Tallgrass Prairie Restoration Legacies, Summary

Beau Larkin

Last updated: 04 December, 2023

Contents

Packages and Libraries

```
packages_needed = c("tidyverse",
                    "png",
                    "knitr",
                    "conflicted")
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")
```

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.

```r
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 == "remnant"), NA),
        yr_since = replace(yr_since, which(field_type == "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. BM = Blue Mounds, FG = Faville Grove, FL = Fermilab, LP = Lake Petite.

The data are 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.

```r
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
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.

Soil collection for fungal communities

We collected soil cores from 10 haphazardly-selected locations in each field. The entire extraction, fragment sequencing, and bioinformatics pipeline is beyond the scope of this summary. Key information here is that a few samples failed to amplify, leaving the sub-samples a little unbalanced.

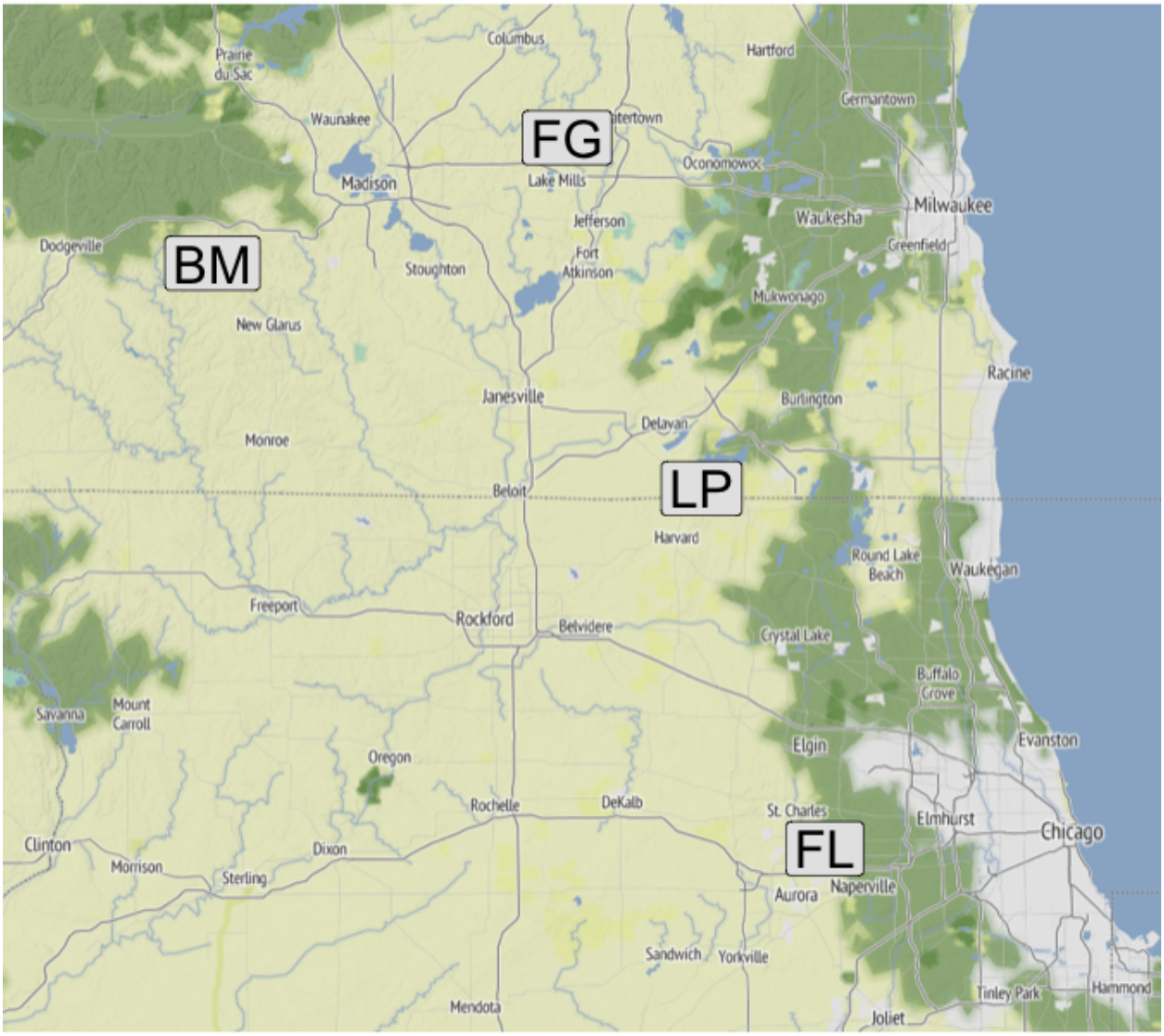

Figure 1: Labels show centroids of regions used for this work.

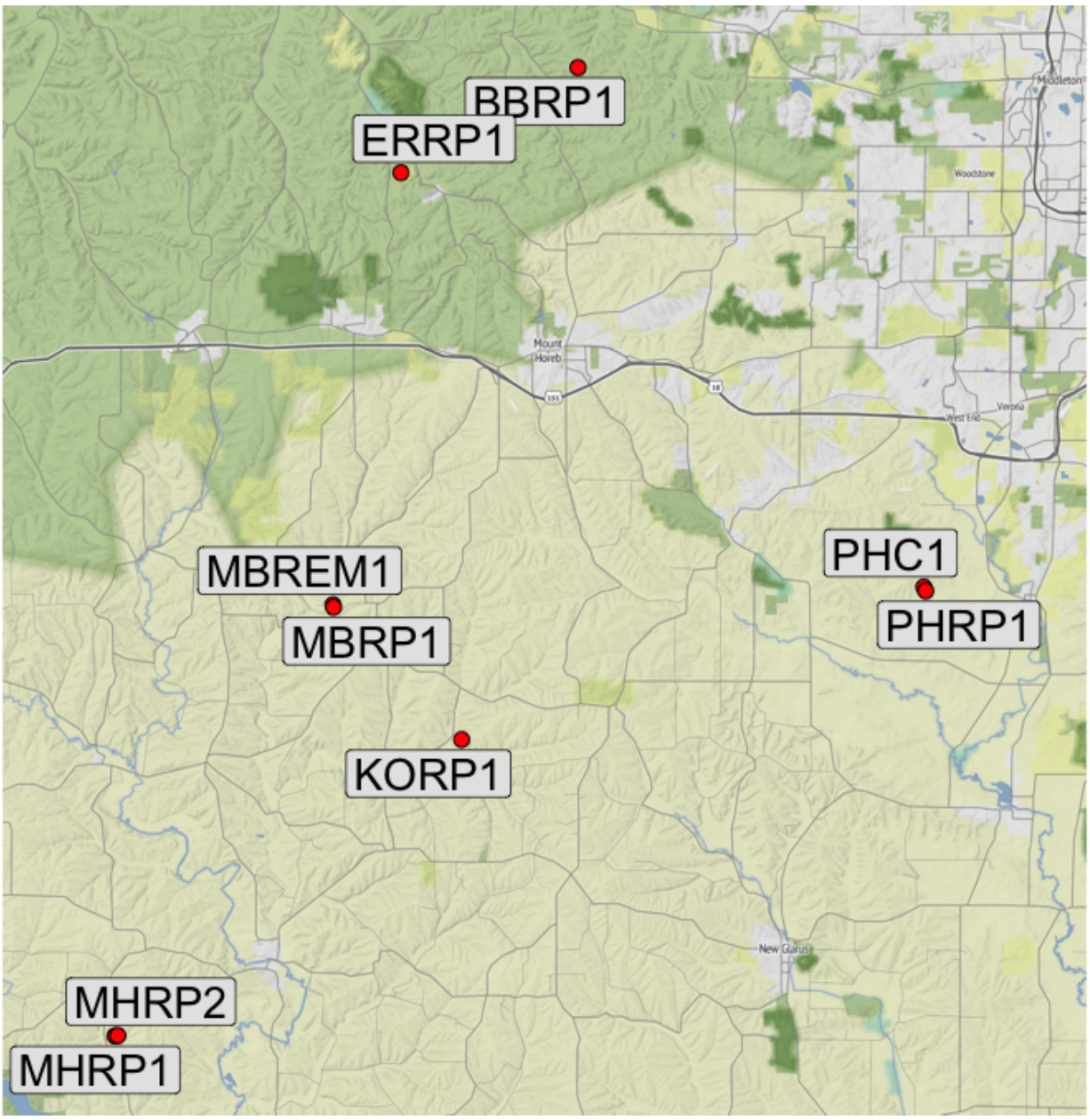

Figure 2: Labels show individual fields in the Blue Mounds region.

This unbalance 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 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.

Finally, after rarefaction it was clear that we had several samples that 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.

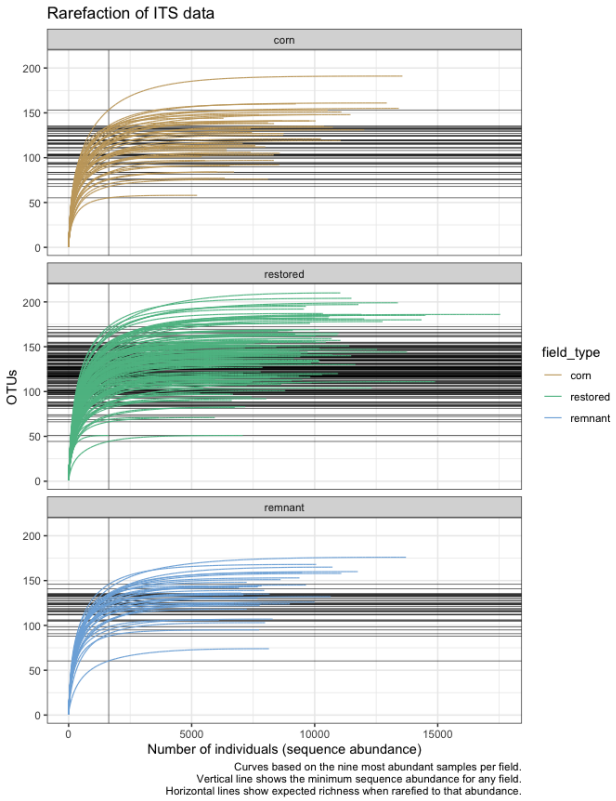

Figure 3: Pre

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

Data sources: Fungi plants metadata

for fungi, subsamples per field? 9, but 8 chosen. Show same for amf

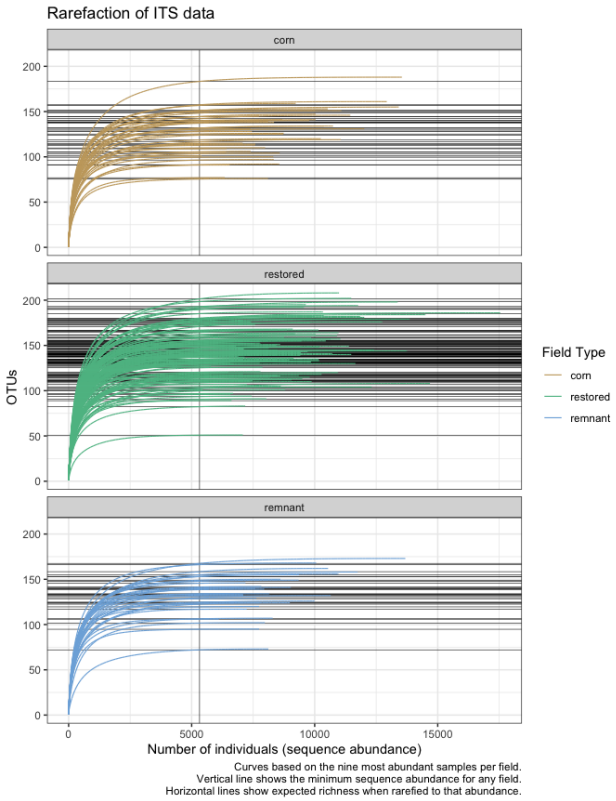

Figure 4: Post

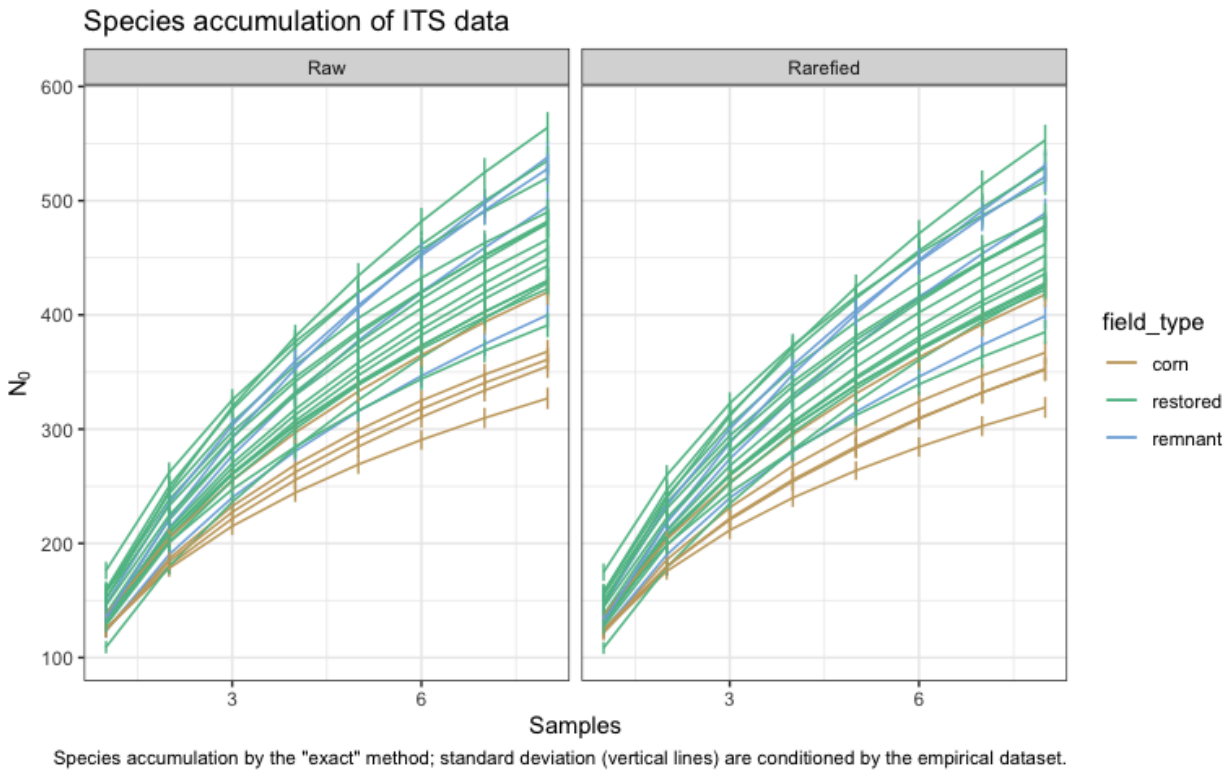

Figure 5: ITS

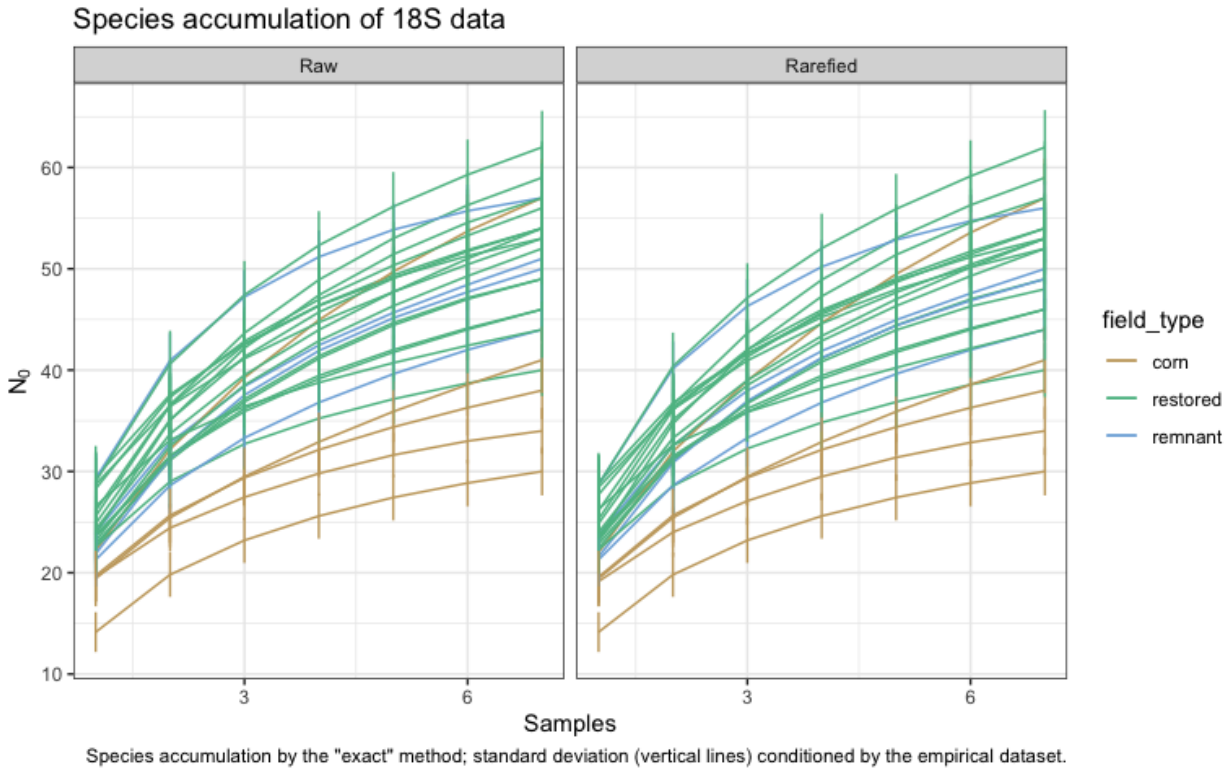

Figure 6: 18S