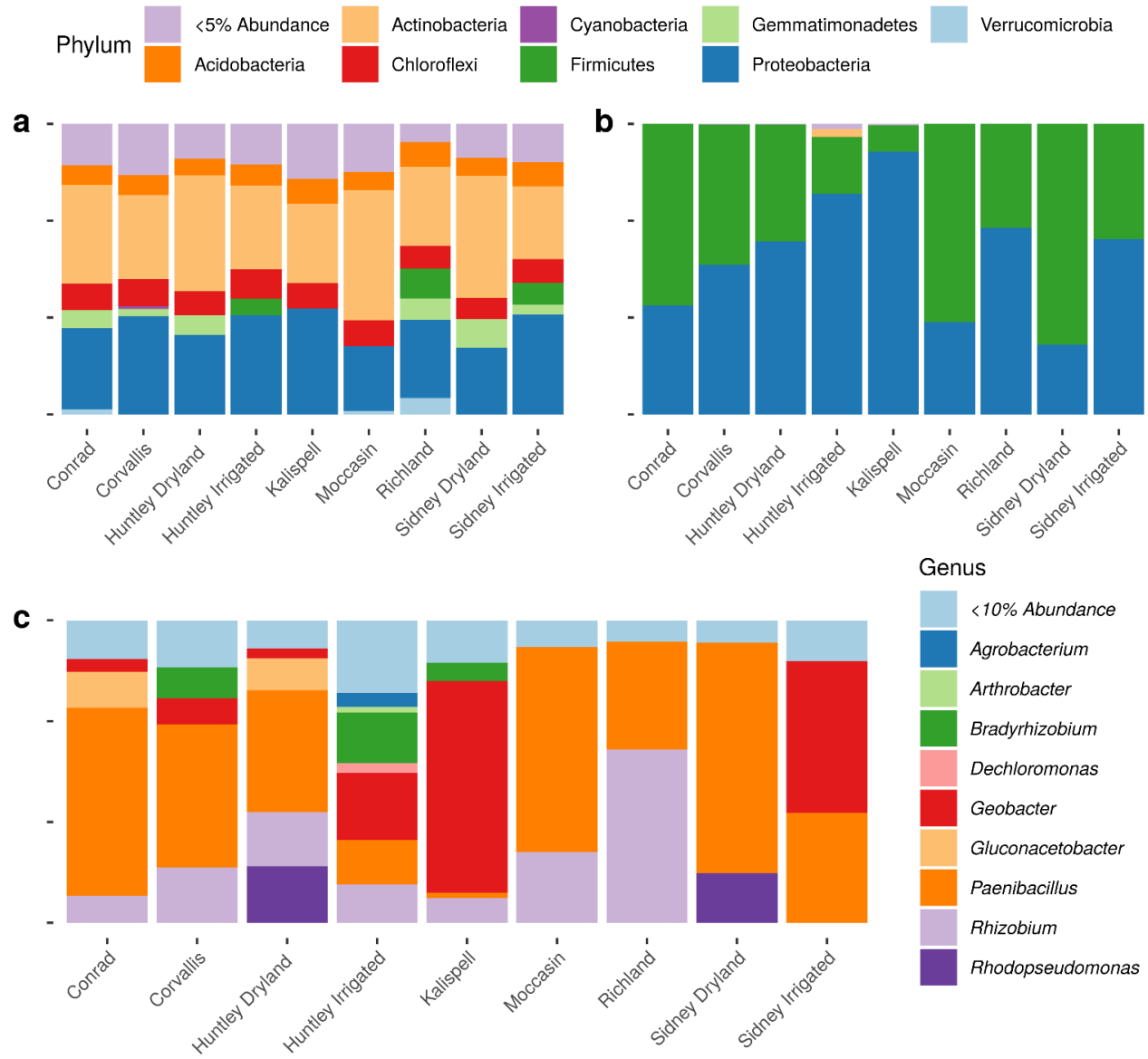
Phylum
<5% Abundance
Acidobacteria
Actinobacteria
Chloroflexi
Cyanobacteria
Firmicutes
Gemmatimonadetes
Proteobacteria
Verrucomicrobia
a
b
c
Conrad
Corvallis
Huntley Dryland
Huntley Irrigated
Kalispell
Moccasin
Richland
Sidney Dryland
Sidney Irrigated
Genus
<10% Abundance
Agrobacterium
Arthrobacter
Bradyrhizobium
Dechloromonas
Geobacter
Gluconacetobacter
Paenibacillus
Rhizobium
Rhodopseudomonas

Figure 2)

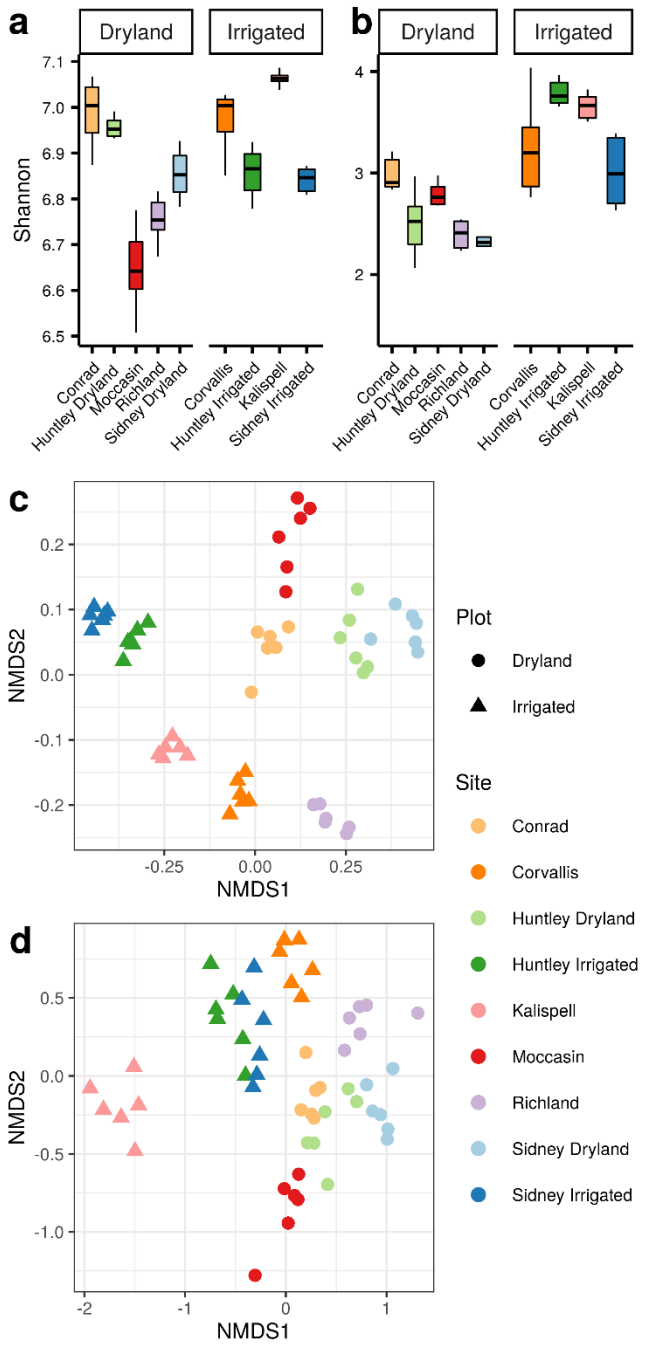
a
Dryland
Irrigated
Shannon
7.1
7.0
6.9
6.8
6.7
6.6
6.5
Conrad
Huntley Dryland
Moccasin
Richland
Sidney Dryland
Corvallis
Huntley Irrigated
Kalispell
Sidney Irrigated
b
Dryland
Irrigated
4
3
2
Conrad
Huntley Dryland
Moccasin
Richland
Sidney Dryland
Corvallis
Huntley Irrigated
Kalispell
Sidney Irrigated
c
NMDS2
0.2
0.1
0.0
-0.1
-0.2
-0.25
0.00
0.25
NMDS1
Plot
Dryland
Irrigated
Site
Conrad
Corvallis
Huntley Dryland
Huntley Irrigated
Kalispell
Moccasin
Richland
Sidney Dryland
Sidney Irrigated
d
NMDS2
0.5
0.0
-0.5
-1.0
-2
-1
0
1
NMDS1

Figure 3)

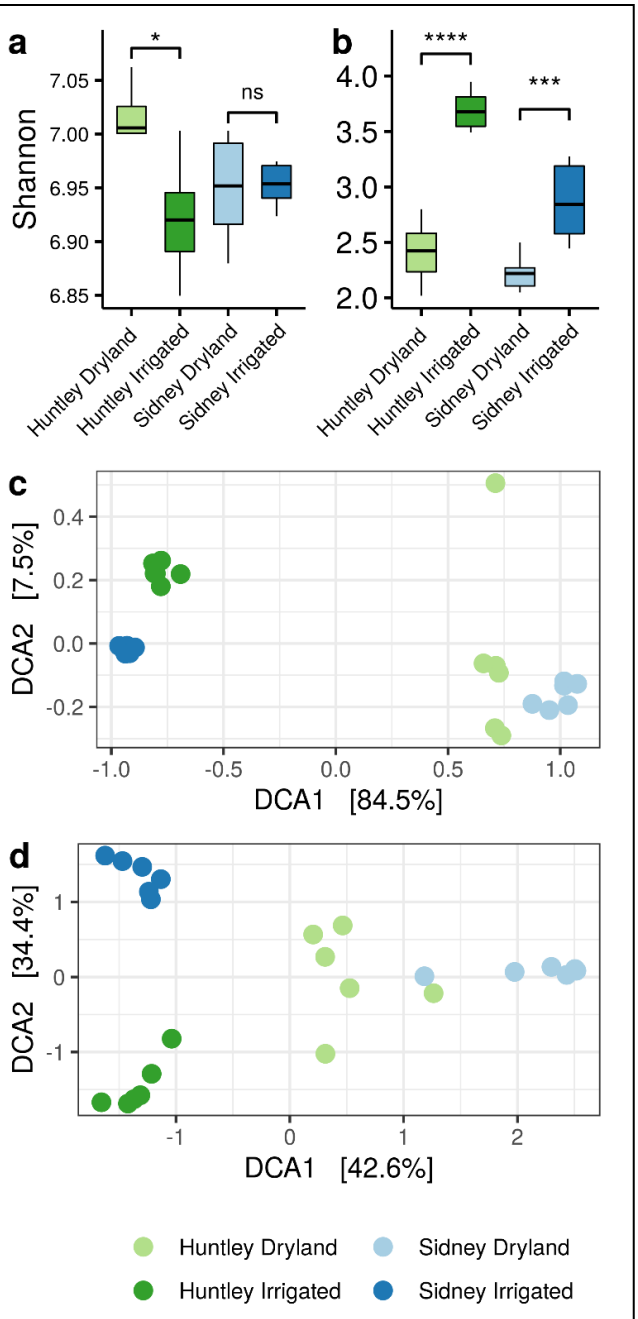
a
b
c
d
Shannon
*
ns
****
***
7.05
7.00
6.95
6.90
6.85
4.0
3.5
3.0
2.5
2.0
Huntley Dryland
Huntley Irrigated
Sidney Dryland
Sidney Irrigated
DCA2 [7.5%]
DCA1 [84.5%]
DCA2 [34.4%]
DCA1 [42.6%]
0.4
0.2
0.0
-0.2
-1.0
-0.5
0.5
1.0
1
0
-1
2

Figure 4)

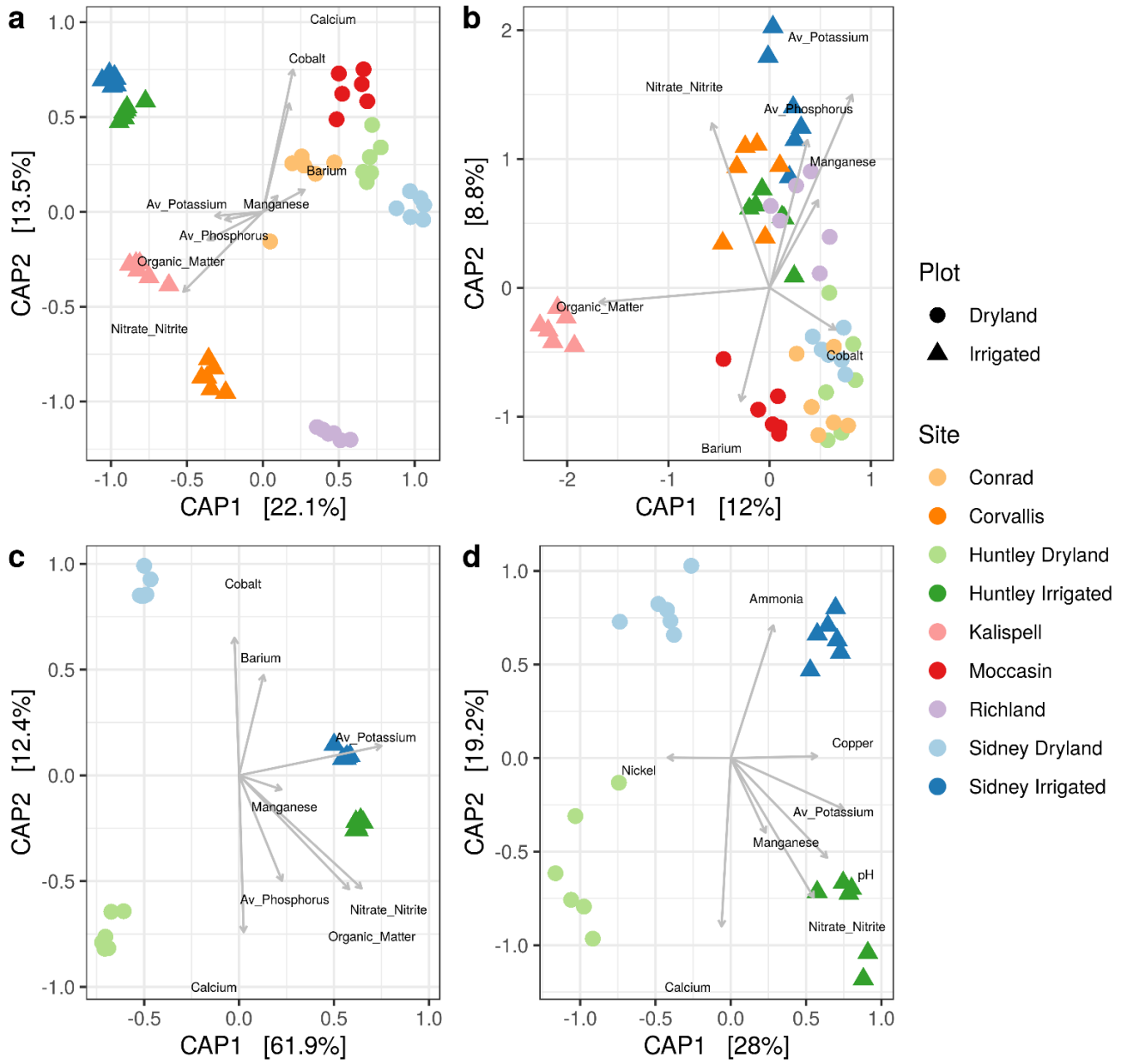
a
Calcium
Cobalt
Barium
Av_Potassium
Manganese
Av_Phosphorus
Organic_Matter
Nitrate_Nitrite
CAP2 [13.5%]
CAP1 [22.1%]
b
Av_Potassium
Nitrate_Nitrite
Av_Phosphorus
Manganese
Organic_Matter
Cobalt
Barium
CAP2 [8.8%]
CAP1 [12%]
c
Cobalt
Barium
Av_Potassium
Manganese
Av_Phosphorus
Nitrate_Nitrite
Organic_Matter
Calcium
CAP2 [12.4%]
CAP1 [61.9%]
d
Ammonia
Copper
Nickel
Av_Potassium
Manganese
pH
Nitrate_Nitrite
Calcium
CAP2 [19.2%]
CAP1 [28%]
Plot
Dryland
Irrigated
Site
Conrad
Corvallis
Huntley Dryland
Huntley Irrigated
Kalispell
Moccasin
Richland
Sidney Dryland
Sidney Irrigated

Figure 5)

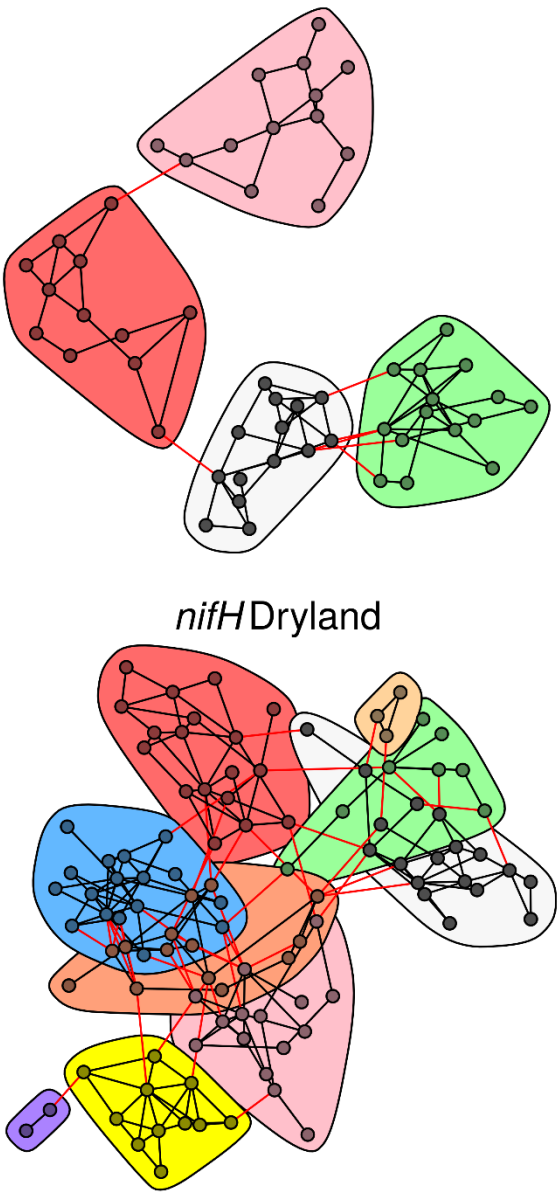

Figure 6)

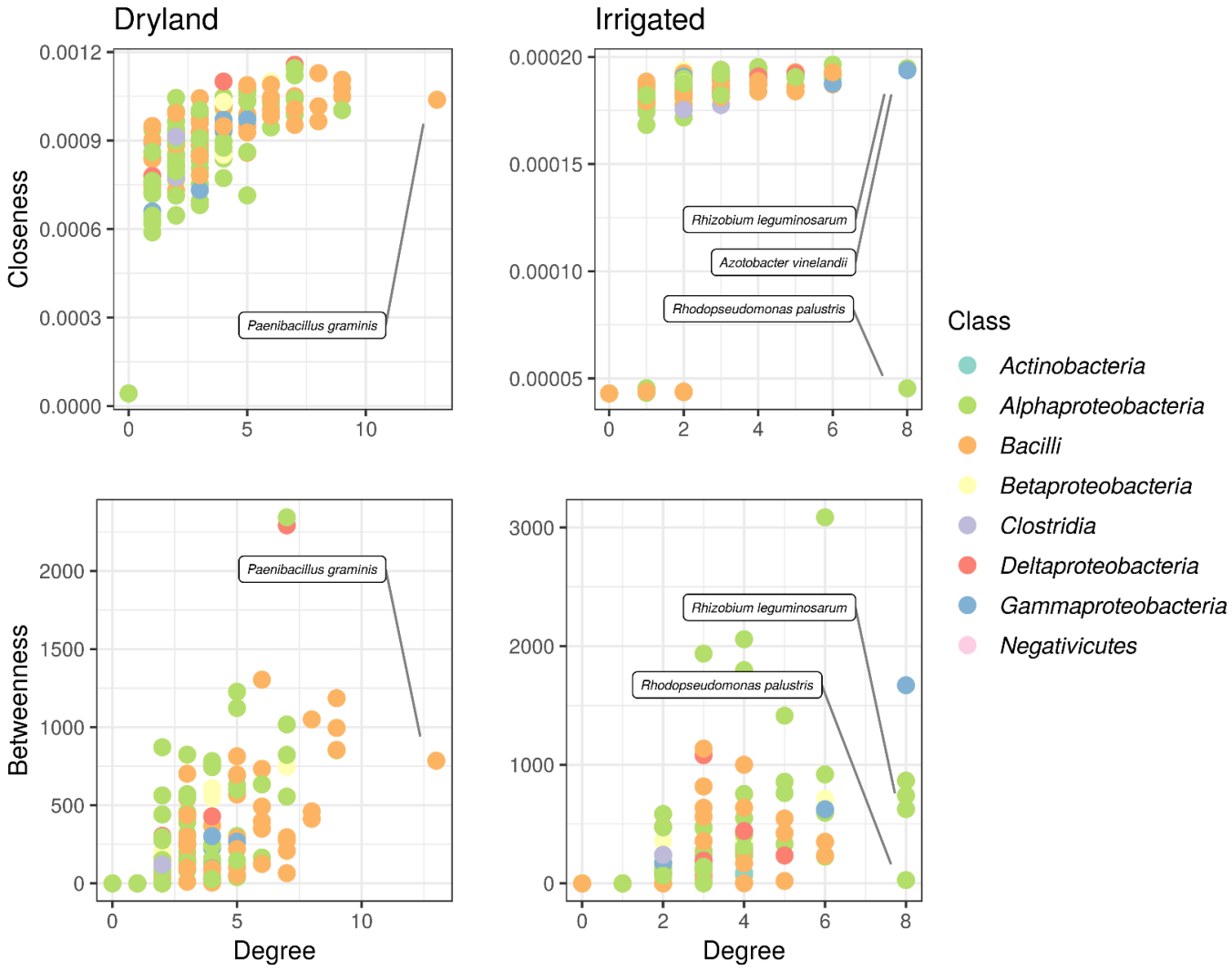
Dryland
Irrigated
Closeness
Betweenness
Degree
Degree
Paenibacillus graminis
Rhizobium leguminosarum
Azotobacter vinelandii
Rhodopseudomonas palustris
Class
Actinobacteria
Alphaproteobacteria
Bacilli
Betaproteobacteria
Clostridia
Deltaproteobacteria
Gammaproteobacteria
Negativicutes

Figure 7)

Supplemental Material for the Manuscript:

"Drivers of diazotroph community structure and co-occurrence in a Northern Great Plains pulse crop rotation system"

Alexander B Alleman[1,2], Yesuf Assen Mohammed[3], Kent A McVay[3], Qasim A Khan[3], Patrick Carr[3], John Miller[3], Zachariah Miller[3], Jessica Torrion[3], Peggy Lamb[3], Florence Mus[1,2], Chengci Chen[3], and John W Peters[1,2]

[1]Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT, USA; [2]Institute of Biological Chemistry, Washington State University, Pullman, WA, USA and [3]Department of Research Centers, Montana State University, Bozeman, MT, USA

Sample ID	Site	ARC	Pea Variety	Plot	Precipitation (in)	Irrigation (in)	Tillage	Previous Crop	Grain Yield (lb/ac)	Elevation (meters)	Latitude	Longitude
JZ030	Kalispell	NWARC	AC Earlystar	Irrigated	9.33	0	Conventional	B	5297	884	48.18	-114.13
JZ031	Kalispell	NWARC	CDC Saffron	Irrigated	9.33	0	Conventional	B	5520	884	48.18	-114.13
JZ032	Kalispell	NWARC	Delta	Irrigated	9.33	0	Conventional	B	5143	884	48.18	-114.13
JZ033	Kalispell	NWARC	DS Admiral	Irrigated	9.33	0	Conventional	B	5699	884	48.18	-114.13
JZ034	Kalispell	NWARC	Majoret	Irrigated	9.33	0	Conventional	B	5024	884	48.18	-114.13
JZ035	Kalispell	NWARC	Navarro	Irrigated	9.33	0	Conventional	B	4364	884	48.18	-114.13
JZ036	Huntley Dryland	SARC	AC Earlystar	Dryland	8.79	0	No_till	CF	877	830	45.922	-108.244
JZ037	Huntley Irrigated	SARC	AC Earlystar	Irrigated	8.79	2.5	No_till	B	1704	830	45.922	-108.244
JZ038	Huntley Dryland	SARC	CDC Saffron	Dryland	8.79	0	No_till	CF	847	830	45.922	-108.244
JZ039	Huntley Irrigated	SARC	CDC Saffron	Irrigated	8.79	2.5	No_till	B	1798	830	45.922	-108.244
JZ040	Huntley Dryland	SARC	Delta	Dryland	8.79	0	No_till	CF	829	830	45.922	-108.244
JZ041	Huntley Irrigated	SARC	Delta	Irrigated	8.79	2.5	No_till	B	1535	830	45.922	-108.244
JZ042	Huntley Dryland	SARC	DS Admiral	Dryland	8.79	0	No_till	CF	716	830	45.922	-108.244
JZ043	Huntley Irrigated	SARC	DS Admiral	Irrigated	8.79	2.5	No_till	B	1911	830	45.922	-108.244
JZ044	Huntley Dryland	SARC	Majoret	Dryland	8.79	0	No_till	CF	693	830	45.922	-108.244
JZ045	Huntley Irrigated	SARC	Majoret	Irrigated	8.79	2.5	No_till	B	1300	830	45.922	-108.244
JZ046	Huntley Dryland	SARC	Navarro	Dryland	8.79	0	No_till	CF	467	830	45.922	-108.244
JZ047	Huntley Irrigated	SARC	Navarro	Irrigated	8.79	2.5	No_till	B	1142	830	45.922	-108.244
JZ078	Moccasin	CARC	AC Earlystar	Dryland	10.73	0	No_till	WW	1285	1293	47.056	-109.95
JZ079	Moccasin	CARC	CDC Saffron	Dryland	10.73	0	No_till	WW	1569	1293	47.056	-109.95
JZ080	Moccasin	CARC	Delta	Dryland	10.73	0	No_till	WW	1405	1293	47.056	-109.95
JZ081	Moccasin	CARC	DS Admiral	Dryland	10.73	0	No_till	WW	1428	1293	47.056	-109.95
JZ082	Moccasin	CARC	Majoret	Dryland	10.73	0	No_till	WW	1265	1293	47.056	-109.95
JZ083	Moccasin	CARC	Navarro	Dryland	10.73	0	No_till	WW	1279	1293	47.056	-109.95
JZ084	Sidney Dryland	EARC	AC Earlystar	Dryland	15.54	0	Conventional	CF	3953	670	47.728	-104.148
JZ085	Sidney Dryland	EARC	CDC Saffron	Dryland	15.54	0	Conventional	CF	4172	670	47.728	-104.148
JZ086	Sidney Dryland	EARC	Delta	Dryland	15.54	0	Conventional	CF	3628	670	47.728	-104.148
JZ087	Sidney Dryland	EARC	DS Admiral	Dryland	15.54	0	Conventional	CF	3591	670	47.728	-104.148
JZ088	Sidney Dryland	EARC	Majoret	Dryland	15.54	0	Conventional	CF	3819	670	47.728	-104.148
JZ089	Sidney Dryland	EARC	Navarro	Dryland	15.54	0	Conventional	CF	3765	670	47.728	-104.148
JZ090	Sidney Irrigated	EARC	AC Earlystar	Irrigated	15.54	5.83	Conventional	SW	4883	670	47.728	-104.148
JZ091	Sidney Irrigated	EARC	CDC Saffron	Irrigated	15.54	5.83	Conventional	SW	4887	670	47.728	-104.148
JZ092	Sidney Irrigated	EARC	Delta	Irrigated	15.54	5.83	Conventional	SW	4352	670	47.728	-104.148
JZ093	Sidney Irrigated	EARC	DS Admiral	Irrigated	15.54	5.83	Conventional	SW	4643	670	47.728	-104.148
JZ094	Sidney Irrigated	EARC	Majoret	Irrigated	15.54	5.83	Conventional	SW	4406	670	47.728	-104.148
JZ095	Sidney Irrigated	EARC	Navarro	Irrigated	15.54	5.83	Conventional	SW	3825	670	47.728	-104.148
JZ096	Richland	EARC	AC Earlystar	Dryland	15.54	0	No_till	CF	5228	899	48.8	-105.42
JZ097	Richland	EARC	CDC Saffron	Dryland	15.54	0	No_till	CF	6043	899	48.8	-105.42
JZ098	Richland	EARC	Delta	Dryland	15.54	0	No_till	CF	5459	899	48.8	-105.42
JZ099	Richland	EARC	DS Admiral	Dryland	15.54	0	No_till	CF	5166	899	48.8	-105.42
JZ100	Richland	EARC	Majoret	Dryland	15.54	0	No_till	CF	4897	899	48.8	-105.42
JZ101	Richland	EARC	Navarro	Dryland	15.54	0	No_till	CF	5769	899	48.8	-105.42
JZ102	Conrad	WTARC	AC Earlystar	Dryland	8.52	0	No_till	CF	4852	1130	48.313	-111.925
JZ103	Conrad	WTARC	DS Admiral	Dryland	8.52	0	No_till	CF	3239	1130	48.313	-111.925
JZ104	Conrad	WTARC	Delta	Dryland	8.52	0	No_till	CF	3933	1130	48.313	-111.925

JZ105	Conrad	WTARC	Majoret	Dryland	8.52	0	No_till	CF	2367	1130	48.313	-111.925
JZ106	Conrad	WTARC	Navarro	Dryland	8.52	0	No_till	CF	4283	1130	48.313	-111.925
JZ107	Conrad	WTARC	CDC Saffron	Dryland	8.52	0	No_till	CF	4367	1130	48.313	-111.925
JZ108	Corvallis	WARC	AC Earlystar	Irrigated	2.57	7	Culti-roller	B	3148	1097	46.327	-114.084
JZ109	Corvallis	WARC	CDC Saffron	Irrigated	2.57	7	Culti-roller	B	3245	1097	46.327	-114.084
JZ110	Corvallis	WARC	Delta	Irrigated	2.57	7	Culti-roller	B	2519	1097	46.327	-114.084
JZ111	Corvallis	WARC	DS Admiral	Irrigated	2.57	7	Culti-roller	B	3005	1097	46.327	-114.084
JZ112	Corvallis	WARC	Majoret	Irrigated	2.57	7	Culti-roller	B	1710	1097	46.327	-114.084

Table S1) Farm Management and Geographic Data. Previous Crop, Barley (B), Chemical Fallow (CF), Winter Wheat (WW), Spring Wheat (SW). Kalispell used sub-surface irrigation methods. Tillage is defined by No Till meaning no tillage has occurred, or Conventional tillage which refers to a minimum tillage and Culti-roller tillage which provides more disturbance using a rotary tiller.

Sample ID	*Organic Matter (%)*	*Moisture Content (%)*	*Nitrate + Nitrite (µg/g)*	*Ammonia (µg/g)*	*Available P (µg/g)*	*Available K (µg/g)*	*Sulfate + Sulfur (µg/g)*	*pH*	*Boron (µg/g)*
JZ030	6.2	14	18	1.9	12	86	2.9	7.5	0.43
JZ031	5.8	15	24	1	12	78	2.2	7.5	0.32
JZ032	5.6	13	27	1.6	11	78	2.7	7.5	0.36
JZ033	5.8	15	16	1.4	8.2	63	2	7.3	0.31
JZ034	5.8	16	20	1.3	9.9	86	2.3	7.4	0.26
JZ035	5.7	17	16	1.5	11	70	1.9	7.4	0.27
JZ036	1.3	8.9	5.3	2.7	12	270	3	7	0.42
JZ037	1.6	16	13	4.1	6.5	210	5.3	7.5	0.57
JZ038	1.3	11	3.3	2.9	7.1	220	3.1	7.3	0.56
JZ039	1.6	15	11	4	5.7	220	5.1	7.7	0.51
JZ040	1.4	11	4.8	4.3	9.9	220	3	7.3	0.38
JZ041	1.6	16	9.6	5	5.1	240	4.1	7.4	0.61
JZ042	1.2	8.4	4.6	3.2	9.8	200	2.3	7.4	0.4
JZ043	1.7	15	12	4.1	6.8	220	4.7	7.5	0.51
JZ044	1.2	13	3.9	2.3	9.7	210	2.5	7.3	0.45
JZ045	1.6	15	15	4.1	6.8	230	5.6	7.5	0.52
JZ046	1.3	9.4	5.2	2.8	9.5	220	3.3	7.2	0.41
JZ047	1.7	16	10	4.9	6.3	230	4.4	7.4	0.64
JZ078	3.6	15	9.1	1.9	6.1	120	2	7.4	0.34
JZ079	3.4	16	7.9	2.2	5.5	130	1.8	7.5	0.48
JZ080	3.7	16	12	3.4	6.9	120	2.8	7.5	0.42
JZ081	3.7	16	9.4	3.1	5.9	120	1.8	7.4	0.38
JZ082	3.7	15	9.4	2.3	6.9	130	2.1	7.5	0.35
JZ083	3.4	16	18	2.6	6.2	140	2.5	7.4	0.42
JZ084	1.9	12	16	1.8	18	130	2.7	7.5	0.34
JZ085	1.9	12	12	1.5	19	130	2.6	7.3	0.4
JZ086	1.9	13	11	2	15	120	3.2	7.3	0.42
JZ087	1.7	11	12	2.2	15	120	2.6	7.4	0.34
JZ088	1.7	13	9.9	2.7	14	140	2.7	7.4	0.37
JZ089	1.8	12	12	2.7	20	150	3.1	7.5	0.33
JZ090	2.8	23	25	1.6	26	400	32	7.8	1.2
JZ091	2.8	21	28	2.8	30	380	23	7.7	1.5
JZ092	2.9	21	24	2	28	410	22	7.8	1.2
JZ093	2.8	22	28	2	29	410	21	7.8	1.6
JZ094	2.6	21	19	1.4	26	350	24	7.9	1.2
JZ095	2.8	21	26	1.9	27	380	17	7.8	1.4
JZ096	3.5	12	18	1.5	16	380	2	6	0.64
JZ097	3.5	12	20	1.4	17	380	2.1	6.2	0.76
JZ098	3.4	12	16	1.1	19	380	1.9	6.2	0.65
JZ099	3.1	11	16	1.2	15	300	1.7	6.1	0.53
JZ100	2.2	11	10	1.6	10	200	1.7	6	0.44
JZ101	2.9	12	19	1.9	9.8	290	1.7	6.2	0.54
JZ102	1.5	12	6.6	0.72	14	230	6.7	7.1	0.36
JZ103	1.9	10	7	0.76	16	160	8.6	7.6	0.41
JZ104	2.1	13	10	1	17	160	9.9	7.6	0.4
JZ105	2	17	8.4	1	16	160	9.4	7.6	0.42
JZ106	2	12	8	1.8	17	190	7.1	7.3	0.39
JZ107	2	17	9.8	0.83	14	160	29	7.7	0.47
JZ108	1.7	11	47	1.2	24	240	1.3	6.7	0.18
JZ109	1.7	11	21	1.1	14	200	1.4	6.5	0.19
JZ110	2	11	29	1.2	15	200	1.6	6.9	0.19
JZ111	1.7	10	32	1.6	12	190	1.3	6.8	0.17
JZ112	1.9	11	70	1.9	17	200	1.5	6.7	0.17
JZ113	1.6	9.4	50	3.4	13	180	1.3	6.7	0.18

Table S2) The Extended Fertility Test. Conducted by the University of Idaho Analytical Sciences Laboratory (https://www.uidaho.edu/cals/analytical-sciences-laboratory).

1

2

Sample ID	*Ba*	*Ca*	*Cr*	*Co*	*Cu*	*Fe*	*Mg*	*Mn*	*Ni*	*P*	*K*	*S*	*V*	*Zn*
JZ030	210	9200	13	6.8	13	14000	6000	330	14	740	2200	460	16	51
JZ031	210	9200	13	5.5	14	14000	6200	320	12	730	2100	450	16	52
JZ032	290	13000	20	8	18	20000	9000	460	22	1000	2900	650	22	76
JZ033	220	8800	13	7.3	13	15000	6200	340	14	700	2200	460	16	52
JZ034	200	8300	12	6.8	13	14000	5900	300	14	690	2000	420	15	51
JZ035	210	8300	13	5.1	13	15000	6200	330	13	740	2200	460	16	52
JZ036	270	12000	50	16	25	35000	10000	690	39	1000	5800	340	72	120
JZ037	180	11000	28	9.7	18	22000	8200	400	26	860	3000	360	41	83
JZ038	180	13000	33	9.6	16	22000	7800	380	27	690	3300	300	47	75
JZ039	170	11000	28	9.4	17	23000	8200	390	26	840	3100	360	42	84
JZ040	160	11000	33	10	16	21000	7000	390	27	720	3200	280	44	74
JZ041	180	11000	30	9.9	18	23000	8500	400	28	860	3100	370	43	85
JZ042	160	13000	30	9.4	14	21000	6800	400	28	630	3100	280	43	69
JZ043	180	11000	29	9.2	18	24000	8600	400	26	890	3200	380	44	87
JZ044	150	9300	28	8.9	14	19000	6200	370	23	560	3000	210	40	64
JZ045	180	12000	30	10	18	23000	8300	410	27	860	3200	360	45	89
JZ046	170	11000	31	9	16	21000	6800	400	25	630	3300	260	44	69
JZ047	180	11000	29	9.3	18	23000	8400	410	26	840	3200	350	42	86
JZ078	270	37000	23	8.5	16	20000	4800	440	30	750	3500	980	45	83
JZ079	280	35000	22	8.8	16	20000	4900	410	26	780	3600	930	45	84
JZ080	270	44000	22	7.7	15	20000	5000	370	29	830	3500	1100	42	84
JZ081	270	30000	23	7.5	16	19000	4900	370	27	770	3600	880	43	83
JZ082	260	29000	22	7.5	15	20000	4600	390	28	720	3600	880	42	82
JZ083	280	20000	24	8.4	16	21000	4700	420	27	730	3900	820	45	89
JZ084	160	25000	33	7.8	15	18000	8000	430	33	570	2200	570	36	66
JZ085	160	29000	31	7.9	15	18000	8300	410	32	630	2100	640	35	69
JZ086	160	29000	23	8.3	16	17000	8500	430	32	600	2100	640	37	67
JZ087	140	23000	32	6.6	14	18000	8400	450	26	590	2100	530	38	59
JZ088	150	27000	31	7.8	14	18000	8000	430	28	590	2200	580	37	63
JZ089	150	16000	28	7.9	14	18000	7000	430	26	590	2300	440	37	61
JZ090	160	26000	26	8	18	21000	11000	530	27	1100	4000	750	36	85
JZ091	160	26000	24	7.8	19	21000	11000	510	26	1200	3800	710	36	84
JZ092	170	26000	25	8.3	19	21000	12000	530	25	1200	4000	750	37	87
JZ093	170	27000	24	7.7	19	21000	12000	520	22	1200	4000	760	38	88
JZ094	160	29000	25	8.1	19	20000	11000	510	27	1100	3800	780	36	84
JZ095	170	28000	30	8.7	19	21000	12000	540	26	1100	4100	760	37	86
JZ096	250	3300	21	7.2	11	16000	3400	560	13	620	3100	320	31	61
JZ097	250	3500	22	7.2	12	15000	3300	540	12	620	3000	330	31	59
JZ098	250	3400	22	7.2	10	15000	3300	580	14	600	3100	330	32	59
JZ099	230	3100	23	6.9	13	15000	3400	530	13	560	2900	290	30	56
JZ100	250	3000	23	7.6	11	16000	3600	550	16	510	2900	260	33	62
JZ101	230	3000	21	7.5	10	15000	3300	490	15	490	2800	300	31	56
JZ102	190	3300	21	6.9	14	18000	4500	320	23	500	4300	220	45	81
JZ103	220	17000	20	6.8	17	18000	5400	270	26	550	3400	450	49	85
JZ104	230	23000	23	7.4	20	18000	5900	250	31	610	3300	570	50	89
JZ105	240	13000	24	8	21	21000	5900	250	31	580	3700	410	57	99
JZ106	230	11000	26	7	19	21000	5800	290	32	570	3900	370	58	100
JZ107	240	14000	23	7.1	18	19000	5500	260	29	560	3700	460	52	89
JZ108	110	4700	13	4.2	7.5	11000	3300	250	7.9	590	2300	220	17	34
JZ109	100	4300	21	3.9	7.5	11000	3000	250	11	540	2100	200	18	31
JZ110	100	4700	16	4.5	7.5	11000	3200	260	8.7	510	2200	220	17	34
JZ111	120	5000	15	4.3	7.5	11000	3100	260	9.2	500	2200	220	17	31
JZ112	110	5100	25	5	7.9	11000	3400	260	14	530	2300	230	18	35
JZ113	110	4700	20	4.1	7.5	11000	3300	240	9.5	550	2200	210	17	32

Table S3) Dissolved Metal Screen. Conducted by the University of Idaho Analytical Sciences Laboratory (https://www.uidaho.edu/cals/analytical-sciences-laboratory). All values are reported in µg/g concentrations.

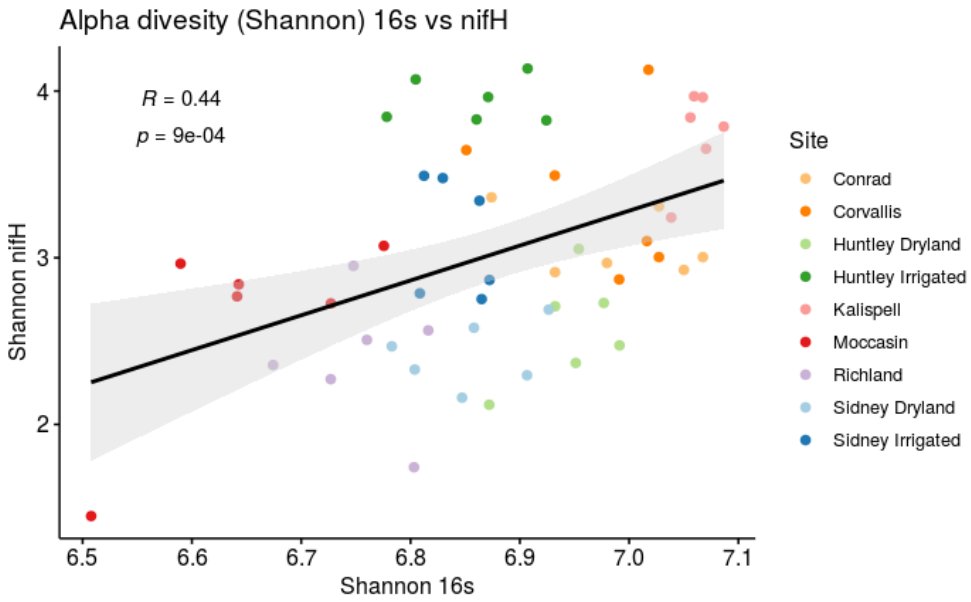

Figure S1) The linear correlation of the diversity of the 16S bacterial community and the *nifH* community across the state of Montana. Grey bar represents the 95% confidence.

Korean Journal of Women Health Nursing | ISSN: 2287-1640 | https://kjwhnb.com/

11

12

Site	Difference	p adj.
Richland-Moccasin	0.141895	0.007
Sidney Irrigated-Moccasin	0.210567	<0.001
Huntley Irrigated-Moccasin	0.220441	<0.001
Sidney Dryland-Moccasin	0.230535	<0.001
Huntley Dryland -Moccasin	0.317364	<0.001
Conrad-Moccasin	0.354823	<0.001
Corvallis-Moccasin	0.358359	<0.001
Kalispell-Moccasin	0.437231	<0.001
Huntley Dryland -Richland	0.175469	<0.001
Conrad-Richland	0.212929	<0.001
Corvallis-Richland	0.216464	<0.001
Kalispell-Richland	0.295336	<0.001
Conrad-Sidney Irrigated	0.144256	0.005
Corvallis-Sidney Irrigated	0.147792	0.004
Kalispell-Sidney Irrigated	0.226663	<0.001
Corvallis-Huntley Irrigated	0.137918	0.010
Kalispell-Huntley Irrigated	0.21679	<0.001
Kalispell-Sidney Dryland	0.206696	<0.001

Table S4) The general bacterial community Shannon diversity index comparing sites with Tukey HSD Multiple comparison of means test. Values only represents significant comparison on Figure 3A in the manuscript.

Site	Difference	p adj.
Conrad-Richland	0.681	0.042
Sidney Irrigated-Richland	0.720	0.026
Corvallis-Richland	0.974	<0.001
Kalispell-Richland	1.343	<0.001
Huntley Irrigated-Richland	1.545	<0.001
Conrad-Sidney Dryland	0.660	0.055
Sidney Irrigated-Sidney Dryland	0.699	0.034
Corvallis-Sidney Dryland	0.953	<0.001
Kalispell-Sidney Dryland	1.322	<0.001
Huntley Irrigated-Sidney Dryland	1.524	<0.001
Corvallis-Huntley Dryland	0.7982	0.009
Kalispell-Huntley Dryland	1.167	<0.001
Huntley Irrigated-Huntley Dryland	1.369	<0.001
Kalispell-Moccasin	1.105	<0.001
Huntley Irrigated-Moccasin	1.307	<0.001
Kalispell-Conrad	0.661	0.054
Huntley Irrigated-Conrad	0.864	0.003
Huntley Irrigated-Sidney Irrigated	0.825	0.006

Table S5) The diazotrophic Shannon diversity index comparing sites with Tukey HSD Multiple comparison of means test. Values only represents significant comparison on Figure 3B in the manuscript.

13

	16S				*nifH*			
	NMDS1	NMDS2	r^2	Pr(>r)	NMDS1	NMDS2	r^2	Pr(>r)
Organic Matter	-0.2851	-0.9584	0.0463	0.3039	-0.5827	-0.8126	**0.4791**	**<0.001**
Moisture Content	0.3594	-0.9331	**0.1496**	**0.015**	0.2163	-0.9763	0.0484	0.287
Nitrate & Nitrite	-0.0526	-0.9986	**0.4311**	**<0.001**	0.0158	0.9998	**0.2201**	**0.001**
Ammonia	-0.1574	0.9875	0.869	0.103	-0.1664	-0.9860	0.0214	0.5805
Available Phosphorus	0.1025	-0.9947	**0.3735**	**<0.001**	0.7747	0.6323	**0.4094**	**<0.001**
Available Potassium	0.4470	-0.8944	**0.2500**	**0.002**	0.9896	-0.1438	**0.4202**	**<0.001**
pH	0.9120	-0.4101	0.0116	0.7416	-0.0216	-0.9997	0.0218	0.5824
Barium	0.2012	0.9795	**0.3251**	**<0.001**	-0.0900	-0.9959	**0.4684**	**<0.001**
Calcium	0.07562	0.9971	0.0695	0.160	0.2922	-0.9563	0.0923	0.084
Cobalt	-0.0417	0.9991	**0.1502**	**0.012**	0.0045	-0.9999	0.0247	0.523
Copper	0.4195	0.9077	0.0829	0.108	0.3109	-0.9504	0.0748	0.138
Iron	0.1520	0.9883	**0.1417**	**0.016**	0.2501	-0.9682	0.0625	0.186
Magnesium	-0.1618	-0.9868	0.0698	0.161	0.2269	-0.9739	0.0438	0.321
Manganese	-0.5538	0.8326	0.0235	0.542	0.1875	-0.9822	0.0048	0.886
Nickel	0.0953	0.9954	**0.2801**	**<0.001**	0.6339	-0.7733	**0.1258**	**0.036**
Phosphorus	0.1857	-0.9825	0.1356	0.023	0.0964	-0.9953	0.0058	0.865
Sulfur	0.2397	0.9708	0.0534	0.243	0.1098	-0.9939	**0.17.00**	**0.009**
Zinc	0.3944	0.9189	**0.2815**	**<0.001**	0.4439	-0.8960	**0.2387**	**0.001**
Pea Variety			0.0692	0.814			0.0449	0.967
Tillage			**0.2532**	**<0.001**			**0.2100**	**<0.001**
Previous Crop			**0.2592**	**<0.001**			**0.3484**	**<0.001**
Season Precipitation	-0.2816	0.9595	0.0412	0.342	0.3009	-0.9536	0.0400	0.354
Irrigation	0.0894	-0.9959	**0.5795**	**<0.001**	0.5402	0.8415	**0.4424**	**<0.001**
Grain Yield	-0.0291	-0.9995	**0.1658**	**0.012**	-0.1033	0.9946	0.0052	0.874
Geographical location			**0.7556**	**<0.001**			**0.8274**	**<0.001**

Table S6) Output of the envfit function showing the proportion of variance explained of the chemical, farm management and geographical variables in relationship to the first and second axis of the Bray-Curtis based NMDS ordination. The r^2 values represent the proportion of variance explained by the ordination of the particular variable. The NMDS1 and NMDS2 values give the directional cosines of the vectors and the p value determines the 'significances' of the fitted vectors and is assessed by 999 permutations of environmental variables. Significant r^2 and p values are in bold.

14

Korean Journal of Women Health Nursing | ISSN: 2287-1640 | https://kjwhnb.com/

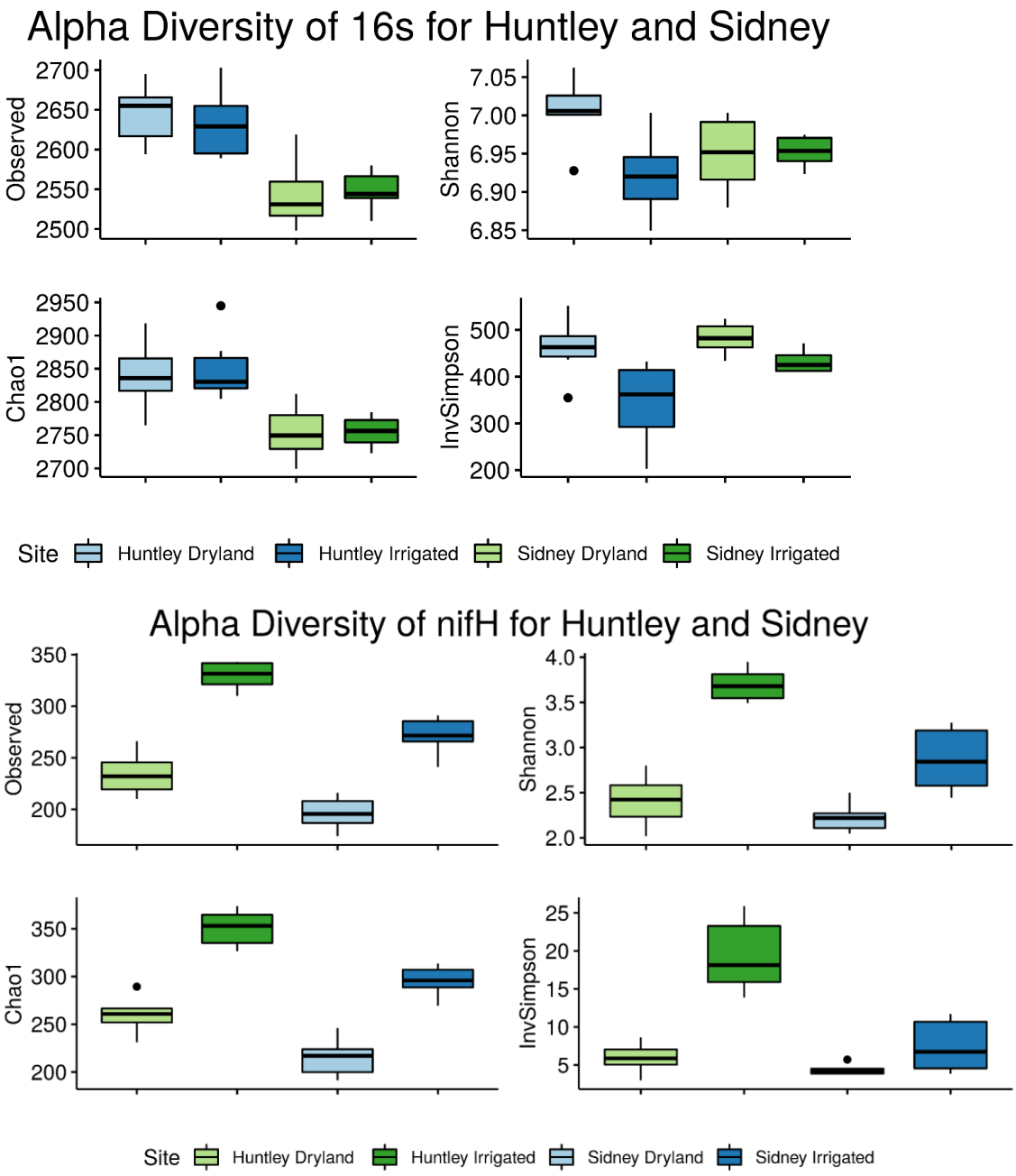

Figure S2) Alpha diversity metrics for the general and diazotrophic bacterial communities.

19

16S

	R2.adj	*AIC*	*F*	*Pr(>F)*
+ Available Phosphorus	0.11444	1125.3	7.8493	0.002
+ Barium	0.20168	1120.6	6.6824	0.002
+ Calcium	0.26223	1117.3	5.1854	0.002
+ Available Potassium	0.32543	1113.4	5.6843	0.002
+ Manganese	0.41798	1106.3	8.7924	0.002
+ Organic Matter	0.46408	1102.7	5.1287	0.002
+ Nitrate Nitrite	0.49021	1100.8	3.4092	0.01
+ Phosphorus	0.51502	1098.9	3.3532	0.006
All variables	0.5263			

nifH

	R2.adj	*AIC*	*F*	*Pr(>F)*
+ Organic Matter	0.09028	158.53	6.0538	0.002
+ Barium	0.14774	156.12	4.345	0.002
+ Available Potassium	0.20662	153.37	4.5984	0.002
+ Manganese	0.26578	150.31	4.8134	0.002
+ Available Phosphorus	0.31551	147.62	4.3473	0.002
+ Cobalt	0.35849	145.21	4.0024	0.002
+ Nitrate Nitrite	0.37945	144.41	2.4499	0.006
All variables	0.39355			

Table S7) Statewide CAPSCALE models explain 52% and 39% for the general and diazotrophic beta diversity respectively. Each model was test for variance inflation factors (VIF) and limited to co-variables with a VIF score less than 10. These variables were run through the function ordiR2step to obtain variance explained percentages.

20

Korean Journal of Women Health Nursing | ISSN: 2287-1640 | https://kjwhnb.com/

21

General bacteria Capsacle model.

	R2.adj	*Df*	*AIC*	*F*	*Pr(>F)*
Moisture Content	0.44777	1	12.9725	19.6494	0.002
Calcium	0.63579	1	3.8663	12.3573	0.002
Available Potassium	0.70493	1	-0.3571	5.9207	0.002
Available Phosphorus	0.732	1	-1.8971	3.0198	0.016
Nitrate Nitrite	0.76856	1	-4.7152	4.0018	0.002
Copper	0.77859	1	-5.1497	1.8149	0.026
All variables	0.7902				

Diazotrophic Capscale model

	R2.adj	*Df*	*AIC*	*F*	*Pr(>F)*
Available Potassium	0.22345	1	609.07	7.6181	0.002
Nitrate Nitrite	0.39195	1	604.08	7.0968	0.002
Copper	0.47121	1	601.56	4.1476	0.002
Calcium	0.53878	1	599.04	3.93	0.002
All variables	0.54973				

Table S8) Statewide CAPSCALE models explain 79% and 55% for the general and diazotrophic beta diversity respectively. Each model was test for variance inflation factors (VIF) and limited to co-variables with a VIF score less than 10. These variables were run through the function ordiR2step to obtain variance explained percentages.

22

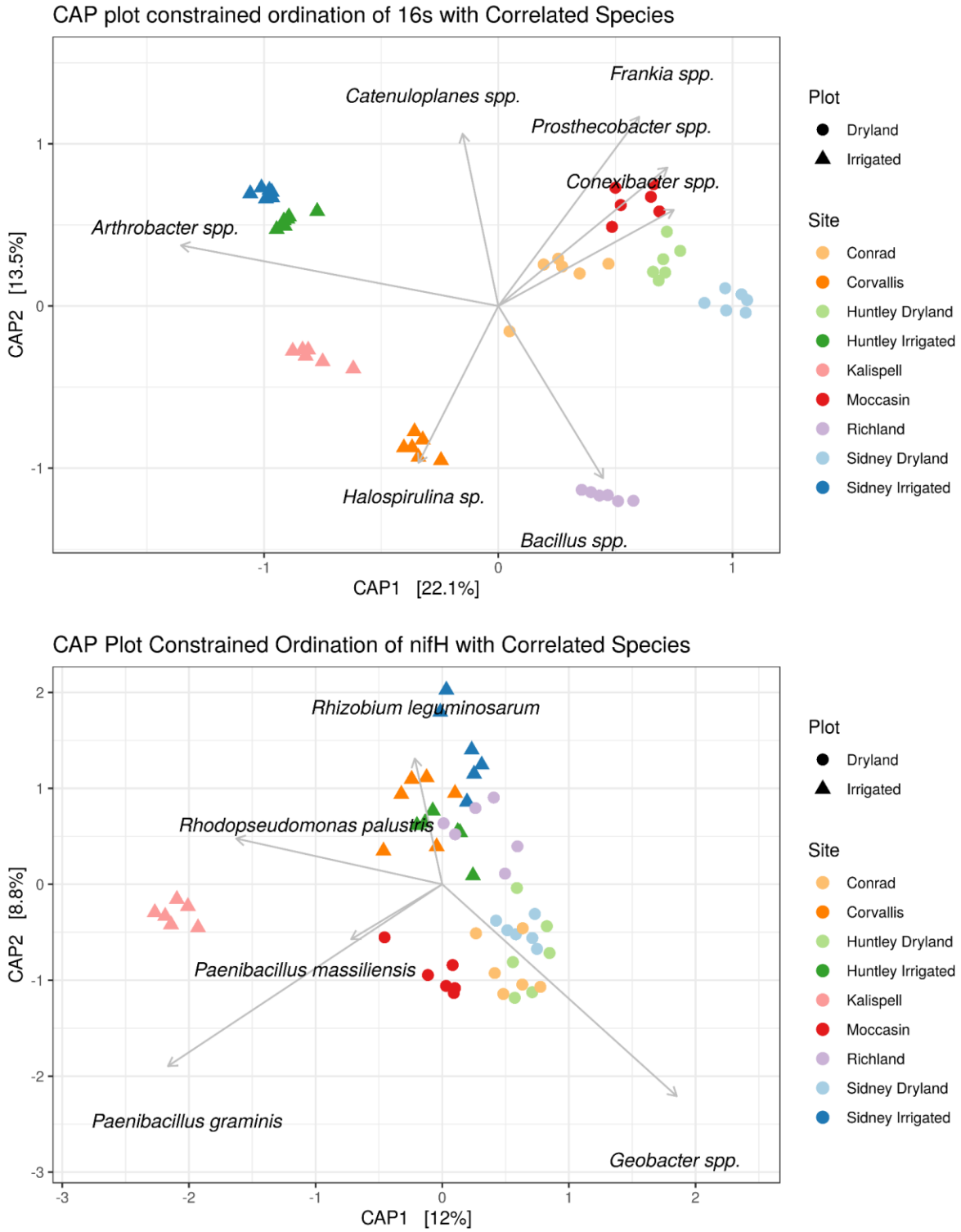

Figure S3) Canonical analysis of principle coordinates (CAP) plots were constructed with the OTUs labels across. OTUs plotted correlate highly ($r^2 > 0.9$) with site location in the CAP space. Large arrow corresponds to higher correlation with the axis direction of the plot. OTUs are labeled with the species names multiple OTUs can correlated with a single species names in the diazotrophs due to the overlap of species names to 3% dissimilar *nifH* genes.

	# nodes	*#edges*	*Avg k*	$\langle d \rangle$	$\langle C_i \rangle$	$\langle CB \rangle$
Dryland	153	274	3.58	5.4	0.23	223
Irrigated	153	204	2.66	6.3	0.22	308
Huntley	153	261	3.41	5.9	0.25	222
Sidney	153	291	3.80	5.2	0.25	225

Table S9) Network topology of irrigated and dryland networks. All nodes are the same but co-occurrence differs depending on irrigation or geography. Average node degree (Avg k), shortest path length $\langle d \rangle$, average centrality $\langle C_i \rangle$, and average betweenness centrality $\langle CB \rangle$, were all measured in the igraph package.

23

24

	Irrigated	*Dryland*
Nodes	48	106
Edges	80	218
Clusters	3	9
Density	0.039	0.7
Transitivity Global	0.269	0.258
Transitivity Average	0.282	0.311
Centralization Degree	0.099	0.084
Centralization Betweenness	0.13	0.22
Eigen Vector Centrality	4.987	6.175
Modularity	0.571	0.647

Table S10) Network properties of irrigated and dryland co-occurrence networks of diazotrophic communities. Network properties of networks shown in Figure 6. These networks removed any node that was not within a community designated by the

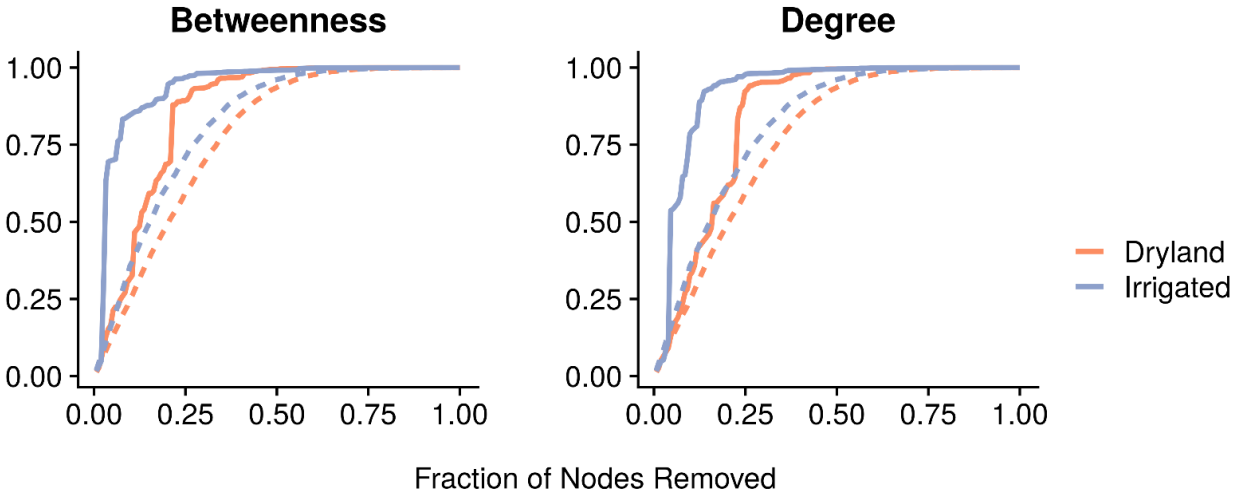

Figure S4) Robustness analysis for diazotrophic co-occurrence networks. Targeted (filled line) and random (dashed lines) for both betweenness centrality and node degree were tested for their robustness. Dryalnd (orange) and irrigated (purple) are effect by node removal at different rates for random, betweenness, and degree node removal.

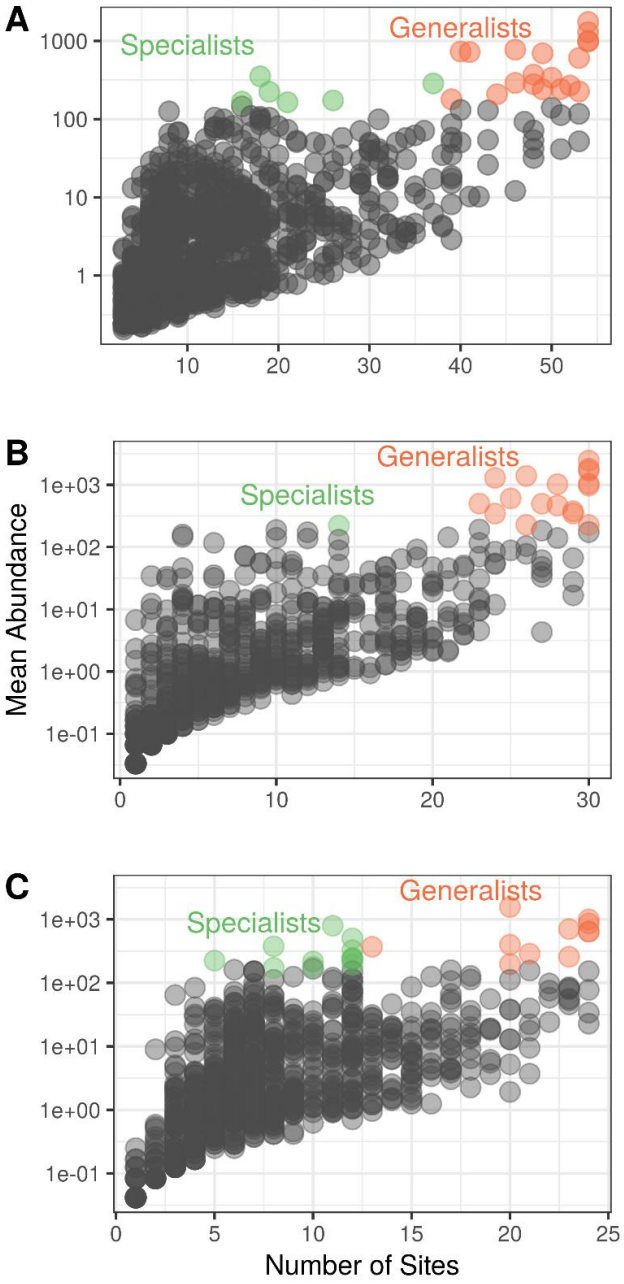

Figure S5) Generalist and Specialist in the diazotroph population. A) Total population of nifH genes across the state of Montana. Generalist OTUs are classified by being the top 5% prevalence and top 2.5% relative abundance. B) All dryland farms across the state. C) All Irrigated plots across the state.

Korean Journal of Women Health Nursing | ISSN: 2287-1640 | https://kjwhnb.com/

57

58

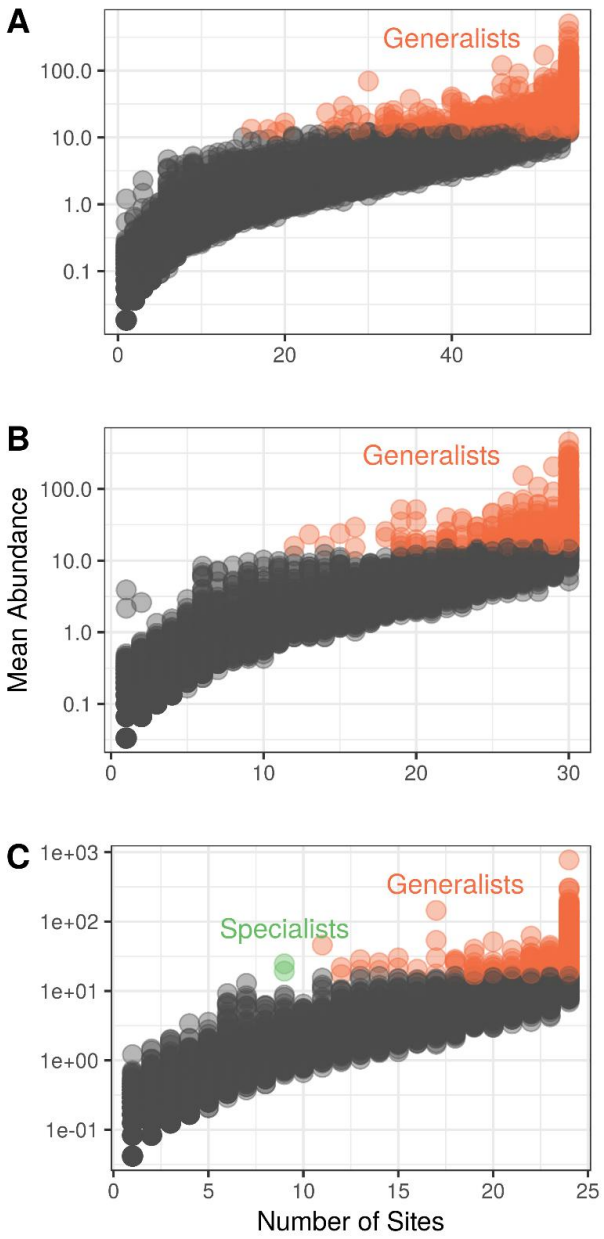

Figure S6) Generalist and Specialist in the 16s population. A) Total population of 16s genes across the state of Montana. Generalist OTUs are classified by being the top 5% prevalence and top 2.5% relative abundance. B) All dryland farms across the state. C) All Irrigated plots across the state.

59

60