

Appendices: R code

Appendices for “Dryland rock detention structures increase herbaceous vegetation cover and stabilize shrub cover over 10 years but do not directly affect soil fertility” by Ossanna et al. (2023). Included are code and analysis used to make figures and full model output. Data can be downloaded from the Zenodo archive under the DOI [10.5281/zenodo.8310363](https://doi.org/10.5281/zenodo.8310363).

Contact: Lia Ossanna, lossanna@arizona.edu

Last updated: 2023-11-28

Appendices:

- Appendix A: Temporal vegetation trends, Figs 3 & S3
- Appendix B: Coefficient of variation, Figs 4 & S54
- Appendix C: Soil fertility, Figs S5-S7
- Appendix D: Structural equation modeling, models 1-4

Package versions:

- R version 4.3.1
- tidyverse 2.0.0
 - dplyr 1.1.2
 - stringr 1.5.0
 - ggplot 3.4.3
 - tidyr 3.2.1
- agricolae 1.3-6
- plotrix 3.8-2
- ggpubr 0.6.0
- rstatix 0.7.2
- scales 1.2.1
- metagenomeSeq 1.42.0, Biobase 2.60.0, BiocGenerics 0.46.0
- vegan 2.6-4
- dada2 1.28.0, ShortRead 1.58.0, Biostrings 2.68.1 used for DADA2 pipeline (code not included here)
- lavaan 0.6-16

Appendix A: Temporal vegetation trends

Setup

```
library(tidyverse)
library(agricolae)
library(plotrix)
library(ggpubr)
library(rstatix)

# Load data -----

notree.all <- read.csv("Herb-and-shrub-cover_2012-2021.csv")
herb.all <- read.csv("Herb-cover_2012-2021.csv")
shrub.all <- read.csv("Shrub-cover_2012-2021.csv")
invasive.all <- read.csv("Invasive-cover_2012-2021.csv")
plant.all <- read.csv("Species-cover_2012-2021.csv")
per.div <- read.csv("Perennial-plant-diversity_2012-2021.csv")

# Functions -----

# Convert columns to factor or date as needed
convert.cols <- function(x) {
  x$year.xaxis <- as.Date(x$year.xaxis)

  group.cols <- c("Sample", "Year", "Treatment")

  x[group.cols] <- lapply(x[group.cols], factor)

  return(x)
}

# Data wrangling -----

notree.all <- convert.cols(notree.all)
herb.all <- convert.cols(herb.all)
shrub.all <- convert.cols(shrub.all)
invasive.all <- convert.cols(invasive.all)
per.div <- convert.cols(per.div)
```

Fig 3a: Grass, forb & shrub cover

```
# Grass, forb, and shrub cover (notree) -----

# Find averages by year
notree.avg <- notree.all %>%
  group_by(Treatment, Year, year.xaxis) %>%
  summarise(mean = mean(Cover),
            SD = sd(Cover),
```

```

SE = std.error(Cover),
.groups = "keep")

# One-way ANOVA for Control
summary(aov(Cover ~ Year, data = filter(notree.all, Treatment == "Control"))) # p = 4.3e-06

```

```

##              Df Sum Sq Mean Sq F value    Pr(>F)
## Year           5   7815   1563.0     7.12 4.3e-06 ***
## Residuals    177   38854    219.5
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

notree.ctrl <- notree.all |>
  filter(Treatment == "Control")
anova.notree.ctrl <- aov(notree.ctrl$Cover ~ notree.ctrl$Year)
hsd.notree.ctrl <- HSD.test(anova.notree.ctrl, trt = "notree.ctrl$Year")
hsd.notree.ctrl$groups

```

```

##      notree.ctrl$Cover groups
## 2021         42.41935      a
## 2012         31.73194     ab
## 2014         31.58750     ab
## 2018         27.71774      b
## 2013         27.71250      b
## 2015         20.89315      b

```

```

# One-way ANOVA for Treated
summary(aov(Cover ~ Year, data = filter(notree.all, Treatment == "Treated"))) # p = 0.00304

```

```

##              Df Sum Sq Mean Sq F value    Pr(>F)
## Year           5   7804   1560.7     3.739 0.00304 **
## Residuals    178   74291    417.4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

notree.trt <- notree.all |>
  filter(Treatment == "Treated")
anova.notree.trt <- aov(notree.trt$Cover ~ notree.trt$Year)
hsd.notree.trt <- HSD.test(anova.notree.trt, trt = "notree.trt$Year")
hsd.notree.trt$groups

```

```

##      notree.trt$Cover groups
## 2018         41.69960      a
## 2021         37.47581     ab
## 2014         31.40927    abc
## 2015         30.33669    abc
## 2012         26.72446     bc
## 2013         21.80029      c

```

```

# Plot with one-way ANOVA letters
notree.ctrl.letters <- hsd.notree.ctrl$groups
notree.ctrl.letters <- notree.ctrl.letters |>
  mutate(Year = rownames(notree.ctrl.letters)) |>
  arrange(Year)
notree.trt.letters <- hsd.notree.trt$groups
notree.trt.letters <- notree.trt.letters |>
  mutate(Year = rownames(notree.trt.letters)) |>
  arrange(Year)

letters.notree2t <- data.frame(x = notree.avg$year.xaxis[1:6],
                              y = c(23, 18, 27, 26, 36.5, 32.5),
                              label = notree.trt.letters$groups)
letters.notree2c <- data.frame(x = notree.avg$year.xaxis[1:6],
                              y = c(35.5, 31.5, 36, 24, 31.5, 47),
                              label = notree.ctrl.letters$groups)

ptext.notree2 <- data.frame(x = as.Date("2019-09-01"),
                            y = 22,
                            label = "ANOVA, p < 0.01")
notree.plot2 <- ggplot(notree.avg, aes(x = year.xaxis, y = mean,
                                       color = Treatment)) +

  geom_line() +
  geom_point() +
  geom_pointrange(aes(ymin = mean - SE, ymax = mean + SE)) +
  xlab(NULL) +
  ylab("Cover (%)") +
  ggtitle("Vegetation cover, 2012-2021") +
  scale_color_manual(values = c("red", "#1F78B4")) +
  theme_bw() +
  theme(legend.position = "bottom") +
  geom_text(data = letters.notree2t,
            mapping = aes(x = x, y = y, label = label),
            color = "#1F78B4",
            size = 3) +
  geom_text(data = letters.notree2c,
            mapping = aes(x = x, y = y, label = label),
            color = "red",
            size = 3,
            fontface = "italic") +
  geom_text(data = ptext.notree2,
            aes(x = x, y = y, label = label),
            color = "gray30",
            size = 2.5) +
  theme(axis.text.x = element_text(color = "black")) +
  theme(plot.margin = margin(t = 0.1, r = 0.1, b = 0.3, l = 0.1, "in")) +
  theme(legend.position = "none")
notree.plot2

```

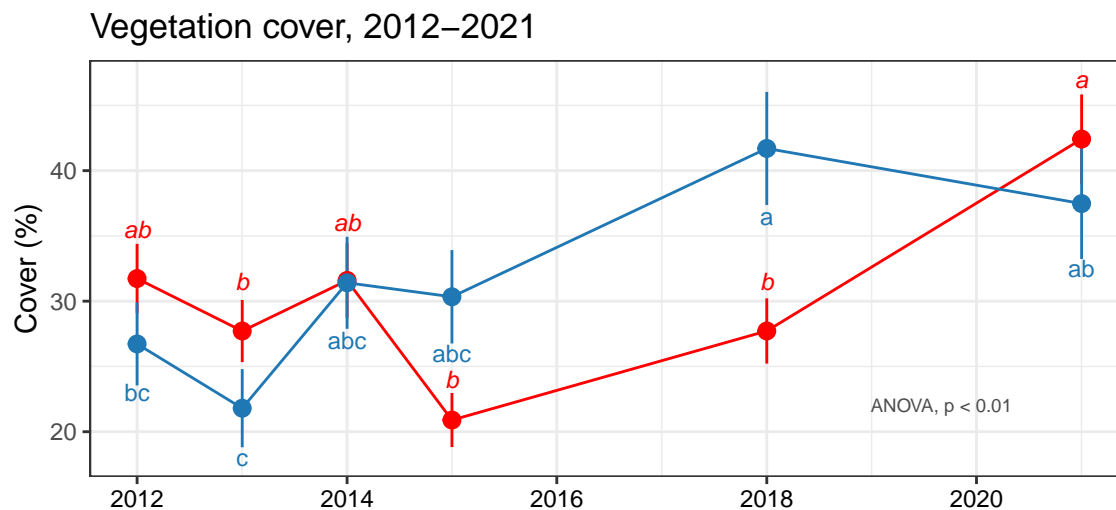


Fig 3b: Herbaceous (grass & forb) cover

```
# Herbaceous cover -----

# Find averages by year
herb.avg <- herb.all %>%
  group_by(Treatment, Year, year.xaxis) %>%
  summarise(mean = mean(Cover),
            SD = sd(Cover),
            SE = std.error(Cover),
            .groups = "keep")

# One-way ANOVA for Control
summary(aov(Cover ~ Year, data = filter(herb.all, Treatment == "Control"))) # 0.00434

##           Df Sum Sq Mean Sq F value Pr(>F)
## Year         5   2821    564.3   3.556 0.00434 **
## Residuals  177  28083    158.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

herb.ctrl <- herb.all |>
  filter(Treatment == "Control")
anova.herb.ctrl <- aov(herb.ctrl$Cover ~ herb.ctrl$Year)
hsd.herb.ctrl <- HSD.test(anova.herb.ctrl, trt = "herb.ctrl$Year")
hsd.herb.ctrl$groups

##      herb.ctrl$Cover groups
## 2021      26.78629      a
## 2014      22.28333     ab
## 2012      20.03472     ab
## 2018      19.89718     ab
```

```
## 2013      17.41528      b
## 2015      14.21169      b
```

```
# One-way ANOVA for Treated
```

```
summary(aov(Cover ~ Year, data = filter(herb.all, Treatment == "Treated"))) # p = 3.77e-10
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Year         5    6847   1369.5    12.17 3.77e-10 ***
## Residuals   178   20035    112.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
herb.trt <- herb.all |>
  filter(Treatment == "Treated")
anova.herb.trt <- aov(herb.trt$Cover ~ herb.trt$Year)
hsd.herb.trt <- HSD.test(anova.herb.trt, trt = "herb.trt$Year")
hsd.herb.trt$groups
```

```
## herb.trt$Cover groups
## 2018      24.489919      a
## 2021      22.201613      ab
## 2014      15.139113      bc
## 2015      12.368952      cd
## 2012      11.436828      cd
## 2013       6.929598      d
```

```
# Plot with one-way ANOVA letters
```

```
herb.ctrl.letters <- hsd.herb.ctrl$groups
herb.ctrl.letters <- herb.ctrl.letters |>
  mutate(Year = rownames(herb.ctrl.letters)) |>
  arrange(Year)
herb.trt.letters <- hsd.herb.trt$groups
herb.trt.letters <- herb.trt.letters |>
  mutate(Year = rownames(herb.trt.letters)) |>
  arrange(Year)

letters.herb3t <- data.frame(x = herb.avg$year.xaxis[1:6],
  y = c(9, 5.5, 12.5, 10, 28.5, 18.3),
  label = herb.trt.letters$groups)
letters.herb3c <- data.frame(x = herb.avg$year.xaxis[1:6],
  y = c(23.5, 20.5, 26.5, 17, 17, 31),
  label = herb.ctrl.letters$groups)

ptext.herb3 <- data.frame(x = as.Date("2019-09-01"),
  y = 8,
  label = "ANOVA, p < 0.001")
herb.plot3 <- ggplot(herb.avg, aes(x = year.xaxis, y = mean,
  color = Treatment)) +
  geom_line() +
  geom_point() +
  geom_pointrange(aes(ymin = mean - SE, ymax = mean + SE)) +
  xlab(NULL) +
```

```

ylab("Cover (%)") +
ggtitle("Herbaceous cover") +
scale_color_manual(values = c("red", "#1F78B4")) +
theme_bw() +
theme(legend.position = "bottom") +
geom_text(data = letters.herb3t,
          mapping = aes(x = x, y = y, label = label),
          color = "#1F78B4",
          size = 3) +
geom_text(data = letters.herb3c,
          mapping = aes(x = x, y = y, label = label),
          color = "red",
          size = 3,
          fontface = "italic") +
geom_text(data = ptext.herb3,
          aes(x = x, y = y, label = label),
          color = "gray30",
          size = 2.5) +
theme(axis.text.x = element_text(color = "black")) +
theme(plot.margin = margin(t = 0.1, r = 0.1, b = 0.3, l = 0.1, "in")) +
theme(legend.position = "none")
herb.plot3

```

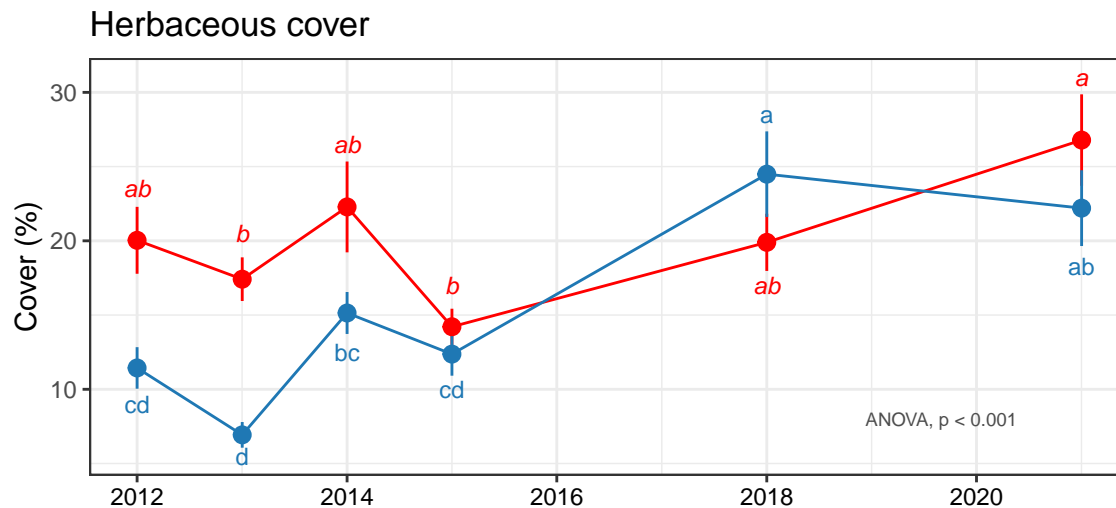


Fig 3c: Shrub cover

```

# Shrub cover -----

# Find averages by year
shrub.avg <- shrub.all %>%
  group_by(Treatment, Year, year.xaxis) %>%
  summarise(mean = mean(Cover),
            SD = sd(Cover),

```

```

      SE = std.error(Cover),
      .groups = "keep")

# One-way ANOVA for Control
summary(aov(Cover ~ Year, data = filter(shrub.all, Treatment == "Control"))) # p = 0.0112

##              Df Sum Sq Mean Sq F value Pr(>F)
## Year          5    1566    313.1   3.062 0.0112 *
## Residuals    177   18100    102.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

shrub.ctrl <- shrub.all |>
  filter(Treatment == "Control")
anova.shrub.ctrl <- aov(shrub.ctrl$Cover ~ shrub.ctrl$Year)
hsd.shrub.ctrl <- HSD.test(anova.shrub.ctrl, trt = "shrub.ctrl$Year")
hsd.shrub.ctrl$groups

##      shrub.ctrl$Cover groups
## 2021      15.633065      a
## 2012      11.697222     ab
## 2013      10.297222     ab
## 2014       9.304167     ab
## 2018       7.820565      b
## 2015       6.681452      b

# One-way ANOVA for Treated
summary(aov(Cover ~ Year, data = filter(shrub.all, Treatment == "Treated"))) # NS, p = 0.982

##              Df Sum Sq Mean Sq F value Pr(>F)
## Year          5     232     46.4   0.141 0.982
## Residuals    178   58465    328.5

# Plot with one-way ANOVA letters
shrub.ctrl.letters <- hsd.shrub.ctrl$groups
shrub.ctrl.letters <- shrub.ctrl.letters |>
  mutate(Year = rownames(shrub.ctrl.letters)) |>
  arrange(Year)

letters.shrub2 <- data.frame(x = shrub.avg$year.xaxis[1:6],
  y = c(9, 8, 7.4, 4.2, 5.5, 11.8),
  label = shrub.ctrl.letters$groups)

ptext.shrub2 <- data.frame(x = as.Date("2019-09-01"),
  y = 6,
  label = "ANOVA, p = 0.982 \nANOVA, p = 0.011")
shrub.plot2 <- ggplot(shrub.avg, aes(x = year.xaxis, y = mean,
  color = Treatment)) +
  geom_line() +
  geom_point() +
  geom_pointrange(aes(ymin = mean - SE, ymax = mean + SE)) +

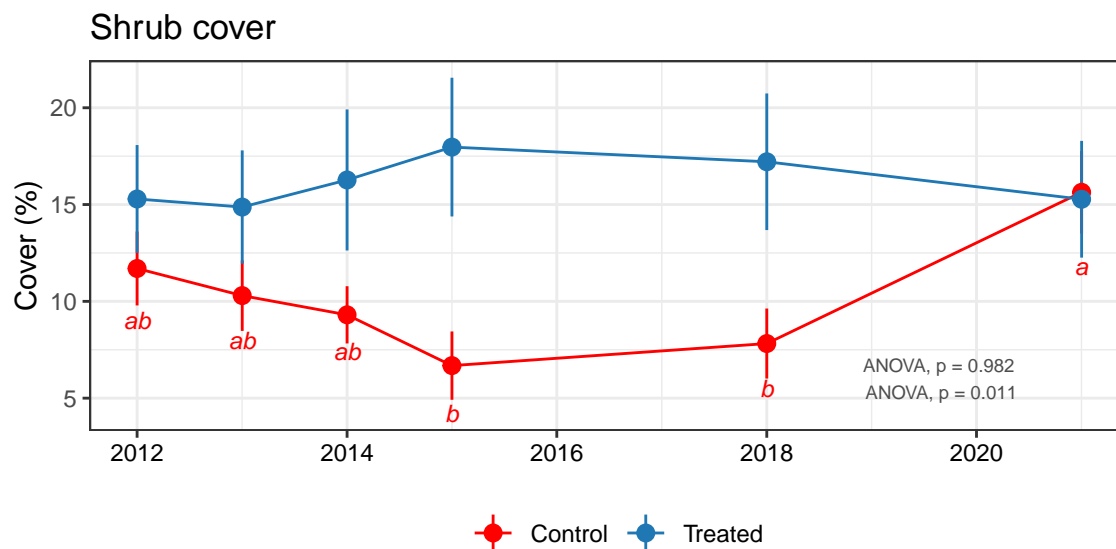
```



```

xlab(NULL) +
ylab("Cover (%)") +
ggtitle("Shrub cover") +
scale_color_manual(values = c("red", "#1F78B4")) +
theme_bw() +
theme(legend.position = "bottom") +
geom_text(data = letters.shrub2,
          mapping = aes(x = x, y = y, label = label),
          color = "red",
          size = 3,
          fontface = "italic") +
geom_text(data = ptext.shrub2,
          aes(x = x, y = y, label = label),
          color = "gray30",
          size = 2.5) +
theme(axis.text.x = element_text(color = "black")) +
theme(plot.margin = margin(t = 0.1, r = 0.1, b = 0, l = 0.1, "in")) +
theme(legend.title = element_blank())
shrub.plot2

```



Combine plots for Fig 3

```

# Combine notree, herb & shrub -----

tiff("figures/2023-12_publish-figures/Fig3_temporal_notree-herb-shrub.tiff", units = "in", height = 8.5)
ggarrange(notree.plot2, herb.plot3, shrub.plot2,
          ncol = 1, nrow = 3,
          labels = c("(A)", "(B)", "(C)"))

dev.off()

```

Supp Fig 3a: Perennial plant species richness

```
# Perennial plant richness -----

# Find averages by year
rich.avg <- per.div %>%
  group_by(Treatment, Year, year.xaxis) %>%
  summarise(mean = mean(rich),
            SD = sd(rich),
            SE = std.error(rich),
            .groups = "keep")

# One-way ANOVA for Control
summary(aov(rich ~ Year, data = filter(per.div, Treatment == "Control"))) # 0.00881

##              Df Sum Sq Mean Sq F value Pr(>F)
## Year           5    92.9  18.587    3.189 0.00881 **
## Residuals    177 1031.7    5.829
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

rich.ctrl <- per.div |>
  filter(Treatment == "Control")
anova.rich.ctrl <- aov(rich.ctrl$rich ~ rich.ctrl$Year)
hsd.rich.ctrl <- HSD.test(anova.rich.ctrl, trt = "rich.ctrl$Year")
hsd.rich.ctrl$groups

##      rich.ctrl$rich groups
## 2012      9.866667      a
## 2013      9.133333     ab
## 2018      8.645161     ab
## 2015      8.612903     ab
## 2014      8.200000     ab
## 2021      7.580645      b

# One-way ANOVA for Treated
summary(aov(rich ~ Year, data = filter(per.div, Treatment == "Treated"))) # p = 0.0516

##              Df Sum Sq Mean Sq F value Pr(>F)
## Year           5     76  15.207    2.247 0.0516 .
## Residuals    178   1204    6.766
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Plot with one-way ANOVA letters
rich.ctrl.letters <- hsd.rich.ctrl$groups
rich.ctrl.letters <- rich.ctrl.letters |>
  mutate(Year = rownames(rich.ctrl.letters)) |>
  arrange(Year)

letters.rich <- data.frame(x = rich.avg$year.xaxis[1:6],
```

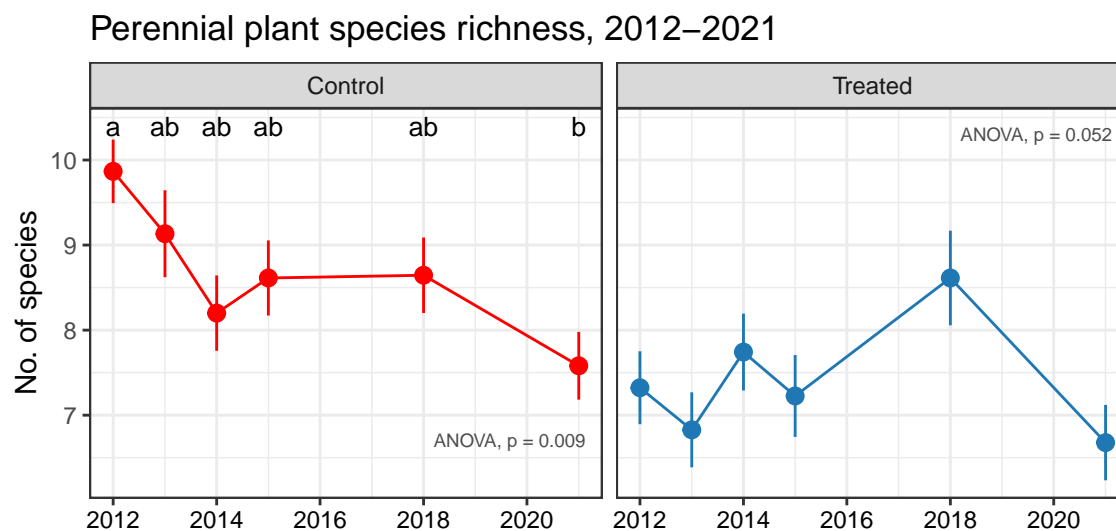
```

y = rep(10.4, 6),
label = rich.ctrl.letters$groups,
Treatment = rep("Control", 6))
ptext.rich <- data.frame(x = rep(as.Date("2019-09-01"), 2),
y = c(6.7, 10.3),
label = c("ANOVA, p = 0.009", "ANOVA, p = 0.052"),
Treatment = c("Control", "Treated"))

rich.plot <- ggplot(rich.avg, aes(x = year.xaxis, y = mean,
group = Treatment,
color = Treatment)) +

geom_line() +
geom_point() +
geom_pointrange(aes(ymin = mean - SE, ymax = mean + SE)) +
facet_wrap(~Treatment) +
xlab(NULL) +
ylab("No. of species") +
ggtitle("Perennial plant species richness, 2012-2021") +
scale_color_manual(values = c("red", "#1F78B4")) +
theme_bw() +
theme(legend.position = "none") +
geom_text(data = letters.rich,
mapping = aes(x = x, y = y, label = label),
color = "black",
size = 3.5) +
geom_text(data = ptext.rich,
aes(x = x, y = y, label = label),
color = "gray30",
size = 2.5) +
theme(axis.text.x = element_text(color = "black")) +
theme(plot.margin = margin(0.1, 0.1, 0.2, 0.1, "in"))
rich.plot

```



Supp Fig 3b: Perennial plant diversity

```
# Perennial plant diversity (Shannon) -----

# Find averages by year
shan.avg <- per.div %>%
  group_by(Treatment, Year, year.xaxis) %>%
  summarise(mean = mean(shan),
            SD = sd(shan),
            SE = std.error(shan),
            .groups = "keep")

# One-way ANOVA for Control
summary(aov(shan ~ Year, data = filter(per.div, Treatment == "Control"))) # p = 0.934

##              Df Sum Sq Mean Sq F value Pr(>F)
## Year          5   0.185  0.03702    0.26  0.934
## Residuals    177  25.197  0.14235

# One-way ANOVA for Treated
summary(aov(shan ~ Year, data = filter(per.div, Treatment == "Treated"))) # p = 0.725

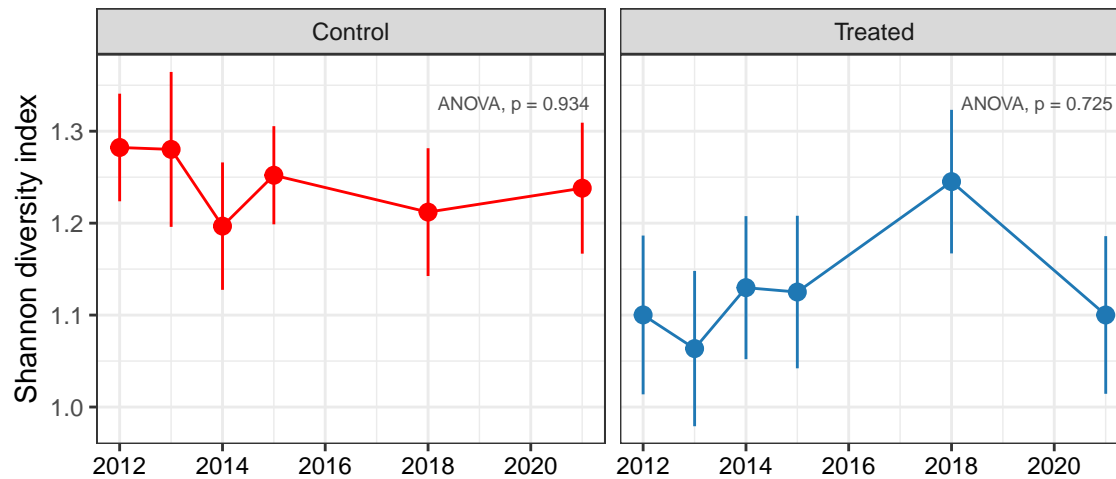
##              Df Sum Sq Mean Sq F value Pr(>F)
## Year          5    0.59  0.1189    0.568  0.725
## Residuals    178  37.27  0.2094

# Plot with one-way ANOVA
ptext.shan <- data.frame(x = rep(as.Date("2019-09-01"), 2),
                        y = c(1.33, 1.33),
                        label = c("ANOVA, p = 0.934", "ANOVA, p = 0.725"),
                        Treatment = c("Control", "Treated"))

shan.plot <- ggplot(shan.avg, aes(x = year.xaxis, y = mean,
                                group = Treatment,
                                color = Treatment)) +

  geom_line() +
  geom_point() +
  geom_pointrange(aes(ymin = mean - SE, ymax = mean + SE)) +
  facet_wrap(~Treatment) +
  xlab(NULL) +
  ylab("Shannon diversity index") +
  ggtitle("Perennial plant diversity") +
  scale_color_manual(values = c("red", "#1F78B4")) +
  theme_bw() +
  theme(legend.position = "none") +
  theme(axis.text.x = element_text(color = "black")) +
  geom_text(data = ptext.shan,
            aes(x = x, y = y, label = label),
            color = "gray30",
            size = 2.5) +
  theme(plot.margin = margin(0.1, 0.1, 0.2, 0.1, "in"))
shan.plot
```

Perennial plant diversity



Combine plots for Supp Fig 3

```
# Combine richness & Shannon -----

tiff("FigS3_temporal_richness-Shannon.tiff", units = "in", height = 5.5, width = 6, res = 300)
ggarrange( rich.plot, shan.plot,
  ncol = 1, nrow = 2,
  labels = c("(A)", "(B)"))

dev.off()
```

Invasive cover and most common species

```
# Average cover by year
invasive.all %>%
  select(-year.axis) |>
  group_by(Treatment, Year) %>%
  summarise(mean = mean(Cover),
    SE = std.error(Cover),
    .groups = "keep") |>
  mutate_if(is.numeric, round, digits = 2)

## 'mutate_if()' ignored the following grouping variables:
## * Columns 'Treatment', 'Year'

## # A tibble: 12 x 4
## # Groups:   Treatment, Year [12]
##   Treatment Year   mean   SE
##   <fct>      <fct> <dbl> <dbl>
## 1 Control   2012    3.34  0.6
```

```
## 2 Control 2013 2.58 0.66
## 3 Control 2014 2.46 0.98
## 4 Control 2015 2.96 0.69
## 5 Control 2018 5.1 1.17
## 6 Control 2021 7.38 1.23
## 7 Treated 2012 3.29 0.84
## 8 Treated 2013 2.18 0.51
## 9 Treated 2014 3.58 0.8
## 10 Treated 2015 5.1 1.08
## 11 Treated 2018 5.17 1.23
## 12 Treated 2021 4.79 0.85
```

```
# Most common species in Control
plant.all |>
  filter(Native == "Invasive",
         Treatment == "Control") |>
  group_by(Common) |>
  summarise(mean = mean(Cover)) |>
  arrange(desc(mean))
```

```
## # A tibble: 4 x 2
##   Common          mean
##   <chr>          <dbl>
## 1 Lehmann lovegrass 4.49
## 2 Stinkgrass       0.833
## 3 Spreading fantails 0.769
## 4 African lovegrass 0.125
```

```
# Most common species in Treated
plant.all |>
  filter(Native == "Invasive",
         Treatment == "Treated") |>
  group_by(Common) |>
  summarise(mean = mean(Cover)) |>
  arrange(desc(mean))
```

```
## # A tibble: 8 x 2
##   Common          mean
##   <chr>          <dbl>
## 1 Lehmann lovegrass 4.75
## 2 Buffelgrass      3.41
## 3 Boer lovegrass   0.625
## 4 African lovegrass 0.531
## 5 Spreading fantails 0.367
## 6 Barnyard         0.225
## 7 Rose Natal grass 0.125
## 8 Stinkgrass       0.125
```

Appendix B: Coefficient of variation

Setup

```
library(tidyverse)
library(car)
library(scales)
library(ggpubr)
```

```
# Load data -----
```

```
notree.all <- read_csv("Herb-and-shrub-cover_2012-2021.csv")
herb.all <- read_csv("Herb-cover_2012-2021.csv")
shrub.all <- read_csv("Shrub-cover_2012-2021.csv")
per.div <- read_csv("Perennial-plant-diversity_2012-2021.csv")
```

Fig 4a: CV of shrub cover

```
# Shrub cover -----
```

```
# Find CV for each sample over time
```

```
shrub.sample <- shrub.all |>
  group_by(Sample, Treatment) |>
  summarise(CV = sd(Cover) / mean(Cover),
            .groups = "keep") # NaNs produced because some have 0 cover and can't divide by 0
```

```
# Replace NaNs with 0
```

```
shrub.sample[1, 3] <- 0
shrub.sample[5, 3] <- 0
shrub.sample[8, 3] <- 0
```

```
# Compare means
```

```
wilcox.test(filter(shrub.sample, Treatment == "Treated")$CV,
             filter(shrub.sample, Treatment == "Control")$CV,
             exact = FALSE) # p = 0.01429
```

```
##
```

```
## Wilcoxon rank sum test with continuity correction
```

```
##
```

```
## data: filter(shrub.sample, Treatment == "Treated")$CV and filter(shrub.sample, Treatment == "Control")$CV
```

```
## W = 306, p-value = 0.01429
```

```
## alternative hypothesis: true location shift is not equal to 0
```

```
# Plot
```

```
letters.shrub <- data.frame(x = c(1, 2),
                           y = c(2.5, 2.5),
                           label = c("a", "b"))
```

```

shrub.plot.cv <- shrub.sample |>
  ggplot(aes(x = Treatment, y = CV)) +
  geom_boxplot(alpha = 0.3,
    outlier.shape = NA,
    aes(fill = Treatment)) +
  geom_jitter(size = 1,
    alpha = 0.8,
    aes(color = Treatment)) +
  scale_color_manual(values = c("red", "#1F78B4")) +
  scale_fill_manual(values = c("red", "#1F78B4")) +
  labs(title = "Shrub cover",
    x = NULL,
    y = "Coefficient of variation") +
  theme_bw() +
  theme(legend.position = "none") +
  scale_y_continuous(labels = percent) +
  theme(axis.text.x = element_text(color = "black")) +
  geom_text(aes(x = 0.95, y = 2.75, label = "Mann-Whitney, \np = 0.014"),
    color = "gray30",
    size = 2.5) +
  geom_text(data = letters.shrub,
    aes(x = x, y = y, label = label),
    color = "black") +
  theme(plot.margin = margin(0.1, 0, 0.1, 0.1, "in")) +
  stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
    width = 0.75, linetype = "dashed") +
  theme(plot.title = element_text(size = 12))
shrub.plot.cv

```

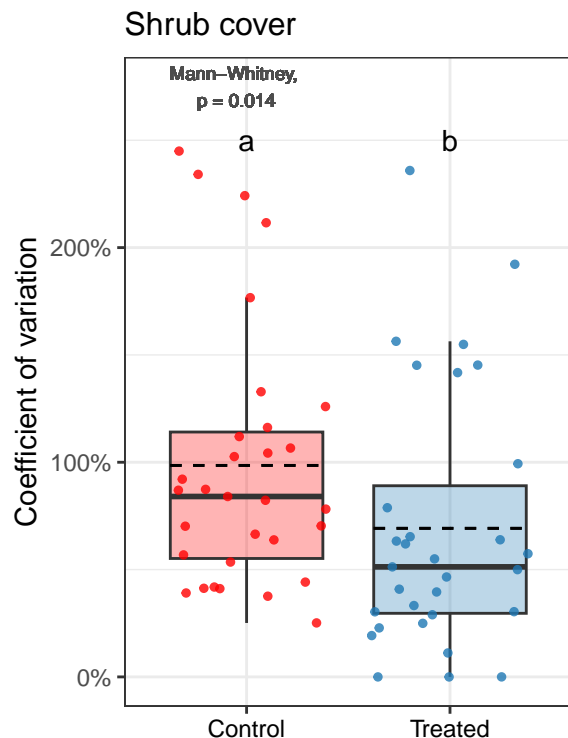


Fig 4b: CV of herbaceous cover

```
# Herb cover -----

# Find CV for each sample over time
herb.sample <- herb.all |>
  group_by(Sample, Treatment) |>
  summarise(CV = sd(Cover) / mean(Cover),
            .groups = "keep")

# Compare means
wilcox.test(filter(herb.sample, Treatment == "Treated")$CV,
            filter(herb.sample, Treatment == "Control")$CV) # NS, p = 0.148

##
## Wilcoxon rank sum exact test
##
## data: filter(herb.sample, Treatment == "Treated")$CV and filter(herb.sample, Treatment == "Control")$CV
## W = 584, p-value = 0.1479
## alternative hypothesis: true location shift is not equal to 0

# Plot
herb.plot.cv <- herb.sample |>
  ggplot(aes(x = Treatment, y = CV)) +
  geom_boxplot(alpha = 0.3,
              outlier.shape = NA,
              aes(fill = Treatment)) +
  geom_jitter(size = 1,
              alpha = 0.8,
              aes(color = Treatment)) +
  scale_color_manual(values = c("red", "#1F78B4")) +
  scale_fill_manual(values = c("red", "#1F78B4")) +
  labs(title = "Herbaceous cover",
       x = NULL,
       y = NULL) +
  theme_bw() +
  theme(legend.position = "none") +
  scale_y_continuous(labels = percent) +
  theme(axis.text.x = element_text(color = "black")) +
  geom_text(aes(x = 0.95, y = 1.18, label = "Mann-Whitney, \np = 0.148"),
            color = "gray30",
            size = 2.5) +
  theme(plot.margin = margin(0.1, 0.1, 0.1, 0.2, "in")) +
  stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
              width = 0.75, linetype = "dashed") +
  theme(plot.title = element_text(size = 12))
herb.plot.cv
```

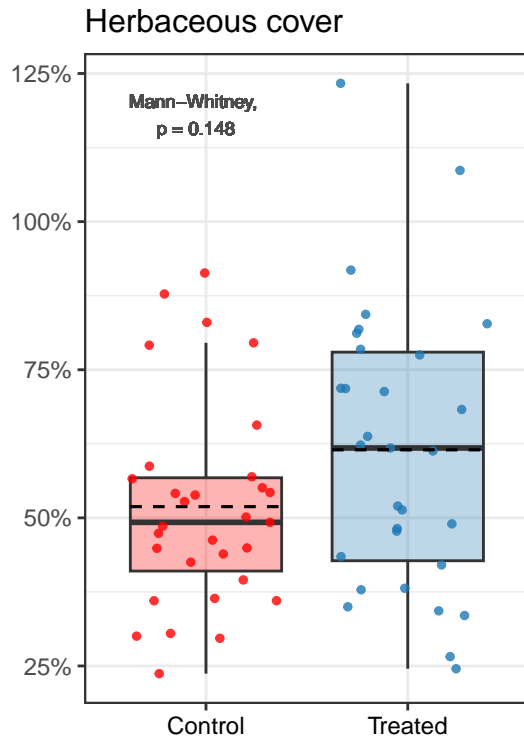


Fig 4c: CV of overall veg cover

```
# Notree cover -----

# Find CV for each sample over time
notree.sample <- notree.all |>
  group_by(Sample, Treatment) |>
  summarise(CV = sd(Cover) / mean(Cover),
    .groups = "keep")

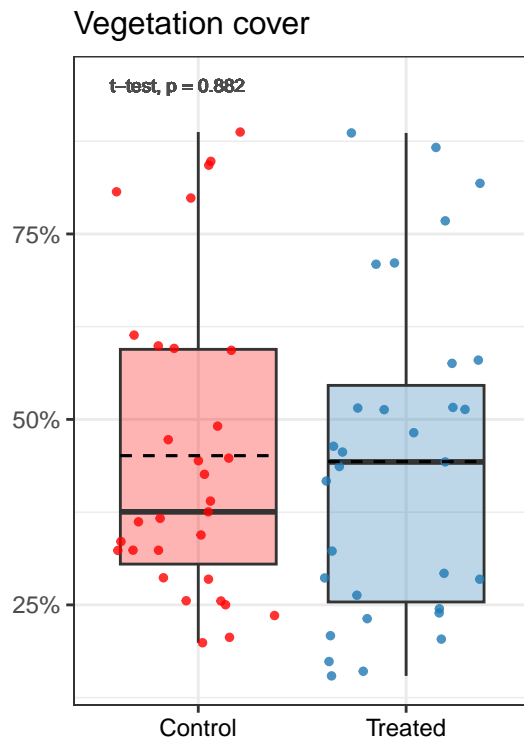
# Compare means
t.test(filter(notree.sample, Treatment == "Treated")$CV,
  filter(notree.sample, Treatment == "Control")$CV) # NS, p = 0.882
```

```
##
## Welch Two Sample t-test
##
## data: filter(notree.sample, Treatment == "Treated")$CV and filter(notree.sample, Treatment == "Control")$CV
## t = -0.14868, df = 59.846, p-value = 0.8823
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.11574119 0.09972676
## sample estimates:
## mean of x mean of y
## 0.4430859 0.4510931
```

```

# Plot
notree.plot.cv <- notree.sample |>
  ggplot(aes(x = Treatment, y = CV)) +
  geom_boxplot(alpha = 0.3,
    outlier.shape = NA,
    aes(fill = Treatment)) +
  geom_jitter(size = 1,
    alpha = 0.8,
    aes(color = Treatment)) +
  scale_color_manual(values = c("red", "#1F78B4")) +
  scale_fill_manual(values = c("red", "#1F78B4")) +
  labs(title = "Vegetation cover",
    x = NULL,
    y = NULL) +
  theme_bw() +
  theme(legend.position = "none") +
  scale_y_continuous(labels = percent) +
  theme(axis.text.x = element_text(color = "black")) +
  geom_text(aes(x = 0.9, y = 0.95, label = "t-test, p = 0.882"),
    color = "gray30",
    size = 2.5) +
  theme(plot.margin = margin(0.1, 0.1, 0.1, 0.2, "in")) +
  stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
    width = 0.75, linetype = "dashed") +
  theme(plot.title = element_text(size = 12))
notree.plot.cv

```



Combine plots for Fig 4

```
# Combine notree, herb, shrub -----

tiff("Fig4_CV_shrub-herb-notree.tiff", units = "in", height = 4, width = 7, res = 1000)
ggarrange(shrub.plot.cv, herb.plot.cv, notree.plot.cv,
          ncol = 3, nrow = 1,
          labels = c("(A)", "(B)", "(C)"))

dev.off()
```

Supp Fig 4a: CV of perennial richness

```
# Richness -----

# Find CV for each sample over time
rich.sample <- per.div |>
  group_by(Sample, Treatment) |>
  summarise(CV = sd(rich) / mean(rich),
            .groups = "keep")

# Compare means
wilcox.test(filter(rich.sample, Treatment == "Treated")$CV,
             filter(rich.sample, Treatment == "Control")$CV) # NS, p = 0.138
```

```
## Warning in wilcox.test.default(filter(rich.sample, Treatment == "Treated")$CV,
## : cannot compute exact p-value with ties
```

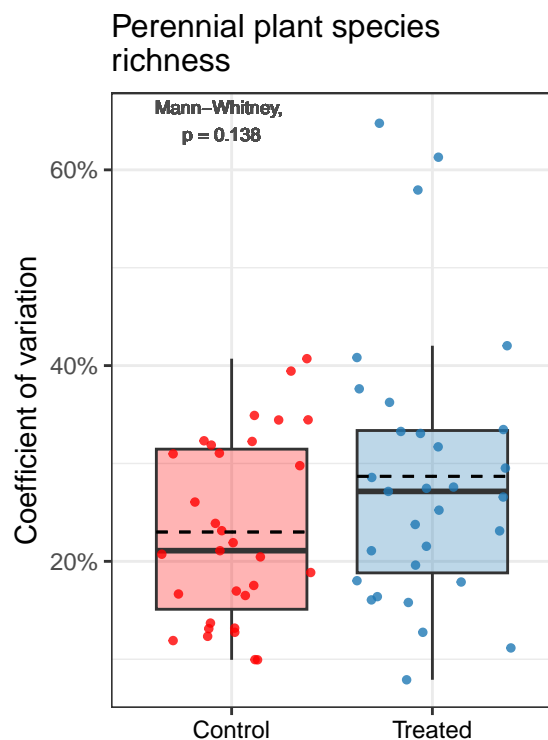
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: filter(rich.sample, Treatment == "Treated")$CV and filter(rich.sample, Treatment == "Control")$CV
## W = 586.5, p-value = 0.1375
## alternative hypothesis: true location shift is not equal to 0
```

```
# Plot
rich.plot.cv <- rich.sample |>
  ggplot(aes(x = Treatment, y = CV)) +
  geom_boxplot(alpha = 0.3,
              outlier.shape = NA,
              aes(fill = Treatment)) +
  geom_jitter(size = 1,
              alpha = 0.8,
              aes(color = Treatment)) +
  scale_color_manual(values = c("red", "#1F78B4")) +
  scale_fill_manual(values = c("red", "#1F78B4")) +
  labs(title = "Perennial plant species richness",
       x = NULL,
       y = "Coefficient of variation") +
  theme_bw() +
```

```

theme(legend.position = "none") +
scale_y_continuous(labels = percent) +
theme(axis.text.x = element_text(color = "black")) +
geom_text(aes(x = 0.95, y = 0.65, label = "Mann-Whitney, \np = 0.138"),
  color = "gray30",
  size = 2.5) +
theme(plot.margin = margin(0.1, 0.1, 0.1, 0.1, "in")) +
stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
  width = 0.75, linetype = "dashed") +
theme(plot.title = element_text(size = 12))
rich.plot.cv

```



Supp Fig 4b: CV of perennial diversity

```

# Shannon -----

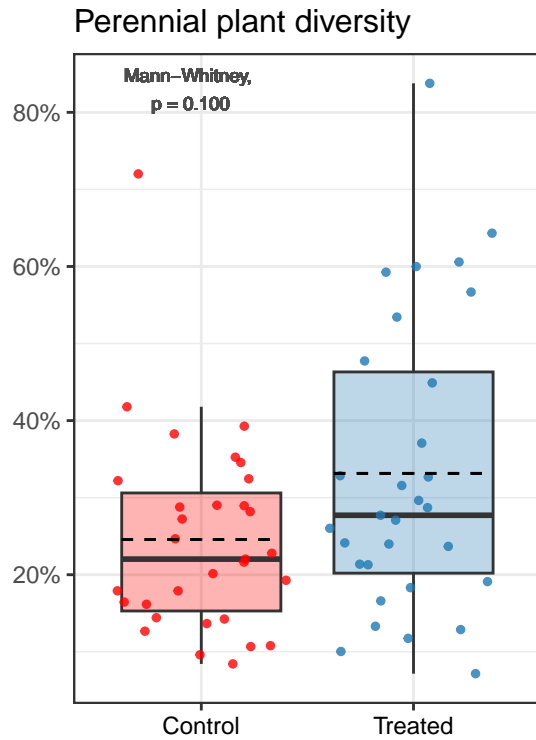
# Find CV for each sample over time
shan.sample <- per.div |>
  group_by(Sample, Treatment) |>
  summarise(CV = sd(shan) / mean(shan),
    .groups = "keep")

# Compare
wilcox.test(filter(shan.sample, Treatment == "Treated")$CV,
  filter(shan.sample, Treatment == "Control")$CV,
  exact = FALSE) # NS, p = 0.100

```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: filter(shan.sample, Treatment == "Treated")$CV and filter(shan.sample, Treatment == "Control")
## W = 598, p-value = 0.09952
## alternative hypothesis: true location shift is not equal to 0
```

```
# Plot
shan.plot.cv <- shan.sample |>
  ggplot(aes(x = Treatment, y = CV)) +
  geom_boxplot(alpha = 0.3,
    outlier.shape = NA,
    aes(fill = Treatment)) +
  geom_jitter(size = 1,
    alpha = 0.8,
    aes(color = Treatment)) +
  scale_color_manual(values = c("red", "#1F78B4")) +
  scale_fill_manual(values = c("red", "#1F78B4")) +
  labs(title = "Perennial plant diversity",
    x = NULL,
    y = NULL) +
  theme_bw() +
  theme(legend.position = "none") +
  scale_y_continuous(labels = percent) +
  theme(axis.text.x = element_text(color = "black")) +
  geom_text(aes(x = 0.95, y = 0.83, label = "Mann-Whitney, \np = 0.100"),
    color = "gray30",
    size = 2.5) +
  theme(plot.margin = margin(0.1, 0.1, 0.1, 0.15, "in")) +
  stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
    width = 0.75, linetype = "dashed") +
  theme(plot.title = element_text(size = 12))
shan.plot.cv
```



Combine plots for Supp Fig 4

```
# Combine richness & Shannon -----

# Supplemental figure
tiff("FigS4_CV_rich-shan.tiff", units = "in", height = 4, width = 5.5, res = 300)
ggarrange(rich.plot.cv, shan.plot.cv,
  ncol = 2, nrow = 1,
  labels = c("(A)", "(B)"))

dev.off()
```

Appendix C: Soil fertility

Setup

```
library(tidyverse)
library(ggpubr)
library(metagenomeSeq)
library(vegan)
```

```
# Load data -----
```

```
barc.asv <- read.table("bac-arc_clean_asv.txt", sep = "\t",
                      header = T, row.names = 1)
fungi.asv <- read.table("fungi_clean_asv.txt",
                      sep = "\t", header = T, row.names = 1)
meta <- read.csv("sequencing_metadata.csv")

dat.2021 <- read.csv("Veg-soil-elev_2021.csv")
```

```
# 16S -----
```

```
# Normalization
```

```
barc.MR <- newMRexperiment(t(barc.asv))
p <- cumNormStat(barc.MR)
```

```
## Default value being used.
```

```
barc.MR <- cumNorm(barc.MR, p = p)
barc.norm <- t(MRcounts(barc.MR, norm = T, log = F))
```

```
# Richness and Shannon
```

```
meta$barc.richness <- specnumber(barc.norm)
meta$barc.shannon <- diversity(barc.norm, index = "shannon")
```

```
# Bray-Curtis distance
```

```
barc.dist <- vegdist(barc.norm, method = "bray")
```

```
# ITS -----
```

```
# Normalization
```

```
fungi.MR <- newMRexperiment(t(fungi.asv))
p <- cumNormStat(fungi.MR)
```

```
## Default value being used.
```

```
fungi.MR <- cumNorm(fungi.MR, p = p)
fungi.norm <- t(MRcounts(fungi.MR, norm = T, log = F))
```

```
# Richness and Shannon
```



```

meta$fungi.richness <- specnumber(fungi.norm)
meta$fungi.shannon <- diversity(fungi.norm, index = "shannon")

# Bray-Curtis distance
fungi.dist <- vegdist(fungi.norm, method = "bray")

```

Fig 5a: Bacteria & archaea NMDS ordination

Note that PERMANOVA model results (adonis2) will vary slightly each time and will not exactly match values published in the paper.

```

# NMDS ordination
barc.nmds <- metaMDS(barcdist, k = 2)

## Run 0 stress 0.1684425
## Run 1 stress 0.1702925
## Run 2 stress 0.1705144
## Run 3 stress 0.1787389
## Run 4 stress 0.1786029
## Run 5 stress 0.1842344
## Run 6 stress 0.1768042
## Run 7 stress 0.1705017
## Run 8 stress 0.1708159
## Run 9 stress 0.1915466
## Run 10 stress 0.1753495
## Run 11 stress 0.1747661
## Run 12 stress 0.1787477
## Run 13 stress 0.1900313
## Run 14 stress 0.1769418
## Run 15 stress 0.168046
## ... New best solution
## ... Procrustes: rmse 0.04031828  max resid 0.2923946
## Run 16 stress 0.1915578
## Run 17 stress 0.1823108
## Run 18 stress 0.1837179
## Run 19 stress 0.1691583
## Run 20 stress 0.1680469
## ... Procrustes: rmse 0.0002821518  max resid 0.001933122
## ... Similar to previous best
## *** Best solution repeated 1 times

```

```
barc.nmds$stress
```

```
## [1] 0.168046
```

```

meta$barc.NMDS1 <- barc.nmds$points[, 1]
meta$barc.NMDS2 <- barc.nmds$points[, 2]

```

```

# PERMANOVA
adonis2(barcdist ~ meta$Treatment)

```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = barc.dist ~ meta$Treatment)
##           Df SumOfSqs      R2      F Pr(>F)
## meta$Treatment  1   0.2559 0.02881 1.7799  0.028 *
## Residual       60   8.6275 0.97119
## Total          61   8.8835 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Plot (using dat.2021 for same NMDS values presented in paper):
barc.nmds.plot.21 <- dat.2021 %>%
  ggplot(aes(x = barc.NMDS1, y = barc.NMDS2, color = Treatment, shape = Treatment)) +
  geom_point() +
  scale_color_manual(values = c("red", "#1F78B4")) +
  theme_bw() +
  labs(x = "Axis 1",
       y = "Axis 2",
       title = "Bacteria & archaea",
       color = "Treatment",
       shape = "Treatment") +
  theme(legend.position = "bottom") +
  theme(plot.margin = margin(0.1, 0.2, 0.1, 0.1, "in")) +
  theme(legend.title = element_blank()) +
  geom_text(aes(x = 0.2, y = -0.55, label = "PERMANOVA, p = 0.029"),
            size = 2.5, color = "gray30") +
  geom_text(aes(x = 0.25, y = -0.65, label = "Stress = 0.168"),
            size = 2.5, color = "gray30") +
  theme(plot.title = element_text(size = 12))
barc.nmds.plot.21
```

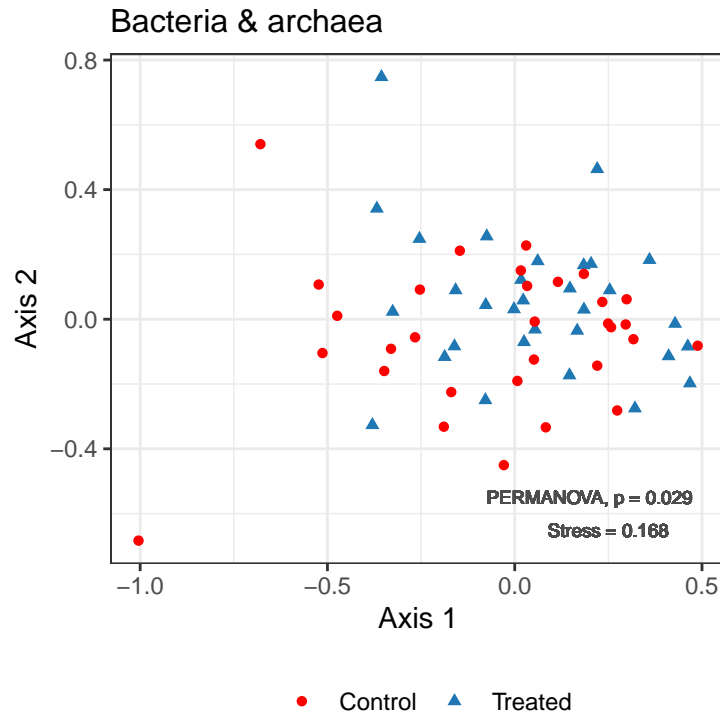


Fig 5b: Fungi NMDS ordination

Note that PERMANOVA model results (`adonis2`) will vary slightly each time and will not exactly match values published in the paper.

```
# NMDS ordination
fungi.nmms <- metaMDS(fungi.dist, k = 2)

## Run 0 stress 0.2461785
## Run 1 stress 0.2710022
## Run 2 stress 0.2512675
## Run 3 stress 0.2533262
## Run 4 stress 0.24328
## ... New best solution
## ... Procrustes: rmse 0.1125388 max resid 0.433694
## Run 5 stress 0.2385107
## ... New best solution
## ... Procrustes: rmse 0.06861351 max resid 0.2199476
## Run 6 stress 0.2474888
## Run 7 stress 0.2448295
## Run 8 stress 0.251045
## Run 9 stress 0.2418675
## Run 10 stress 0.2529092
## Run 11 stress 0.256877
## Run 12 stress 0.2456739
## Run 13 stress 0.2444618
## Run 14 stress 0.2429744
## Run 15 stress 0.2412761
```

```
## Run 16 stress 0.2499875
## Run 17 stress 0.2455412
## Run 18 stress 0.2406315
## Run 19 stress 0.2415854
## Run 20 stress 0.2425369
## *** Best solution was not repeated -- monoMDS stopping criteria:
##      1: no. of iterations >= maxit
##     19: stress ratio > sratmax
```

```
fungi.nmds$stress
```

```
## [1] 0.2385107
```

```
meta$fungi.NMDS1 <- fungi.nmds$points[, 1]
meta$fungi.NMDS2 <- fungi.nmds$points[, 2]
```

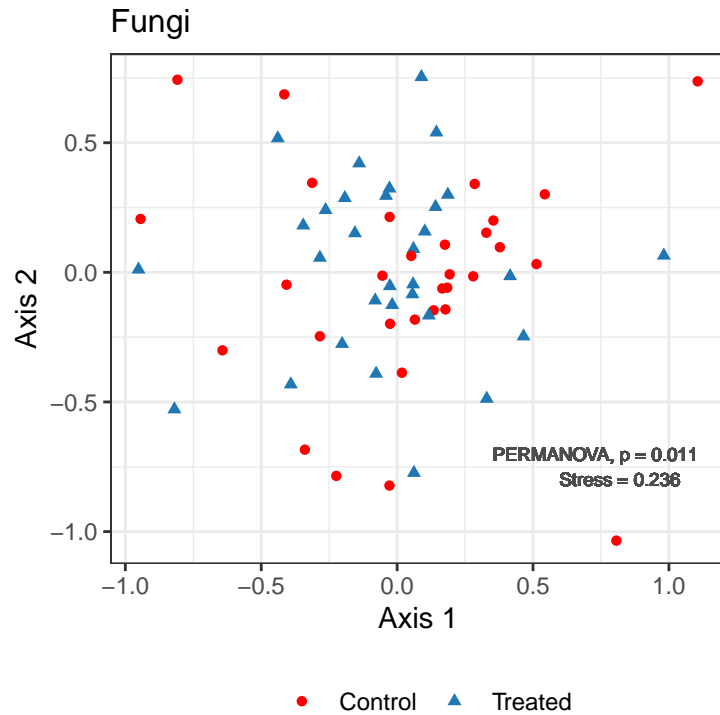
```
# PERMANOVA
```

```
adonis2(fungi.dist ~ meta$Treatment)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = fungi.dist ~ meta$Treatment)
##              Df SumOfSqs      R2      F Pr(>F)
## meta$Treatment  1   0.4511 0.02371 1.457  0.01 **
## Residual       60  18.5783 0.97629
## Total          61  19.0294 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Plot (using dat.2021 for same NMDS values presented in paper):
```

```
fungi.nmds.plot.21 <- dat.2021 %>%
  ggplot(aes(x = fungi.NMDS1, y = fungi.NMDS2, color = Treatment, shape = Treatment)) +
  geom_point() +
  scale_color_manual(values = c("red", "#1F78B4")) +
  theme_bw() +
  labs(x = "Axis 1",
       y = "Axis 2",
       title = "Fungi",
       color = "Treatment",
       shape = "Treatment") +
  theme(legend.position = "bottom") +
  theme(legend.title = element_blank()) +
  theme(plot.margin = margin(0.1, 0.2, 0.1, 0.1, "in")) +
  geom_text(aes(x = 0.73, y = -0.7, label = "PERMANOVA, p = 0.011"),
            size = 2.5, color = "gray30") +
  geom_text(aes(x = 0.82, y = -0.8, label = "Stress = 0.236"),
            size = 2.5, color = "gray30") +
  theme(plot.title = element_text(size = 12))
fungi.nmds.plot.21
```



Combine plots for Fig 5

```
# Combine NMDS -----

tiff("Fig5_Soil-NMDS.tiff", height = 4, width = 7, units = "in", res = 1000)
ggarrange(barc.nmds.plot.21, fungi.nmds.plot.21,
  nrow = 1, ncol = 2,
  labels = c("(A)", "(B)"),
  common.legend = TRUE, legend = "bottom")
dev.off()
```

Supp Fig 5a: Total soil nitrogen

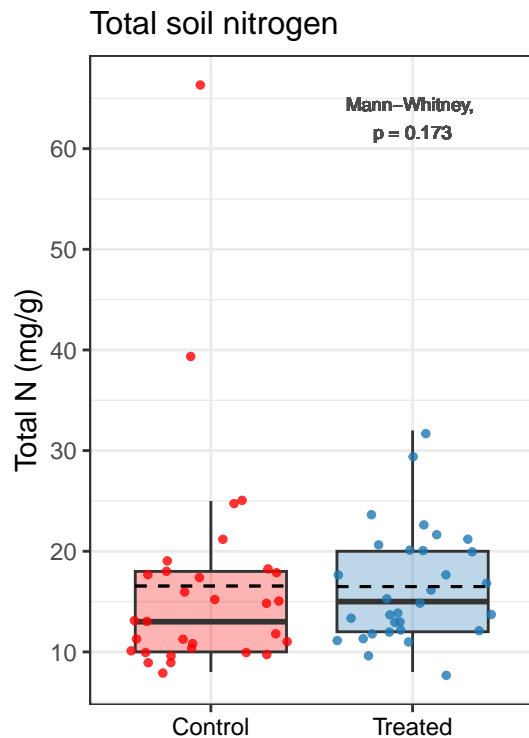
```
# Total N -----

# Mann-Whitney
wilcox.test(filter(dat.2021, Treatment == "Control")$TN_ppt,
  filter(dat.2021, Treatment == "Treated")$TN_ppt,
  exact = FALSE) # p-value = 0.1731
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: filter(dat.2021, Treatment == "Control")$TN_ppt and filter(dat.2021, Treatment == "Treated")$
```

```
## W = 383.5, p-value = 0.1731
## alternative hypothesis: true location shift is not equal to 0
```

```
# Plot
tn.plot.21 <- dat.2021 |>
  ggplot(aes(x = Treatment, y = TN_ppt)) +
  geom_boxplot(alpha = 0.3,
    outlier.shape = NA,
    aes(fill = Treatment)) +
  geom_jitter(size = 1,
    alpha = 0.8,
    aes(color = Treatment)) +
  scale_color_manual(values = c("red", "#1F78B4")) +
  scale_fill_manual(values = c("red", "#1F78B4")) +
  labs(title = "Total soil nitrogen",
    x = NULL,
    y = "Total N (mg/g)") +
  theme_bw() +
  theme(legend.position = "none") +
  theme(axis.text.x = element_text(color = "#000000")) +
  theme(plot.margin = margin(0.1, 0.2, 0.1, 0.1, "in")) +
  geom_text(aes(x = 2, y = 63, label = "Mann-Whitney, \np = 0.173"),
    color = "gray30",
    size = 2.5) +
  stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
    width = 0.75, linetype = "dashed") +
  theme(plot.title = element_text(size = 12))
tn.plot.21
```



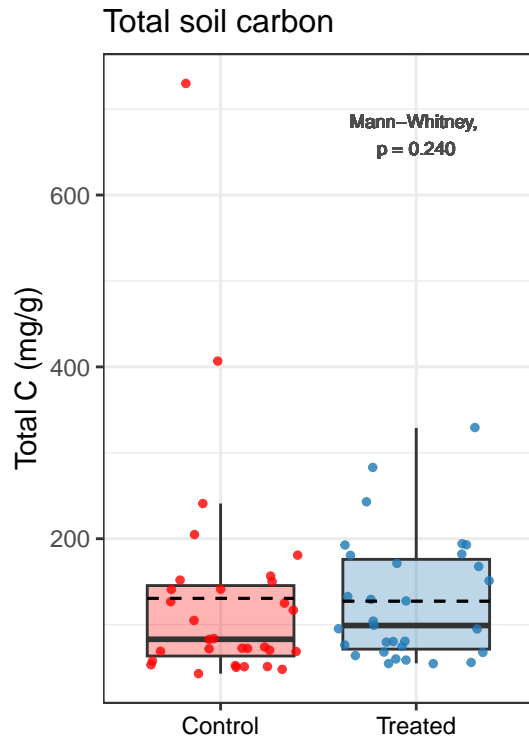
Supp Fig 5b: Total soil carbon

```
# Total C -----

# Mann-Whitney
wilcox.test(filter(dat.2021, Treatment == "Control")$TC_ppt,
            filter(dat.2021, Treatment == "Treated")$TC_ppt,
            exact = FALSE) # p-value = 0.2397

##
## Wilcoxon rank sum test with continuity correction
##
## data: filter(dat.2021, Treatment == "Control")$TC_ppt and filter(dat.2021, Treatment == "Treated")$
## W = 396.5, p-value = 0.2397
## alternative hypothesis: true location shift is not equal to 0

# Plot
tc.plot.21 <- dat.2021 |>
  ggplot(aes(x = Treatment, y = TC_ppt)) +
  geom_boxplot(alpha = 0.3,
               outlier.shape = NA,
               aes(fill = Treatment)) +
  geom_jitter(size = 1,
              alpha = 0.8,
              aes(color = Treatment)) +
  scale_color_manual(values = c("red", "#1F78B4")) +
  scale_fill_manual(values = c("red", "#1F78B4")) +
  labs(title = "Total soil carbon",
       x = NULL,
       y = "Total C (mg/g)") +
  theme_bw() +
  theme(legend.position = "none") +
  theme(axis.text.x = element_text(color = "#000000")) +
  theme(plot.margin = margin(0.1, 0.2, 0.1, 0.1, "in")) +
  geom_text(aes(x = 2, y = 670, label = "Mann-Whitney, \np = 0.240"),
            color = "gray30",
            size = 2.5) +
  stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
              width = 0.75, linetype = "dashed") +
  theme(plot.title = element_text(size = 12))
tc.plot.21
```



Supp Fig 5c: Soil organic matter

```
# Organic matter -----
# Mann-Whitney
wilcox.test(filter(dat.2021, Treatment == "Control")$OM_perc,
             filter(dat.2021, Treatment == "Treated")$OM_perc) # p-value = 0.4332
```

```
##
## Wilcoxon rank sum exact test
##
## data: filter(dat.2021, Treatment == "Control")$OM_perc and filter(dat.2021, Treatment == "Treated")$OM_perc
## W = 424, p-value = 0.4332
## alternative hypothesis: true location shift is not equal to 0
```

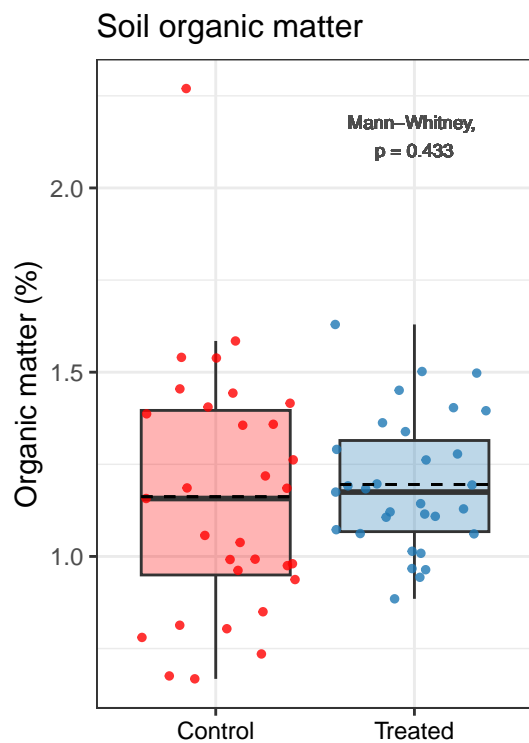
```
# Plot
om.plot.21 <- dat.2021 |>
ggplot(aes(x = Treatment, y = OM_perc)) +
  geom_boxplot(alpha = 0.3,
               outlier.shape = NA,
               aes(fill = Treatment)) +
  geom_jitter(size = 1,
              alpha = 0.8,
              aes(color = Treatment)) +
  scale_color_manual(values = c("red", "#1F78B4")) +
  scale_fill_manual(values = c("red", "#1F78B4")) +
```



```

labs(title = "Soil organic matter",
     x = NULL,
     y = "Organic matter (%)") +
theme_bw() +
theme(legend.position = "none") +
theme(axis.text.x = element_text(color = "#000000")) +
theme(plot.margin = margin(0.1, 0.2, 0.1, 0.1, "in")) +
geom_text(aes(x = 2, y = 2.14, label = "Mann-Whitney, \np = 0.433"),
         color = "gray30",
         size = 2.5) +
stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
         width = 0.75, linetype = "dashed") +
theme(plot.title = element_text(size = 12))
om.plot.21

```



Combine plots for Supp Fig 5

```

# Combine soil chem -----

# TN, TC, OM
tiff("FigS5_Soil-chem.tiff", units = "in", height = 4, width = 7, res = 300)
ggarrange(tn.plot.21, tc.plot.21, om.plot.21,
         ncol = 3, nrow = 1,
         labels = c("(A)", "(B)", "(C)"))

dev.off()

```

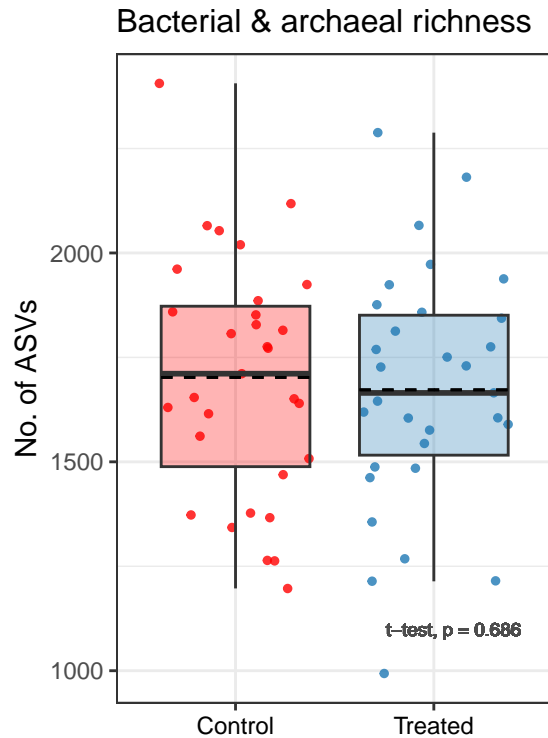
Supp Fig 6a: Bacterial & archaeal richness

```
# Barc richness -----

# T-test
t.test(filter(dat.2021, Treatment == "Control")$barc.richness,
        filter(dat.2021, Treatment == "Treated")$barc.richness) # NS, p = 0.686

##
## Welch Two Sample t-test
##
## data: filter(dat.2021, Treatment == "Control")$barc.richness and filter(dat.2021, Treatment == "Treated")$barc.richness
## t = 0.40607, df = 59.999, p-value = 0.6861
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -116.6407 176.0601
## sample estimates:
## mean of x mean of y
## 1702.065 1672.355

# Plot
barc.rich.plot.21 <- dat.2021 |>
  ggplot(aes(Treatment, barc.richness)) +
  geom_jitter(aes(color = Treatment),
              alpha = 0.8,
              size = 1) +
  geom_boxplot(aes(fill = Treatment),
               alpha = 0.3,
               outlier.shape = NA) +
  xlab(NULL) +
  ylab("No. of ASVs") +
  ggtitle("Bacterial & archaeal richness") +
  scale_color_manual(values = c("red", "#1F78B4")) +
  scale_fill_manual(values = c("red", "#1F78B4")) +
  theme_bw() +
  theme(legend.position = "none") +
  theme(axis.text.x = element_text(color = "#000000")) +
  geom_text(aes(x = 2.1, y = 1100, label = "t-test, p = 0.686"),
            color = "gray30",
            size = 2.5) +
  theme(plot.margin = margin(0.1, 0.1, 0.1, 0.1, "in")) +
  stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
              width = 0.75, linetype = "dashed") +
  theme(plot.title = element_text(size = 11.5))
barc.rich.plot.21
```



Supp Fig 6b: Fungal richness

```
# Fungi richness -----
# T-test
t.test(filter(dat.2021, Treatment == "Control")$fungi.richness,
       filter(dat.2021, Treatment == "Treated")$fungi.richness) # NS,  $p = 0.938$ 
```

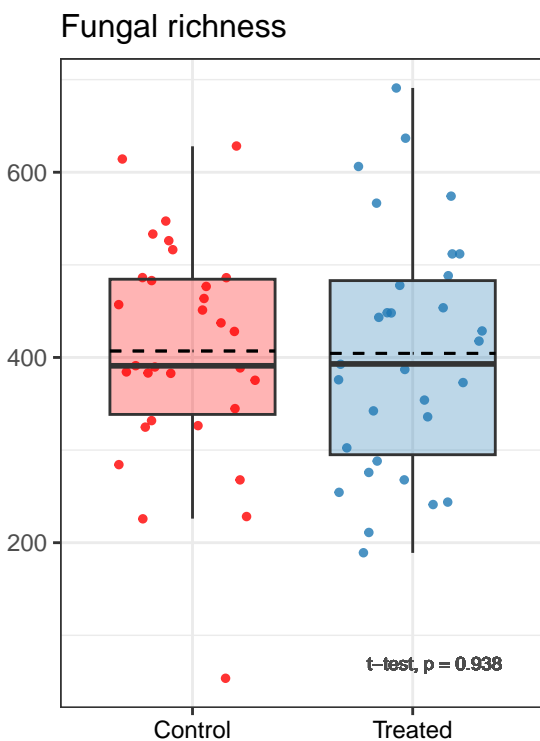
```
##
## Welch Two Sample t-test
##
## data: filter(dat.2021, Treatment == "Control")$fungi.richness and filter(dat.2021, Treatment == "Treated")$fungi.richness
## t = 0.078, df = 59.653, p-value = 0.9381
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -61.22215 66.18989
## sample estimates:
## mean of x mean of y
## 406.9677 404.4839
```

```
# Plot
fungi.rich.plot.21 <- dat.2021 %>%
  ggplot(aes(Treatment, fungi.richness)) +
  geom_jitter(aes(color = Treatment),
             alpha = 0.8,
             size = 1) +
```

```

geom_boxplot(aes(fill = Treatment),
             alpha = 0.3,
             outlier.shape = NA) +
xlab(NULL) +
ylab(NULL) +
ggtitle("Fungal richness") +
scale_color_manual(values = c("red", "#1F78B4")) +
scale_fill_manual(values = c("red", "#1F78B4")) +
theme_bw() +
theme(legend.position = "none") +
theme(axis.text.x = element_text(color = "#000000")) +
geom_text(aes(x = 2.1, y = 70, label = "t-test, p = 0.938"),
          color = "gray30",
          size = 2.5) +
theme(plot.margin = margin(0.1, 0.1, 0.1, 0.1, "in")) +
stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
            width = 0.75, linetype = "dashed") +
theme(plot.title = element_text(size = 12))
fungi.rich.plot.21

```



Combine plots for Supp Fig 6

```

# Combine soil richness -----
tiff("FigS6_Soil-richness.tiff", units = "in", height = 4, width = 5.5, res = 300)
ggarrange(barc.rich.plot.21, fungi.rich.plot.21,

```

```

ncol = 2, nrow = 1,
labels = c("(A)", "(B)"))

dev.off()

```

Supp Fig 7a: Chemoheterotrophic bacteria & archaea

```

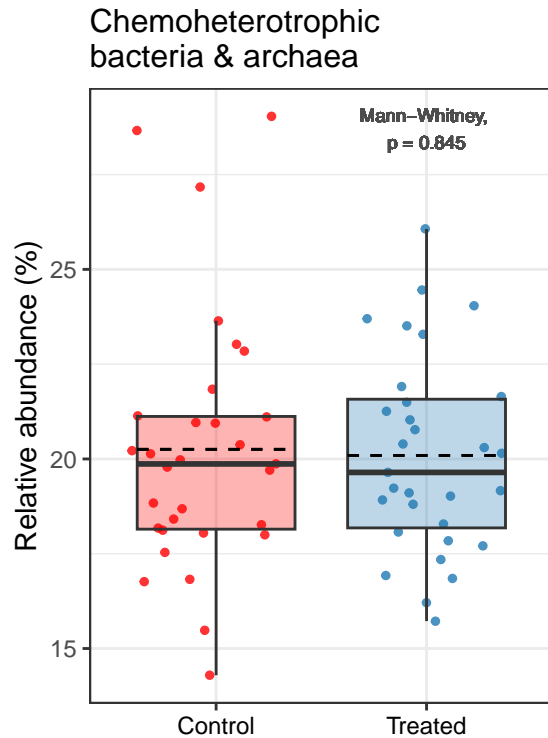
# Chemoheterotrophs -----

# Mann-Whitney
wilcox.test(filter(dat.2021, Treatment == "Control")$chemoheterotrophy_perc,
             filter(dat.2021, Treatment == "Treated")$chemoheterotrophy_perc) # p = 0.8449

##
## Wilcoxon rank sum exact test
##
## data: filter(dat.2021, Treatment == "Control")$chemoheterotrophy_perc and filter(dat.2021, Treatment == "Treated")$chemoheterotrophy_perc
## W = 466, p-value = 0.8449
## alternative hypothesis: true location shift is not equal to 0

# Plot
chemohet.plot.21 <- dat.2021 %>%
  ggplot(aes(Treatment, chemoheterotrophy_perc)) +
  geom_jitter(aes(color = Treatment),
              alpha = 0.8,
              size = 1) +
  geom_boxplot(aes(fill = Treatment),
               alpha = 0.3,
               outlier.shape = NA) +
  xlab(NULL) +
  ylab("Relative abundance (%)") +
  ggtitle("Chemoheterotrophic \nbacteria & archaea") +
  scale_color_manual(values = c("red", "#1F78B4")) +
  scale_fill_manual(values = c("red", "#1F78B4")) +
  theme_bw() +
  theme(legend.position = "none") +
  theme(axis.text.x = element_text(color = "#000000")) +
  theme(plot.margin = margin(0.1, 0.1, 0.1, 0.1, "in")) +
  geom_text(aes(x = 2, y = 28.7, label = "Mann-Whitney, \np = 0.845"),
            color = "gray30",
            size = 2.5) +
  stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
              width = 0.75, linetype = "dashed") +
  theme(plot.title = element_text(size = 12))
chemohet.plot.21

```



Supp Fig 7b: Nitrogen-cycling bacteria & archaea

```
# N-cyclers -----
# Mann-Whitney
wilcox.test(filter(dat.2021, Treatment == "Control")$n.cycler_perc,
             filter(dat.2021, Treatment == "Treated")$n.cycler_perc) # p-value = 0.5854
```

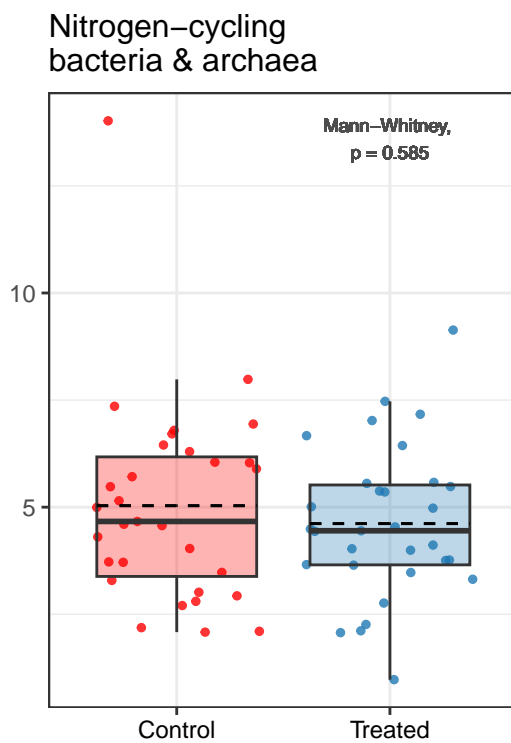
```
##
## Wilcoxon rank sum exact test
##
## data: filter(dat.2021, Treatment == "Control")$n.cycler_perc and filter(dat.2021, Treatment == "Treated")$n.cycler_perc
## W = 520, p-value = 0.5854
## alternative hypothesis: true location shift is not equal to 0
```

```
# Plot
ncycler.plot.21 <- dat.2021 %>%
  ggplot(aes(Treatment, n.cycler_perc)) +
  geom_jitter(aes(color = Treatment),
              alpha = 0.8,
              size = 1) +
  geom_boxplot(aes(fill = Treatment),
               alpha = 0.3,
               outlier.shape = NA) +
  xlab(NULL) +
  ylab(NULL) +
```

```

ggtitle("Nitrogen-cycling \nbacteria & archaea") +
scale_color_manual(values = c("red", "#1F78B4")) +
scale_fill_manual(values = c("red", "#1F78B4")) +
theme_bw() +
theme(legend.position = "none") +
theme(axis.text.x = element_text(color = "#000000")) +
theme(plot.margin = margin(0.1, 0.1, 0.1, 0.25, "in")) +
geom_text(aes(x = 2, y = 13.6, label = "Mann-Whitney, \np = 0.585"),
          color = "gray30",
          size = 2.5) +
stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
            width = 0.75, linetype = "dashed") +
theme(plot.title = element_text(size = 12))
ncycler.plot.21

```



Supp Fig 7c: Saprotrophic fungi

```

# Saprotrophs -----

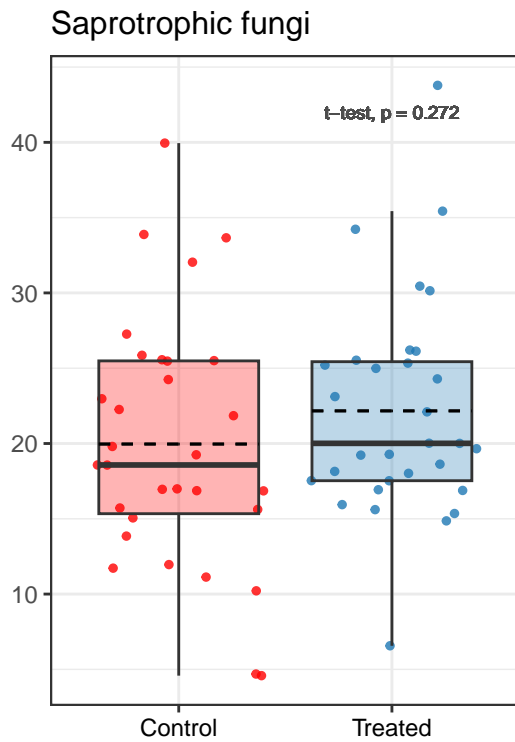
# T-test
t.test(filter(dat.2021, Treatment == "Control")$saprotroph,
        filter(dat.2021, Treatment == "Treated")$saprotroph) # NS, p = 0.272

##
## Welch Two Sample t-test
##

```

```
## data: filter(dat.2021, Treatment == "Control")$saprotroph and filter(dat.2021, Treatment == "Treated")$saprotroph
## t = -1.1097, df = 59.068, p-value = 0.2716
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -6.168384  1.767276
## sample estimates:
## mean of x mean of y
## 19.96490 22.16546
```

```
# Plot
sapro.plot.21 <- dat.2021 %>%
  ggplot(aes(Treatment, saprotroph)) +
  geom_jitter(aes(color = Treatment),
              alpha = 0.8,
              size = 1) +
  geom_boxplot(aes(fill = Treatment),
               alpha = 0.3,
               outlier.shape = NA) +
  xlab(NULL) +
  ylab(NULL) +
  ggtitle("Saprotrophic fungi") +
  scale_color_manual(values = c("red", "#1F78B4")) +
  scale_fill_manual(values = c("red", "#1F78B4")) +
  theme_bw() +
  theme(legend.position = "none") +
  theme(axis.text.x = element_text(color = "#000000")) +
  theme(plot.margin = margin(0.1, 0.1, 0.1, 0.25, "in")) +
  geom_text(aes(x = 2, y = 42, label = "t-test, p = 0.272"),
            color = "gray30",
            size = 2.5) +
  stat_summary(fun = mean, geom = "errorbar", aes(ymax = after_stat(y), ymin = after_stat(y)),
              width = 0.75, linetype = "dashed") +
  theme(plot.title = element_text(size = 12))
sapro.plot.21
```

Combine plots for Supp Fig 7

```
# Combine FAPROTAX and FUNGuild -----

tiff("FigS7_Soil-functional.tiff", units = "in", height = 4, width = 7, res = 300)
ggarrange(chemohet.plot.21, ncycler.plot.21, sapro.plot.21,
  ncol = 3, nrow = 1,
  labels = c("(A)", "(B)", "(C)"))

dev.off()
```

Appendix D: Structural equation modeling

Setup

```
library(lavaan)
library(tidyverse)

# Load data -----

dat.2021 <- read.csv("Veg-soil-elev_2021.csv")

# Data wrangling -----

# Add Control/Treated as binary variable and select only variables needed for SEM
sem.dat.unscaled <- dat.2021 |>
  mutate(rocks = case_when(
    Treatment == "Control" ~ 0,
    Treatment == "Treated" ~ 1)) |>
  select(Sample, rocks, notree, notree.18, herb, herb.18, tree, perveg.richness, perveg.shannon,
    TN_log, TC_log, OM_log, barc.richness, fungi.richness,
    chemoheterotrophy_log, n.cycler_log, saprotroph)

# Center and scale continuous variables
sem.dat <- sem.dat.unscaled |>
  mutate(rocks = as.character(rocks),
    Sample = as.character(Sample)) |>
  mutate_if(is.numeric, scale) |>
  mutate(rocks = as.numeric(rocks))
```

Latent variables

Soil microbiome

```
lvmod.soimic <- '
  # latent variable model
  soil_microbe =~ barc.richness + fungi.richness + chemoheterotrophy_log + n.cycler_log + saprotroph
'

fit.soimic <- sem(lvmod.soimic, data = sem.dat)
summary(fit.soimic, fit.measures = TRUE, standardized = TRUE)

## lavaan 0.6.16 ended normally after 18 iterations
##
##   Estimator                      ML
##   Optimization method          NLMINB
##   Number of model parameters    10
##
##   Number of observations        62
##
## Model Test User Model:
```

```

##
## Test statistic 2.723
## Degrees of freedom 5
## P-value (Chi-square) 0.743
##
## Model Test Baseline Model:
##
## Test statistic 52.070
## Degrees of freedom 10
## P-value 0.000
##
## User Model versus Baseline Model:
##
## Comparative Fit Index (CFI) 1.000
## Tucker-Lewis Index (TLI) 1.108
##
## Loglikelihood and Information Criteria:
##
## Loglikelihood user model (H0) -412.677
## Loglikelihood unrestricted model (H1) -411.315
##
## Akaike (AIC) 845.354
## Bayesian (BIC) 866.625
## Sample-size adjusted Bayesian (SABIC) 835.162
##
## Root Mean Square Error of Approximation:
##
## RMSEA 0.000
## 90 Percent confidence interval - lower 0.000
## 90 Percent confidence interval - upper 0.125
## P-value H_0: RMSEA <= 0.050 0.800
## P-value H_0: RMSEA >= 0.080 0.134
##
## Standardized Root Mean Square Residual:
##
## SRMR 0.044
##
## Parameter Estimates:
##
## Standard errors Standard
## Information Expected
## Information saturated (h1) model Structured
##
## Latent Variables:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## soil_microbe =~
## barc.richness 1.000 0.853 0.860
## fungi.richness 0.342 0.168 2.033 0.042 0.291 0.294
## chmhttrtrphy_lg 0.158 0.164 0.962 0.336 0.135 0.136
## n.cycler_log 0.892 0.241 3.699 0.000 0.761 0.767
## saprotroph 0.461 0.172 2.675 0.007 0.393 0.397
##
## Variances:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all

```

```
##      .barc.richness      0.256      0.184      1.392      0.164      0.256      0.260
##      .fungi.richness    0.899      0.165      5.448      0.000      0.899      0.914
##      .chmhttrtrphy_lg   0.966      0.174      5.545      0.000      0.966      0.982
##      .n.cycler_log       0.405      0.160      2.538      0.011      0.405      0.412
##      .saprotroph         0.829      0.156      5.321      0.000      0.829      0.843
##      soil_microbe        0.728      0.247      2.951      0.003      1.000      1.000
```

```
modindices(fit.soimic, sort = TRUE, minimum.value = 3.5)
```

```
## [1] lhs      op      rhs      mi      epc      sepc.lv  sepc.all sepc.nox
## <0 rows> (or 0-length row.names)
```

```
# No paths to add.
```

Soil chemistry

- Soil chemistry does not do well as a latent variable because TN, TC, and OM are collinear and must be modeled separately.

```
lvmod.soichem <- '
# latent variable model
soil_chem =~ TN_log + OM_log
'
fit.soichem <- sem(lvmod.soichem, data = sem.dat)
```

```
## Warning in lav_model_vcov(lavmodel = lavmodel, lavsamplestats = lavsamplestats, : lavaan WARNING:
##      Could not compute standard errors! The information matrix could
##      not be inverted. This may be a symptom that the model is not
##      identified.
```

```
summary(fit.soichem)
```

```
## lavaan 0.6.16 ended normally after 13 iterations
##
##      Estimator                      ML
##      Optimization method           NLMINB
##      Number of model parameters      4
##
##      Number of observations           62
##
## Model Test User Model:
##
##      Test statistic                  NA
##      Degrees of freedom              -1
##      P-value (Unknown)              NA
##
## Parameter Estimates:
##
##      Standard errors                Standard
##      Information                    Expected
##      Information saturated (h1) model Structured
```

```
##
## Latent Variables:
##           Estimate Std.Err z-value P(>|z|)
##   soil_chem =~
##     TN_log      1.000
##     OM_log      0.874      NA
##
## Variances:
##           Estimate Std.Err z-value P(>|z|)
##     .TN_log      0.139      NA
##     .OM_log      0.338      NA
##     soil_chem    0.845      NA
```

Model 1

Model for Supp Fig 8.

```
mod1 <- '
# latent variables
soil_microbe =~ barc.richness + fungi.richness + chemoheterotrophy_log + n.cycler_log + saprotroph

# structure
notree ~ rocks + notree.18 + tree
OM_log ~ rocks
soil_microbe ~ rocks
notree.18 ~ rocks

# covariance
OM_log ~~ soil_microbe
OM_log ~~ notree
soil_microbe ~~ notree
'

fit1 <- sem(mod1, data = sem.dat)
summary(fit1, fit.measures = TRUE, standardized = TRUE)
```

```
## lavaan 0.6.16 ended normally after 27 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of model parameters      22
##
##      Number of observations          62
##
## Model Test User Model:
##
##      Test statistic                  26.242
##      Degrees of freedom              30
##      P-value (Chi-square)            0.663
##
## Model Test Baseline Model:
##
##      Test statistic                  124.316
##      Degrees of freedom              44
```

```

##      P-value                                0.000
##
## User Model versus Baseline Model:
##
##      Comparative Fit Index (CFI)                1.000
##      Tucker-Lewis Index (TLI)                  1.069
##
## Loglikelihood and Information Criteria:
##
##      Loglikelihood user model (H0)              -650.724
##      Loglikelihood unrestricted model (H1)       -637.603
##
##      Akaike (AIC)                              1345.448
##      Bayesian (BIC)                             1392.245
##      Sample-size adjusted Bayesian (SABIC)      1323.026
##
## Root Mean Square Error of Approximation:
##
##      RMSEA                                0.000
##      90 Percent confidence interval - lower     0.000
##      90 Percent confidence interval - upper     0.080
##      P-value H_0: RMSEA <= 0.050              0.829
##      P-value H_0: RMSEA >= 0.080              0.049
##
## Standardized Root Mean Square Residual:
##
##      SRMR                                0.076
##
## Parameter Estimates:
##
##      Standard errors                        Standard
##      Information                          Expected
##      Information saturated (h1) model      Structured
##
## Latent Variables:
##      Estimate  Std.Err  z-value  P(>|z|)  Std.lv  Std.all
##      soil_microbe =~
##      barc.richness      1.000
##      fungi.richness     0.361    0.165    2.194    0.028    0.305    0.308
##      chmhttrtrphy_lg    0.192    0.164    1.167    0.243    0.162    0.163
##      n.cycler_log        0.908    0.194    4.675    0.000    0.768    0.774
##      saprotroph          0.442    0.165    2.680    0.007    0.374    0.377
##
## Regressions:
##      Estimate  Std.Err  z-value  P(>|z|)  Std.lv  Std.all
##      notree ~
##      rocks      -0.615    0.239   -2.567    0.010   -0.615   -0.307
##      notree.18    0.535    0.107    5.011    0.000    0.535    0.530
##      tree       -0.037    0.106   -0.347    0.728   -0.037   -0.036
##      OM_log ~
##      rocks        0.249    0.250    0.996    0.319    0.249    0.126
##      soil_microbe ~
##      rocks       -0.106    0.237   -0.446    0.656   -0.125   -0.063
##      notree.18 ~

```

```
##      rocks          0.674    0.237    2.844    0.004    0.674    0.340
##
## Covariances:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## .soil_microbe ~~
## .OM_log          0.370    0.129    2.861    0.004    0.439    0.446
## .notree ~~
## .OM_log          0.332    0.115    2.877    0.004    0.332    0.393
## .soil_microbe ~~
## .notree          0.213    0.107    1.991    0.046    0.252    0.294
##
## Variances:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## .barc.richness    0.269    0.138    1.955    0.051    0.269    0.274
## .fungi.richness   0.891    0.164    5.442    0.000    0.891    0.905
## .chmhttrtrphy_lg 0.958    0.173    5.535    0.000    0.958    0.973
## .n.cycler_log     0.395    0.129    3.067    0.002    0.395    0.401
## .saprotroph        0.844    0.157    5.367    0.000    0.844    0.858
## .notree           0.740    0.133    5.568    0.000    0.740    0.737
## .OM_log           0.968    0.174    5.568    0.000    0.968    0.984
## .notree.18        0.870    0.156    5.568    0.000    0.870    0.885
## .soil_microbe     0.712    0.213    3.349    0.001    0.996    0.996
```

```
modindices(fit1, sort = TRUE, minimum.value = 3.5)
```

```
##      lhs op   rhs   mi   epc sepc.lv sepc.all sepc.nox
## 50      saprotroph ~~ OM_log 4.331 -0.213 -0.213 -0.235 -0.235
## 43 chemoheterotrophy_log ~~ OM_log 3.616 0.203 0.203 0.211 0.211
```

```
# OM already covaries with soil microbiome latent variable;
# does not make sense to add paths for saprotrophs or chemoheterotrophs.
```

Model 2

Model for Supp Fig 9.

```
mod2 <- '
# latent variables
soil_microbe =~ barc.richness + fungi.richness + chemoheterotrophy_log + n.cycler_log + saprotroph

# structure
notree ~ rocks + notree.18 + tree
TN_log ~ rocks
soil_microbe ~ rocks
notree.18 ~ rocks

# covariance
TN_log ~~ soil_microbe
TN_log ~~ notree
soil_microbe ~~ notree
'

fit2 <- sem(mod2, data = sem.dat)
summary(fit2, fit.measures = TRUE, standardized = TRUE)
```

```

## lavaan 0.6.16 ended normally after 32 iterations
##
## Estimator ML
## Optimization method NLMINB
## Number of model parameters 22
##
## Number of observations 62
##
## Model Test User Model:
##
## Test statistic 31.847
## Degrees of freedom 30
## P-value (Chi-square) 0.375
##
## Model Test Baseline Model:
##
## Test statistic 164.607
## Degrees of freedom 44
## P-value 0.000
##
## User Model versus Baseline Model:
##
## Comparative Fit Index (CFI) 0.985
## Tucker-Lewis Index (TLI) 0.978
##
## Loglikelihood and Information Criteria:
##
## Loglikelihood user model (H0) -633.381
## Loglikelihood unrestricted model (H1) -617.457
##
## Akaike (AIC) 1310.762
## Bayesian (BIC) 1357.559
## Sample-size adjusted Bayesian (SABIC) 1288.340
##
## Root Mean Square Error of Approximation:
##
## RMSEA 0.032
## 90 Percent confidence interval - lower 0.000
## 90 Percent confidence interval - upper 0.103
## P-value H_0: RMSEA <= 0.050 0.595
## P-value H_0: RMSEA >= 0.080 0.165
##
## Standardized Root Mean Square Residual:
##
## SRMR 0.079
##
## Parameter Estimates:
##
## Standard errors Standard
## Information Expected
## Information saturated (h1) model Structured
##
## Latent Variables:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all

```



```
## soil_microbe =~
##   barc.richness      1.000                0.808    0.815
##   fungi.richness     0.455    0.167    2.719    0.007    0.367    0.370
##   chmhttrtrphy_lg    0.284    0.169    1.679    0.093    0.229    0.231
##   n.cycler_log        0.948    0.163    5.802    0.000    0.766    0.772
##   saprotroph          0.464    0.167    2.777    0.005    0.375    0.378
##
## Regressions:
##           Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## notree ~
##   rocks      -0.651    0.237   -2.744    0.006   -0.651   -0.315
##   notree.18    0.593    0.099    6.011    0.000    0.593    0.570
##   tree       -0.033    0.098   -0.336    0.737   -0.033   -0.032
## TN_log ~
##   rocks      0.177    0.251    0.706    0.480    0.177    0.089
## soil_microbe ~
##   rocks     -0.102    0.230   -0.444    0.657   -0.126   -0.063
## notree.18 ~
##   rocks      0.674    0.237    2.844    0.004    0.674    0.340
##
## Covariances:
##           Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## .soil_microbe ~~
##   .TN_log      0.602    0.145    4.157    0.000    0.747    0.756
## .notree ~~
##   .TN_log      0.454    0.123    3.696    0.000    0.454    0.532
## .soil_microbe ~~
##   .notree      0.214    0.104    2.054    0.040    0.266    0.308
##
## Variances:
##           Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## .barc.richness  0.331    0.100    3.310    0.001    0.331    0.337
## .fungi.richness 0.849    0.157    5.398    0.000    0.849    0.863
## .chmhttrtrphy_lg 0.931    0.169    5.508    0.000    0.931    0.947
## .n.cycler_log    0.397    0.103    3.867    0.000    0.397    0.404
## .saprotroph      0.843    0.156    5.389    0.000    0.843    0.857
## .notree          0.746    0.134    5.568    0.000    0.746    0.699
## .TN_log          0.976    0.175    5.568    0.000    0.976    0.992
## .notree.18       0.870    0.156    5.568    0.000    0.870    0.885
## .soil_microbe    0.650    0.184    3.529    0.000    0.996    0.996
```

```
modindices(fit2, sort = TRUE, minimum.value = 3.5)
```

```
##           lhs op           rhs    mi    epc sepc.lv sepc.all sepc.nox
## 43 chemoheterotrophy_log ~~ TN_log 6.485 0.207    0.207    0.217    0.217
## 29      barc.richness ~~ n.cycler_log 4.776 0.303    0.303    0.835    0.835
## 38      fungi.richness ~~ TN_log 4.552 0.170    0.170    0.187    0.187
## 39      fungi.richness ~~ notree.18 4.096 0.224    0.224    0.260    0.260
```

```
# TN already covaries with soil microbiome latent variable;
# does not make sense to add path.
# Prior veg (notree.18) is too far removed to have a plausible
```

Model 3

Model for Supp Fig 10.

```
mod3 <- '
  # latent variables
  soil_microbe =~ barc.richness + fungi.richness + chemoheterotrophy_log +
  n.cyclor_log + saprotroph

  # structure
  herb ~ rocks + herb.18 + tree
  OM_log ~ rocks
  soil_microbe ~ rocks
  herb.18 ~ rocks

  # covariance
  OM_log ~~ soil_microbe
  OM_log ~~ herb
  soil_microbe ~~ herb
'

fit3 <- sem(mod3, data = sem.dat)
summary(fit3, fit.measures = TRUE, standardized = TRUE)
```

```
## lavaan 0.6.16 ended normally after 27 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of model parameters    22
##
##      Number of observations        62
##
## Model Test User Model:
##
##      Test statistic                21.703
##      Degrees of freedom            30
##      P-value (Chi-square)          0.865
##
## Model Test Baseline Model:
##
##      Test statistic                113.947
##      Degrees of freedom            44
##      P-value                       0.000
##
## User Model versus Baseline Model:
##
##      Comparative Fit Index (CFI)    1.000
##      Tucker-Lewis Index (TLI)      1.174
##
## Loglikelihood and Information Criteria:
##
##      Loglikelihood user model (H0)  -653.639
##      Loglikelihood unrestricted model (H1) -642.787
##
##      Akaike (AIC)                  1351.278
```

```

## Bayesian (BIC) 1398.075
## Sample-size adjusted Bayesian (SABIC) 1328.856
##
## Root Mean Square Error of Approximation:
##
## RMSEA 0.000
## 90 Percent confidence interval - lower 0.000
## 90 Percent confidence interval - upper 0.052
## P-value H_0: RMSEA <= 0.050 0.946
## P-value H_0: RMSEA >= 0.080 0.011
##
## Standardized Root Mean Square Residual:
##
## SRMR 0.069
##
## Parameter Estimates:
##
## Standard errors Standard
## Information Expected
## Information saturated (h1) model Structured
##
## Latent Variables:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## soil_microbe =~
## barc.richness 1.000 0.850 0.857
## fungi.richness 0.351 0.163 2.151 0.031 0.298 0.301
## chmhttrtrphy_lg 0.194 0.163 1.194 0.233 0.165 0.167
## n.cycler_log 0.899 0.189 4.762 0.000 0.764 0.770
## saprotroph 0.442 0.163 2.703 0.007 0.375 0.378
##
## Regressions:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## herb ~
## rocks -0.485 0.234 -2.070 0.038 -0.485 -0.250
## herb.18 0.391 0.102 3.839 0.000 0.391 0.400
## tree -0.102 0.106 -0.965 0.335 -0.102 -0.104
## OM_log ~
## rocks 0.249 0.250 0.996 0.319 0.249 0.126
## soil_microbe ~
## rocks -0.106 0.238 -0.443 0.658 -0.124 -0.062
## herb.18 ~
## rocks 0.334 0.248 1.347 0.178 0.334 0.169
##
## Covariances:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## .soil_microbe ~~
## .OM_log 0.371 0.130 2.867 0.004 0.438 0.445
## .herb ~~
## .OM_log 0.349 0.118 2.954 0.003 0.349 0.405
## .soil_microbe ~~
## .herb 0.268 0.112 2.404 0.016 0.316 0.361
##
## Variances:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all

```

```
##      .barc.richness      0.261      0.135      1.936      0.053      0.261      0.266
##      .fungi.richness     0.895      0.164      5.451      0.000      0.895      0.910
##      .chmhttrtrphy_lg    0.957      0.173      5.535      0.000      0.957      0.972
##      .n.cycler_log       0.400      0.126      3.172      0.002      0.400      0.407
##      .saprotroph         0.843      0.157      5.369      0.000      0.843      0.857
##      .herb               0.766      0.138      5.568      0.000      0.766      0.813
##      .OM_log             0.968      0.174      5.568      0.000      0.968      0.984
##      .herb.18            0.956      0.172      5.568      0.000      0.956      0.972
##      .soil_microbe       0.720      0.212      3.401      0.001      0.996      0.996
```

```
modindices(fit3, sort = TRUE, minimum.value = 3.5)
```

```
##      lhs op      rhs      mi      epc sepc.lv sepc.all sepc.nox
## 50 saprotroph ~~ OM_log 4.195 -0.21  -0.21  -0.232  -0.232
```

```
# OM already covaries with soil microbiome latent variable;
# does not make sense to add path for saprotrophs.
```

Model 4

Model for Fig 6.

```
mod4 <- '
# latent variables
soil_microbe =~ barc.richness + fungi.richness + chemoheterotrophy_log +
n.cycler_log + saprotroph

# structure
herb ~ rocks + herb.18 + tree
TN_log ~ rocks
soil_microbe ~ rocks
herb.18 ~ rocks

# covariance
TN_log ~~ soil_microbe
TN_log ~~ herb
soil_microbe ~~ herb
'

fit4 <- sem(mod4, data = sem.dat)
summary(fit4, fit.measures = TRUE, standardized = TRUE)
```

```
## lavaan 0.6.16 ended normally after 30 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of model parameters          22
##
##      Number of observations          62
##
## Model Test User Model:
##
##      Test statistic                  25.702
```

```

## Degrees of freedom          30
## P-value (Chi-square)       0.690
##
## Model Test Baseline Model:
##
## Test statistic             148.339
## Degrees of freedom         44
## P-value                    0.000
##
## User Model versus Baseline Model:
##
## Comparative Fit Index (CFI)      1.000
## Tucker-Lewis Index (TLI)        1.060
##
## Loglikelihood and Information Criteria:
##
## Loglikelihood user model (H0)     -638.442
## Loglikelihood unrestricted model (H1) -625.591
##
## Akaike (AIC)                  1320.884
## Bayesian (BIC)                 1367.681
## Sample-size adjusted Bayesian (SABIC) 1298.463
##
## Root Mean Square Error of Approximation:
##
## RMSEA                        0.000
## 90 Percent confidence interval - lower 0.000
## 90 Percent confidence interval - upper 0.077
## P-value H_0: RMSEA <= 0.050      0.847
## P-value H_0: RMSEA >= 0.080      0.042
##
## Standardized Root Mean Square Residual:
##
## SRMR                        0.066
##
## Parameter Estimates:
##
## Standard errors              Standard
## Information                  Expected
## Information saturated (h1) model Structured
##
## Latent Variables:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## soil_microbe =~
##   barc.richness    1.000
##   fungi.richness   0.451    0.167    2.711    0.007    0.366    0.369
##   chmhttrtrphy_lg  0.273    0.168    1.623    0.105    0.222    0.224
##   n.cycler_log     0.940    0.163    5.764    0.000    0.763    0.770
##   saprotroph       0.467    0.166    2.810    0.005    0.379    0.383
##
## Regressions:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## herb ~
##   rocks           -0.495    0.233   -2.121    0.034   -0.495   -0.253

```

```

##      herb.18          0.413    0.098    4.236    0.000    0.413    0.418
##      tree            -0.106    0.101   -1.043    0.297   -0.106   -0.107
##      TN_log ~
##      rocks           0.177    0.251    0.706    0.480    0.177    0.089
##      soil_microbe ~
##      rocks           -0.102    0.231   -0.439    0.660   -0.125   -0.063
##      herb.18 ~
##      rocks           0.334    0.248    1.347    0.178    0.334    0.169
##
## Covariances:
##              Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      .soil_microbe ~~
##      .TN_log          0.604    0.145    4.162    0.000    0.745    0.754
##      .herb ~~
##      .TN_log          0.442    0.123    3.589    0.000    0.442    0.512
##      .soil_microbe ~~
##      .herb            0.256    0.108    2.374    0.018    0.316    0.362
##
## Variances:
##              Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      .barc.richness    0.325    0.101    3.225    0.001    0.325    0.330
##      .fungi.richness   0.850    0.157    5.397    0.000    0.850    0.863
##      .chmhttrtrphy_lg  0.935    0.170    5.511    0.000    0.935    0.950
##      .n.cycler_log     0.401    0.104    3.873    0.000    0.401    0.408
##      .saprotroph       0.840    0.156    5.383    0.000    0.840    0.854
##      .herb             0.765    0.137    5.568    0.000    0.765    0.798
##      .TN_log           0.976    0.175    5.568    0.000    0.976    0.992
##      .herb.18          0.956    0.172    5.568    0.000    0.956    0.972
##      .soil_microbe     0.656    0.185    3.541    0.000    0.996    0.996

```

```
modindices(fit4, sort = TRUE, minimum.value = 3.5)
```

```

##              lhs op              rhs    mi    epc sepc.lv sepc.all sepc.nox
## 38      fungi.richness ~~      TN_log 7.073 0.220    0.220    0.241    0.241
## 29      barc.richness ~~ n.cycler_log 4.802 0.309    0.309    0.855    0.855
## 43 chemoheterotrophy_log ~~      TN_log 4.418 0.177    0.177    0.186    0.186

```

```

# TN already covaries with soil microbiome latent variable;
# does not make sense to add path for chemoheterotrophs or fungal richness.
# Effect of soil microbiome is modeled as a single latent variable,
# so bacterial/archeal richness and N-cyclers do not need to covary.

```