Data for:

Nicholas S. Marzolf

Updated March 13, 2023

Introduction

RMarkdown file to accompany *On the breakdown of woody debris across a groundwater gradient in Neotropical streams, Costa Rica,* submitted to the journal *Freshwater Science.*

To recreate this HTML, load the data files into a folder named 'Data' at the same path as the .rmd file, open the .rmd file, and click 'Knit'.

Load packages

```
# data manipulation
library(tidyverse)
library(readxl)
library(dplyr)
library(forcats)
library(purrr)

# plotting
library(ggplot2)
library(ggeffects)
library(grid)
library(ggpubr)
library(ggrepel)
library(ggcorrplot)
library(lemon)
library(signs)

# statistics
library(car)
library(drc)
library(nlme)
library(nlstools)
library(qpcR)

# spatial
library(sf)
library(raster)
library(spData)
library(cowplot)
```

```
library(leaflet)
library(spDataLarge)

# community
library(vegan)
library(pander)
library(lattice)
library(permute)
```

Load data

```
chem <- readxl::read_excel('Data/LT_CWD_datasheet.xlsx',
                           sheet = 'Chemistry')

cwd <- readxl::read_excel('Data/LT_CWD_datasheet.xlsx',
                          sheet = 'Sheet1')

table_s1 <- readxl::read_excel('Data/LT_CWD_datasheet.xlsx',
                               sheet = 'Wood Morphology')
readr::write_csv(table_s1,
                 'Data/table_s1.csv')
```

Stream chemistry for 5 sites

```
chem <- chem %>%
  dplyr::rename(srp = `SRP (ug/L)`,
                no3_n = `NO3-N (ug/L)`,
                nh4_n = `NH4-N (ug/L)`) %>%
  dplyr::mutate(site = case_when(Site == 'Arb' ~ 'Arboleda-30',
                                 Site == 'Sur30' ~ 'Sura-30',
                                 Site == 'Tito60' ~ 'Saltito-60',
                                 Site == 'Piper' ~ 'Piper-30',
                                 Site == 'Tac' ~ 'Taconazo-30'))

# calculate mean of all measurements
chem_sum <- chem %>%
  dplyr::group_by(site) %>%
  dplyr::summarise(dplyr::across(.cols = 3:8,
                                 .fns = mean,
                                 na.rm = TRUE))

# calculate SD from stream chemistry data
chem_sd <- chem %>%
  dplyr::group_by(site) %>%
  dplyr::summarise(dplyr::across(.cols = 3:8,
                                 .fns = sd,
                                 na.rm = TRUE))
```

```
# create object that sorts sites by decreasing mean conductivity
sites <- chem_sum %>%
  dplyr::group_by(site) %>%
  dplyr::summarise(mean_cond = mean(Cond, na.rm = TRUE)) %>%
  dplyr::arrange(desc(mean_cond))

# create a characeter vector of sites in order
sites <- as.character(sites$site)


# calculate DIN and N:P ratio
chem_sum <- chem_sum %>%
  dplyr::mutate(din = no3_n + nh4_n,
                n_p = (din/14.0067)/(srp/30.973762))


# re-level the site factor
chem_sum$site <- forcats::fct_relevel(chem_sum$site,
                                      sites)

table1 <- dplyr::arrange(chem_sum,
                         desc(Cond))
table1
```

```
## # A tibble: 5 x 9
##   site           srp no3_n nh4_n    pH  Cond  Temp   din    n_p
##   <fct>        <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  <dbl>
## 1 Arboleda-30 134.    193.  37.6  6.22 260.   24.8  231.   3.82
## 2 Sura-30      80.8   259   75.5  6.53 146.   24.8  334.   9.16
## 3 Saltito-60   32.1   211.  46.2  6.25  94.3  24.2  257.  17.7
## 4 Piper-30      7.37  292  113.   5.75  23.6  25.3  405. 122.
## 5 Taconazo-30   6.02  164.  52.4  5.73  18.4  24.8  216.  79.4
```

Figure 1: Map

```
streams_gps <- readr::read_csv('C:/Users/Nick Marzolf/Desktop/NCSU/STREAMS/La Selva GIS data/LTREB Data/
```

```
## Rows: 17 Columns: 4
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (1): Site
## dbl (3): Lat, Long, Elevation
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
cwd_coords <- streams_gps %>%
  dplyr::mutate(hyphen_site = gsub(' ', '-', Site)) %>%
  dplyr::mutate(Site = ifelse(Site == 'Piper',
                              'Piper-30',
```

```
                              hyphen_site),
               Elevation = ifelse(Site == 'Piper-30',
                                  30,
                                  Elevation)) %>%
  dplyr::filter(Site %in% sites) %>%
  dplyr::select(-hyphen_site) %>%
  sf::st_as_sf(.,
               coords = c('Long', 'Lat'),
               crs = st_crs(4326))

cwd_coords$Site <- forcats::fct_relevel(cwd_coords$Site,
                                        c('Arboleda-30', 'Sura-30', 'Saltito-60',
                                          'Piper-30', 'Taconazo-30'))

# La Selva boundary
lsbs <- sf::st_read(dsn = 'Data/Spatial/laselvaboundary.shp')
```

```
## Reading layer 'laselvaboundary' from data source
##   'C:\Users\Nick Marzolf\Desktop\NCSU\STREAMS\Projects\Long-term Wood decomp\LTCWD\Data\Spatial\lasel
##   using driver 'ESRI Shapefile'
## Simple feature collection with 4 features and 8 fields
## Geometry type: POLYGON
## Dimension:     XY
## Bounding box:  xmin: 823381.8 ymin: 1151045 xmax: 828917.3 ymax: 1156986
## Projected CRS: WGS 84 / UTM zone 16N
```

```
# Streams at La Selva shapefile
streams <- sf::st_read(dsn = 'Data/Spatial/streamsclip.shp')
```

```
## Reading layer 'streamsclip' from data source
##   'C:\Users\Nick Marzolf\Desktop\NCSU\STREAMS\Projects\Long-term Wood decomp\LTCWD\Data\Spatial\strea
##   using driver 'ESRI Shapefile'
## Simple feature collection with 521 features and 10 fields
## Geometry type: MULTILINESTRING
## Dimension:     XY
## Bounding box:  xmin: 823758 ymin: 1151440 xmax: 828668.2 ymax: 1156895
## Projected CRS: WGS 84 / UTM zone 16N
```

```
# shapefile for Costa Rica
cr <- world %>%
  dplyr::filter(name_long == 'Costa Rica')

# make site map, with 5 streams
sites_map <- ggplot()+
  geom_sf(data = lsbs,
          fill = 'white',
          color = 'black')+
  geom_sf(data = streams,
          color = 'blue')+
  geom_sf(data = cwd_coords,
          size = 5,
          aes(color = Site))+
```

```
  coord_sf(xlim = c(-84.05, -83.98),
           ylim = c(10.395, 10.45),
           crs = 4326)+
  scale_color_viridis_d()+
  ggspatial::annotation_scale(location = "bl",
                              bar_cols = c("black", "white"))+
  ggspatial::annotation_north_arrow(location = "tl",
                                    which_north = "true",
                                    style = ggspatial::north_arrow_minimal(line_col = "black"))

# create a centroid for La Selva
centroid <- sf::st_centroid(lsbs[1,])

# create the inset map
inset <- ggplot()+
  geom_sf(data = cr)+
  geom_sf(data = centroid)+
  scale_x_continuous(breaks = -84)+
  scale_y_continuous(breaks = 10)

# make the full figure
fig1 <- ggdraw()+
  draw_plot(sites_map)+
  draw_plot(inset, x = .6, y = 0.1, height = 0.25, width = 0.25)+
  theme()

# save the figure
ggsave(plot = fig1,
       filename = 'Figures/fig1_ggplot.png',
       dpi = 600,
       width = 9.85, height = 7)
```

Calculate wood decomposition rates

```
# begin cleaning data
cwd <- cwd %>%
  dplyr::mutate(site = case_when(Site == 'Arb' ~ 'Arboleda-30',
                                 Site == 'Sur30' ~ 'Sura-30',
                                 Site == 'Tito60' ~ 'Saltito-60',
                                 Site == 'Piper' ~ 'Piper-30',
                                 Site == 'Tac' ~ 'Taconazo-30')) %>%
  dplyr::filter(Flag == 0) %>%
  dplyr::select(site,
                month = `Collection Month`,
                rep = Rep,
                init_mass = `initial CWD mass (g)`,
                dry_mass = `CWD Pack Dry Mass (g)`,
                init_den = `init wood density (g/cm3)`,
                fin_den = `final wood density (g/cm3)`) %>%
  dplyr::filter(month < 24)
```

```r
# re-level site factor based on conductivity
cwd$site <- forcats::fct_relevel(cwd$site,
                                 sites)

# calculate % dry mass remaining
cwd_calc <- cwd %>%
  dplyr::filter(site != 'Sac') %>%
  dplyr::mutate(percent_mass = (dry_mass/init_mass)*100)

# dry mass of each pack to be used in the bug analysis
final_dry_mass <- cwd_calc %>%
  dplyr::select(site, month, rep, dry_mass) %>%
  dplyr::filter(month != 0)

# calculate decay rates using ANCOVA
k_cwd_int <- lm(data = cwd_calc,
                log(percent_mass) ~ month * site)

summary(k_cwd_int)
```

```
##
## Call:
## lm(formula = log(percent_mass) ~ month * site, data = cwd_calc)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.41381 -0.13079  0.01234  0.12742  0.39291
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              4.45289    0.07116  62.577  < 2e-16 ***
## month                   -0.06407    0.01154  -5.551 7.41e-07 ***
## siteSura-30              0.07091    0.10063   0.705  0.48383
## siteSaltito-60          -0.02906    0.10098  -0.288  0.77457
## sitePiper-30            -0.17348    0.10098  -1.718  0.09112 .
## siteTaconazo-30          0.12792    0.10937   1.170  0.24695
## month:siteSura-30        0.04731    0.01632   2.898  0.00529 **
## month:siteSaltito-60     0.04650    0.01637   2.841  0.00619 **
## month:sitePiper-30       0.04914    0.01637   3.002  0.00395 **
## month:siteTaconazo-30    0.04704    0.01683   2.795  0.00703 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.193 on 58 degrees of freedom
## Multiple R-squared:  0.539,  Adjusted R-squared:  0.4675
## F-statistic: 7.536 on 9 and 58 DF,  p-value: 3.288e-07
```

```r
anova(k_cwd_int)
```

```
## Analysis of Variance Table
##
## Response: log(percent_mass)
##            Df  Sum Sq Mean Sq F value    Pr(>F)
```

```
## month       1 0.91680 0.91680 24.6079 6.478e-06 ***
## site        4 1.10744 0.27686  7.4312 6.658e-05 ***
## month:site  4 0.50252 0.12563  3.3721   0.01505 *
## Residuals  58 2.16087 0.03726
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
car::Anova(k_cwd_int, type = 'III')
```

```
## Anova Table (Type III tests)
##
## Response: log(percent_mass)
##              Sum Sq Df   F value    Pr(>F)
## (Intercept) 145.894  1 3915.9413 < 2.2e-16 ***
## month         1.148  1   30.8088 7.408e-07 ***
## site          0.348  4    2.3354   0.06609 .
## month:site    0.503  4    3.3721   0.01505 *
## Residuals     2.161 58
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

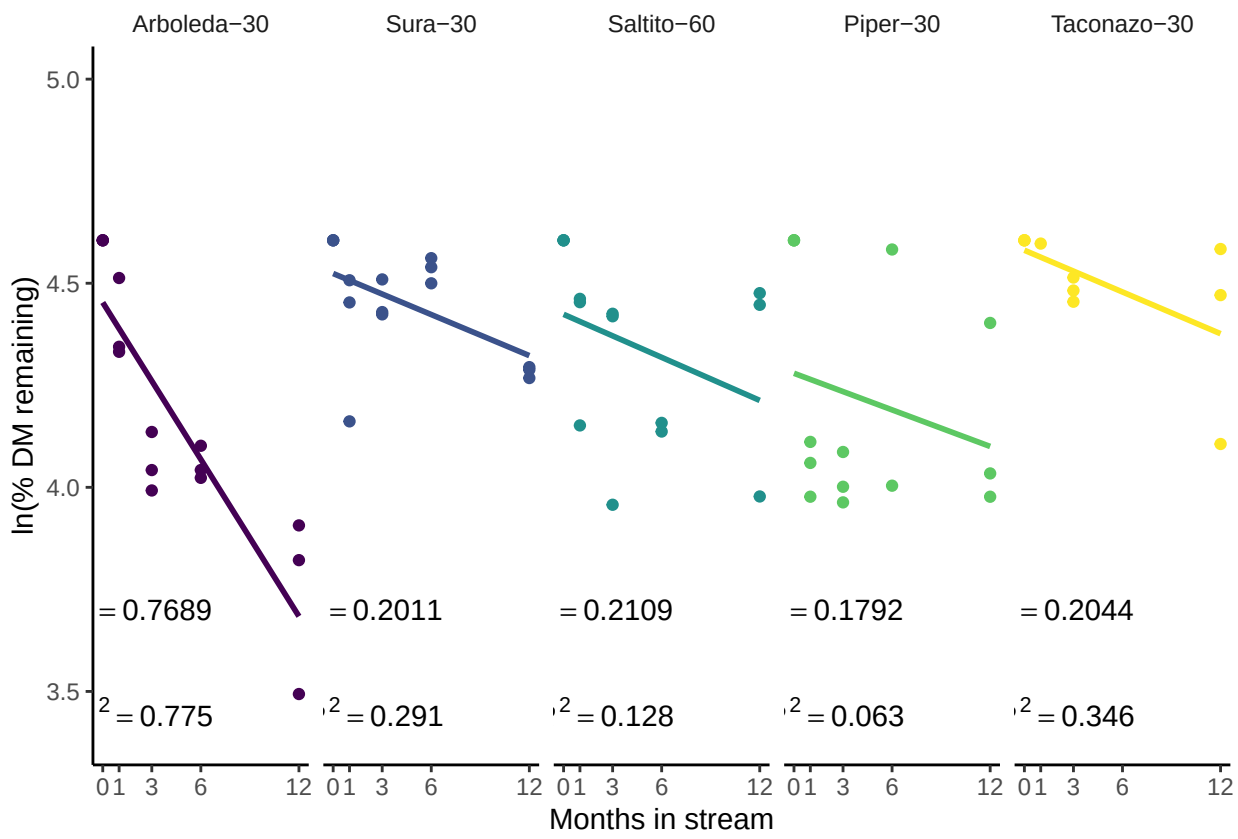
```
table2 <- cwd_calc %>%
  dplyr::filter(site != 'Sac') %>%
  dplyr::group_by(site) %>%
  dplyr::summarise(
    int = round((coef(lm(log(percent_mass) ~ month))[1]), 3),
    k_yr = (round((coef(lm(log(percent_mass) ~ month))[2])*12, 4))*-1,
    error = summary(lm(log(percent_mass) ~ month))$coefficient[3],
    #df = summary(lm(log(percent_mass) ~ month))$fstatistic,
    r2 = round(summary(lm(log(percent_mass) ~ month))$r.squared, 3),
    p = anova(lm(log(percent_mass) ~ month))$'Pr(>F)'[1])
table2
```

```
## # A tibble: 5 x 6
##   site          int  k_yr  error    r2         p
##   <fct>       <dbl> <dbl>  <dbl> <dbl>     <dbl>
## 1 Arboleda-30  4.45 0.769 0.0591 0.775 0.0000151
## 2 Sura-30      4.52 0.201 0.0447 0.291 0.0379
## 3 Saltito-60   4.42 0.211 0.0818 0.128 0.209
## 4 Piper-30     4.28 0.179 0.103  0.063 0.386
## 5 Taconazo-30  4.58 0.204 0.0561 0.346 0.0735
```

```
cwd_plot <- dplyr::left_join(cwd_calc,
                             table2,
                             by = 'site')

fig2 <- ggplot(data = cwd_plot,
               aes(y = log(percent_mass),
                   x = month))+
  geom_point(aes(color = site))+
  geom_smooth(aes(color = site),
              method = 'lm',
```

```
              se = FALSE)+
  lims(y = c(3.4, 5))+
  geom_text(data = table2 %>%
              dplyr::mutate(label = paste("italic(k) == ", k_yr)),
            aes(x = 2.5, y = 3.7,
                label = label),
            parse = TRUE)+
  geom_text(data = table2 %>%
              dplyr::mutate(label = paste("italic('R'^2) == ", r2)),
            aes(x = 2.5,
                y = 3.45,
                label = label),
            parse = TRUE,
            inherit.aes = FALSE)+
  facet_grid(. ~ site)+
  labs(x = 'Months in stream',
       y = 'ln(% DM remaining)')+
  scale_color_viridis_d()+
  scale_x_continuous(breaks = c(0, 1, 3, 6, 12))+
  theme_classic()+
  theme(legend.position = 'none',
        panel.grid = element_blank(),
        strip.background = element_blank())
fig2
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
ggsave(plot = fig2,
       'Figures/fig2.png',
       dpi = 600,
       width = 10,height = 3)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
cwd_calc %>%
  group_by(site, month) %>%
  summarise(mean_per_mass = mean(percent_mass, na.rm = TRUE))
```

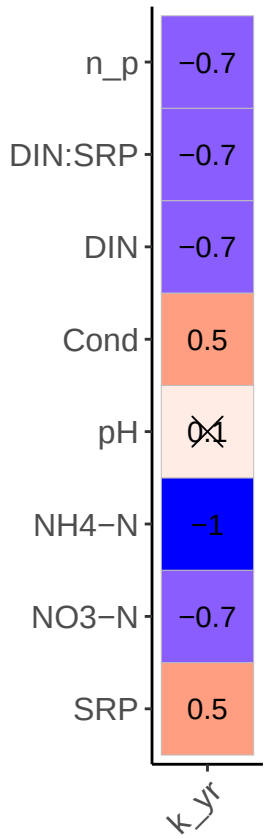
```
## `summarise()` has grouped output by 'site'. You can override using the
## `.groups` argument.
```

```
## # A tibble: 24 x 3
## # Groups:   site [5]
##    site        month mean_per_mass
##    <fct>       <dbl>         <dbl>
##  1 Arboleda-30     0         100
##  2 Arboleda-30     1          81.4
##  3 Arboleda-30     3          57.9
##  4 Arboleda-30     6          57.8
##  5 Arboleda-30    12          42.8
##  6 Sura-30         0         100
```

```
##  7 Sura-30         1          80.2
##  8 Sura-30         3          86.1
##  9 Sura-30         6          93.1
## 10 Sura-30        12          72.5
## # i 14 more rows
```

Merge chemistry with decay rates

```
merged <- table2 %>%
  dplyr::select(site, k_yr, error) %>%
  dplyr::right_join(chem_sum, 'site') %>%
  # dplyr::select(-Temp) %>%
  tidyr::pivot_longer(srp:n_p)

names_long <- c(`Cond` = 'Conductivity (µS/cm)',
                `din` = 'DIN (µg/L)',
                `n_p` = 'DIN:SRP',
                `nh4_n` = 'NH4-N (µg/L)',
                `no3_n` = 'NO3-N (µg/L)',
                `pH` = 'pH',
                `srp` = 'SRP (µg/L)')

wide <- table2 %>%
  dplyr::select(site, k_yr, error) %>%
  dplyr::right_join(chem_sum, 'site')

cor <- cor(wide %>%
             dplyr::select(-site, -error) %>%
             mutate_if(is.character, as.numeric),
           method = 'spearman')

cor_sub <- cor[-c(2:nrow(cor)), 2:ncol(cor), drop = FALSE]
rownames(cor_sub)[1] <- 'k_yr'
colnames(cor_sub)[1:7] <- c('SRP', 'NO3-N', 'NH4-N', 'pH', 'Cond', 'DIN', 'DIN:SRP')

pmat <- ggcorrplot::cor_pmat(cor,
                             'spearman',
                             alternative = 'two.sided')
pmat_sub <- pmat[-c(2:nrow(pmat)), 2:ncol(pmat),drop = FALSE]

ggcorrplot(cor_sub, lab = TRUE,
           ggtheme = theme_classic(),show.legend = FALSE,p.mat = pmat_sub)
```



Macroinvertebrate analysis

```
data <- readxl::read_excel('Data/Samples_Nick_ana_2024.xlsx',
                           sheet = 'Data') %>%
  dplyr::mutate(Stream = case_when(Stream == 'Arboleda 30' ~ 'Arboleda-30',
                                   Stream == 'Sura 30' ~ 'Sura-30',
                                   Stream == 'Saltito 60' ~ 'Saltito-60',
                                   Stream == 'Piper' ~ 'Piper-30',
                                   Stream == 'Taconazo 30' ~ 'Taconazo-30'),
                functional_group = case_when(`Functional group` == "Predator" ~ 'Predators',
                                             `Functional group` == "Filters" ~ 'Filterers',
                                             `Functional group` == "Collectors-Gatherers" ~ "Collector-
                                             `Functional group` == "Shredder plant tissue" ~ "Shredder p
                                             `Functional group` == "Shredder detritus" ~ "Shredder detri
                                             `Functional group` == "Scrapers" ~ "Scrapers",
                                             `Functional group` == "Piercers" ~ "Piercers",
                                             `Functional group` == "Miners" ~ "Miners",
                                             `Functional group` == "Shredder detritus" ~ "Shredder detri
                                             `Functional group` == "Endoparasite" ~ "Endoparasite"))

data$Stream <- forcats::fct_relevel(data$Stream,
                                    c('Arboleda-30', 'Sura-30', 'Saltito-60',
                                      'Piper-30', 'Taconazo-30'))
```

```r
# clean the data
tax_data_clean <- data %>%
  dplyr::mutate(month = as.numeric(gsub(".*?([0-9]+).*", "\\1", Sample)),
                rep = gsub("\\d+", "", Sample)) %>%
  dplyr::rename(ffg = functional_group) %>%
  dplyr::filter(QAQC == 0,
                month != 24)

fig_s1 <- tax_data_clean %>%
  filter(!is.na(ffg)) %>%
  group_by(ffg) %>%
  summarise(total_ffg = sum(Total, na.rm = TRUE)) %>%
  ggplot(.,
         aes(x = ffg,
             y = total_ffg))+
  geom_bar(stat = 'identity')+
  geom_label(aes(label = total_ffg),nudge_y = 0.15)+
  scale_y_log10()+
  labs(x = element_blank(),
       y = 'Total Abundance')+
  coord_flip()+
  theme(axis.text.x = element_text(size = 8))

ggsave(plot = fig_s1,
       'Figures/fig_s1.png',
       dpi = 600,
       height = 4, width = 9)

tot_macros <- sum(tax_data_clean$Total)
threshold <- tot_macros*0.01

# main FFGS: at least 1% of all FFGs identified
main_ffgs <- tax_data_clean %>%
  dplyr::group_by(ffg) %>%
  dplyr::summarise(n = sum(Total)) %>%
  dplyr::filter(n > threshold) %>%
  dplyr::pull(ffg)
```

```r
final_dry_mass <- readr::read_csv('Data/final_dry_mass.csv')
```

```
## Rows: 53 Columns: 4
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (2): site, rep
## dbl (2): month, dry_mass
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
final_dry_mass$site <- dplyr::recode_factor(final_dry_mass$site,
                                            Arb = 'Arboleda-30',
                                            Sur30 = 'Sura-30',
```

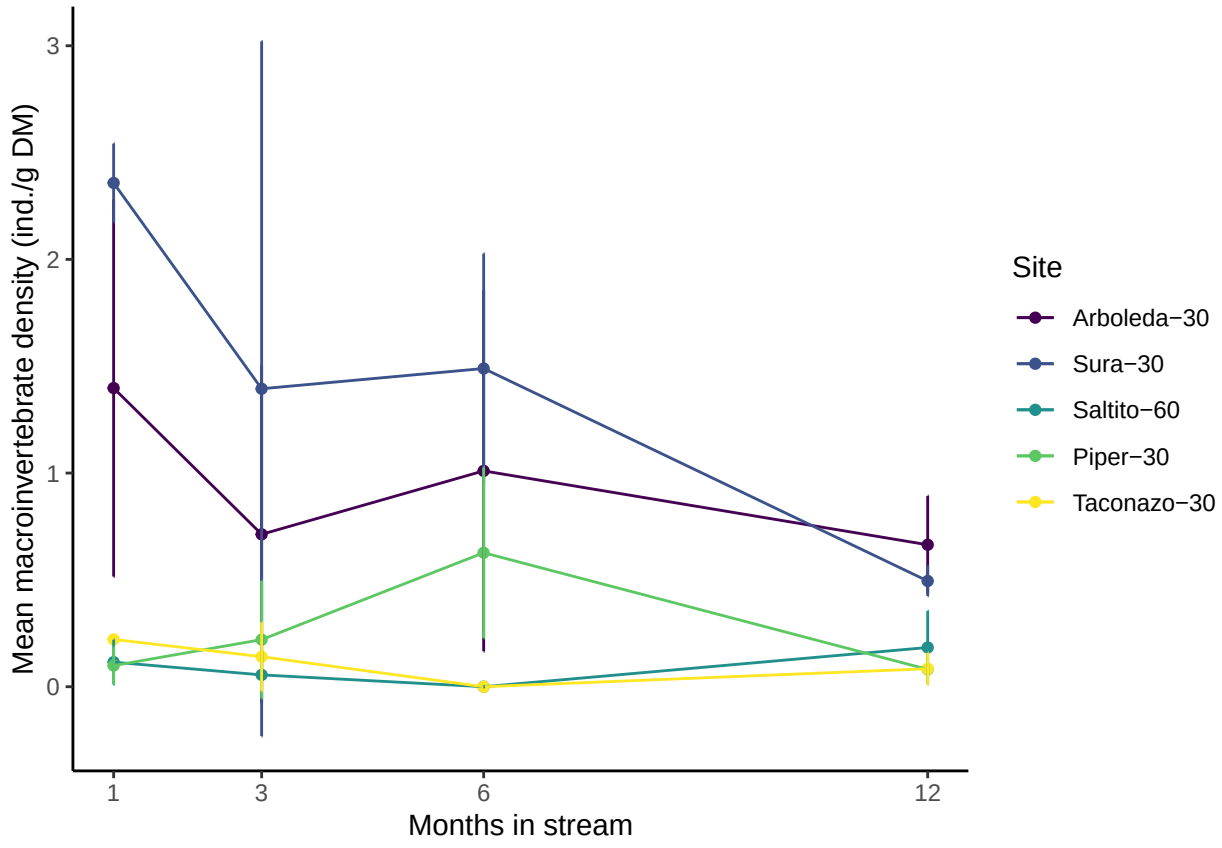
```
                                          Tito60 = 'Saltito-60',
                                          Piper = 'Piper-30',
                                          Tac = 'Taconazo-30')


macro_density <- tax_data_clean %>%
  dplyr::filter(!is.na(order)) %>%                              # remove NAs from Order
  dplyr::select(site = Stream, Family, month, rep, Total) %>%     # get the necessary columns
  dplyr::group_by(site, month, rep) %>%                         # and do the grouping
  dplyr::summarise(total_family = sum(Total)) %>%
  dplyr::left_join(final_dry_mass,
                   by = c('site', 'month', 'rep')) %>%
  dplyr::mutate(abund_per_dm = total_family/dry_mass)
```

```
## `summarise()` has grouped output by 'site', 'month'. You can override using the
## `.groups` argument.
```

```
macro_density_sum <- macro_density %>%
  group_by(site, month) %>%
  summarise(mean_den = mean(abund_per_dm, na.rm = TRUE),
            se_den = sd(abund_per_dm, na.rm = TRUE)/length(mean_den))
```

```
## `summarise()` has grouped output by 'site'. You can override using the
## `.groups` argument.
```

```
fig3 <- ggplot(macro_density_sum %>%
                 mutate(across(everything(), ~ ifelse(is.nan(.), 0, .))),
               aes(x = month,
                   y = mean_den,
                   color = site,
                   group = site))+
  geom_point()+
  geom_line()+
  geom_errorbar(aes(ymin = mean_den - se_den,
                    ymax = mean_den + se_den),
                width = 0)+
  scale_x_continuous(breaks = c(0, 1, 3, 6, 12))+
  labs(y = 'Mean macroinvertebrate density (ind./g DM)',
       x = 'Months in stream')+
  scale_color_viridis_d(name = 'Site')+
  theme(axis.title = element_text(size = 12),
        panel.grid = element_blank())+
  theme_classic()
fig3
```




```
ggsave(plot = fig3,
       'Figures/fig3.png',
       dpi = 600,
       width = 6, height = 4)


macro_density_mod <- macro_density %>%
  dplyr::mutate(log_abund = log10(abund_per_dm)) %>%
  dplyr::filter(is.finite(log_abund))

macro_lm <- lm(data = macro_density_mod %>%
                 dplyr::filter(is.finite(abund_per_dm)),
               log_abund ~ site * factor(month))

summary(macro_lm)
```

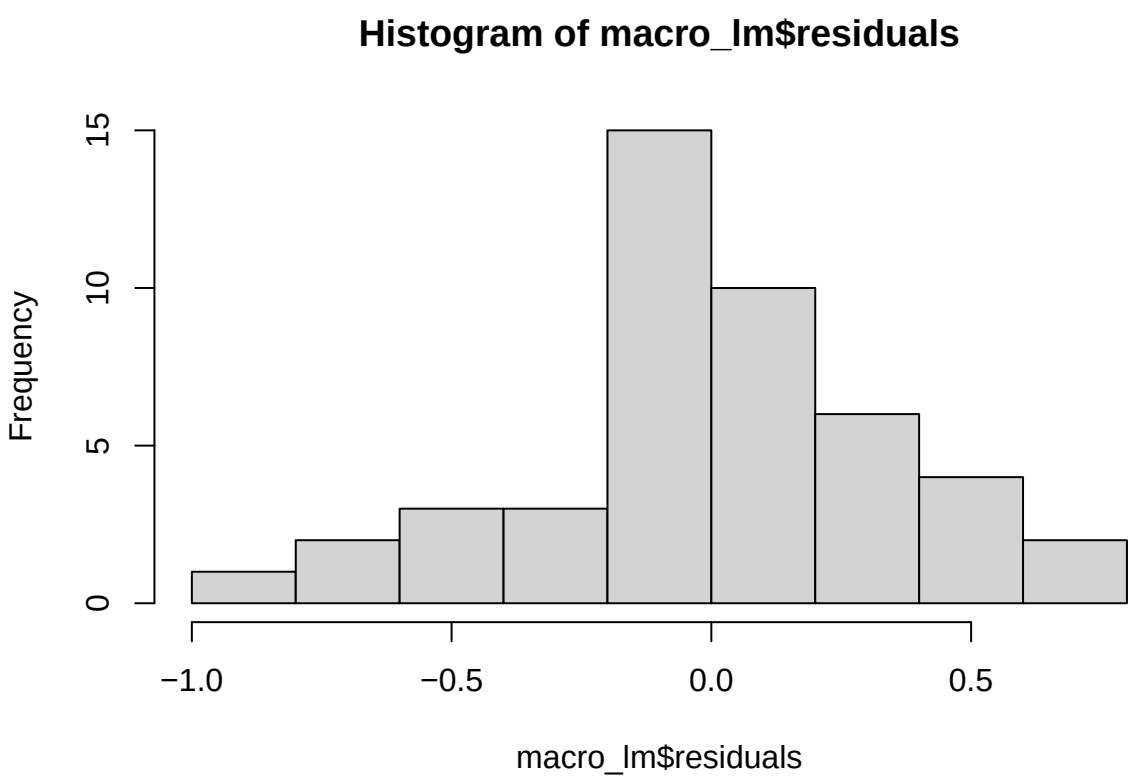
```
##
## Call:
## lm(formula = log_abund ~ site * factor(month), data = macro_density_mod %>%
##     dplyr::filter(is.finite(abund_per_dm)))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.89963 -0.13689 -0.00345  0.20023  0.62367
##
```

```
## Coefficients: (2 not defined because of singularities)
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                      0.093209   0.263905   0.353  0.72659
## siteSura-30                      0.278458   0.373218   0.746  0.46182
## siteSaltito-60                  -0.864065   0.417270  -2.071  0.04771 *
## sitePiper-30                    -1.200646   0.417270  -2.877  0.00759 **
## siteTaconazo-30                 -0.748172   0.527809  -1.418  0.16737
## factor(month)3                  -0.467030   0.373218  -1.251  0.22116
## factor(month)6                  -0.376372   0.373218  -1.008  0.32187
## factor(month)12                 -0.287686   0.373218  -0.771  0.44727
## siteSura-30:factor(month)3       0.002746   0.527809   0.005  0.99589
## siteSaltito-60:factor(month)3    0.123994   0.590109   0.210  0.83510
## sitePiper-30:factor(month)3      0.593267   0.590109   1.005  0.32334
## siteTaconazo-30:factor(month)3   0.004342   0.646432   0.007  0.99469
## siteSura-30:factor(month)6       0.159771   0.527809   0.303  0.76435
## siteSaltito-60:factor(month)6          NA         NA      NA       NA
## sitePiper-30:factor(month)6      1.231915   0.590109   2.088  0.04605 *
## siteTaconazo-30:factor(month)6         NA         NA      NA       NA
## siteSura-30:factor(month)12     -0.392632   0.527809  -0.744  0.46314
## siteSaltito-60:factor(month)12   0.198692   0.590109   0.337  0.73885
## sitePiper-30:factor(month)12     0.297927   0.559826   0.532  0.59880
## siteTaconazo-30:factor(month)12 -0.372137   0.646432  -0.576  0.56943
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4571 on 28 degrees of freedom
## Multiple R-squared:  0.6796, Adjusted R-squared:  0.4851
## F-statistic: 3.494 on 17 and 28 DF,  p-value: 0.001663
```

```
anova(macro_lm)
```

```
## Analysis of Variance Table
##
## Response: log_abund
##                    Df Sum Sq Mean Sq F value    Pr(>F)
## site                4 9.7470 2.43674 11.6625 1.063e-05 ***
## factor(month)       3 1.2218 0.40727  1.9492    0.1446
## site:factor(month) 10 1.4418 0.14418  0.6901    0.7249
## Residuals          28 5.8502 0.20894
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
hist(macro_lm$residuals)
```




```
(agricolae::HSD.test(macro_lm, trt = 'site'))
```

```
## $statistics
##     MSerror Df       Mean        CV
##   0.2089371 28 -0.5067934 -90.19384
##
## $parameters
##    test name.t ntr StudentizedRange alpha
##   Tukey   site   5           4.1203  0.05
##
## $means
##               log_abund       std  r        se        Min         Max
## Arboleda-30 -0.18956334 0.4697270 12 0.1319524 -1.1827949  0.38288383
## Piper-30    -0.88585069 0.5085039  9 0.1523655 -1.5797836 -0.04139269
## Saltito-60  -0.91486584 0.2944634  6 0.1866088 -1.2833012 -0.51648452
## Sura-30      0.03136597 0.3699442 12 0.1319524 -0.6345892  0.51224243
## Taconazo-30 -1.13603949 0.5886647  7 0.1727662 -2.0813473 -0.49398056
##                    Q25         Q50        Q75
## Arboleda-30 -0.3458852 -0.05760492  0.1306942
## Piper-30    -1.1949766 -1.02015403 -0.4623980
## Saltito-60  -1.1385314 -0.89888760 -0.7296380
## Sura-30     -0.2647593  0.07216143  0.3516841
## Taconazo-30 -1.4294868 -1.05744874 -0.7302632
##
## $comparison
```

```
## NULL
##
## $groups
##                log_abund groups
## Sura-30      0.03136597      a
## Arboleda-30 -0.18956334      a
## Piper-30    -0.88585069      b
## Saltito-60  -0.91486584      b
## Taconazo-30 -1.13603949      b
##
## attr(,"class")
## [1] "group"
```

```
emmeans::emmeans(macro_lm, ~site*month)
```

```
##  site         month  emmean    SE df lower.CL upper.CL
##  Arboleda-30      1  0.0932 0.264 28   -0.447    0.634
##  Sura-30          1  0.3717 0.264 28   -0.169    0.912
##  Saltito-60       1 -0.7709 0.323 28   -1.433   -0.109
##  Piper-30         1 -1.1074 0.323 28   -1.770   -0.445
##  Taconazo-30      1 -0.6550 0.457 28   -1.591    0.281
##  Arboleda-30      3 -0.3738 0.264 28   -0.914    0.167
##  Sura-30          3 -0.0926 0.264 28   -0.633    0.448
##  Saltito-60       3 -1.1139 0.323 28   -1.776   -0.452
##  Piper-30         3 -0.9812 0.323 28   -1.643   -0.319
##  Taconazo-30      3 -1.1177 0.264 28   -1.658   -0.577
##  Arboleda-30      6 -0.2832 0.264 28   -0.824    0.257
##  Sura-30          6  0.1551 0.264 28   -0.386    0.696
##  Saltito-60       6  nonEst    NA NA       NA       NA
##  Piper-30         6 -0.2519 0.323 28   -0.914    0.410
##  Taconazo-30      6  nonEst    NA NA       NA       NA
##  Arboleda-30     12 -0.1945 0.264 28   -0.735    0.346
##  Sura-30         12 -0.3087 0.264 28   -0.849    0.232
##  Saltito-60      12 -0.8598 0.323 28   -1.522   -0.198
##  Piper-30        12 -1.0972 0.264 28   -1.638   -0.557
##  Taconazo-30     12 -1.3148 0.264 28   -1.855   -0.774
##
## Confidence level used: 0.95
```

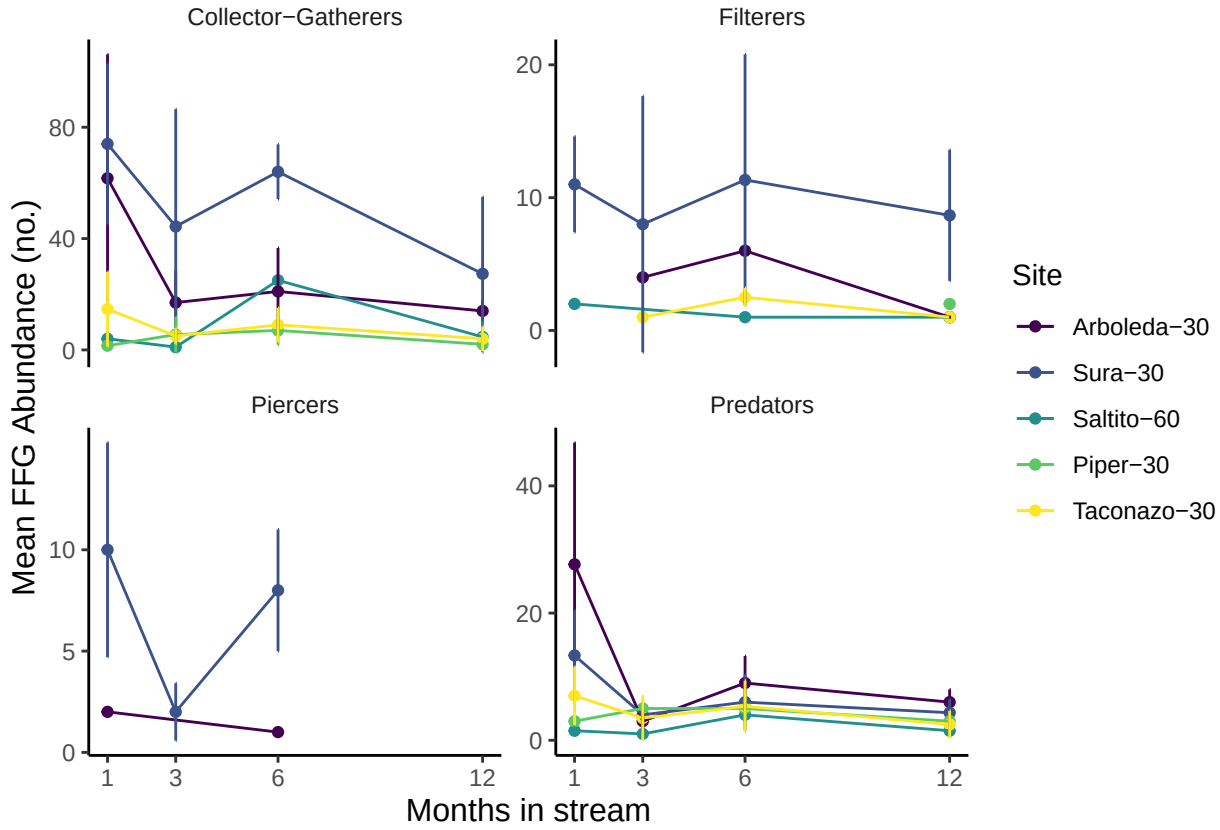
```
macro_density_ffg <- tax_data_clean %>%
  dplyr::filter(!is.na(order)) %>%                              # remove NAs from Order
  dplyr::select(site = Stream, ffg, month, rep, Total) %>%     # get the necessary columns
  dplyr::group_by(site, month, ffg, rep) %>%                        # and do the grouping
  dplyr::summarise(total_family = sum(Total))
```

```
## `summarise()` has grouped output by 'site', 'month', 'ffg'. You can override
## using the `.groups` argument.
```

```
macro_density_sum_ffg <- macro_density_ffg %>%
  group_by(site, month, ffg) %>%
  summarise(mean_abund = mean(total_family, na.rm = TRUE),
            se_abund = sd(total_family, na.rm = TRUE)/length(mean_abund))
```

```
## `summarise()` has grouped output by 'site', 'month'. You can override using the
## `.groups` argument.
```

```
fig4 <- ggplot(macro_density_sum_ffg %>%
                 dplyr::filter(!is.na(ffg),
                               ffg %in% main_ffgs),
               aes(x = month,
                   y = mean_abund,
                   color = site))+
  geom_point()+
  geom_line()+
  geom_errorbar(aes(ymin = mean_abund - se_abund,
                    ymax = mean_abund + se_abund),
                width = 0)+
  # geom_boxplot()+
  facet_wrap(ffg ~ .,
             scales = 'free_y')+
  scale_color_viridis_d(name = 'Site')+
  scale_x_continuous(breaks = c(1, 3, 6, 12))+
  scale_y_continuous(breaks = function(y) {
    max_value <- max(y, na.rm = TRUE)
    tick_marks <- seq(-0, round(max_value*.75, -1), length.out = 3)
    return(tick_marks)})+
  labs(y = 'Mean FFG Abundance (no.)',
       x = 'Months in stream')+
  theme_classic()+
  theme(axis.title = element_text(size = 12),
        panel.grid = element_blank(),
        strip.background = element_blank(),
        axis.line.x = element_line(color = 'black'))
fig4
```




```
# fig4_b <- shift_legend2(fig4_b)

ggsave(plot = fig4,
       'Figures/fig4.png',
       dpi = 600,
       width = 6, height = 5)
```

```
tax_data_clean_family_log <- tax_data_clean %>%
  dplyr::filter(!is.na(Family)) %>%                                  # remove NAs from Order
  dplyr::select(Stream, Family, month, rep, Total) %>%    # get the necessary columns
  dplyr::group_by(Stream, month, Family) %>%          # and do the grouping
  dplyr::summarise(mean_family = mean(Total, na.rm = TRUE)) %>%   # sum by order in each possible group
  dplyr::mutate(log_total = log10(1 + mean_family)) %>%          # log10 + 1 transform data
  dplyr::select(-mean_family) %>%
  tidyr::pivot_wider(names_from = Family,                            # pivot data
                     values_from = log_total,
                     values_fill = 0)
```

```
## `summarise()` has grouped output by 'Stream', 'month'. You can override using
## the `.groups` argument.
```

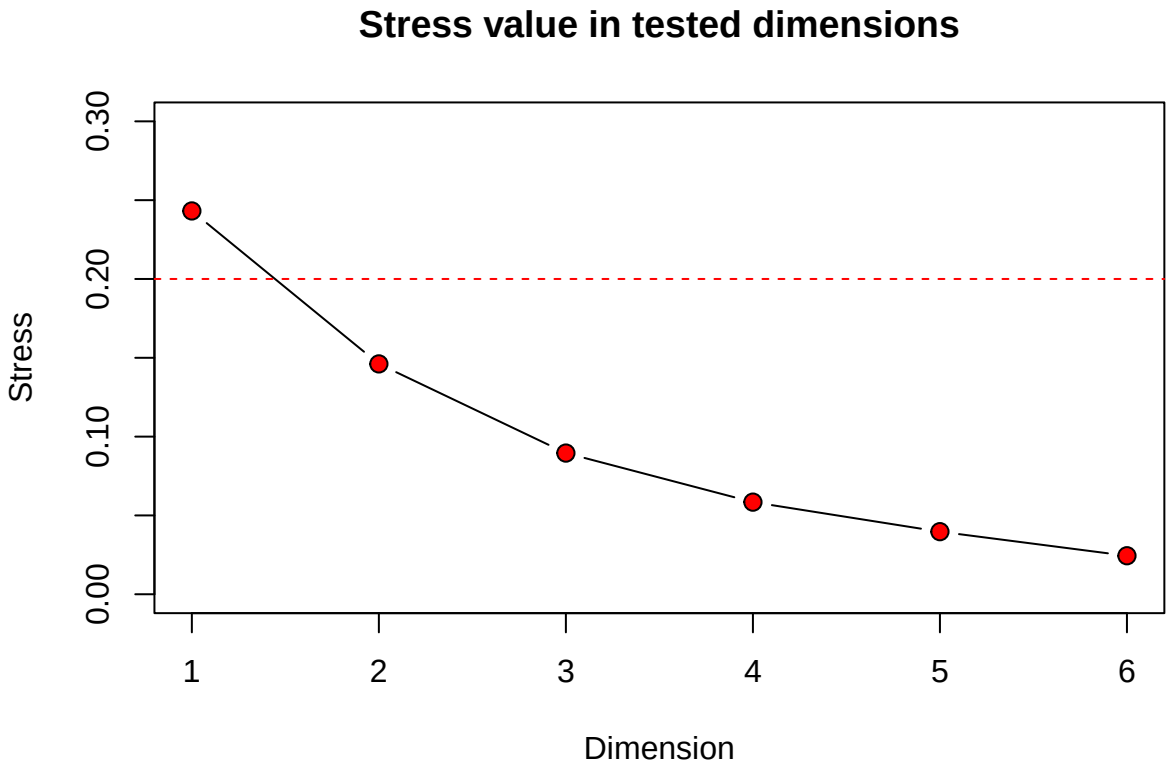
```
tax_data_log_matrix_fam <- tax_data_clean_family_log[,-c(1:2)]
tax_data_log_meta_fam <- tax_data_clean_family_log[,c(1:2)]

nmds_family_log_2d <- vegan::metaMDS(tax_data_log_matrix_fam,
```

```
                                    distance = 'bray',
                                    k = 2,
                                    autotransform = FALSE)
```

```
## Run 0 stress 0.1466003
## Run 1 stress 0.1566809
## Run 2 stress 0.1573421
## Run 3 stress 0.1537507
## Run 4 stress 0.1534708
## Run 5 stress 0.1537933
## Run 6 stress 0.155521
## Run 7 stress 0.1538881
## Run 8 stress 0.1526783
## Run 9 stress 0.1514922
## Run 10 stress 0.1527885
## Run 11 stress 0.1552138
## Run 12 stress 0.1466008
## ... Procrustes: rmse 0.0004490266  max resid 0.001619462
## ... Similar to previous best
## Run 13 stress 0.1534708
## Run 14 stress 0.1535372
## Run 15 stress 0.1461227
## ... New best solution
## ... Procrustes: rmse 0.0385572  max resid 0.1471893
## Run 16 stress 0.1561415
## Run 17 stress 0.1517733
## Run 18 stress 0.151375
## Run 19 stress 0.1537981
## Run 20 stress 0.1573421
## *** Best solution was not repeated -- monoMDS stopping criteria:
##     20: stress ratio > sratmax
```

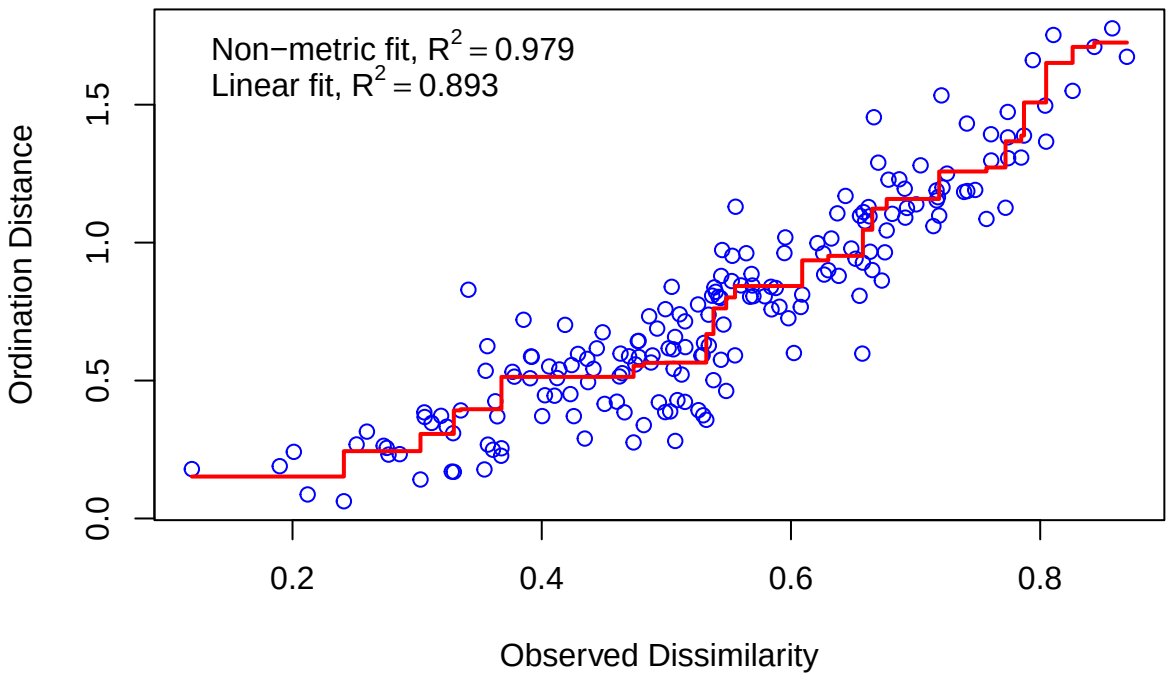
```
stress_fam_log_2d <- round(nmds_family_log_2d$stress, 3)

# evaluate the NMDS via screeplot and Shepard plot
scree <- goeveg::screeplot_NMDS(tax_data_log_matrix_fam,
                                distance = 'bray',
                                autotransform = FALSE)
```

```
##   |                                                                      |
```



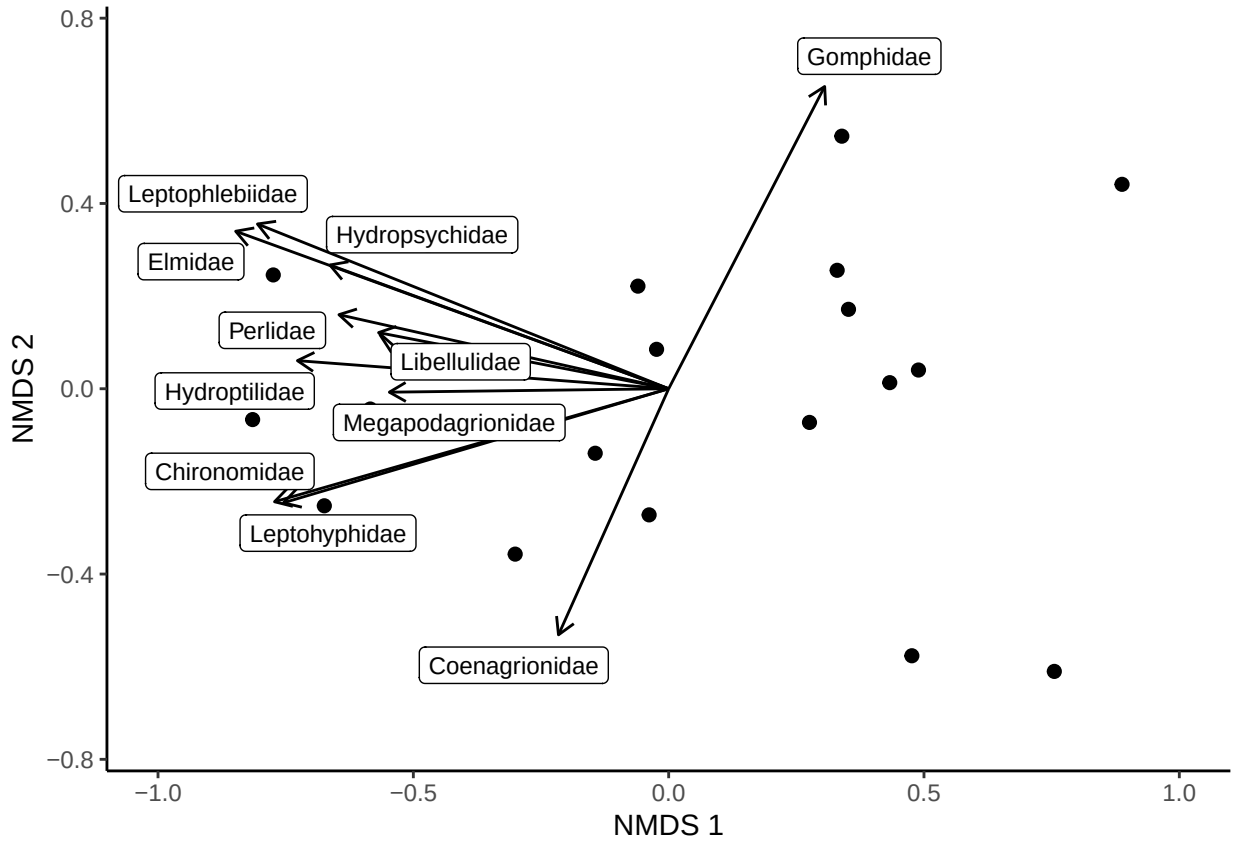

```
##     1     2     3     4     5     6
## 0.243 0.146 0.090 0.058 0.040 0.024
```

```
df_scree <- tibble::rownames_to_column(data.frame(scree))

scree_plot <- ggplot(df_scree,
                     aes(x = rowname,
                         y = scree))+
  geom_bar(stat = 'identity')+
  geom_label(aes(label = scree))+
  labs(x = 'Dimensions',
       y = 'Stress')+
  theme_classic()

shepard <- data.frame(vegan::stressplot(nmds_family_log_2d))
```



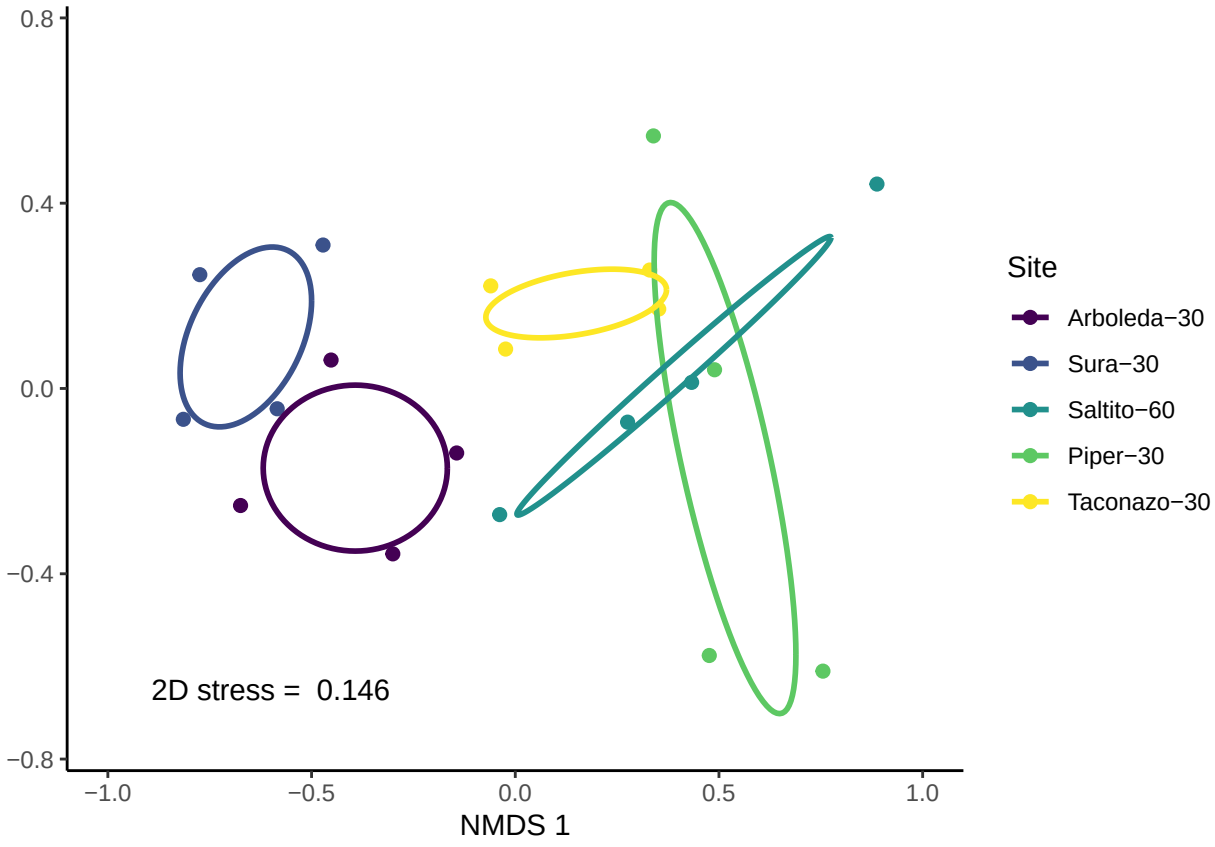

```
shepard_plot <- ggplot(shepard)+
  geom_point(aes(x = x, y = y))+
  geom_line(aes(x = x, y = yf),
            color = 'red')+
  geom_label(x = 0.2, y = 1.5,
             label = 'Non-metric fit R2 = 0.979',
             label.size = 0)+
  geom_label(x = 0.2, y = 1.2,
             label = 'Linear fit R2 = 0.893',
             label.size = 0)+
  labs(x = 'Observed dissimilarity',
       y = 'Ordination distance')+
  theme_classic()+
  theme(panel.grid = element_blank())

fig_s2 <- ggpubr::ggarrange(scree_plot,
                            shepard_plot,
                            labels = 'AUTO',
                            ncol = 2,
                            widths = c(1, 2))

ggsave(plot = fig_s2,
       'Figures/fig_s2.png',
       dpi = 600,
       width = 11, height = 5)
```

```r
# pick the log transformed 2d NMDS (14.6% stress)

nmds_family_out <- data.frame(x = nmds_family_log_2d$points[,1],
                              y = nmds_family_log_2d$points[,2])

nmds_family_out <- cbind(tax_data_log_meta_fam,
                         nmds_family_out)

fit_fam <- (vegan::envfit(nmds_family_log_2d,
                          tax_data_log_matrix_fam,
                          perm = 9999))

scrs_fam <- data.frame(vegan::scores(fit_fam, 'vectors'))

scrs_fam$pvals <- fit_fam$vectors$pvals

scrs_fam_sig <- subset(scrs_fam, pvals <= 0.05)

scrs_fam_sig$env.variables <- row.names(scrs_fam_sig)


plot_nmds_family_scrs <- ggplot(nmds_family_out,
                                aes(x = x, y = y))+
  geom_point(size = 2)+
  geom_segment(data = scrs_fam_sig,
               aes(x = 0, xend = NMDS1,
                   y = 0, yend = NMDS2),
               arrow = arrow(length = unit(0.25, 'cm')),
               color = 'black')+
  ggrepel::geom_label_repel(data = scrs_fam_sig,
                            aes(NMDS1, NMDS2,
                                label = env.variables),
                            size = 3)+
  labs(x = 'NMDS 1',
       y = 'NMDS 2')+
  scale_x_continuous(limits = c(-1,1),
                     labels = signs_format(accuracy = 0.1))+
  scale_y_continuous(limits = c(-0.75, 0.75),
                     labels = signs_format(accuracy = 0.1))+
  theme_classic()+
  theme(panel.grid = element_blank())
plot_nmds_family_scrs
```



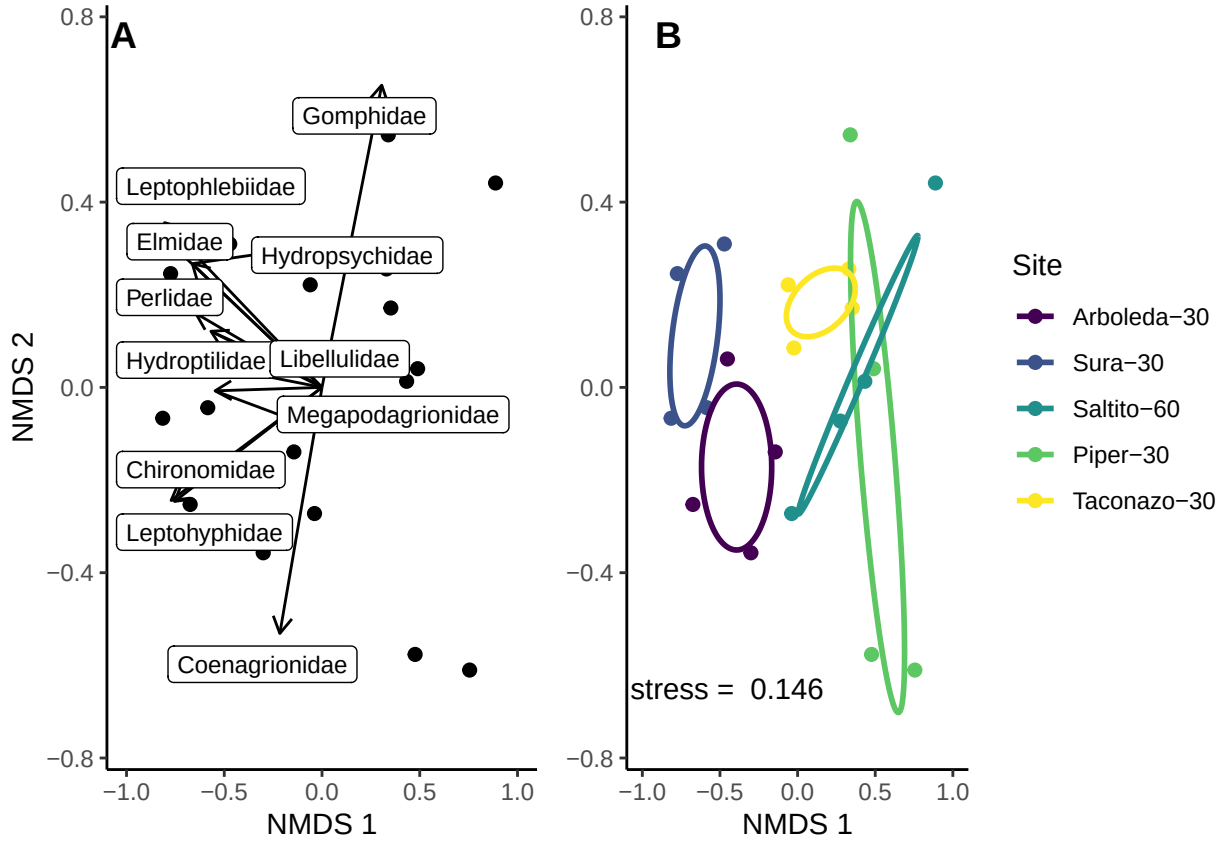

```
source('Code/veganCovEllipse.R')
streams <- data.frame()
for(i in unique(nmds_family_out$Stream)){
  streams <- rbind(streams,
                   cbind(
                     as.data.frame(
                       with(nmds_family_out[nmds_family_out$Stream == i,],
                            veganCovEllipse(cov.wt(cbind(x, y),
                                                   wt = rep(1/length(x),
                                                            length(x)))$cov,
                                            center = c(mean(x),
                                                       mean(y)
                                            )
                            )
                       )
                     ),
                     Stream = i)
  )
} # end for loop

plot_nmds_streams <- ggplot(data = nmds_family_out,
                            aes(x = x, y = y))+
  geom_point(aes(color = Stream),
             size = 2)+
  geom_path(data = streams,
            linewidth = 1,
```

```
            aes(x = x, y = y, color = Stream))+
  labs(x = "NMDS 1",
       y = "NMDS 2")+
  scale_x_continuous(limits = c(-1,1),
                     labels = signs_format(accuracy = 0.1))+
  scale_y_continuous(limits = c(-0.75, 0.75),
                     labels = signs_format(accuracy = 0.1))+
  scale_color_viridis_d(name = 'Site')+
  geom_label(label = paste('2D stress = ', stress_fam_log_2d),
             x = -0.6,
             y = -0.65,
             label.size = 0)+
  theme_classic()+
  theme(panel.grid = element_blank(),
        legend.background = element_blank(),
        axis.title.y = element_blank())
plot_nmds_streams
```




```
fig5 <- ggpubr::ggarrange(plot_nmds_family_scrs,
                          plot_nmds_streams,
                          align = 'h',
                          widths = c(1,1.25),
                          labels = 'AUTO',
                          label.x = c(0.18,0.13),
                          label.y = 0.98)
```

```
fig5
```




```
ggsave(plot = fig5,
       'Figures/fig5.png',
       dpi = 600,
       width = 11.5,height = 4)

# Permanova
# The goal of this test is to tell you if there are significant differences in your response variables
permanova <- vegan::adonis2(tax_data_log_matrix_fam ~ Stream * month,
                            data = tax_data_log_meta_fam,
                            permutations = 9999,
                            method = 'bray',
                            by = 'terms')
permanova
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 9999
##
## vegan::adonis2(formula = tax_data_log_matrix_fam ~ Stream * month, data = tax_data_log_meta_fam, perm
##              Df SumOfSqs      R2      F Pr(>F)
## Stream        4  1.56809 0.52929 4.1208 0.0001 ***
## month         1  0.13925 0.04700 1.4637 0.1965
```

```
## Stream:month  4  0.30394 0.10259 0.7987 0.7212
## Residual     10  0.95133 0.32111
## Total        19  2.96261 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# SIMPER
sim <- with(tax_data_log_meta_fam,
            vegan::simper(tax_data_log_matrix_fam, Stream))

summary(sim)
```

```
##
## Contrast: Arboleda-30_Sura-30
##
##                   average      sd   ratio     ava     avb cumsum     p
## Hydroptilidae     0.04465 0.03069 1.45510 0.19450 0.61820  0.104 0.097 .
## Perlidae          0.03581 0.02526 1.41810 0.07530 0.45040  0.187 0.083 .
## Hydropsychidae    0.03418 0.01836 1.86130 0.34510 0.74720  0.267 0.929
## Leptohyphidae     0.02984 0.01978 1.50890 0.50550 0.72550  0.336 0.989
## Chironomidae      0.02956 0.01523 1.94090 0.77320 0.87940  0.405 0.995
## Leptophlebiidae   0.02352 0.01574 1.49400 0.47580 0.67660  0.460 0.996
## Ceratopogonidae   0.02224 0.01647 1.35050 0.22580 0.28650  0.511 0.923
## Caenidae          0.02172 0.01577 1.37770 0.28650 0.07530  0.562 0.554
## Polycentropodidae 0.02159 0.02297 0.94020 0.00000 0.23860  0.612 0.059 .
## Elmidae           0.02049 0.01095 1.87160 0.83920 1.03930  0.660 1.000
## Libellulidae      0.02037 0.02188 0.93080 0.07530 0.25000  0.707 0.120
## Simuliidae        0.01795 0.01907 0.94160 0.00000 0.20600  0.749 0.073 .
## Baetidae          0.01705 0.01981 0.86070 0.19450 0.00000  0.789 0.347
## Coenagrionidae    0.01705 0.01409 1.20960 0.22580 0.07530  0.828 0.715
## Gyrinidae         0.01450 0.01517 0.95540 0.15050 0.00000  0.862 0.146
## Megapodagrionidae 0.01384 0.01469 0.94250 0.07530 0.15050  0.894 0.181
## Empididae         0.01344 0.01409 0.95390 0.00000 0.15050  0.925 0.475
## Scirtidae         0.00983 0.01342 0.73250 0.07530 0.07530  0.948 0.166
## Leptoceridae      0.00775 0.01398 0.55430 0.00000 0.07530  0.967 0.602
## Calamoceratidae   0.00723 0.01306 0.55410 0.07530 0.00000  0.983 0.403
## Planariidae       0.00719 0.01295 0.55480 0.00000 0.07530  1.000 0.375
## Gomphidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Gordiidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Contrast: Arboleda-30_Saltito-60
##
##                   average      sd   ratio     ava     avb cumsum     p
## Elmidae           0.07900 0.03646 2.16650 0.83920 0.36930  0.140 0.163
## Leptophlebiidae   0.07653 0.03517 2.17620 0.47580 0.00000  0.275 0.028 *
## Chironomidae      0.05650 0.02820 2.00340 0.77320 0.51340  0.375 0.236
## Caenidae          0.04161 0.02883 1.44320 0.28650 0.11930  0.449 0.039 *
## Leptohyphidae     0.04114 0.03879 1.06060 0.50550 0.22580  0.521 0.838
## Hydropsychidae    0.03934 0.03466 1.13510 0.34510 0.22580  0.591 0.841
## Coenagrionidae    0.03661 0.02385 1.53460 0.22580 0.00000  0.656 0.058 .
## Ceratopogonidae   0.03354 0.02877 1.16590 0.22580 0.15050  0.715 0.563
## Baetidae          0.03076 0.03740 0.82260 0.19450 0.00000  0.770 0.004 **
```

```
## Gyrinidae         0.02775 0.02900 0.95720 0.15050 0.00000  0.819 0.002 **
## Hydroptilidae     0.02695 0.02808 0.95990 0.19450 0.00000  0.866 0.575
## Gomphidae         0.02628 0.02819 0.93220 0.00000 0.15050  0.913 0.362
## Calamoceratidae   0.01383 0.02492 0.55480 0.07530 0.00000  0.937 0.008 **
## Libellulidae      0.00885 0.01588 0.55730 0.07530 0.00000  0.953 0.499
## Megapodagrionidae 0.00885 0.01588 0.55730 0.07530 0.00000  0.969 0.490
## Perlidae          0.00885 0.01588 0.55730 0.07530 0.00000  0.984 0.897
## Scirtidae         0.00885 0.01588 0.55730 0.07530 0.00000  1.000 0.245
## Polycentropodidae 0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Simuliidae        0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Empididae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Planariidae       0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Leptoceridae      0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Gordiidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Contrast: Arboleda-30_Piper-30
##
##                   average      sd   ratio     ava     avb cumsum     p
## Elmidae           0.11845 0.04991 2.37350 0.83920 0.15050  0.179 0.002 **
## Leptophlebiidae   0.06748 0.04382 1.53990 0.47580 0.07530  0.282 0.098 .
## Leptohyphidae     0.06456 0.03297 1.95820 0.50550 0.07530  0.379 0.084 .
## Hydropsychidae    0.05495 0.04068 1.35080 0.34510 0.07530  0.462 0.219
## Chironomidae      0.04841 0.03522 1.37470 0.77320 0.46110  0.536 0.498
## Caenidae          0.04474 0.02722 1.64350 0.28650 0.00000  0.604 0.014 *
## Ceratopogonidae   0.04351 0.02789 1.55980 0.22580 0.36930  0.669 0.265
## Baetidae          0.03148 0.03827 0.82270 0.19450 0.00000  0.717 0.002 **
## Coenagrionidae    0.03107 0.02624 1.18370 0.22580 0.07530  0.764 0.120
## Gyrinidae         0.02848 0.02966 0.96010 0.15050 0.00000  0.807 0.003 **
## Hydroptilidae     0.02750 0.02860 0.96170 0.19450 0.00000  0.849 0.554
## Leptoceridae      0.02518 0.04572 0.55080 0.00000 0.17470  0.887 0.408
## Calamoceratidae   0.01418 0.02552 0.55590 0.07530 0.00000  0.909 0.011 *
## Empididae         0.01351 0.02468 0.54750 0.00000 0.07530  0.929 0.421
## Gomphidae         0.01084 0.01969 0.55080 0.00000 0.07530  0.946 0.933
## Libellulidae      0.00900 0.01614 0.55760 0.07530 0.00000  0.959 0.494
## Megapodagrionidae 0.00900 0.01614 0.55760 0.07530 0.00000  0.973 0.504
## Perlidae          0.00900 0.01614 0.55760 0.07530 0.00000  0.986 0.898
## Scirtidae         0.00900 0.01614 0.55760 0.07530 0.00000  1.000 0.225
## Polycentropodidae 0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Simuliidae        0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Planariidae       0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Gordiidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Contrast: Arboleda-30_Taconazo-30
##
##                   average      sd   ratio     ava     avb cumsum     p
## Leptohyphidae     0.06749 0.02101 3.21300 0.50550 0.00000  0.130 0.055 .
## Elmidae           0.05222 0.02824 1.84900 0.83920 0.47970  0.230 0.757
## Leptophlebiidae   0.05000 0.03250 1.53800 0.47580 0.28260  0.326 0.588
## Ceratopogonidae   0.04086 0.03070 1.33100 0.22580 0.40050  0.404 0.316
## Caenidae          0.03864 0.02328 1.66000 0.28650 0.00000  0.478 0.065 .
```

```
## Chironomidae      0.03675 0.02222 1.65400 0.77320 0.63870  0.549 0.896
## Hydropsychidae    0.03294 0.02603 1.26600 0.34510 0.25700  0.612 0.945
## Gordiidae         0.02882 0.03161 0.91200 0.00000 0.19450  0.667 0.133
## Baetidae          0.02704 0.03223 0.83900 0.19450 0.00000  0.719 0.029 *
## Coenagrionidae    0.02688 0.02235 1.20200 0.22580 0.07530  0.771 0.273
## Hydroptilidae     0.02408 0.02501 0.96300 0.19450 0.00000  0.817 0.705
## Gyrinidae         0.02402 0.02488 0.96600 0.15050 0.00000  0.863 0.024 *
## Gomphidae         0.02096 0.02215 0.94700 0.00000 0.15050  0.903 0.604
## Perlidae          0.01431 0.01939 0.73800 0.07530 0.07530  0.931 0.805
## Calamoceratidae   0.01197 0.02146 0.55800 0.07530 0.00000  0.954 0.135
## Libellulidae      0.00807 0.01444 0.55900 0.07530 0.00000  0.969 0.623
## Megapodagrionidae 0.00807 0.01444 0.55900 0.07530 0.00000  0.985 0.577
## Scirtidae         0.00807 0.01444 0.55900 0.07530 0.00000  1.000 0.409
## Polycentropodidae 0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Simuliidae        0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Empididae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Planariidae       0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Leptoceridae      0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Contrast: Sura-30_Saltito-60
##
##                   average      sd   ratio     ava     avb cumsum     p
## Leptophlebiidae   0.07984 0.01426 5.60000 0.67660 0.00000  0.120 0.026 *
## Elmidae           0.07939 0.01541 5.15000 1.03930 0.36930  0.240 0.127
## Hydroptilidae     0.06702 0.04391 1.52600 0.61820 0.00000  0.341 0.004 **
## Hydropsychidae    0.06303 0.02445 2.57700 0.74720 0.22580  0.436 0.062 .
## Leptohyphidae     0.06076 0.03319 1.83100 0.72550 0.22580  0.527 0.157
## Perlidae          0.05012 0.03376 1.48400 0.45040 0.00000  0.603 0.005 **
## Chironomidae      0.04977 0.03313 1.50200 0.87940 0.51340  0.678 0.470
## Polycentropodidae 0.02897 0.03106 0.93300 0.23860 0.00000  0.721 0.004 **
## Libellulidae      0.02548 0.02969 0.85800 0.25000 0.00000  0.760 0.013 *
## Ceratopogonidae   0.02425 0.02136 1.13500 0.28650 0.15050  0.796 0.876
## Simuliidae        0.02380 0.02490 0.95600 0.20600 0.00000  0.832 0.004 **
## Megapodagrionidae 0.01890 0.02002 0.94400 0.15050 0.00000  0.861 0.018 *
## Gomphidae         0.01885 0.02000 0.94300 0.00000 0.15050  0.889 0.662
## Empididae         0.01784 0.01867 0.95600 0.15050 0.00000  0.916 0.339
## Caenidae          0.01745 0.02260 0.77200 0.07530 0.11930  0.942 0.678
## Leptoceridae      0.01079 0.01940 0.55600 0.07530 0.00000  0.959 0.420
## Scirtidae         0.00974 0.01748 0.55700 0.07530 0.00000  0.973 0.145
## Planariidae       0.00974 0.01748 0.55700 0.07530 0.00000  0.988 0.010 **
## Coenagrionidae    0.00810 0.01453 0.55800 0.07530 0.00000  1.000 0.922
## Baetidae          0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Gyrinidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Calamoceratidae   0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Gordiidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Contrast: Sura-30_Piper-30
##
##                   average      sd   ratio     ava     avb cumsum     p
## Elmidae           0.10786 0.02607 4.13700 1.03930 0.15050  0.149 0.005 **
```

```
## Hydropsychidae    0.08135 0.02132 3.81600 0.74720 0.07530  0.262 0.003 **
## Leptohyphidae     0.07875 0.03093 2.54600 0.72550 0.07530  0.370 0.013 *
## Leptophlebiidae   0.07303 0.02391 3.05400 0.67660 0.07530  0.471 0.054 .
## Hydroptilidae     0.06808 0.04451 1.53000 0.61820 0.00000  0.565 0.002 **
## Chironomidae      0.05236 0.02745 1.90700 0.87940 0.46110  0.638 0.400
## Perlidae          0.05094 0.03427 1.48600 0.45040 0.00000  0.708 0.004 **
## Polycentropodidae 0.02950 0.03161 0.93300 0.23860 0.00000  0.749 0.001 ***
## Libellulidae      0.02586 0.03011 0.85900 0.25000 0.00000  0.785 0.018 *
## Leptoceridae      0.02524 0.03090 0.81700 0.07530 0.17470  0.820 0.415
## Simuliidae        0.02421 0.02530 0.95700 0.20600 0.00000  0.853 0.001 ***
## Megapodagrionidae 0.01925 0.02038 0.94400 0.15050 0.00000  0.879 0.016 *
## Ceratopogonidae   0.01910 0.01615 1.18200 0.28650 0.36930  0.906 0.956
## Empididae         0.01828 0.01927 0.94800 0.15050 0.07530  0.931 0.319
## Coenagrionidae    0.01355 0.01844 0.73500 0.07530 0.07530  0.950 0.870
## Scirtidae         0.00992 0.01779 0.55700 0.07530 0.00000  0.964 0.123
## Planariidae       0.00992 0.01779 0.55700 0.07530 0.00000  0.977 0.006 **
## Caenidae          0.00823 0.01475 0.55800 0.07530 0.00000  0.989 0.898
## Gomphidae         0.00820 0.01484 0.55300 0.00000 0.07530  1.000 0.955
## Baetidae          0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Gyrinidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Calamoceratidae   0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Gordiidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Contrast: Sura-30_Taconazo-30
##
##                   average      sd   ratio     ava     avb cumsum     p
## Leptohyphidae     0.07887 0.02362 3.33900 0.72550 0.00000  0.135 0.010 **
## Hydroptilidae     0.06147 0.04031 1.52500 0.61820 0.00000  0.240 0.008 **
## Elmidae           0.05963 0.01146 5.20300 1.03930 0.47970  0.341 0.657
## Hydropsychidae    0.05319 0.02071 2.56800 0.74720 0.25700  0.432 0.296
## Leptophlebiidae   0.04507 0.03241 1.39100 0.67660 0.28260  0.509 0.772
## Perlidae          0.04262 0.02874 1.48300 0.45040 0.07530  0.582 0.019 *
## Chironomidae      0.03568 0.02244 1.59000 0.87940 0.63870  0.643 0.923
## Polycentropodidae 0.02625 0.02788 0.94200 0.23860 0.00000  0.688 0.019 *
## Libellulidae      0.02351 0.02738 0.85900 0.25000 0.00000  0.728 0.060 .
## Simuliidae        0.02166 0.02267 0.95600 0.20600 0.00000  0.765 0.019 *
## Gordiidae         0.02165 0.02368 0.91400 0.00000 0.19450  0.802 0.304
## Ceratopogonidae   0.02025 0.01831 1.10600 0.28650 0.40050  0.836 0.934
## Megapodagrionidae 0.01709 0.01796 0.95100 0.15050 0.00000  0.865 0.067 .
## Empididae         0.01624 0.01691 0.96000 0.15050 0.00000  0.893 0.388
## Gomphidae         0.01598 0.01679 0.95200 0.00000 0.15050  0.920 0.760
## Coenagrionidae    0.01193 0.01610 0.74100 0.07530 0.07530  0.941 0.879
## Leptoceridae      0.00964 0.01727 0.55800 0.07530 0.00000  0.957 0.489
## Scirtidae         0.00879 0.01575 0.55800 0.07530 0.00000  0.972 0.252
## Planariidae       0.00879 0.01575 0.55800 0.07530 0.00000  0.987 0.130
## Caenidae          0.00744 0.01333 0.55900 0.07530 0.00000  1.000 0.921
## Baetidae          0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Gyrinidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Calamoceratidae   0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Contrast: Saltito-60_Piper-30
##
##                   average      sd   ratio     ava     avb cumsum     p
## Elmidae           0.07103 0.05900 1.20380 0.36930 0.15050  0.140 0.320
## Ceratopogonidae   0.06988 0.06205 1.12630 0.15050 0.36930  0.277 0.006 **
## Chironomidae      0.06391 0.05177 1.23440 0.51340 0.46110  0.403 0.103
## Hydropsychidae    0.05534 0.04699 1.17780 0.22580 0.07530  0.512 0.210
## Leptohyphidae     0.05502 0.04763 1.15510 0.22580 0.07530  0.620 0.298
## Gomphidae         0.05153 0.05732 0.89900 0.15050 0.07530  0.721 0.013 *
## Leptoceridae      0.04146 0.07514 0.55180 0.00000 0.17470  0.803 0.042 *
## Caenidae          0.02967 0.05356 0.55390 0.11930 0.00000  0.861 0.263
## Empididae         0.02650 0.04881 0.54280 0.00000 0.07530  0.913 0.075 .
## Coenagrionidae    0.02616 0.04816 0.54320 0.00000 0.07530  0.965 0.270
## Leptophlebiidae   0.01786 0.03236 0.55180 0.00000 0.07530  1.000 0.999
## Baetidae          0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Hydroptilidae     0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Libellulidae      0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Megapodagrionidae 0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Perlidae          0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Scirtidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Gyrinidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Calamoceratidae   0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Polycentropodidae 0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Simuliidae        0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Planariidae       0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Gordiidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Contrast: Saltito-60_Taconazo-30
##
##                   average      sd   ratio     ava     avb cumsum     p
## Leptophlebiidae   0.06402 0.07033 0.91020 0.00000 0.28260  0.142 0.135
## Chironomidae      0.06359 0.04595 1.38370 0.51340 0.63870  0.283 0.101
## Ceratopogonidae   0.06261 0.05573 1.12350 0.15050 0.40050  0.422 0.024 *
## Leptohyphidae     0.04965 0.03018 1.64500 0.22580 0.00000  0.532 0.529
## Gordiidae         0.04838 0.05262 0.91950 0.00000 0.19450  0.639 0.004 **
## Gomphidae         0.03616 0.03871 0.93400 0.15050 0.15050  0.719 0.092 .
## Hydropsychidae    0.03549 0.04011 0.88470 0.22580 0.25700  0.798 0.919
## Elmidae           0.03348 0.02569 1.30330 0.36930 0.47970  0.872 0.998
## Caenidae          0.02381 0.04272 0.55740 0.11930 0.00000  0.925 0.478
## Coenagrionidae    0.01793 0.03249 0.55180 0.00000 0.07530  0.965 0.693
## Perlidae          0.01584 0.02863 0.55340 0.00000 0.07530  1.000 0.761
## Baetidae          0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Hydroptilidae     0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Libellulidae      0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Megapodagrionidae 0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Scirtidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Gyrinidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Calamoceratidae   0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Polycentropodidae 0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Simuliidae        0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Empididae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Planariidae       0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
```

```
## Leptoceridae      0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Contrast: Piper-30_Taconazo-30
##
##                   average      sd   ratio     ava     avb cumsum     p
## Elmidae           0.08479 0.05310 1.59670 0.15050 0.47970  0.171 0.076 .
## Leptophlebiidae   0.06670 0.06650 1.00310 0.07530 0.28260  0.305 0.098 .
## Chironomidae      0.05850 0.03975 1.47180 0.46110 0.63870  0.423 0.217
## Hydropsychidae    0.05275 0.04041 1.30550 0.07530 0.25700  0.530 0.298
## Gordiidae         0.05001 0.05401 0.92600 0.00000 0.19450  0.630 0.002 **
## Gomphidae         0.03556 0.03710 0.95850 0.07530 0.15050  0.702 0.089 .
## Leptoceridae      0.03446 0.06181 0.55750 0.17470 0.00000  0.772 0.218
## Coenagrionidae    0.02850 0.03846 0.74090 0.07530 0.07530  0.829 0.208
## Ceratopogonidae   0.02816 0.02893 0.97310 0.36930 0.40050  0.886 0.761
## Leptohyphidae     0.02018 0.03628 0.55620 0.07530 0.00000  0.926 0.999
## Empididae         0.02018 0.03628 0.55620 0.07530 0.00000  0.967 0.271
## Perlidae          0.01631 0.02938 0.55500 0.00000 0.07530  1.000 0.773
## Baetidae          0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Caenidae          0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Hydroptilidae     0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Libellulidae      0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Megapodagrionidae 0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Scirtidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Gyrinidae         0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Calamoceratidae   0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Polycentropodidae 0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Simuliidae        0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## Planariidae       0.00000 0.00000     NaN 0.00000 0.00000  1.000    NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Permutation: free
## Number of permutations: 999
```

```
simper_use <- do.call(rbind, summary(sim)) %>%
  tibble::rownames_to_column() %>%
  dplyr::rowwise() %>%
  dplyr::mutate(comparison = strsplit(rowname, "[.]")[[1]][1],
                family = strsplit(rowname, "[.]")[[1]][2]) %>%
  dplyr::select(-rowname)

simper_sig <- simper_use %>%
  dplyr::filter(p <= 0.05) %>%
  dplyr::group_by(comparison, family) %>%
  dplyr::summarise(n_sig = length(p),
                   mean_avg = mean(average, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'comparison'. You can override using the
## `.groups` argument.
```

```
table_s2 <- simper_use %>%
  dplyr::filter(p <= 0.05) %>%
```

```
  dplyr::group_by(comparison, family)

table_s2[,-c(8:9)] <- round(table_s2[,-c(8:9)],
                            digits = 4)
readr::write_csv(table_s2,
                 'Data/table_s2.csv')


# ANOSIM
anosim <- vegan::anosim(tax_data_log_matrix_fam,
                        grouping = tax_data_log_meta_fam$Stream,
                        distance = 'bray',
                        permutations = 9999)
summary(anosim)
```

```
##
## Call:
## vegan::anosim(x = tax_data_log_matrix_fam, grouping = tax_data_log_meta_fam$Stream,      permutations
## Dissimilarity: bray
##
## ANOSIM statistic R: 0.5483
##       Significance: 2e-04
##
## Permutation: free
## Number of permutations: 9999
##
## Upper quantiles of permutations (null model):
##   90%   95% 97.5%   99%
## 0.132 0.180 0.223 0.274
##
## Dissimilarity ranks between and within classes:
##             0%   25%   50%    75% 100%   N
## Between      2 60.25 107.5 149.25  190 160
## Arboleda-30  6 20.75  46.5  58.00   76   6
## Sura-30      7 11.25  17.0  36.25   47   6
## Saltito-60   1 18.25  32.5  86.50  136   6
## Piper-30    59 76.50  93.5 109.75  154   6
## Taconazo-30  4 14.25  38.0  68.50   89   6
```

```
anosim_r <- anosim$statistic
anosim_p <- anosim$signif
```

Session Info

```
pander(sessionInfo())
```

R version 4.2.3 (2023-03-15 ucrt)

Platform: x86_64-w64-mingw32/x64 (64-bit)

locale: *LC_COLLATE=English_United States.utf8*, *LC_CTYPE=English_United States.utf8*, *LC_MONETARY=English_United States.utf8*, *LC_NUMERIC=C* and *LC_TIME=English_United States.utf8*

attached base packages: *grid*, *stats*, *graphics*, *grDevices*, *utils*, *datasets*, *methods* and *base*

other attached packages: *pander(v.0.6.5)*, *vegan(v.2.6-8)*, *lattice(v.0.20-45)*, *permute(v.0.9-7)*, *spDataLarge(v.2.0.9)*, *leaflet(v.2.2.2)*, *cowplot(v.1.1.3)*, *spData(v.2.3.3)*, *raster(v.3.6-26)*, *sp(v.2.1-4)*, *sf(v.1.0-17)*, *qpcR(v.1.4-1)*, *Matrix(v.1.6-3)*, *robustbase(v.0.99-4-1)*, *rgl(v.1.3.1)*, *minpack.lm(v.1.2-4)*, *nlstools(v.2.1-0)*, *nlme(v.3.1-162)*, *drc(v.3.0-1)*, *MASS(v.7.3-58.2)*, *car(v.3.1-3)*, *carData(v.3.0-5)*, *signs(v.0.1.2)*, *lemon(v.0.4.9)*, *ggcorrplot(v.0.1.4.1)*, *ggrepel(v.0.9.6)*, *ggpubr(v.0.6.0)*, *ggeffects(v.1.7.1)*, *readxl(v.1.4.3)*, *lubridate(v.1.9.3)*, *forcats(v.1.0.0)*, *stringr(v.1.5.1)*, *dplyr(v.1.1.3)*, *purrr(v.1.0.2)*, *readr(v.2.1.5)*, *tidyr(v.1.3.1)*, *tibble(v.3.2.1)*, *ggplot2(v.3.5.1)* and *tidyverse(v.2.0.0)*

loaded via a namespace (and not attached): *spam(v.2.10-0)*, *backports(v.1.5.0)*, *Hmisc(v.5.1-3)*, *systemfonts(v.1.1.0)*, *plyr(v.1.8.9)*, *splines(v.4.2.3)*, *AlgDesign(v.1.2.1.1)*, *crosstalk(v.1.2.1)*, *TH.data(v.1.1-2)*, *digest(v.0.6.33)*, *htmltools(v.0.5.8.1)*, *fansi(v.1.0.5)*, *checkmate(v.2.3.2)*, *magrittr(v.2.0.3)*, *cluster(v.2.1.4)*, *tzdb(v.0.4.0)*, *vroom(v.1.6.5)*, *sandwich(v.3.1-1)*, *timechange(v.0.3.0)*, *colorspace(v.2.1-0)*, *textshaping(v.0.4.0)*, *xfun(v.0.45)*, *crayon(v.1.5.3)*, *jsonlite(v.1.8.9)*, *survival(v.3.5-3)*, *zoo(v.1.8-12)*, *glue(v.1.6.2)*, *gtable(v.0.3.5)*, *emmeans(v.1.10.4)*, *maps(v.3.4.2)*, *DEoptimR(v.1.1-3)*, *abind(v.1.4-8)*, *scales(v.1.3.0)*, *mvtnorm(v.1.3-1)*, *DBI(v.1.2.3)*, *rstatix(v.0.7.2)*, *Rcpp(v.1.0.13)*, *plotrix(v.3.8-4)*, *htmlTable(v.2.4.3)*, *viridisLite(v.0.4.2)*, *xtable(v.1.8-4)*, *units(v.0.8-5)*, *foreign(v.0.8-84)*, *bit(v.4.5.0)*, *proxy(v.0.4-27)*, *dotCall64(v.1.1-1)*, *Formula(v.1.2-5)*, *htmlwidgets(v.1.6.4)*, *pkgconfig(v.2.0.3)*, *farver(v.2.1.2)*, *nnet(v.7.3-18)*, *utf8(v.1.2.4)*, *tidyselect(v.1.2.1)*, *labeling(v.0.4.3)*, *rlang(v.1.1.2)*, *ggspatial(v.1.1.9)*, *reshape2(v.1.4.4)*, *munsell(v.0.5.1)*, *cellranger(v.1.1.0)*, *tools(v.4.2.3)*, *cli(v.3.6.1)*, *generics(v.0.1.3)*, *broom(v.1.0.7)*, *evaluate(v.1.0.0)*, *fastmap(v.1.1.1)*, *yaml(v.2.3.10)*, *ragg(v.1.3.3)*, *knitr(v.1.48)*, *bit64(v.4.5.2)*, *compiler(v.4.2.3)*, *rstudioapi(v.0.16.0)*, *e1071(v.1.7-16)*, *ggsignif(v.0.6.4)*, *stringi(v.1.8.4)*, *highr(v.0.11)*, *fields(v.16.3)*, *classInt(v.0.4-10)*, *vctrs(v.0.6.4)*, *pillar(v.1.9.0)*, *lifecycle(v.1.0.4)*, *goeveg(v.0.7.5)*, *estimability(v.1.5.1)*, *data.table(v.1.15.4)*, *insight(v.0.20.4)*, *agricolae(v.1.3-7)*, *R6(v.2.5.1)*, *KernSmooth(v.2.23-20)*, *gridExtra(v.2.3)*, *codetools(v.0.2-19)*, *gtools(v.3.9.5)*, *withr(v.3.0.1)*, *multcomp(v.1.4-26)*, *mgcv(v.1.8-42)*, *parallel(v.4.2.3)*, *hms(v.1.1.3)*, *terra(v.1.7-55)*, *rpart(v.4.1.19)*, *coda(v.0.19-4.1)*, *class(v.7.3-21)*, *rmarkdown(v.2.28)* and *base64enc(v.0.1-3)*