# Tallgrass Restoration Legacies, Summary

Beau Larkin

Last updated: 06 December, 2023

## Contents



## Packages

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

# Introduction

The tallgrass prairie biome in North America is nearly extinct, creating a tremendous desire for restoration or recreation of this critical habitat. In the upper Midwest, interested landowners have attempted to convert previous agricultural fields (corn or soybeans, usually) to prairie for decades. The consequences of this practice for soil biota are largely unknown. Agricultural fields 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 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?

# 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

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.

## Design

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.

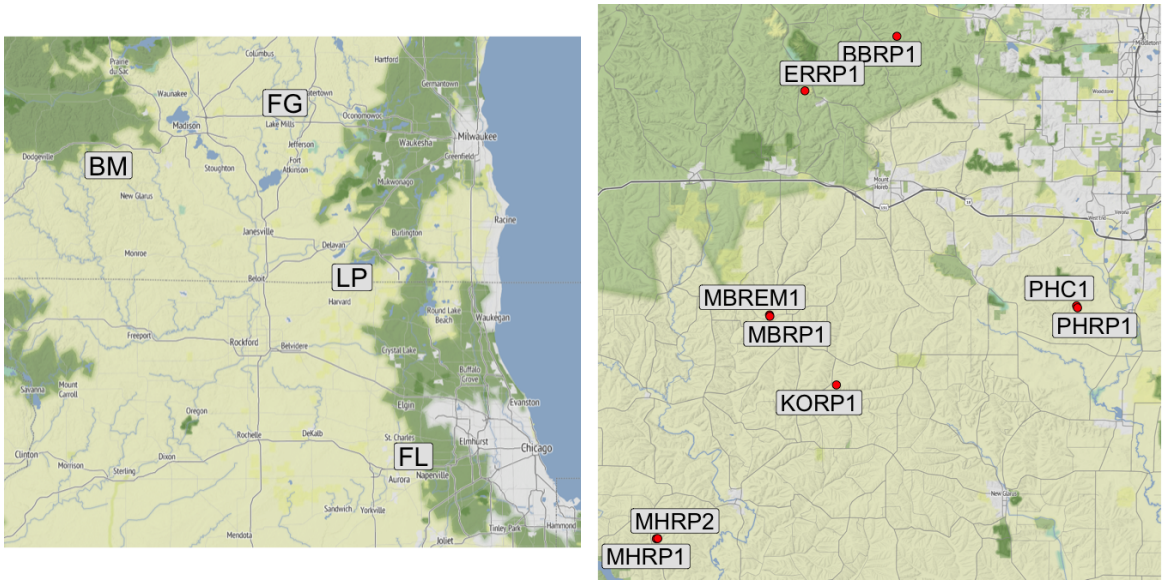


Figure 1: (Left) Labels show centroids of regions used for this work. BM = Blue Mounds, FG = Faville Grove, FL = Fermilab, LP = Lake Petite. (Right) Labels show individual fields in the Blue Mounds region.

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 with corn or remnant fields. 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 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

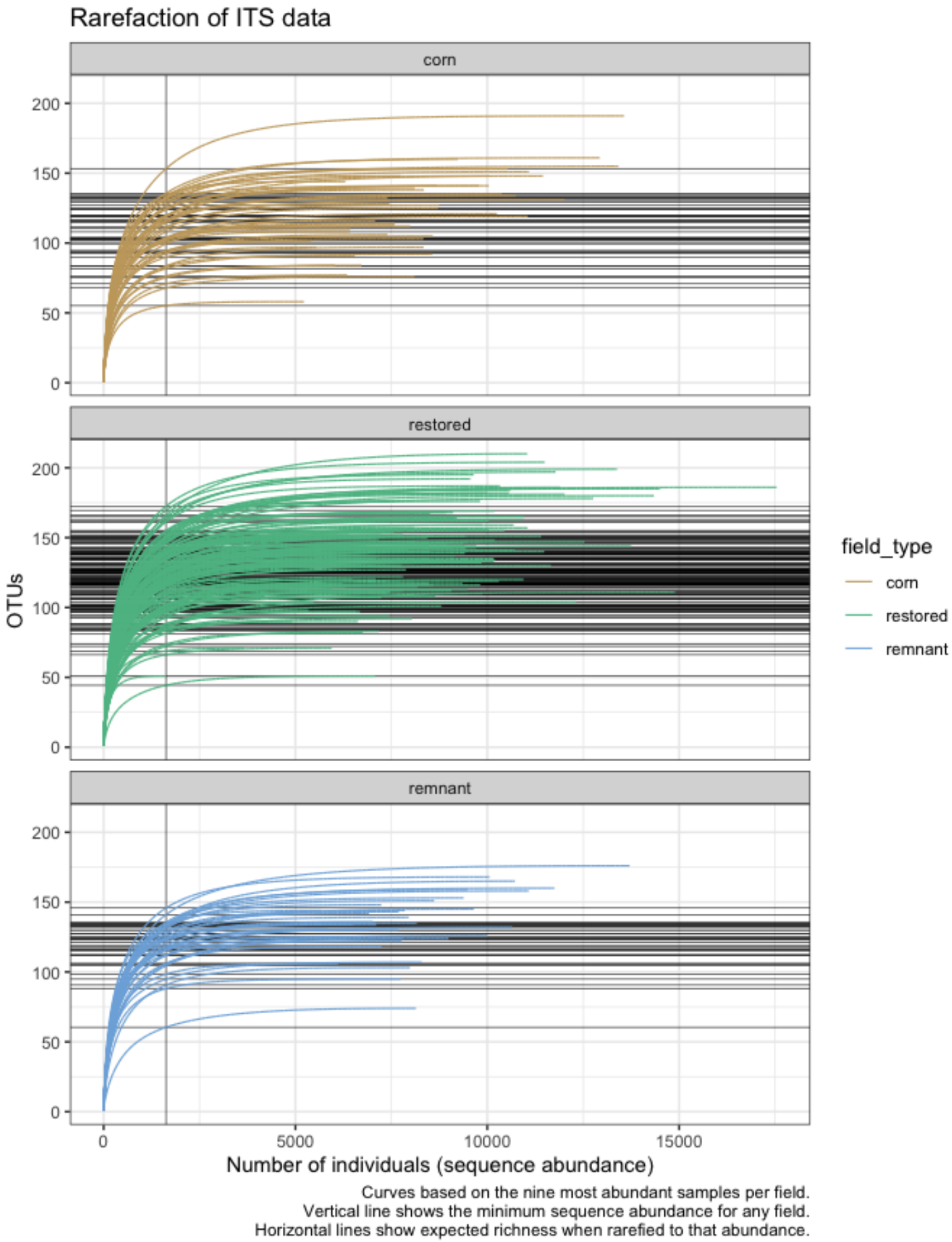
Rarefaction of ITS data
corn
restored
remnant
OTUs
0
50
100
150
200
0
5000
10000
15000
Number of individuals (sequence abundance)
field_type
corn
restored
remnant
Curves based on the nine most abundant samples per field.
Vertical line shows the minimum sequence abundance for any field.
Horizontal lines show expected richness when rarefied to that abundance.

Figure 2: Pre

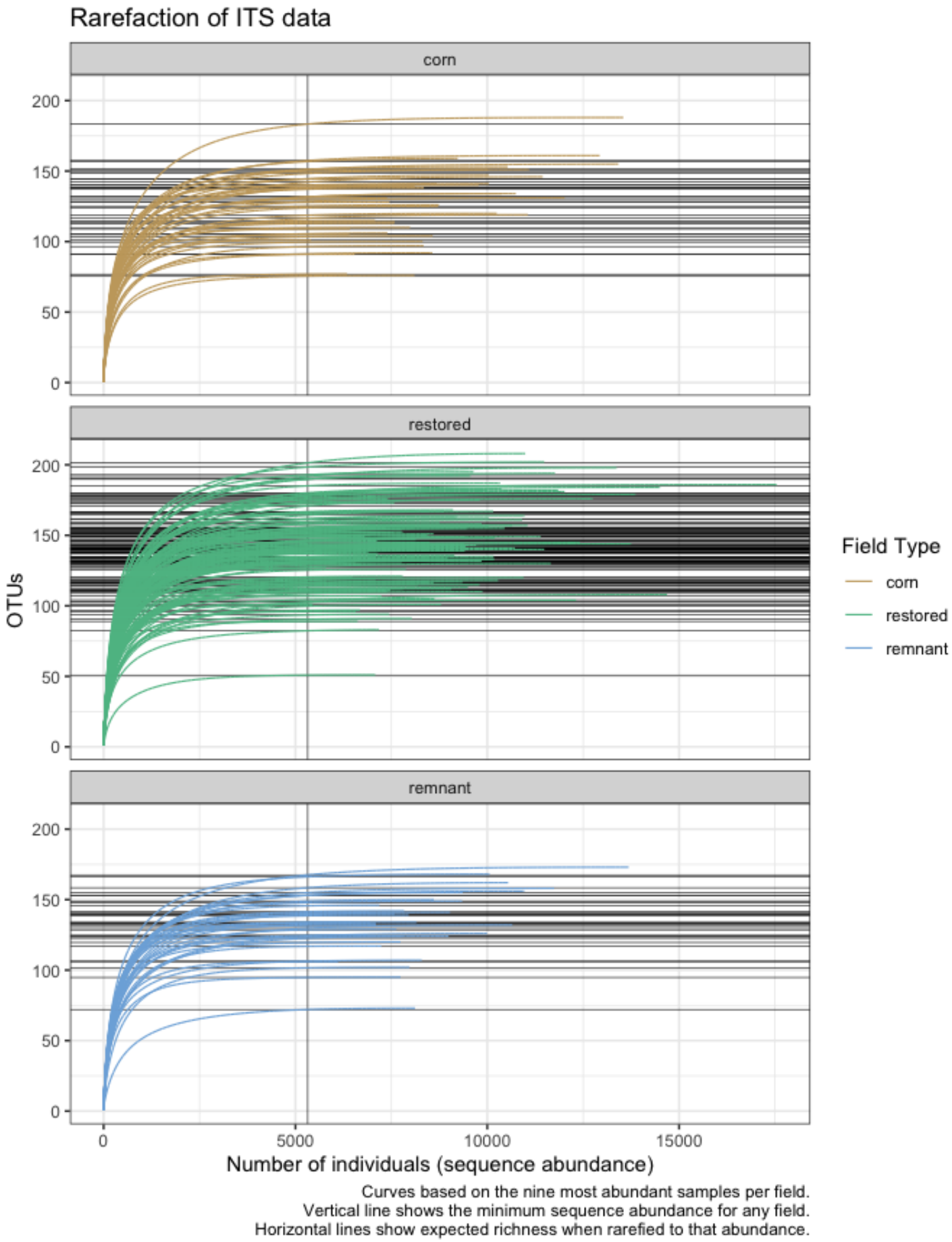


Figure 3: Post