Tallgrass Restoration Legacies, Summary

Beau Larkin, coauthors L. Bullington, P. Ramsey, & Y. Lekberg

Last updated: 12 December, 2023

Contents

1 Introduction

The tallgrass prairie biome in North America is nearly extinct, creating a tremendous desire for restoration of this critical habitat. In the upper Midwest, interested landowners have attempted to convert previous agricultural fields (corn or soybeans, usually) to something that looks like prairie for decades. The consequences of this practice for soil biota are largely unknown. Agricultural practices filter fungal and bacterial communities for species that can tolerate extreme nutrient loads, frequent tillage, and poor plant diversity. When agricultural practices cease, and native perennial plants are introduced, soil biotic communities are again filtered by the new regime and rapidly change. What is the endpoint of these new communities? Do they begin to resemble communities in remnant prairie (the reference condition), or do they form novel communities, indicating a change of state in the system due to disturbance associated with agriculture? Are these communities' fates similar across distant regions, or is each context dependent?

2 Description

The goal is to present results and discuss whether a path forward with these data exists. If so, we will determine the strategy that presents the best story, organization, and interpretation of these results.

3 Methods

3.1 Sites

We visited fields in southern Wisconsin and at Fermilab, just west of Chicago, Illinois. Site descriptions and photos are available in this field note. The survey followed an unbalanced complete block design. Corn, restored, and remnant fields are compared, with at least one of each field type in each block. I have called blocks "regions". We collected samples and data from four regions (Fig 1, left).
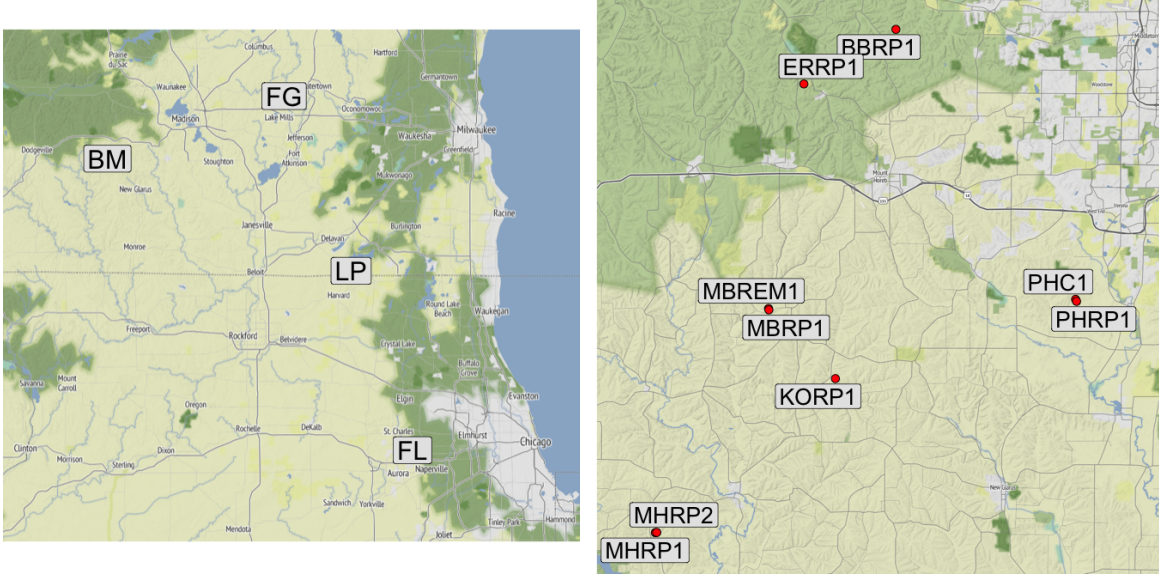

Figure 1: (Left) Labels show centroids of regions used for this work. BM = Blue Mounds, FG = Faville Grove, FL = Fermilab, LP = Lake Petite. (Right) Labels show individual fields in the Blue Mounds region.

3.2 Design

The design is unbalanced because there are more restored fields than corn or remnant (table 1), and the count of restored fields differs among regions. In all but one case, only single corn and remnant fields were available in each region. This means that we only have replication to separate field types when using the entire block design.

Table 1: Count of fields by type and region: BM = Blue Mounds, FG = Faville Grove, FL = Fermilab, LP = Lake Petite

	corn	restored	remnant
BM	1	7	1
FG	1	1	1
FL	2	6	1
LP	1	2	1

A feature of our design is that restored fields vary in time since restoration, setting up the potential for a chronosequence in addition to the contrasts among corn, restored, or remnant fields. The rules for establishing a chronosequence are strict. We cannot call fields from all regions a chronosequence. The

Blue Mounds region, which has seven restored fields in a relatively small geographic area, is our only opportunity to take advantage different field ages (Fig 1, right), but we will likely have to call this a "pseudochronosequence" and avoid some inferences. Supporting our case in the Blue Mounds, mantel tests (not shown) failed to find correlations between soil variables and pairwise distance, giving us a little confidence that we're avoiding systemic bias.

3.3 Soil Fungal Communities

We collected soil cores from 10 haphazardly-selected locations in each field. Genomic DNA was extracted, and the lab and bioinformatics pipeline delivered community data from ITS or 18S regions clustered as 97% similar OTUs or 99% similar SVs. In preliminary work, inferences made with SVs were weaker but not qualitatively different. I continued with OTUs only.

Originally, I had planned to average the sequence abundance per OTU in each field. In part, I did this for stastistical design reasons. Since fields are replicates, the samples in fields aren't independent. Also, explanatory and response data exist as one data point per field, meaning that they should not be expanded to apply singly to each sample. Finally, I didn't have the tools to account for this design in permutation tests. With available tools, I could only choose one "nesting" factor, but to include samples, I'd have to account for regions and fields as blocking elements when testing the design element of field type. Luckily, I've recently found a new tool, the function `how()` from package `permute`, that allows for more complicated permutation designs. I decided to proceed along parallel analysis paths: one with sample-level data, and one with field-level data.

Creating the sample-level data took some work. A few samples had failed to amplify, leaving some fields represented with 9 samples instead of 10. This unbalance, which is normally not a problem, became unacceptable for a couple of reasons. First, the permutations I had planned using function `how()` require balance at the plot (i.e., field) level. Also, some samples contained very few sequences compared with the others. Rarefying the entire dataset to these tiny depths would result in the loss of many admittedly low-abundance OTUs. To acheive a balanced count of samples per field and retain a higher rarefaction threshold, I used an iterative process to remove samples with few sequences, choosing a rarefaction cutoff near the plateau of OTU recovery at depth (Fig 2). This ended up being 8 samples for ITS and 7 for 18S datasets. These sample sequence values were then rarefied and used as-is for sample-level data, or, when analyzing at the field level, sample sequence values were summed and rarefied. The resulting data files recovered many more OTUs, but didn't change any major interpretation of the data. The situation with 18S data is similar, but at a much lower total sequencing depth (not shown).

Species (OTU) accumulation was performed on field-level sums of sequences. Most fields were undersampled (Fig 3), particularly remnant and restored fields (ITS) and corn and remnant fields (18S).

3.4 Plant data

Plant community data resulted from two independent surveys. In the Wisconsin regions, haphazard transects were established and 10 meter frames placed, with percent cover estimated for all species, resulting in a dataset with plant composition. In Fermi, relevé methods were used, resulting in presence/absence data only. Plant metadata were assembled from multiple sources and include plant traits and natural history details. These metadata aren't used with fields from Fermi because it doesn't make sense to think of the presence/absence of traits only. More information about the plant data, including preliminary analysis, is available in this report.
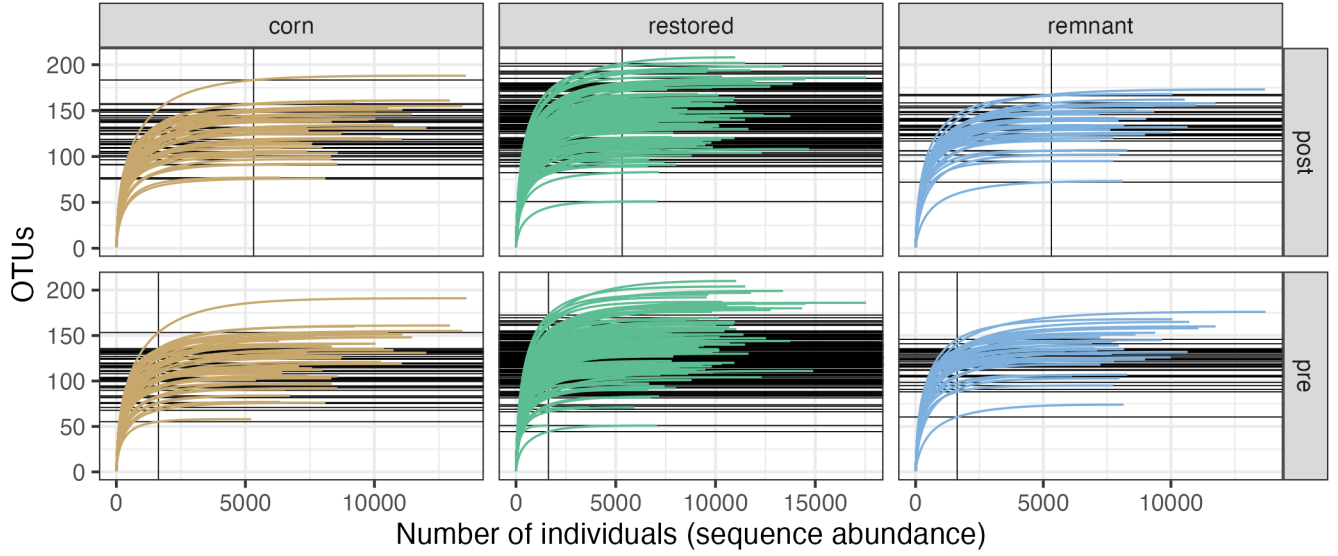
Figure 2: Individual-based rarefaction of ITS OTUs in subsamples when nine subsamples per field are retained (pre, top row) or when eight subsamples are retained (post, bottom row). Vertical lines show the minimum sequencing depth available across samples; this is the depth to which the entire set would be rarefied. Horizontal lines show the number of OTUs that would be recovered from each sample at that rarefaction depth.
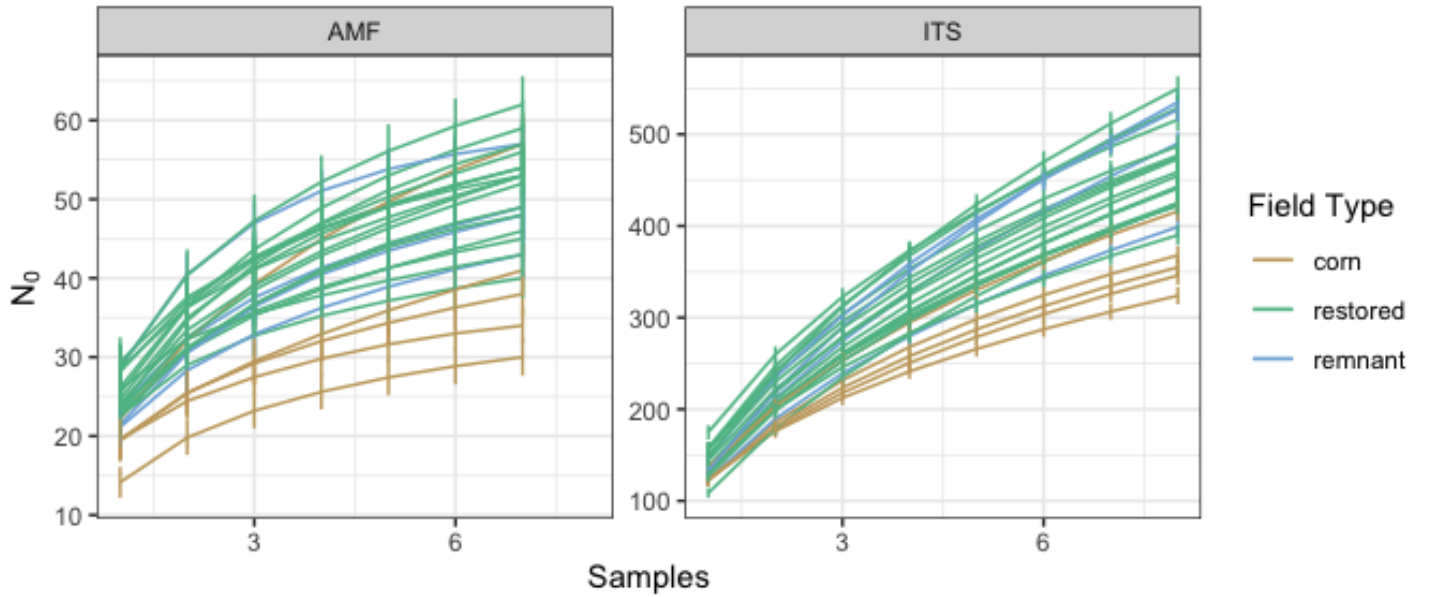

Figure 3: Species accumulations for rarefied sums of field sequences for ITS and 18S datasets. Species accumulation by the exact method; standard deviation (vertical lines) conditioned by the empirical dataset.

3.5   Soil data

In the field, soil was pooled from 10 haphazardly-selected locations, mixed, and sampled once for soil chemical analysis. Soil data includes abiotic macro and micronutrients, organic carbon, and properties like pH. Average precipitation was determined for each field using PRISM climate data and is included with the soil data for constrained analyses. More information is available here.

3.6   Design/site data

- Field type: corn, restored, remnant
- Field age: years since restoration (NA for corn and remnant)
- Region: blocks

3.7   Response data

Also taken from the pooled soil in each field, one sample was taken for analysis of Water Stable Aggregates (I don't know who did this), and one for quantification of microbial biomass.

3.8   Data sources, summary

- Fungal genomic data, ITS and 18S, OTU clusters
- Plant community data, composition in Wisconsin and presence/absence in Fermi
- Plant traits and natural history

    - Native status, photosystem, functional guild, lifecycle...

- Soil properties
- Fungal biomass (PLFA, NLFA)
- Water stable aggregates
- Site metadata and design

4   Results

4.1   Fungal communities

Fungal communities in cornfields stand apart from remnants and restored fields restored fields (Fig 4). As the age of restored fields increases, their distance from corn grows, and remnants lie in an intermediate position among restored fields. The high dimensionality here keeps the explanatory power of axes 1 & 2 low. Clusters based on field type were significant based on a permanova (with regions as blocks and subsamples held constant within fields, the replicate, table 2). However, a pairwise post-hoc permutation failed to separate remnant fields from any other type (table 3). This is due partially to a low number of replicates, and it's also a problem with dispersion. The remnants are as distant from each other as they are from corn. I also attempted this using field-level data (not shown), and the p values were somewhat improved, but remnants still failed to separate from corn or restored fields.
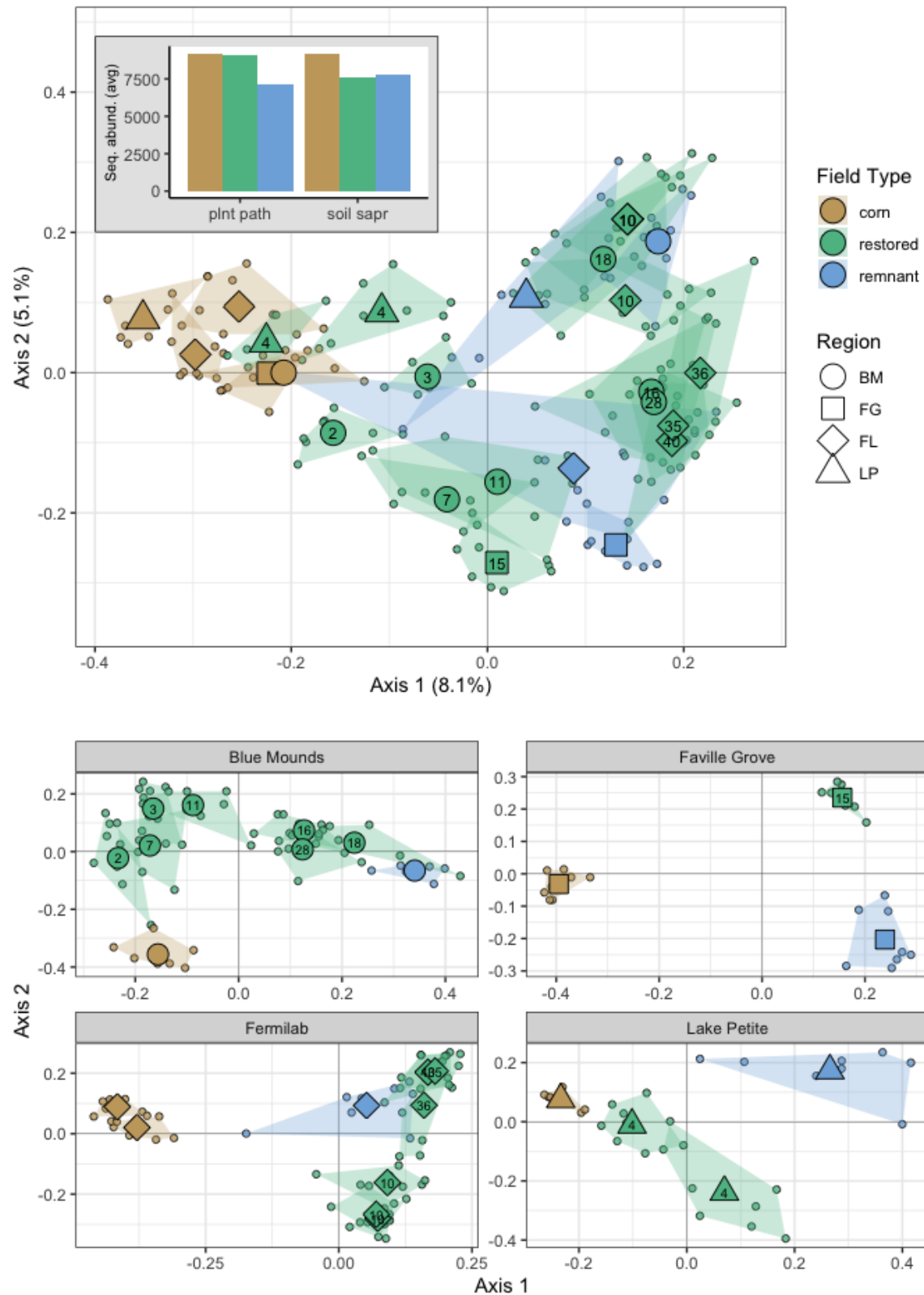
Figure 4: The top panel is a PCoA ordination of samples in fields, ITS dataset. Small circular icons show locations of samples and hulls enclose fields. Each field is represented with a large icon on the field centroid, where icon shapes identify the field region and color identifies the field type. Text on the restored field icons shows years since restoration. The inset column chart displays average sequence abundance across all fields for the three most abundant fungal guilds. The lower four panels show independent db-PCoAs for each region.

Table 2: Permutation test under reduced model. Blocks: region, samples not permuted (not independent), number of permutations: 1999

Df	SumOfSqs	R2	F	Pr(>F)
2	5.758	0.081	8.676	0
197	65.371	0.919	NA	NA
199	71.129	1.000	NA	NA

Table 3: Pairwise tests under reduced model. Blocks: region, samples not permuted (not independent), number of permutations: 1999, p values corrected after multiple tests using the Benjamini & Hochberg correction.

group1	group2	R2	F_value	df1	df2	p_value	p_value_adj
restored	corn	0.066	11.692	1	166	0.002	0.006
restored	remnant	0.018	2.827	1	158	0.220	0.220
corn	remnant	0.134	10.854	1	70	0.143	0.214

AMF communities (18S data) showed qualitatively similar patterns to the ITS results (Fig 5), but with weaker support (tables 4 & 5).

Table 4: Permutation test under reduced model. Blocks: region, samples not permuted (not independent), number of permutations: 1999

Df	SumOfSqs	R2	F	Pr(>F)
2	5.400	0.109	10.571	0.002
172	43.931	0.891	NA	NA
174	49.330	1.000	NA	NA

Table 5: Pairwise tests under reduced model. Blocks: region, samples not permuted (not independent), number of permutations: 1999, p values corrected after multiple tests using the Benjamini & Hochberg correction.

group1	group2	R2	F_value	df1	df2	p_value	p_value_adj
restored	corn	0.057	8.787	1	145	0.001	0.003
restored	remnant	0.008	1.072	1	138	0.999	0.999
corn	remnant	0.111	7.616	1	61	0.143	0.214

4.2 Fungal diversity

Species richness, Shannon's, and Simpson's diversity are lowest in cornfields, and the trend is an increase in all indices through restored and remnant fields (Fig 6). These differences are significant for Richness
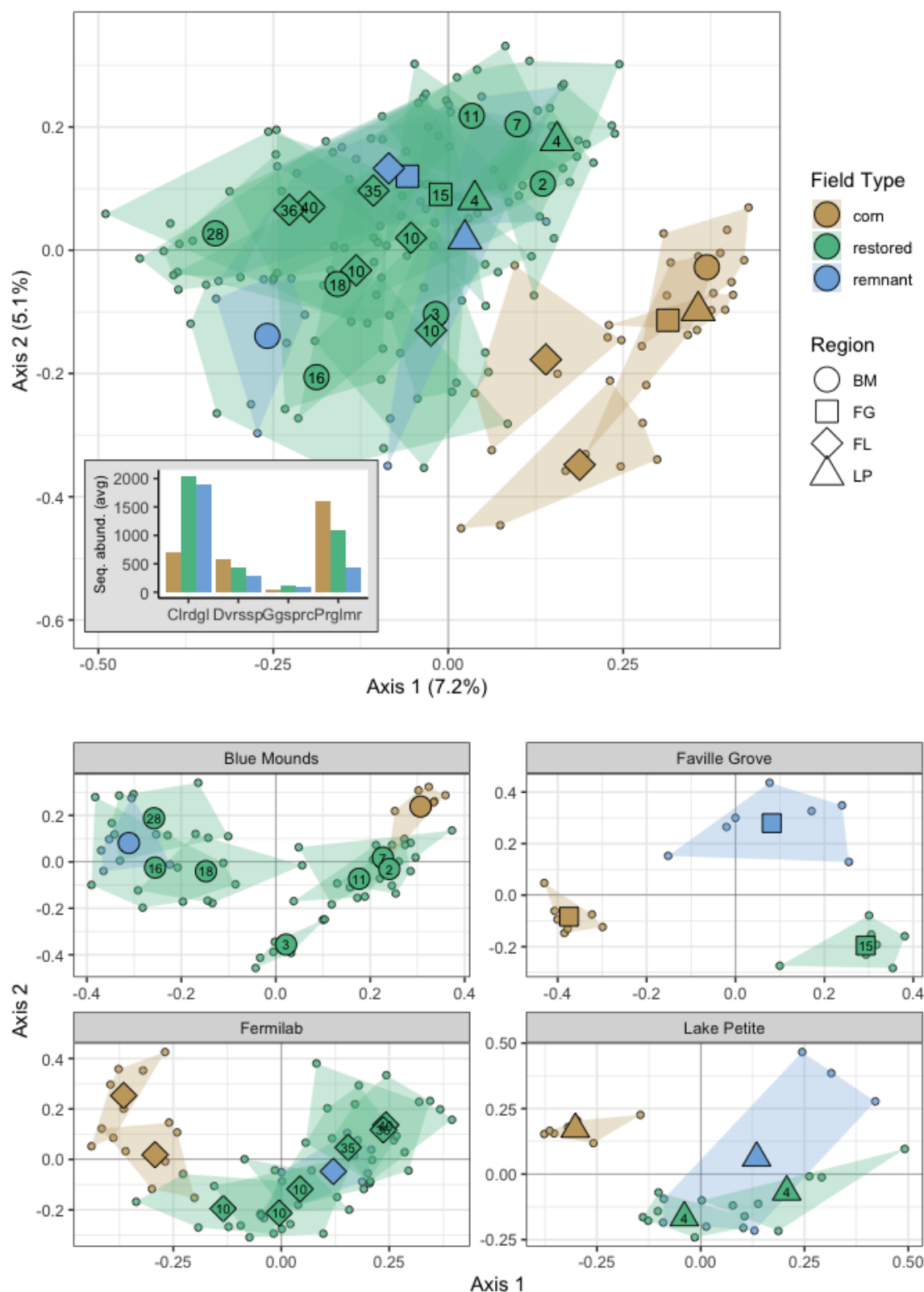
Figure 5: The top panel is a PCoA ordination of samples in fields, 18S dataset. Small circular icons show locations of samples and hulls enclose fields. Each field is represented with a large icon on the field centroid, where icon shapes identify the field region and color identifies the field type. Text on the restored field icons shows years since restoration. The inset column chart displays average sequence abundance across all fields for the most variable taxa (Claroideoglomeraceae, Diversisporaceae, Gigasporaceae, and Paraglomeraceae) The lower four panels show independent db-PCoAs for each region.
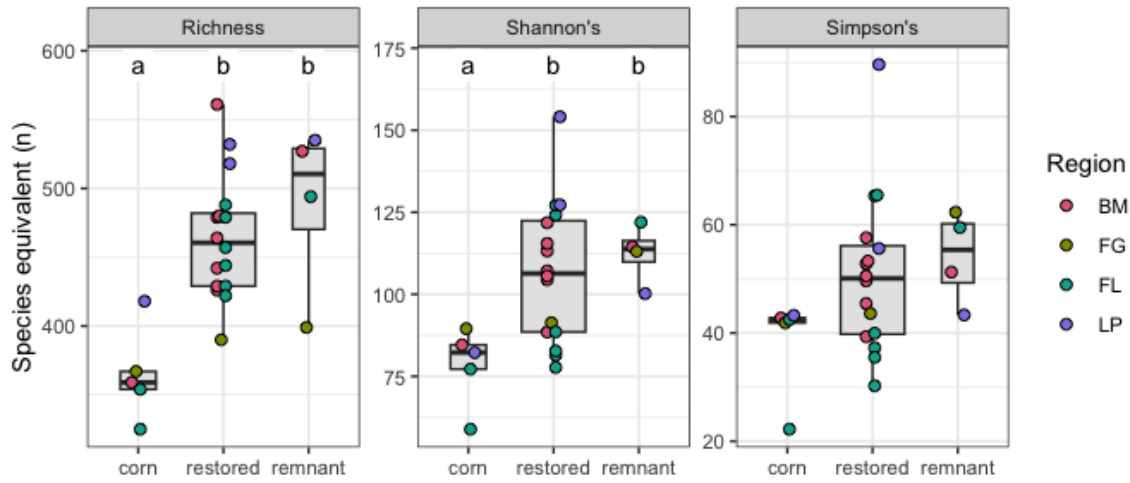
Figure 6: Counts of ITS OTUs are shown for richness and Shannon's and Simpson's diversity in all field types. Differences in means among field types were tested with mixed-effects linear models where region was coded as a random effect. Post-hoc comparisions are Tukey contrasts and letters show significant differences at p<0.05.

and Shannon's diversity, with cornfields being significnatly lower than either remnants or restored fields, which were not significantly different. Index values were more variable for restored fields than for other field types, which insults prerequisite needs for equal variance among groups.

Trends in diversity were similar in Blue Mounds, Faville Grove, and Fermilab, but at Lake Petite, Shannon's and Simpson's diversity indices were much higher than in cornfields or remnants (not shown), indicating that a sizeable percentage of OTUs were in relatively equal abundance there compared with other sites. This may partially be because the restored fields at Lake Petite are both young and haven't had time to develop the long tail of rarer species which characterizes most older habitats and tends to lower Shannon's and Simpson's index values.
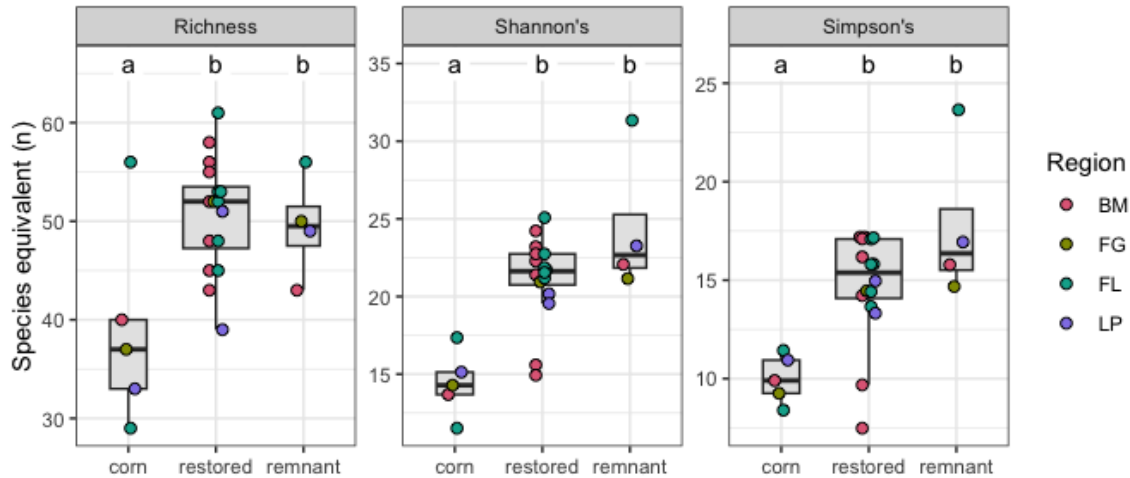

Figure 7: Counts of AMF OTUs are shown for richness and Shannon's and Simpson's diversity in all field types. Differences in means among field types were tested with mixed-effects linear models where region was coded as a random effect. Post-hoc comparisions are Tukey contrasts and letters show significant differences at p<0.05.

With AMF, trends in diversity largely matched those of soil fungi in general. Species richness showed a mild interaction, where it was highest in restored fields for Blue Mounds and Faville Grove, and in remnants

for Fermilab and Lake Petite. Shannon's and Simpson's indices increased from corn through restored and remnant fields for all regions.

For both soil fungi and AMF, diversity indices were significantly lower for corn but not between restored and remnant fields, suggesting a similarity between the latter.

4.3 Fungal guilds

Table 6: Average sequence abundances per field in each field type are shown for the top six guilds. A linear mixed model was used to separate means in field types and letters indicate significant differences based on Tukey's post-hoc test of that model at p<0.05.

guild	corn	restored	remnant	total
unidentified	25995	31580	34912	92487
plant pathogen	9226	9086	7182	25493
soil saprotroph	9145	7634	7796	24574
wood saprotroph	3700 (a)	2880 (ab)	1774 (b)	8354
dung saprotroph	3436	1998	1115	6549
litter saprotroph	2518	1765	1296	5579

Pathogen and saprotroph guilds are the most numerous in this dataset (table 6). Compared with the sequence abundance in the unidentified group, plant pathogens and soil saprotrophs are abundant enough to feel somewhat confident about in terms of coverage. I would have predicted soil saprotrophs to be lowest in cornfields (which have the least organic material), increasing in restored fields, and highest in remnants. With pathogens, I would have predicted the highest abundance in cornfields (due to the plant monoculture and high nutrients), lowest in restored fields (many plant species and less time for specialists to accumulate), and intermediate in remnants (high plant diversity and the presence of antagonistic fungi, but perennial plants in place to be attacked for decades). Instead, we found that the abundance of most guilds is highest in cornfields, intermediate in restored fields, and lowest in remnants.

At the level of taxonomic order, composition of soil saprotrophs in corn fields looks different from restored and remnant fields (Fig 8, left). Among well-represented orders, Agarics increase from corn to remnant, Cystofilobasidiales are prevalent only in corn, and Filobasidiales, often human pathogens, decline from corn to remnant. Geoglossales, which are associated with woodlands, moss, grasslands, probably don't like tillage or lack of substrate in cornfields and aren't found there. The order Hypocreales only occurs in restored and remnant fields. OTUs identified as Mortierellales are abundant in corn and restored fields, less so in remnants, with different OTUs associated with different field types. The contrast among field types with plant pathogens isn't obvious (Fig 8, right). The orders Glomerales, Hypocreales, and Pleosporales were the most abundant.

4.3.1 Trends in regions

Within regions, trends in abundance were also apparent for soil saprotrophs and plant pathogens, but these trends weren't consistent among regions (Fig 9). The abundance of soil saprotrophs in restored fields falls between corn and remnant fields in Blue Mounds and Fermilab, indicating that restored fields become more like remnants, but these trends are in opposite directions. In Lake Petite, the abundance of soil
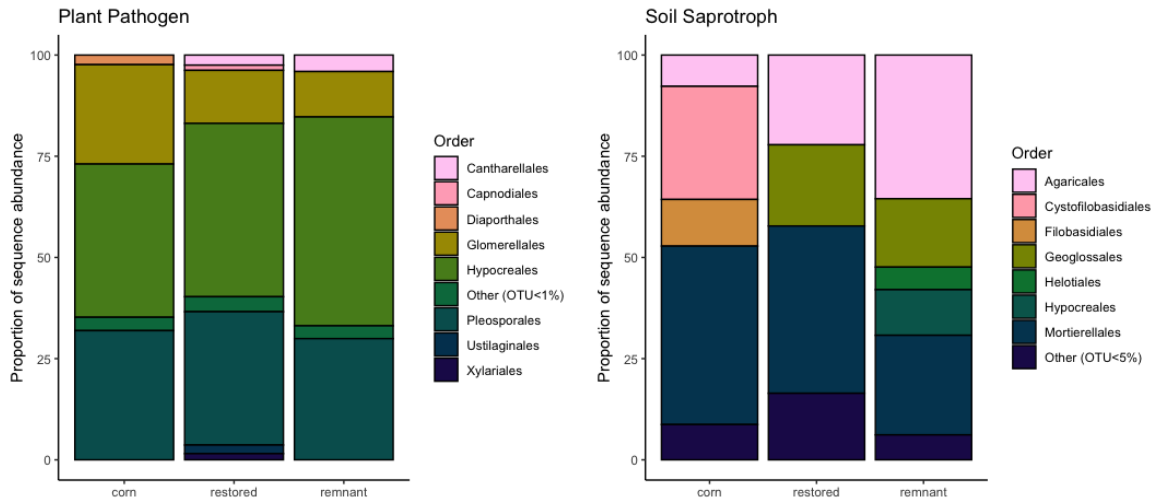
Figure 8: Composition of soil saprotroph (left) and plant pathogen (right) orders are shown across field types. Unidentified taxa were not included, and low abundance OTUs were filtered, see legends for the cutoff percentage.

saprotrophs is much lower than in corn or remnant fields, indicating little resemblance to either. In Faville Grove, soil saprotroph abundance is similar in all field types. Pathogen abundance is similarly inconsistent. In Blue Mounds and Faville Grove, pathogens spike in remnants and are lowest in remnants. At Fermilab, pathogens are instead lowest in restored fields. In Lake Petite, pathogens are highest in cornfields (which I would have predicted across the board), and decline from restored to remnant fields.
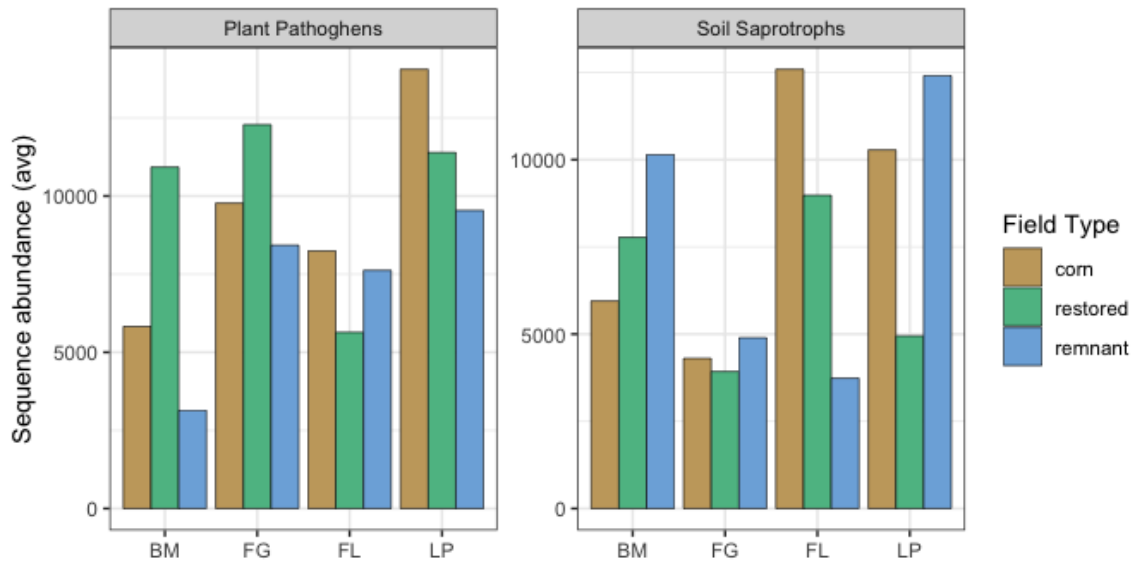

Figure 9: Soil saprotroph and pathogen guilds across field types, regional trends. Values are average sequence abundance in fields. No symbols of variance are shown because corn and remnant fields have only one replicate in most regions.

4.3.2 Trends over time

Contrary to predictions, plant pathogens show a strong decline from young to old restorations in the Blue Mounds region (Fig 10, left). Pathogen abundance reaches a low of 5,233 sequences in the 28 year old
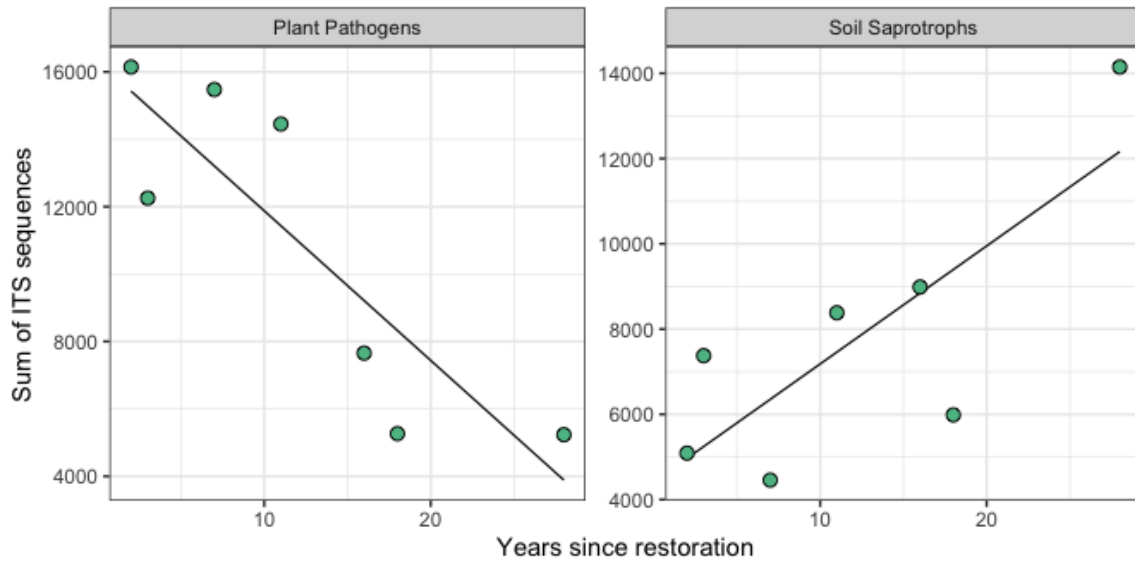
Figure 10: Icons show sums of ITS sequences in each restored field from the Blue Mounds region. Plant pathogen abundance decines with years since restoration (left, $R^2_{Adj} = 0.75$, $p = 0.013$). Soil saprotroph abundance increases with years since restoration (right, $R^2_{Adj} = 0.54$, $p = 0.037$), although leverage of the most extreme value is apparent.

field, which is lower than we found in the cornfield (5,822), though not as low as we found in the remnant (3,104). Aligned with predictions, soil saprotrophs increased with time in Blue Mounds' restored fields (Fig 10, right). Again, this increase is trending in the right direction, away from the cornfield's abundance of 5,950, reaching a high of 14,147 in the 28 year old field, which exceeds the remnant's abundance of 10,145.

4.3.3 Indicator species

OTUs from saprotroph and pathogen guilds were also heavily represented in an analysis to find indicator species for each field type. Note that the indicator species analysis is flawed because of the unbalanced number of fields for each field type. It's easier for a species to be detected in five cornfields than in sixteen restored fields, for example. I'd argue that as a basic filter to find interesting species, this analysis is acceptable if we don't try to call these species "indicators" of a particular habitat.

4.3.4 Natural history

I filtered indicator species to those which were identified to at least the genus level. Saprotrophs associated with cornfields tended to be disturbance-tolerant extremophiles, in particular, tolerating very cold environments. Some are known to produce suites of extracellular enzymes, exploting various nutrient sources (*Tausonia pullulans*) and making recalcitrant nutrients available to plants (*Mortierella minutissima*, which is used as an agricultural inoculant). *Pseudaleuria sp.* was also associated with cornfields. Species in this genus are suspected to increase crop yields and protect plants from disease. The only saprotroph associated with restored fields was another species in the genus *Mortierella*. Remnants hosted few indicator species. OTUs identified as *Gymnostellatospora sp.* and *Striaticonidium cinctum* appear to have little natural history information. An OTU in the genus *Clavaria* is a coral fungus, associated with woodland, savanna, or old-field habitats worldwide.

12

Not surprisingly, most pathogenic taxa associated with cornfields are known agricultural pathogens, including *Corynespora cassiicola*, *Curvularia sp.*, *Plectosphaerella cucumerina*, and *Setophoma terrestris*. *Microdochium colombiense*, a snow mold of rye and other cereal grains, is also an extreme cold tolerant species. In restored fields, an indicator pathogen was a *Pseudopithomyces sp.*, which is often found in anthropogenic environments and is a known wheat pathogen.

Natural history details about indicator fungi could make this paper more compelling, but the best value for such an analysis here is to show whether restored fields look more like cornfields, more like remnants, or more like something else entirely. But only 36% of these OTUs were assigned to a guild and 46% were assigned to a genus. It's likely that many saprotrophs and pathogens exist in these data that are opaque to our analysis. We'd have to assume that these unassigned taxa are distributed equally across field types, but this assumption is likely invalid because agricultural systems have received much more attention. Recovery of crop-associated fungi in restored fields probably reflects this bias as much as it shows a similarity with cornfield fungi. With few taxa identified in remnant fields, there's not much to compare with restored fields. Finally, very abundant OTUs tended to be commonly distributed, but all of these indicator OTUs occur in low abundance. Are they relevant to ecological processes in these fields? It's hard to say.

If natural history details are included, I think the most important case studies will come from these taxa:

1. Mortierella

   - Cornfields mostly, lots known about them, they have interesting physiology

2. Stagnospora

   - Aligns restored and remnant

3. Pseudopithomyces

   - Aligns corn and restored

4. Clavaria

   - Aligns restored and remnant

5. Microscypha/DSE

   - proposed mycorrhizae, in all remnants and 4 restored

4.4   AMF taxa

Table 7: Average 18S sequence abundances per field in each field type are shown for the top five families. A linear mixed model was used to separate means in field types and letters indicate significant differences based on Tukey's post-hoc test of that model at p<0.05.

family	corn	restored	remnant	total
Glomeraceae	15066	14225	15266	44557
Claroideoglomeraceae	702 (a)	2050 (b)	1892 (ab)	4644
Paraglomeraceae	1595	1082	429	3106
Diversisporaceae	588	434	287	1309
Gigasporaceae	39	109	105	253

Within the 18S dataset (AMF), taxa in the *Glomeraceae* are most abundant (7). As far as I know, these tend to be generalist taxa and are widespread. In this case, they do not differ in abundance among field types. Abundances fall off quickly for other taxa, with *Claroideoglomeraceae* and *Paraglomeraceae* an order of magnitude lower than the *Glomeraceae*, and *Gigasporaceae* abundance only 0.5% as much[1]. Differences among corn and restored fields are significant for the *Claroideoglomeraceae*, where they are lowest in corn and highest in restored fields (7). The abundance of these taxa in remnants doesn't separate due to large within-group variance, but the mean for remnants is closer to restored fields. Despite their small overall abundance, taxa in the *Gigasporaceae* showed a strong, increasing relationship with years since restoration in the Blue Mounds' restored fields (11).
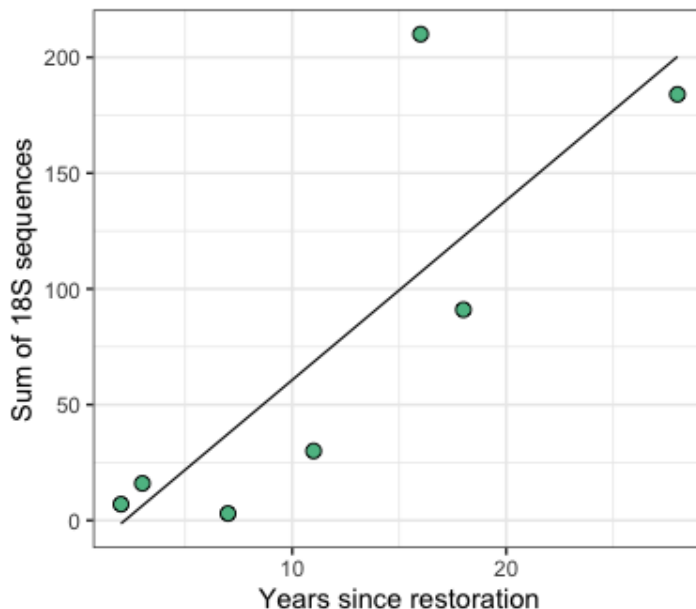

Figure 11: Icons show sums of 18S sequences in each restored field from the Blue Mounds region. The abundance of taxa in the *Gigasporaceae* increase with years since restoration ($R^2_{Adj} = 0.61$, $p = 0.023$).

4.5 Environmental constraints on fungal communities

We want to know how environmental/explanatory variables like soil chemistry, precipitation, and plant communities affect change in soil microbes among sites. But which sites? Constraining variables would undoubtedly find that soil nutrients explained differences between corn and restored or remnant fields, but this is hardly an interesting result. More worthwhile would be to constrain on restored fields and test whether years since restoration aligns fields in a relatively linear manner. Discovering if other explanatory variables constrain the data in a similar direction to years since restoration would be a wonderful outcome, because we'd learn what it is in the soil abiotic environment that's correlated with microbial community change.

Since the Blue Mounds restored fields are the most amenable to chronosequence methodology, I started there. With only seven fields, the available number of permutations was low, and with axes defined by very weak effects already, no significant constraining variables were found. Although reviewers may object, I decided that the only way forward was to use restored fields from other regions, breaking the rules of chronosequence analysis. Since the regions' soils and other environmental characteristics are so different, I parsed some out as covariables in a partial db-RDA (covariables: pH, SO4, Zn, Fe, Mn, Cu, Ca, Mg, Na,

---

[1]The switchgrass fields at Fermilab could not be included because no plant data exists for these fileds.

precipitation). I retained the soil nutrients N, P, and K as test variables because they were undoubtedly applied to cornfields and would decrease rapidly after agricultural practice ceased. I kept soil organic matter (OM) as a test variable too because it should be low in cornfields and increase after tillage ceases and mycotrophic, perennial plant species establish. I also included plant traits or plant community axes as test variables. Finally, I used years since restoration (yr_since) as a test variable. For the analysis, I first ran forward selection on regions in Wisconsin only because that's where we have plant composition data. Then, I ran forward selection with explanatory variables on all regions, using plant presence/absence axes only[2].

4.6 Constrained analysis

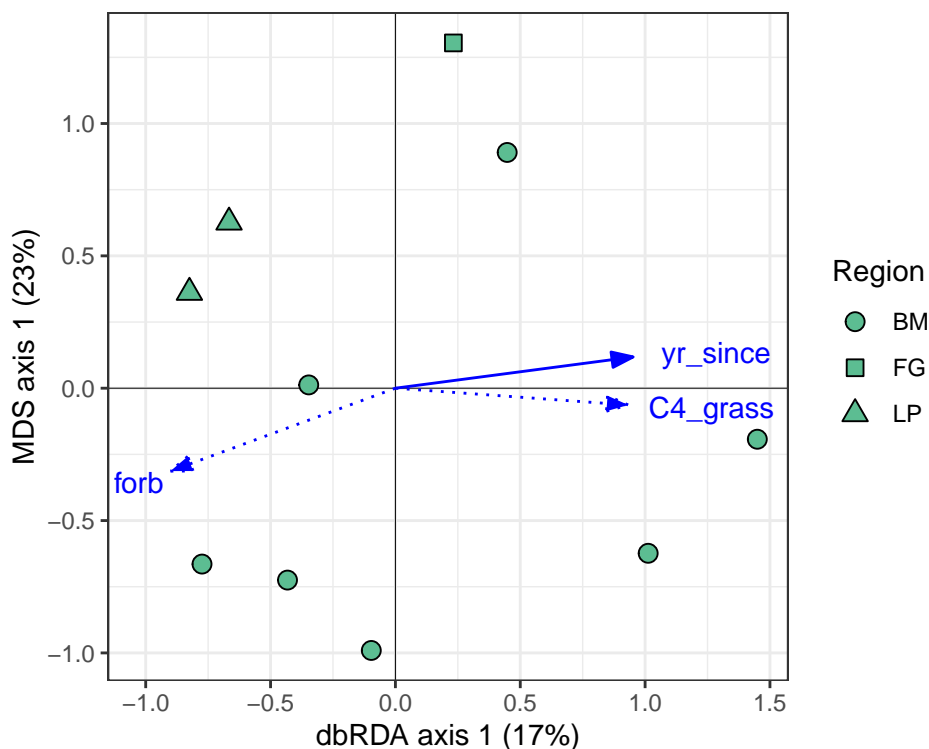### 4.6.1 Wisconsin regions, plant traits data

Figure 12: Results of a distance-based redundancy analysis of soil fungal community composition in restored fields in Wisconsin regions. Icons represent the constrained locations of fields. Solid blue arrows show the direction and strength of associations between communities and explanatory variables. Dotted blue arrows show correlations with strong, but non-significant explanatory variables. Axis 1 is significant based on 1999 permutations at p=0.011.

4.6.2 All regions, plant presence/absence data

4.6.3 Plants constraint...

4.7 Responses

---

[2]The switchgrass fields at Fermilab could not be included because no plant data exists for these fileds.
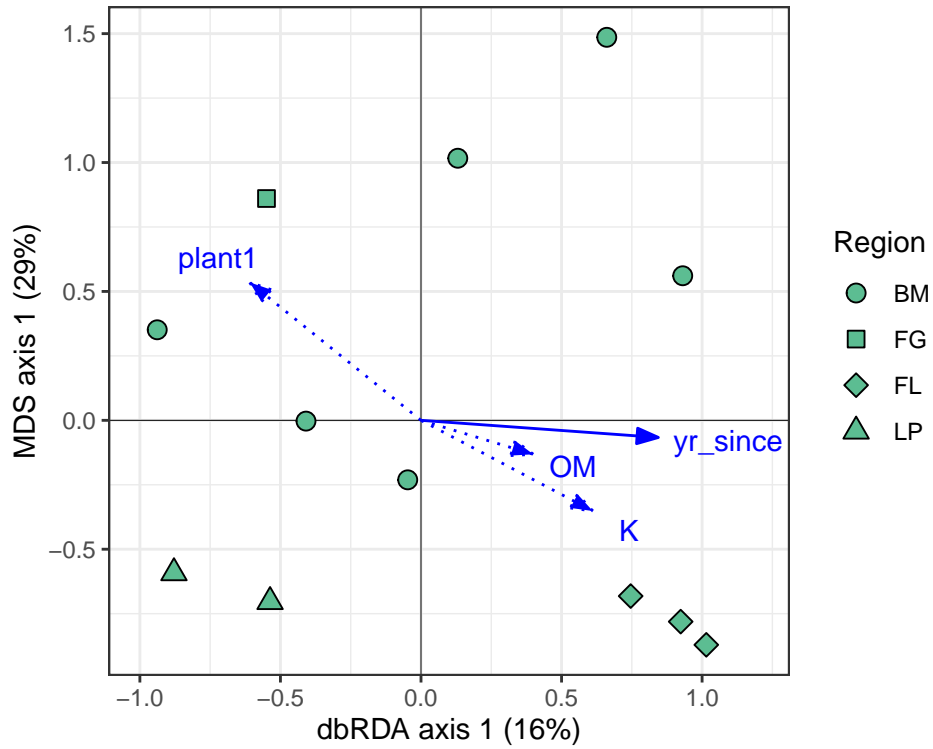
Figure 13: Results of a distance-based redundancy analysis of soil fungal community composition in restored fields from all regions. Icons represent the constrained locations of fields. Solid blue arrows show the direction and strength of associations between communities and explanatory variables. Dotted blue arrows show correlations with strong, but non-significant explanatory variables. Axis 1 is significant based on 1999 permutations at p=0.008.