Tallgrass Prairie Restoration Legacies, Summary

Beau Larkin

Last updated: 05 December, 2023

Contents

Packages and Libraries

```
packages_needed = c("tidyverse",
                    "png",
                    "knitr",
                    "conflicted",
                    "gridExtra")
packages_installed = packages_needed %in% rownames(installed.packages())
```

```
if (any(!packages_installed)) {
    install.packages(packages_needed[!packages_installed])
}
```

```
for (i in 1:length(packages_needed)) {
    library(packages_needed[i], character.only = T)
}
```

```
conflict_prefer("filter", "dplyr")
conflict_prefer("select", "dplyr")
```

Description

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

Data

For this summary, I'll pull as many objects as possible from existing files to reduce the number of interspersed code chunks. A few quick new analyses will be necessary, though. Data are loaded here.

```
sites <-
    read_csv(paste0(getwd(), "/clean_data/sites.csv"), show_col_types = FALSE) %>%
    mutate(
        field_type = factor(
            field_type,
            ordered = TRUE,
            levels = c("corn", "restored", "remnant")),
        yr_since = replace(yr_since, which(field_type %in% c("remnant", "corn")), NA)) %>%
    select(-lat, -long, -yr_restore, -yr_rank) %>%
    arrange(field_key)
```

Methods

Sites and design

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" so far. We collected samples and data from four regions, shown on the map below.

The design is unbalanced because there are more restored fields than corn or remnant. 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.

```
kable(table(sites$region, sites$field_type),
      format = "pandoc",
      caption = "Count of fields by type and region:\nBM = Blue Mounds, FG = Faville Grove,\nFL = Fermi
```

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

	corn	restored	remnant
LP	1	2	1

The rules for establishing a chronosequence are strict. We cannot call fields from all regions a chronosequence. With seven restored fields in a small geographic area, the Blue Mounds fields are our best bet for this, but we will likely have to call this a "pseudochronosequence" and avoid some inferences. 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.

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 OTUs or SVs. In preliminary work, inferences made with SVs were weaker but not qualitatively different. I continued with OTUs only.

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, I had planned to use permutation tests in ordinations, and application function `how()` from package `permute` requires balance at the plot (i.e., field) level. Second, I wanted to summarize the subsamples at the field level because fields are our replicates. Any comparisons with field metadata, where we have one data point per field, would be pseudoreplicated when regressed against subsample level data.

Finally, after rarefaction it was clear that several samples contained very few sequences compared with the others. I used an iterative process to remove subsamples with few sequences, choosing a rarefication cutoff near the plateau of OTU recovery at depth (see figures). This ended up being 8 subsamples for ITS and 7 for 18S datasets. These subsample sequence values were then rarefied and used as-is, or, when reporting at the field level, subsample sequence values were summed and rarefied.

This process didn't remove many OTUs and doesn't change any major interpretation.

Even after this process, it's clear that fungal communities were undersampled.

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.

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.

Design/site data

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

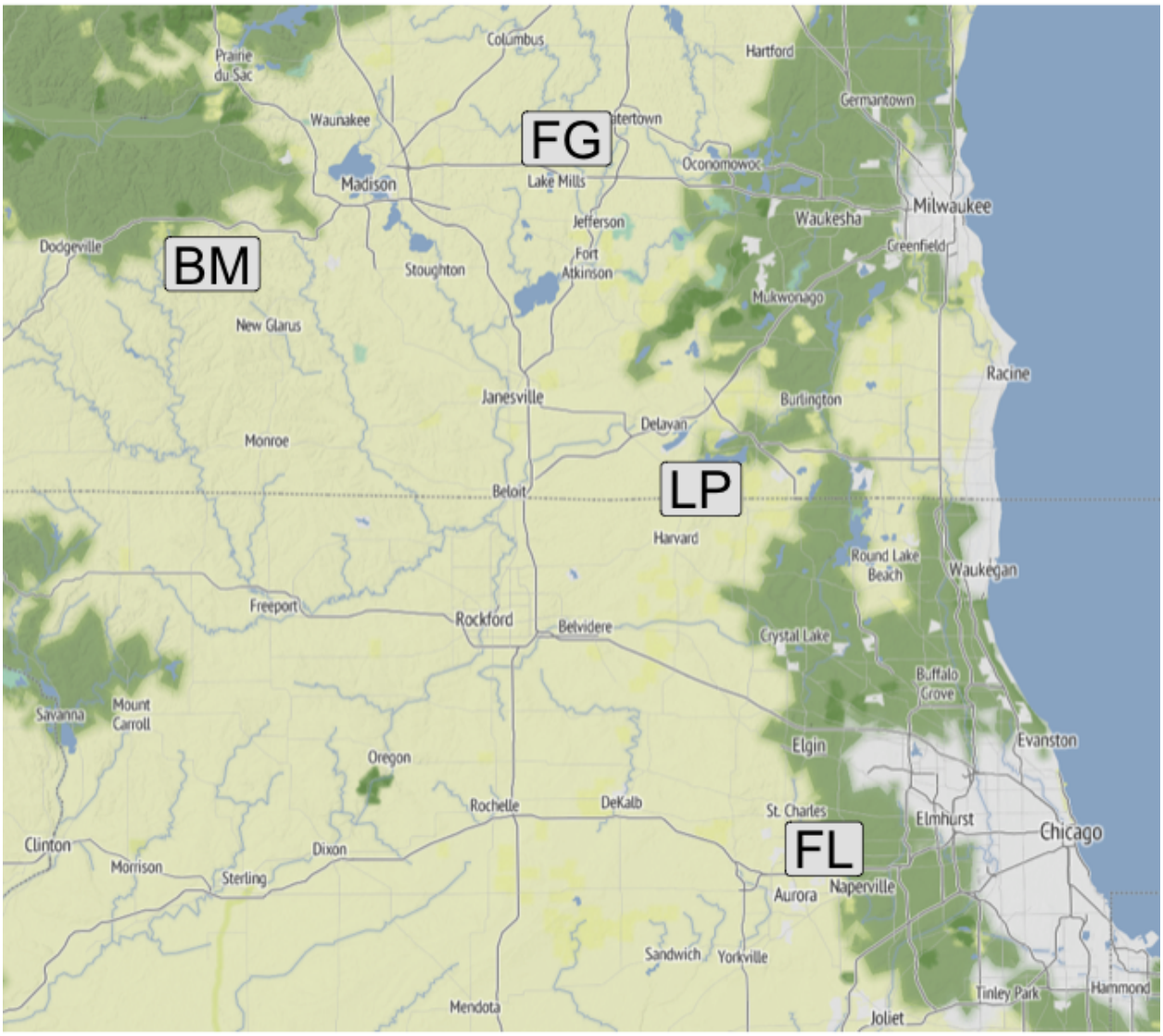

Figure 1: Labels show centroids of regions used for this work. BM = Blue Mounds, FG = Faville Grove, FL = Fermilab, LP = Lake Petite.

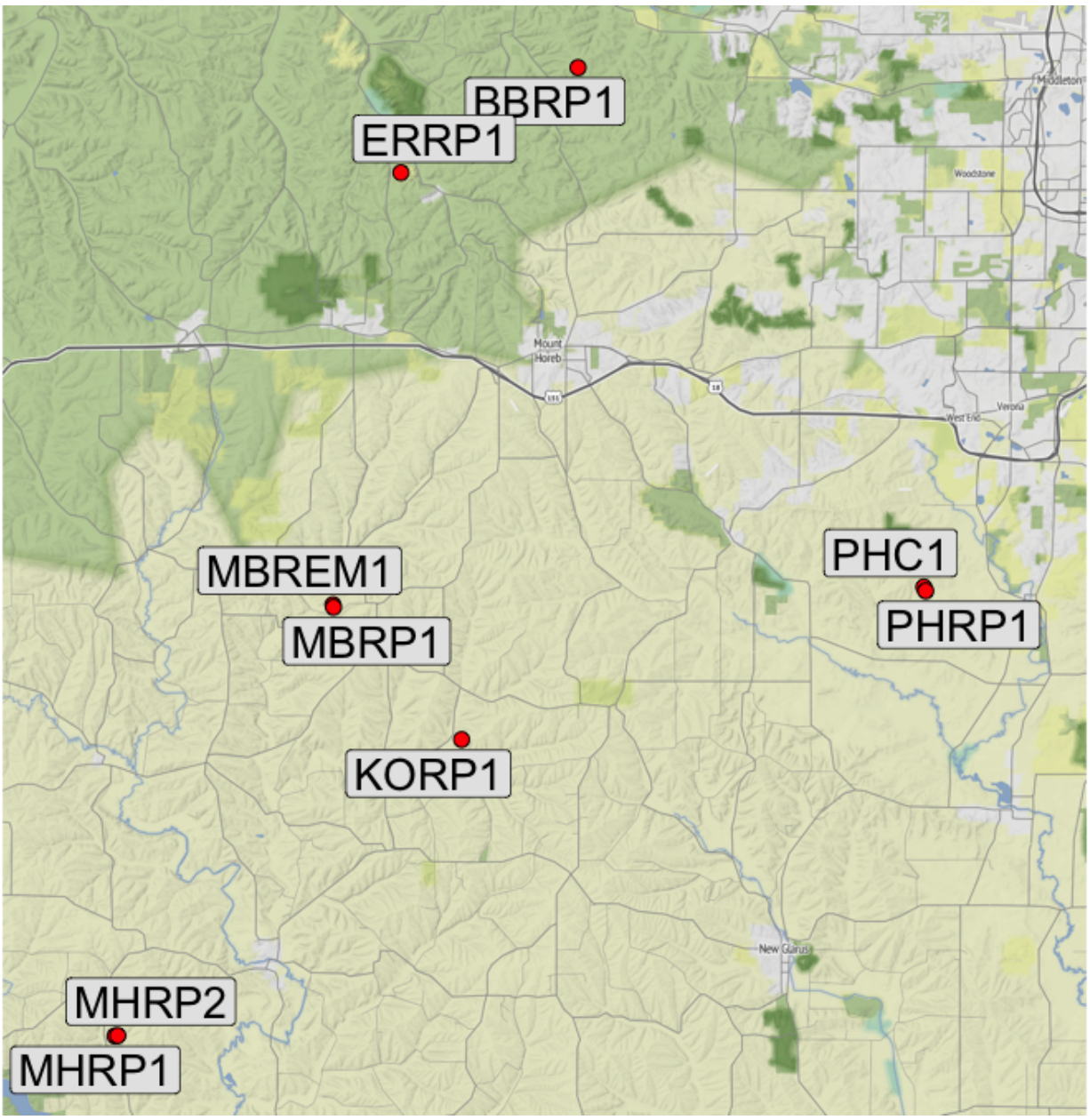

Figure 2: Labels show individual fields in the Blue Mounds region.

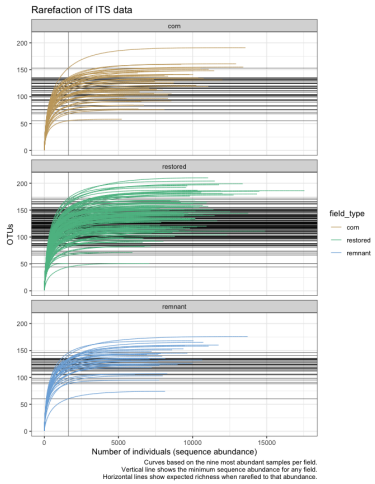

Figure 3: Pre

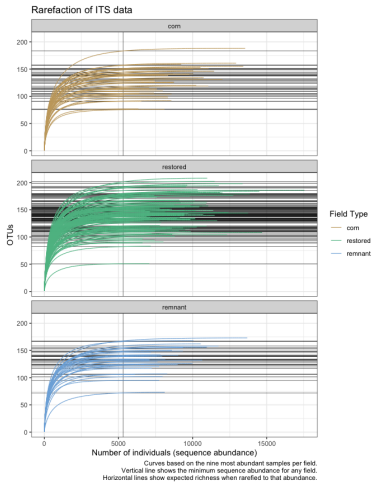

Figure 4: Post

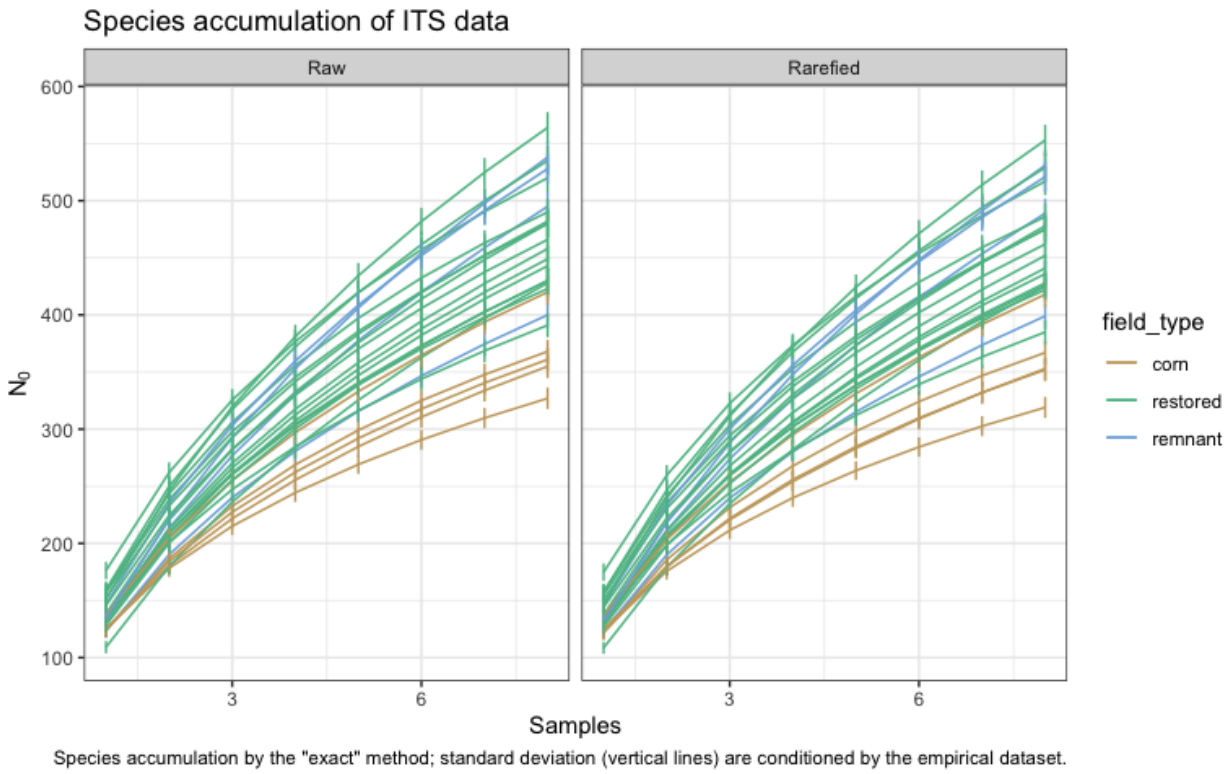

Figure 5: ITS

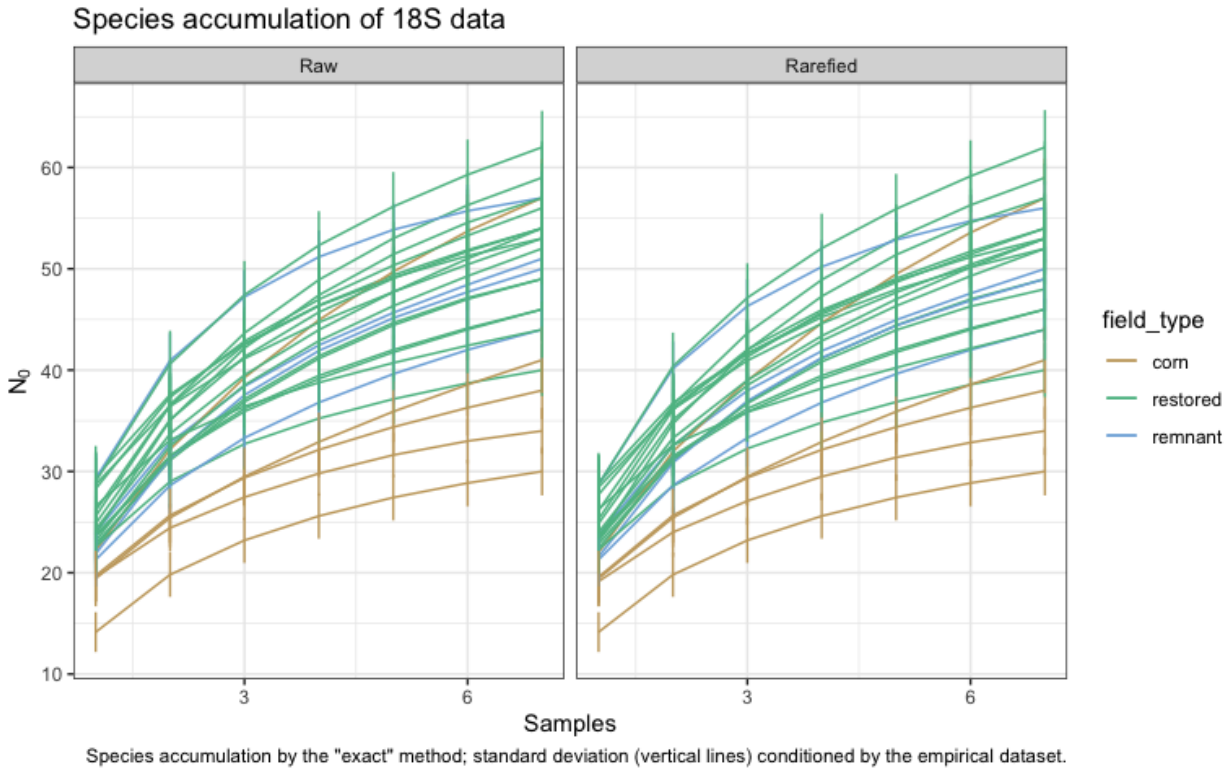

Figure 6: 18S

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 these), and one for quantification of microbial biomass. ## 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
- Soil properties
- Fungal biomass
- Water stable aggregates
- Site metadata and design

Results

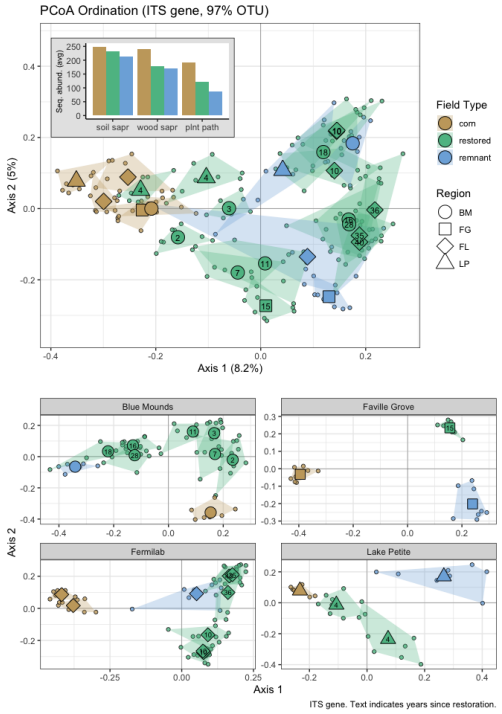

Figure 7: ITS-based fungal communities, most abundant guilds inset.

```
# kable(read_csv("microbial_communities_files/pcoa_its_eig.csv", show_col_types = FALSE), format = "pan
```

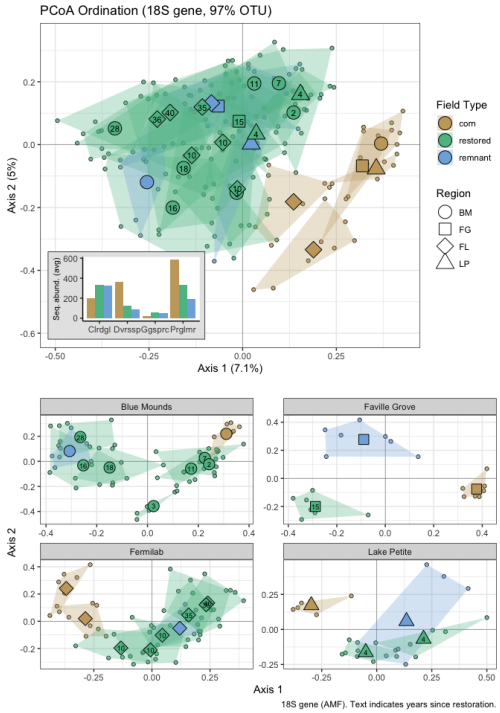

Figure 8: 18S-based fungal communities, most important families inset.