6. Worksheet: Among Site (Beta) Diversity – Part 1

Anna Lennon

05 February, 2025

OVERVIEW

In this worksheet, we move beyond the investigation of within-site α-diversity. We will explore β-diversity, which is defined as the diversity that occurs among sites. This requires that we examine the compositional similarity of assemblages that vary in space or time.

After completing this exercise you will know how to:

1. formally quantify β-diversity
2. visualize β-diversity with heatmaps, cluster analysis, and ordination
3. test hypotheses about β-diversity using multivariate statistics

Directions:

1. In the Markdown version of this document in your cloned repo, change "Student Name" on line 3 (above) with your name.
2. Complete as much of the worksheet as possible during class.
3. Use the handout as a guide; it contains a more complete description of data sets along with examples of proper scripting needed to carry out the exercises.
4. Answer questions in the worksheet. Space for your answers is provided in this document and is indicated by the ">" character. If you need a second paragraph be sure to start the first line with ">". You should notice that the answer is highlighted in green by RStudio (color may vary if you changed the editor theme).
5. Before you leave the classroom, **push** this file to your GitHub repo.
6. For the assignment portion of the worksheet, follow the directions at the bottom of this file.
7. When you are done, **Knit** the text and code into a PDF file.
8. After Knitting, submit the completed exercise by creating a **pull request** via GitHub. Your pull request should include this file (**6.BetaDiversity_1_Worksheet.Rmd**) with all code blocks filled out and questions answered) and the PDF output of `Knitr` (**6.BetaDiversity_1_Worksheet.pdf**).

The completed exercise is due on **Wednesday, February 5th, 2025 before 12:00 PM (noon)**.

1) R SETUP

Typically, the first thing you will do in either an R script or an RMarkdown file is setup your environment. This includes things such as setting the working directory and loading any packages that you will need.

In the R code chunk below, please provide the code to:

1) Clear your R environment,
2) Print your current working directory,
3) Set your working directory to your `Week3-Beta/` folder folder, and
4) Load the `vegan` R package (be sure to install first if you have not already).

```
rm(list=ls())
getwd()
```

```
## [1] "/cloud/project/QB2025_ALennon/Week3-Beta"
```

```
#setwd("cloud/project/AB2025_ALennon/Week3-Beta")
```

2) LOADING DATA

Load dataset

In the R code chunk below, do the following:

1. load the `doubs` dataset from the `ade4` package, and
2. explore the structure of the dataset.

```
# note, please do not print the dataset when submitting
# package.list <- c('vegan', 'ade4', 'viridis', 'gplots', 'indicspecies')
# for (package in package.list) {
# if (!require(package, character.only = TRUE, quietly = TRUE)) {
# install.packages(package)
#install.packages("lattice")  # Install lattice if not installed
library(lattice)
#library(package, character.only = TRUE)}}
library("vegan")
```

```
## Loading required package: permute
```

```
library("ade4")
library("viridis")
```

```
## Loading required package: viridisLite
```

```
library("gplots")
```

```
##
## Attaching package: 'gplots'
```

```
## The following object is masked from 'package:stats':
##
##     lowess
```

```
library("indicspecies")
library("tidyverse")
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.0.2
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become error
```

```
#install.packages("BiodiversityR")
#library("BiodiversityR")
```

```
data(doubs)
#print(doubs)
str(doubs, max.leve=1)
```

```
## List of 4
##  $ env    :'data.frame': 30 obs. of  11 variables:
##  $ fish   :'data.frame': 30 obs. of  27 variables:
##  $ xy     :'data.frame': 30 obs. of  2 variables:
##  $ species:'data.frame': 27 obs. of  4 variables:
```

```
doubs.se.df <- doubs$env
#print(doubs.se.df)
doubs.ss.df <- doubs$fish
#print(doubs.ss.df)
doubs.xy.df <-doubs$xy
```

Question 1: Describe some of the attributes of the `doubs` dataset.

a. How many objects are in `doubs`?
b. How many fish species are there in the `doubs` dataset?
c. How many sites are in the `doubs` dataset?

Answer 1a: There are 30 objects in "doubs" ***Answer 1b***: There are 27 fish species contained in "doubs" ***Answer 1c***: There are 30 sires in "doubs"

Visualizing the Doubs River Dataset

Question 2: Answer the following questions based on the spatial patterns of richness (i.e., α-diversity) and Brown Trout (*Salmo trutta*) abundance in the Doubs River.

a. How does fish richness vary along the sampled reach of the Doubs River?
b. How does Brown Trout (*Salmo trutta*) abundance vary along the sampled reach of the Doubs River?
c. What do these patterns say about the limitations of using richness when examining patterns of biodiversity?

Answer 2a:In general, species richness is highest downstream and continually decreases going further >upstream. There are two main clusters of high (>20) richness in the downstream section with three sampling sites with low (<20) in between these downstream clusters. ***Answer 2b***:When focused on the Brown Trout, abundance is greatest in the upstream area, which is where richness is lowest. Contrastly, the downstream area has extremely low aundance of trout with most N-values equal to 0 which is where the downstream richness is greatest. ***Answer 2c***:The comparison of richness to Brown Trout abundance displays how richness is limited to looking at large scale population patterns and may not reflect the ecological patterns of an indvidual species.

3) QUANTIFYING BETA-DIVERSITY

In the R code chunk below, do the following:

1. write a function (`beta.w()`) to calculate Whittaker's β-diversity (i.e., β_w) that accepts a site-by-species matrix with optional arguments to specify pairwise turnover between two sites, and
2. use this function to analyze various aspects of β-diversity in the Doubs River.

```
#original function
# beta.w <- function(site.by.species = ""){
#   SbyS.pa <- decostand(site.by.species, method = "pa")
#   S <- ncol(SbyS.pa[,which(colSums(SbyS.pa) > 0)])
```

```
#   a.bar <- mean(specnumber(SbyS.pa))
#   b.w <- round(S/a.bar, 3)
#   return(b.w)
#}
#pairwise Bw
beta.w <- function(site.by.species = "", sitenum1 = "", sitenum2 = "",
    pairwise = FALSE){
    if (pairwise == TRUE){
      if (sitenum1 == "" | sitenum2 == "") {
        print("Error: please specify sites to compare")
        return(NA)}
      site1 = site.by.species[sitenum1,]
      site2 = site.by.species[sitenum2,]
      site1 = subset(site1, select = site1 > 0)
      site2 = subset(site2, select = site2 > 0)
      gamma = union(colnames(site1), colnames(site2))
      s     = length(gamma)
      a.bar = mean(c(specnumber(site1), specnumber(site2)))
      b.w   = round(s/a.bar - 1, 3)
      return(b.w)
    }
  else{
      SbyS.pa <- decostand(site.by.species, method = "pa")
      S <- ncol(SbyS.pa[,which(colSums(SbyS.pa) > 0)])
      a.bar <- mean(specnumber(SbyS.pa))
      b.w <- round(S/a.bar, 3)
      return(b.w)
  }
}

##################
site1.doubs <-doubs.ss.df[1,]
site2.doubs <-doubs.ss.df[2,]
site10.doubs <-doubs.ss.df[10,]

beta.w(site1.doubs, site2.doubs)
```

```
## numeric(0)
```

```
beta.w(site10.doubs, site2.doubs)
```

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

```
beta.w(site10.doubs, site1.doubs)
```

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

Question 3: Using your `beta.w()` function above, answer the following questions:

a. Describe how local richness (α) and turnover (β) contribute to regional (γ) fish diversity in the Doubs.
b. Is the fish assemblage at site 1 more similar to the one at site 2 or site 10?
c. Using your understanding of the equation $\beta_w = \gamma/\alpha$, how would your interpretation of β change if we instead defined beta additively (i.e., $\beta = \gamma - \alpha$)?

Answer 3aGamma diversity, under Whittaker's Beta-diversity, is the result of how many more times a region is more diverse (variation and turnover- Beta diversity) multiplied by the average richness at each site. ***Answer 3b***: Site 1 is more similar to Site 2 with a Beta-diversity equal

to 0 compared to Site1-Site10 which was 1. Therefore, Site1-Site2 have minimum B-diversity or variation.

Answer 3c:Through additive partitioning, which combines alpha diversity and beta diversity, gamma diversity can be calculated. Additive paritioning transforms Beta diversity to determine how many more species exists at the regional level which is added to the local richness. Therefore, the local species richness combined with the turnover combined contribute to the reginal diversity, which logically makes senese that reginal diversity is a combination of local diversity and how many more species exist on the regional level compared to the local sites.

The Resemblance Matrix

In order to quantify β-diversity for more than two samples, we need to introduce a new primary ecological data structure: the **Resemblance Matrix**.

Question 4: How do incidence- and abundance-based metrics differ in their treatment of rare species?

Answer 4: Incidence-based measures incorportate rare species in both the first and second assemblages compared to abundance-based measures that looks at the overall abundance of each species at each site. Incidence-based measures take more care to rare species compared to the abundance-based.

In the R code chunk below, do the following:

1. make a new object, `fish`, containing the fish abundance data for the Doubs River,
2. remove any sites where no fish were observed (i.e., rows with sum of zero),
3. construct a resemblance matrix based on Sørensen's Similarity ("fish.ds"), and
4. construct a resemblance matrix based on Bray-Curtis Distance ("fish.db").

```
fish <-doubs$fish
fish <-fish[-8,] # remove 0 observation
#str(fish)
library("vegan")
#Sorensen

fish.ds <- vegdist(fish, method = "bray", binary = TRUE)

print(fish.ds)
```

```
##            1          2          3          4          5          6          7
## 2  0.50000000
## 3  0.60000000 0.14285714
## 4  0.77777778 0.45454545 0.33333333
## 5  0.83333333 0.57142857 0.46666667 0.15789474
## 6  0.81818182 0.53846154 0.42857143 0.11111111 0.04761905
## 7  0.66666667 0.25000000 0.33333333 0.38461538 0.37500000 0.33333333
## 9  1.00000000 0.50000000 0.55555556 0.38461538 0.37500000 0.33333333 0.40000000
## 10 0.71428571 0.33333333 0.40000000 0.28571429 0.29411765 0.25000000 0.09090909
## 11 0.71428571 0.33333333 0.40000000 0.42857143 0.52941176 0.50000000 0.27272727
## 12 0.71428571 0.33333333 0.40000000 0.42857143 0.52941176 0.50000000 0.27272727
## 13 0.71428571 0.33333333 0.40000000 0.57142857 0.64705882 0.62500000 0.45454545
## 14 0.81818182 0.53846154 0.42857143 0.33333333 0.42857143 0.40000000 0.46666667
## 15 0.83333333 0.57142857 0.60000000 0.36842105 0.36363636 0.33333333 0.37500000
## 16 0.88888889 0.70000000 0.61904762 0.36000000 0.28571429 0.25925926 0.54545455
## 17 0.91304348 0.76000000 0.69230769 0.46666667 0.39393939 0.37500000 0.62962963
## 18 0.91666667 0.76923077 0.70370370 0.48387097 0.41176471 0.39393939 0.64285714
## 19 1.00000000 0.84615385 0.77777778 0.54838710 0.41176471 0.45454545 0.71428571
## 20 1.00000000 0.84000000 0.76923077 0.53333333 0.39393939 0.43750000 0.70370370
```

```
## 21 1.00000000 0.84615385 0.77777778 0.54838710 0.41176471 0.45454545 0.71428571
## 22 1.00000000 0.92000000 0.84615385 0.60000000 0.45454545 0.50000000 0.77777778
## 23 1.00000000 1.00000000 1.00000000 0.81818182 0.71428571 0.69230769 0.75000000
## 24 1.00000000 1.00000000 1.00000000 0.75000000 0.68421053 0.66666667 0.84615385
## 25 1.00000000 1.00000000 0.83333333 0.62500000 0.36842105 0.44444444 0.69230769
## 26 1.00000000 0.91666667 0.84000000 0.58620690 0.43750000 0.48387097 0.76923077
## 27 1.00000000 0.92000000 0.84615385 0.60000000 0.45454545 0.50000000 0.77777778
## 28 1.00000000 0.92000000 0.84615385 0.60000000 0.45454545 0.50000000 0.77777778
## 29 0.92592593 0.79310345 0.73333333 0.52941176 0.40540541 0.44444444 0.67741935
## 30 1.00000000 1.00000000 0.92000000 0.65517241 0.50000000 0.54838710 0.84615385
##             9         10         11         12         13         14         15
## 2
## 3
## 4
## 5
## 6
## 7
## 9
## 10 0.45454545
## 11 0.45454545 0.33333333
## 12 0.45454545 0.33333333 0.00000000
## 13 0.63636364 0.50000000 0.16666667 0.16666667
## 14 0.60000000 0.37500000 0.25000000 0.25000000 0.25000000
## 15 0.50000000 0.29411765 0.29411765 0.29411765 0.29411765 0.14285714
## 16 0.54545455 0.47826087 0.56521739 0.56521739 0.56521739 0.33333333 0.28571429
## 17 0.62962963 0.57142857 0.57142857 0.57142857 0.57142857 0.37500000 0.33333333
## 18 0.64285714 0.58620690 0.58620690 0.58620690 0.58620690 0.39393939 0.35294118
## 19 0.64285714 0.65517241 0.79310345 0.79310345 0.79310345 0.57575758 0.52941176
## 20 0.62962963 0.64285714 0.78571429 0.78571429 0.85714286 0.62500000 0.57575758
## 21 0.64285714 0.65517241 0.79310345 0.79310345 0.86206897 0.63636364 0.58823529
## 22 0.70370370 0.71428571 0.85714286 0.85714286 0.92857143 0.68750000 0.63636364
## 23 0.50000000 0.77777778 0.77777778 0.77777778 1.00000000 0.84615385 0.85714286
## 24 0.69230769 0.71428571 0.85714286 0.85714286 1.00000000 0.77777778 0.78947368
## 25 0.69230769 0.57142857 0.85714286 0.85714286 1.00000000 0.66666667 0.68421053
## 26 0.69230769 0.70370370 0.85185185 0.85185185 0.92592593 0.67741935 0.62500000
## 27 0.70370370 0.71428571 0.85714286 0.85714286 0.92857143 0.68750000 0.63636364
## 28 0.70370370 0.71428571 0.85714286 0.85714286 0.92857143 0.68750000 0.63636364
## 29 0.67741935 0.62500000 0.68750000 0.68750000 0.68750000 0.50000000 0.45945946
## 30 0.76923077 0.77777778 0.92592593 0.92592593 1.00000000 0.74193548 0.68750000
##            16         17         18         19         20         21         22
## 2
## 3
## 4
## 5
## 6
## 7
## 9
## 10
## 11
## 12
## 13
## 14
## 15
## 16
```

```
## 17 0.12820513
## 18 0.15000000 0.02222222
## 19 0.25000000 0.15555556 0.13043478
## 20 0.28205128 0.18181818 0.15555556 0.02222222
## 21 0.30000000 0.20000000 0.17391304 0.04347826 0.02222222
## 22 0.33333333 0.22727273 0.20000000 0.06666667 0.04545455 0.02222222
## 23 0.80000000 0.76000000 0.76923077 0.76923077 0.76000000 0.76923077 0.76000000
## 24 0.68000000 0.60000000 0.54838710 0.48387097 0.46666667 0.48387097 0.46666667
## 25 0.60000000 0.60000000 0.54838710 0.48387097 0.46666667 0.48387097 0.46666667
## 26 0.36842105 0.25581395 0.22727273 0.09090909 0.06976744 0.04545455 0.02325581
## 27 0.33333333 0.22727273 0.20000000 0.06666667 0.04545455 0.02222222 0.00000000
## 28 0.33333333 0.22727273 0.20000000 0.06666667 0.04545455 0.02222222 0.00000000
## 29 0.25581395 0.12500000 0.10204082 0.06122449 0.08333333 0.06122449 0.08333333
## 30 0.36842105 0.25581395 0.22727273 0.09090909 0.06976744 0.04545455 0.02325581
##            23         24         25         26         27         28         29
## 2
## 3
## 4
## 5
## 6
## 7
## 9
## 10
## 11
## 12
## 13
## 14
## 15
## 16
## 17
## 18
## 19
## 20
## 21
## 22
## 23
## 24 0.45454545
## 25 0.45454545 0.37500000
## 26 0.75000000 0.44827586 0.44827586
## 27 0.76000000 0.46666667 0.46666667 0.02325581
## 28 0.76000000 0.46666667 0.46666667 0.02325581 0.00000000
## 29 0.79310345 0.52941176 0.52941176 0.10638298 0.08333333 0.08333333
## 30 0.75000000 0.44827586 0.44827586 0.04761905 0.02325581 0.02325581 0.10638298
```

```r
# Bray-Curtis
fish.db <- vegdist(fish, method = "bray")
fish.db <- vegdist(fish, method = "bray", upper = TRUE, diag = TRUE)
#print(fish.db)
```

Question 5: Using the distance matrices from above, answer the following questions:

a. Does the resemblance matrix (`fish.db`) represent similarity or dissimilarity? What information in the resemblance matrix led you to arrive at your answer?
b. Compare the resemblance matrices (`fish.db` or `fish.ds`) you just created. How does the choice of the Sørensen or Bray-Curtis distance influence your interpretation of site (dis)similarity?

Answer 5a: fish.db generally represents similarity. The only zero values (fill dissimilarity) were the diagonals, while the rest of the values were above zero. While some sites are more similar than others (range [.0978, 1.00]), they all generally trend towards aimilarity. ***Answer 5b***: While both matrixes are similar, there are a few key differences that are due to how Sorensen and Bray determine similarity/dissimilarity. Relative abundance is not accounted for in Sorensen, an incidence based measure that emphasizes similarity due to shared species, compared to Bray-Curtis which is an abundance-based measure. Therefore, Bray-Curtis is more commonly used to find similarity. If two sites do not have many shared species but are similar for other reasons, Sorensen will measure them as more dissimilar compared to Bray-Curits.

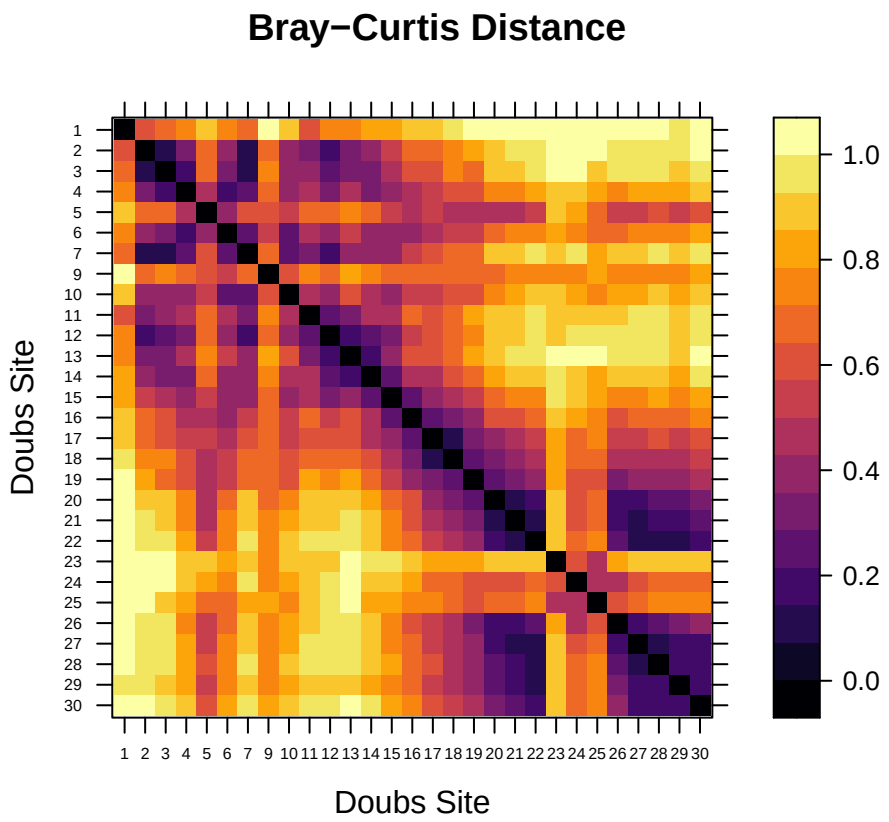
4) VISUALIZING BETA-DIVERSITY

A. Heatmaps

In the R code chunk below, do the following:

1. define a color palette,
2. define the order of sites in the Doubs River, and
3. use the `levelplot()` function to create a heatmap of fish abundances in the Doubs River.

```
# install.packages("viridis")
# install.packages("viridisLite")
#palette(viridis)
order <-rev(attr(fish.db, "Labels"))
levelplot(as.matrix(fish.db)[, order], aspect = "iso", col.regions = inferno,
          xlab = "Doubs Site", ylab = "Doubs Site", scales = list(cex = 0.5),
          main = "Bray-Curtis Distance")
```

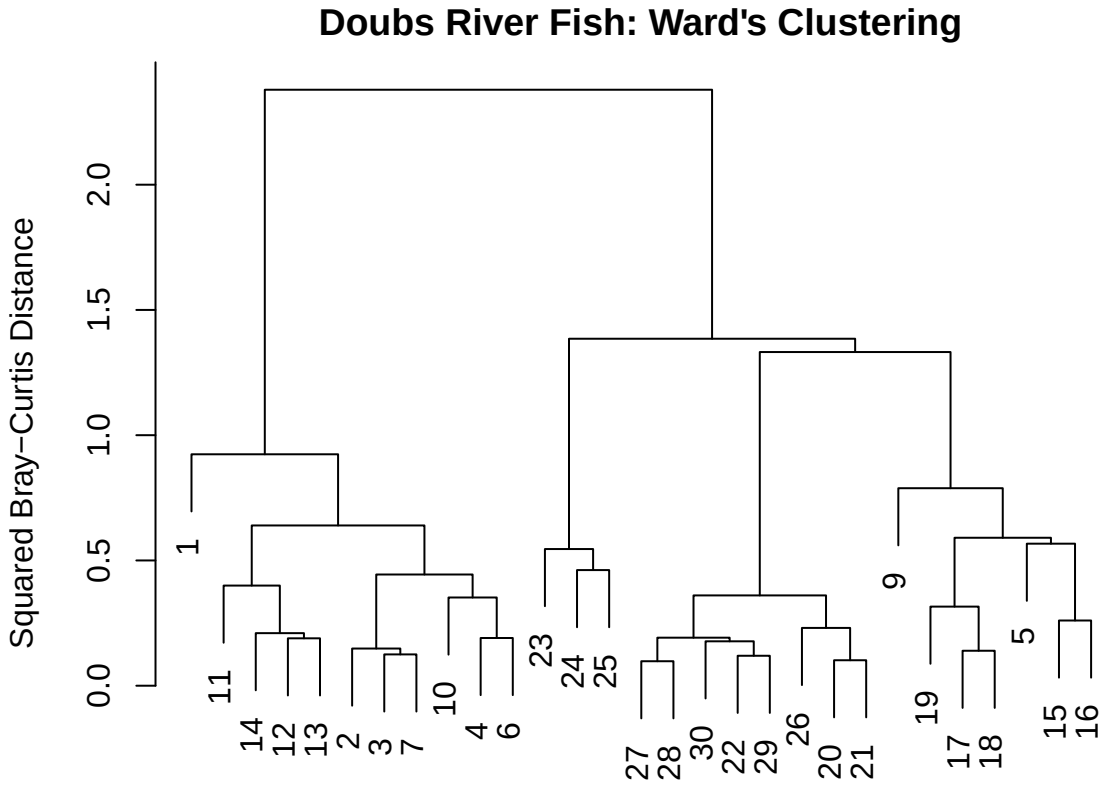


B. Cluster Analysis

In the R code chunk below, do the following:

1. perform a cluster analysis using Ward's Clustering, and
2. plot your cluster analysis (use either `hclust` or `heatmap.2`).

```
fish.ward <- hclust(fish.db, method = "ward.D2")

par(mar = c(1, 5, 2, 2) + 0.1)
plot(fish.ward, main = "Doubs River Fish: Ward's Clustering",
     ylab = "Squared Bray-Curtis Distance")
```



Question 6: Based on cluster analyses and the introductory plots that we generated after loading the data, develop an ecological hypothesis for fish diversity the `doubs` data set?

> ***Answer 6***: Given the preliminary data, early sites contrasted to later sites have greater fish diversity due to variable environmental conditions.

C. Ordination

Principal Coordinates Analysis (PCoA)

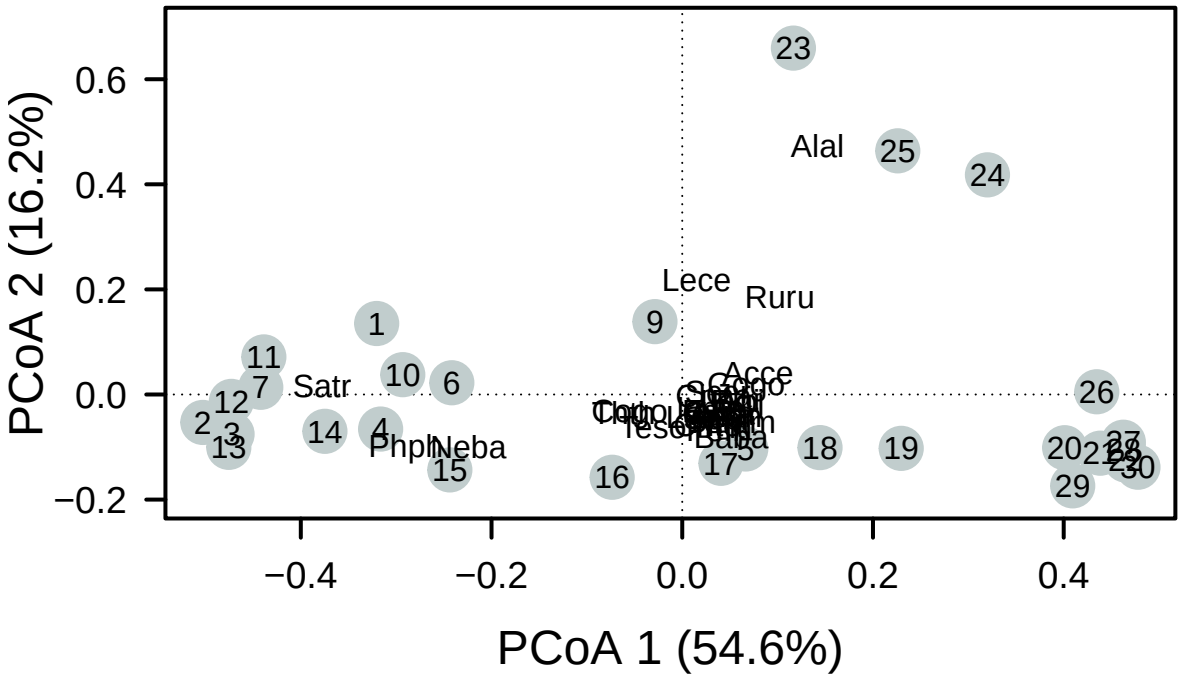
In the R code chunk below, do the following:

1. perform a Principal Coordinates Analysis to visualize beta-diversity
2. calculate the variation explained by the first three axes in your ordination
3. plot the PCoA ordination,
4. label the sites as points using the Doubs River site number, and
5. identify influential species and add species coordinates to PCoA plot.

```
`add.spec.scores.class` <-
  function(ordi,comm,method="cor.scores",multi=1,Rscale=F,scaling="1") {
```

```
  ordiscores <- scores(ordi,display="sites")
  n <- ncol(comm)
  p <- ncol(ordiscores)
  specscores <- array(NA,dim=c(n,p))
  rownames(specscores) <- colnames(comm)
  colnames(specscores) <- colnames(ordiscores)
  if (method == "cor.scores") {
    for (i in 1:n) {
      for (j in 1:p) {specscores[i,j] <- cor(comm[,i],ordiscores[,j],method="pearson")}
    }
  }
  if (method == "wa.scores") {specscores <- wascores(ordiscores,comm)}
  if (method == "pcoa.scores") {
    rownames(ordiscores) <- rownames(comm)
    eigenv <- ordi$eig
    accounted <- sum(eigenv)
    tot <- 2*(accounted/ordi$GOF[2])-(accounted/ordi$GOF[1])
    eigen.var <- eigenv/(nrow(comm)-1)
    neg <- length(eigenv[eigenv<0])
    pos <- length(eigenv[eigenv>0])
    tot <- tot/(nrow(comm)-1)
    eigen.percen <- 100*eigen.var/tot
    eigen.cumpercen <- cumsum(eigen.percen)
    constant <- ((nrow(comm)-1)*tot)^0.25
    ordiscores <- ordiscores * (nrow(comm)-1)^-0.5 * tot^-0.5 * constant
    p1 <- min(p, pos)
    for (i in 1:n) {
      for (j in 1:p1) {
        specscores[i,j] <- cor(comm[,i],ordiscores[,j])*sd(comm[,i])/sd(ordiscores[,j])
        if(is.na(specscores[i,j])) {specscores[i,j]<-0}
      }
    }
    if (Rscale==T && scaling=="2") {
      percen <- eigen.var/tot
      percen <- percen^0.5
      ordiscores <- sweep(ordiscores,2,percen,"/")
      specscores <- sweep(specscores,2,percen,"*")
    }
    if (Rscale==F) {
      specscores <- specscores / constant
      ordiscores <- ordi$points
    }
    ordi$points <- ordiscores
    ordi$eig <- eigen.var
    ordi$eig.percen <- eigen.percen
    ordi$eig.cumpercen <- eigen.cumpercen
    ordi$eigen.total <- tot
    ordi$R.constant <- constant
    ordi$Rscale <- Rscale
    ordi$scaling <- scaling
  }
  specscores <- specscores * multi
  ordi$cproj <- specscores
```

```
    return(ordi)
  }


fish.pcoa <- cmdscale(fish.db, eig = TRUE, k = 3)

#First 3 axes
var1 <- round(fish.pcoa$eig[1] / sum(fish.pcoa$eig), 3) * 100
var2 <- round(fish.pcoa$eig[2] / sum(fish.pcoa$eig), 3) * 100
var3 <- round(fish.pcoa$eig[3] / sum(fish.pcoa$eig), 3) * 100
sum.eig <- sum(var1, var2, var3)

#Plot
plot(fish.pcoa$points[ ,1], fish.pcoa$points[ ,2], ylim = c(-0.2, 0.7),
     xlab = paste("PCoA 1 (", var1, "%)", sep = ""),
     ylab = paste("PCoA 2 (", var2, "%)", sep = ""),
     pch = 16, cex = 2.0, type = "n", cex.lab = 1.5,
     cex.axis = 1.2, axes = FALSE)

# Add Axes
axis(side = 1, labels = T, lwd.ticks = 2, cex.axis = 1.2, las = 1)
axis(side = 2, labels = T, lwd.ticks = 2, cex.axis = 1.2, las = 1)
abline(h = 0, v = 0, lty = 3)
box(lwd = 2)

points(fish.pcoa$points[ ,1], fish.pcoa$points[ ,2],
       pch = 19, cex = 3, bg = "gray", col = "azure3")
text(fish.pcoa$points[ ,1], fish.pcoa$points[ ,2],
     labels = row.names(fish.pcoa$points))

#Influential Species
fishREL <- fish
  for(i in 1:nrow(fish)){
    fishREL[i, ] = fish[i, ] / sum(fish[i, ])
  }

library("vegan")

fish.pcoa <- add.spec.scores.class(fish.pcoa,fishREL, method = "pcoa.scores")
text(fish.pcoa$cproj[,1], fish.pcoa$cproj[,2],
     labels = row.names(fish.pcoa$cproj), col = "black")
```



In the R code chunk below, do the following:

1. identify influential species based on correlations along each PCoA axis (use a cutoff of 0.70), and
2. use a permutation test (999 permutations) to test the correlations of each species along each axis.

```
spe.corr <- add.spec.scores.class(fish.pcoa, fishREL, method = "cor.scores")$cproj
corrcut  <- 0.7
imp.spp  <- spe.corr[abs(spe.corr[, 1]) >= corrcut | abs(spe.corr[, 2]) >= corrcut, ]
print(imp.spp)
```

```
##            Dim1       Dim2        Dim3
## Phph -0.8674640 -0.1699316 -0.12463098
## Neba -0.7674114 -0.1855678 -0.36963830
## Rham  0.8088751 -0.4192567  0.14136301
## Legi  0.8201759 -0.1701803  0.12423941
## Cyca  0.7595122 -0.4442926  0.17313658
## Abbr  0.7704744 -0.3452714  0.29277803
## Acce  0.7635195  0.2155765  0.10288179
## Blbj  0.8118483 -0.1324698  0.25581178
## Alal  0.4471283  0.8119843 -0.05167131
## Anan  0.7974122 -0.3918972  0.20944968
```

```
# Permutation Test
fit <- envfit(fish.pcoa, fishREL, perm = 999)
```

Question 7: Address the following questions about the ordination results of the `doubs` data set:

a. Describe the grouping of sites in the Doubs River based on fish community composition.
b. Generate a hypothesis about which fish species are potential indicators of river quality.

Answer 7a: There are two main clusters presented on the PCoA plot with indvidual sites scattered throughout the graph. There seems to be a rough grouping of lower number sites (1-14) and a site of upper numbers (20-30). This could be reflective of upstream versus down stream.
Answer 7b: Given the high correlation coefficients along the PCoA (<.80), PhpH, Blbj, Rham, and Legi would all be strong canidates for indicator species. It should be noted that PhpH has a

negative correlation but still is a strong idicator that if their species is abundant than the river quailty is poor.

SYNTHESIS

Load the dataset from that you and your partner are using for the team project. Use one of the tools introduced in the beta diversity module to visualize your data. Describe any interesting patterns and identify a hypothesis is relevant to the principles of biodiversity. >Question: What is the influence of the type of mycorrhizal association on tree species abundance when exposed to the pressure of invasive species? Hypothesis: Mycorrhizal association will influence tree species abundance in plots containing invasive species. Prediction: In particular, plots with invasive pressures will have reduced biodiversity compared to plots without invasive species given the ability for invasive species to outcompete native species. when looking at mycorrhizal association, AM trees will have greater abundance in comparison to EMC species.

This is having so much trouble switching directories. I am not sure what do do- it is in text format while I try to figure out why it will not switch.

getwd() setwd("/cloud/project/QB_biodiversity_project_EH") # Load necessary libraries library(vegan) library(tibble) library(lattice)

Read the CSV file

tree <- read.csv("TREE.csv") #print(tree)

tree.species.df <- data.frame(Plot_ID = tree$PLOT, Species_ID = tree$SPCD) #print(tree.species.df)

tree.ss.df <- as.data.frame.matrix(table(tree.species.df$Plot_ID, tree.species.df$Species_ID)) tree.ss.df <- rownames_to_column(tree.ss.df, var = "Plot_ID") tree.ss.df <-tree.ss.df[1:100,1:100] #print(tree.ss.df)

tree.species.only.df <- tree.ss.df[, -1] print(tree.species.only.df)

tree.species.only.df.L <-tree.species.df

#Bray-Curtis

tree.db <-vegdist(tree.species.only.df, method = "bray")

order <-rev(attr(tree.db, "Labels")) levelplot(as.matrix(tree.db)[,order], aspect = "iso", col.regions =inferno, xlab = "Plot", ylab = "Plot", scales =list(cex =.5), main = "Bray-Curtis Distance"))