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

Dustin Brewer; Z620: Quantitative Biodiversity, Indiana University

05 February, 2019

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 today, it is *imperative* that you **push** this file to your GitHub repo, at whatever stage you are. Ths will enable you to pull your work onto your own computer.
6. When you have completed the worksheet, **Knit** the text and code into a single PDF file by pressing the `Knit` button in the RStudio scripting panel. This will save the PDF output in your '8.BetaDiversity' folder.
7. After Knitting, please submit the worksheet by making a **push** to your GitHub repo and then create a **pull request** via GitHub. Your pull request should include this file (**8.BetaDiversity_1_Worksheet.Rmd**) with all code blocks filled out and questions answered) and the PDF output of `Knitr` (**8.BetaDiversity_1_Worksheet**

The completed exercise is due on **Wednesday, February 6th, 2019 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, provide the code to:

1. clear your R environment,
2. print your current working directory,
3. set your working directory to your "*/8.BetaDiversity*" folder, and
4. load the `vegan` R package (be sure to install if needed).

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

```
## [1] "C:/Users/dusti/GitHub/QB2019_Brewer/2.Worksheets/8.BetaDiversity"
```

```
package.list <- c('vegan', 'ade4', 'viridis', 'gplots', 'BiodiversityR', 'indicspecies')
  for (package in package.list) {
    if (!require(package, character.only = TRUE, quietly = TRUE)) {
      install.packages(package)
      library(package, character.only=TRUE)
    }
  }
```

```
## This is vegan 2.5-3
```

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

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

```
## BiodiversityR 2.11-1: Use command BiodiversityRGUI() to launch the Graphical User Interface;
## to see changes use BiodiversityRGUI(changeLog=TRUE, backward.compatibility.messages=TRUE)
```

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, pleae do not print the dataset when submitting
data(doubs)
str(doubs$env)
```

```
## 'data.frame':    30 obs. of  11 variables:
##  $ dfs: num  3 22 102 185 215 324 268 491 705 990 ...
##  $ alt: num  934 932 914 854 849 846 841 792 752 617 ...
##  $ slo: num  6.18 3.43 3.64 3.5 3.18 ...
##  $ flo: num  84 100 180 253 264 286 400 130 480 1000 ...
##  $ pH : num  79 80 83 80 81 79 81 81 80 77 ...
##  $ har: num  45 40 52 72 84 60 88 94 90 82 ...
##  $ pho: num  1 2 5 10 38 20 7 20 30 6 ...
##  $ nit: num  20 20 22 21 52 15 15 41 82 75 ...
##  $ amm: num  0 10 5 0 20 0 0 12 12 1 ...
##  $ oxy: num  122 103 105 110 80 102 111 70 72 100 ...
##  $ bdo: num  27 19 35 13 62 53 22 81 52 43 ...
```

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***: It appears that there are 4 objects, 'species', 'xy', 'fish', and 'env' ***Answer 1b***: 27 species ***Answer 1c***: 30 sites

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***: Richness tends to increase as one moves from site 1 (upstream) to site 30 (downstream). ***Answer 2b***: Brown trout abundance varies distinctly, with none detected in sites 19-30, and between 1 and 5 individuals found at most of sites 1-18.
> ***Answer 2c***: Some limitations of using richness when examining patterns of diversity are that sites which have the same amount of species but do not have have similiar abundance of species are represented the same. Reporting diversity seems to be an important supplement to reporting richness.

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.

```
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)
}

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)
}
}

beta.w(doubs$fish)
```

```
## [1] 2.16
```

```
beta.w(doubs$fish, 1, 2, pairwise = TRUE)
```

```
## [1] 0.5
```

```
beta.w(doubs$fish, 1, 10, pairwise = TRUE)
```

```
## [1] 0.714
```

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 3a: Movement between sites appears to occur to some degree, as some sites share species. However, my interpretation is that regional diversity is 2.16 times more than the average richness at each site. Thus, it appears that processes are occuring which keep all of the sites from having the same species (perhaps a dam?).
Answer 3b: Sites 1 and 2 are more similiar than sites 1 and 10 in regard to fish assemblages.
Answer 3c: If beta were defined additively, then the value would indicate not how many more times speciose the regional species pool is than the average site, but rather how many more species the regional pool has than the average site.

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: Incidnence metrices treat rare species the same as any other species, but abundance-based metrics account for rarity.

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 SÃ¸rensen's Similarity ("fish.ds"), and
4. construct a resemblance matrix based on Bray-Curtis Distance ("fish.db").

```
fish <- doubs$fish
fish <- fish[-8, ]
fish.dj <- vegdist(fish, method = "jaccard", binary = TRUE)

fish.db <- vegdist(fish, method = "bray")


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

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 SÃ¸rensen or Bray-Curtis distance influence your interpretation of site (dis)similarity?

Answer 5a: The matrix represents dissimilarity. The bigger the #, the more dissimilar. I was able to determine this because when site 1 is compared with site 1, the value is '0.' ***Answer 5b***: It was a challenge for me to compare the matrices visually, but my impression is that using a Sorenson Index might tend to consider a pair of sites more similiar than the Bray-Curtis would because the Sorenson does not consider abundnace, only incidence.

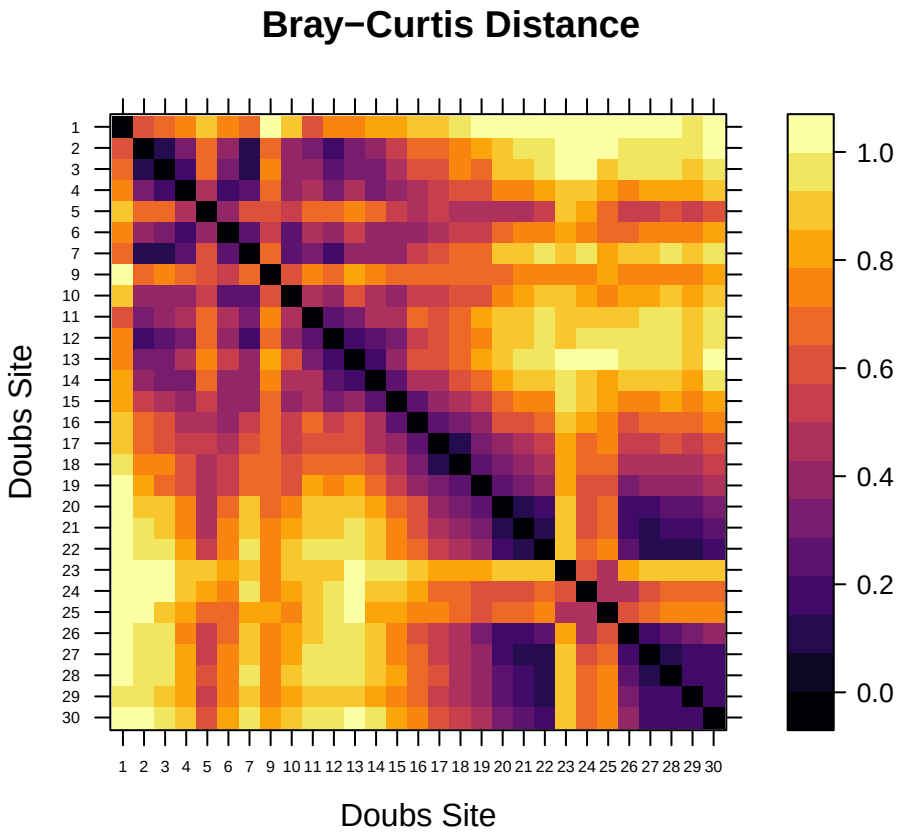
4) VISUALIZING BETA-DIVERSITY

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.

```
order <- rev(attr(fish.db, "Labels"))

heatmap <- 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")

heatmap
```

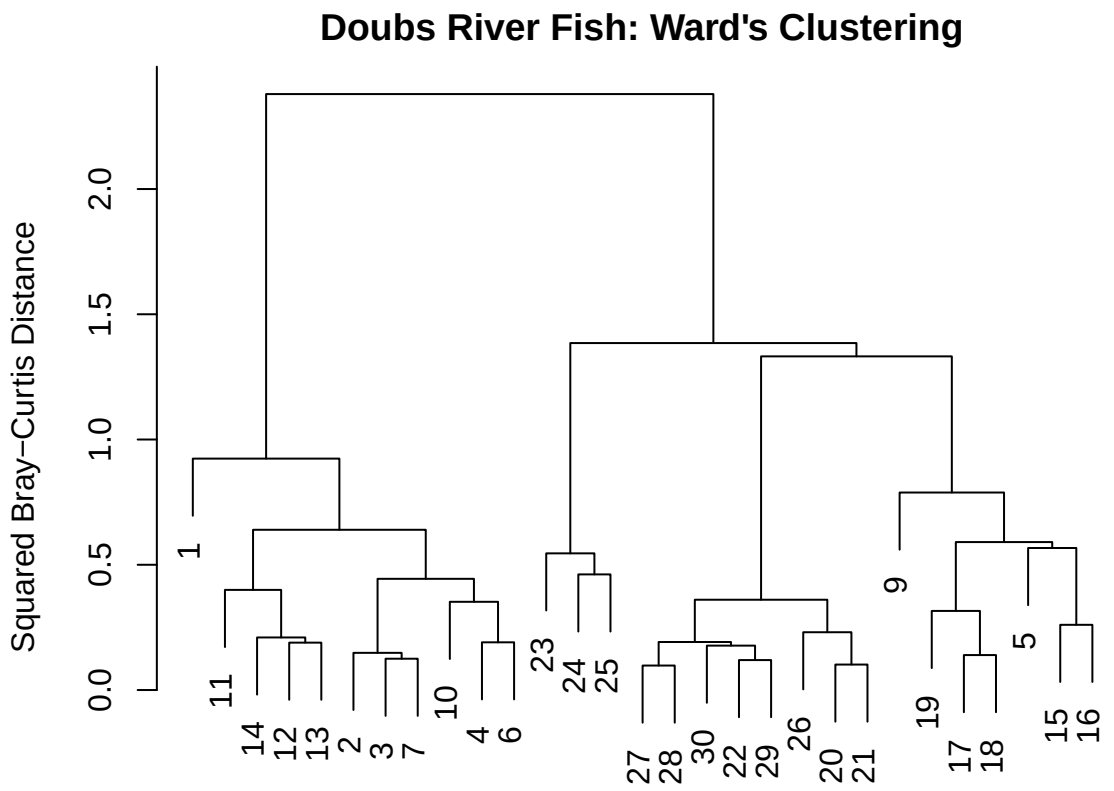


B. Cluster Analysis

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***: Hypothesis: sites are more similar that are closer to each other than are sites that are further from each other.

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.

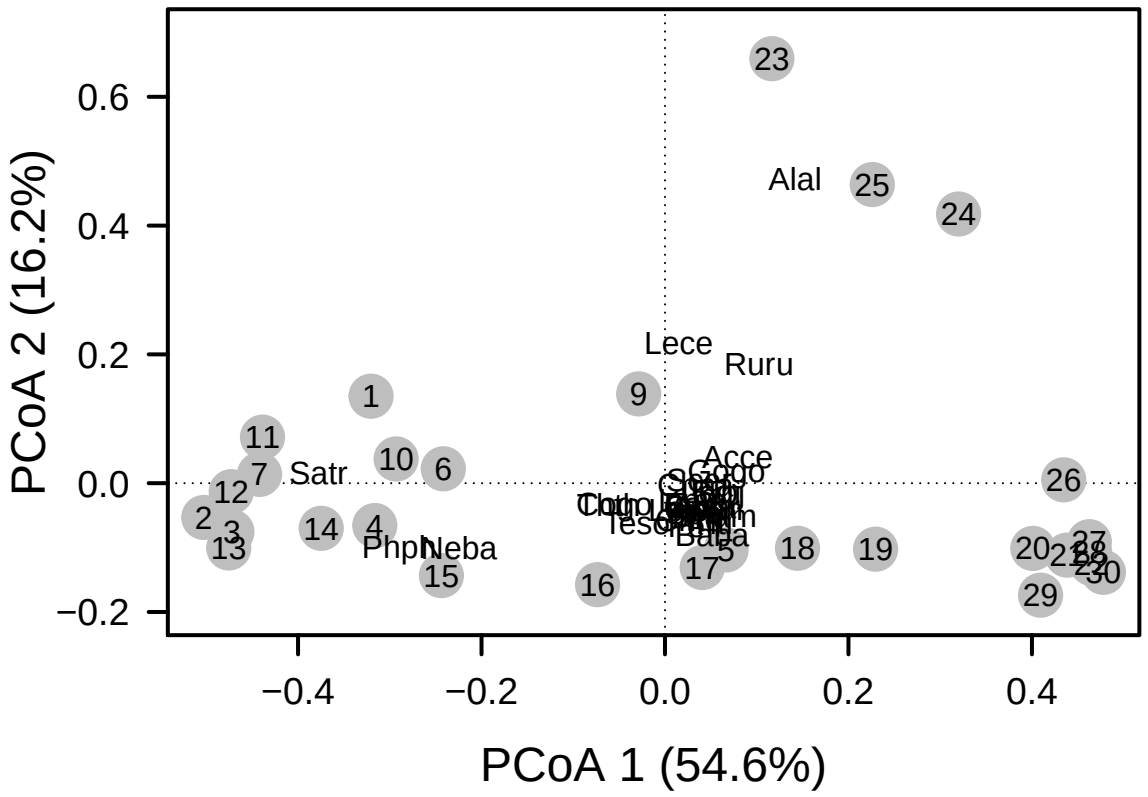
```
fish.pcoa <- cmdscale(fish.db, eig = TRUE, k = 3)
fish.pcoa
```

```
## $points
##           [,1]         [,2]         [,3]
## 1  -0.32043291  0.135094526  0.610444135
```

```
## 2  -0.50271272 -0.053204485  0.115084582
## 3  -0.47207593 -0.075509864  0.025191639
## 4  -0.31629787 -0.065793925 -0.054743001
## 5   0.06659050 -0.103685307 -0.061632012
## 6  -0.24164704  0.022330383 -0.117689415
## 7  -0.44197058  0.013721347 -0.019578989
## 9  -0.02864444  0.138633466 -0.189715336
## 10 -0.29305689  0.037682458 -0.202079336
## 11 -0.43875065  0.071482520  0.153768001
## 12 -0.47279103 -0.013160075 -0.006637071
## 13 -0.47550783 -0.100573921  0.079082050
## 14 -0.37472705 -0.070111200 -0.076607528
## 15 -0.24370479 -0.143518516 -0.131459333
## 16 -0.07348942 -0.157547743 -0.149621439
## 17  0.04061679 -0.130989703 -0.133635222
## 18  0.14445084 -0.101088883 -0.146573604
## 19  0.22960480 -0.102509909 -0.156586724
## 20  0.40100791 -0.101317037 -0.022574780
## 21  0.43837569 -0.111705508  0.048669123
## 22  0.46516232 -0.124748909  0.094770962
## 23  0.11656630  0.659234875 -0.006196177
## 24  0.32004280  0.418009171 -0.030531755
## 25  0.22601066  0.463713151 -0.110983530
## 26  0.43466174  0.005166762  0.006723518
## 27  0.46255346 -0.090842861  0.086824252
## 28  0.46286016 -0.106465286  0.108084941
## 29  0.40942763 -0.173891275  0.123971141
## 30  0.47787757 -0.138404251  0.164230911
##
## $eig
##  [1]  3.695331e+00  1.098472e+00  7.104740e-01  4.149729e-01  3.045604e-01
##  [6]  1.917884e-01  1.569703e-01  1.319099e-01  1.294251e-01  8.667896e-02
## [11]  4.615780e-02  3.864487e-02  2.745800e-02  1.306508e-02  7.087896e-03
## [16]  4.039469e-03  1.300594e-03 -9.346500e-17 -3.534426e-05 -3.940676e-03
## [21] -8.956051e-03 -1.461149e-02 -1.598905e-02 -2.145686e-02 -3.017013e-02
## [26] -3.432671e-02 -3.760052e-02 -6.037087e-02 -6.880061e-02
##
## $x
## NULL
##
## $ac
## [1] 0
##
## $GOF
## [1] 0.7484133 0.7798263
```

```
explainvar1 <- round(fish.pcoa$eig[1] / sum(fish.pcoa$eig), 3) * 100
explainvar2 <- round(fish.pcoa$eig[2] / sum(fish.pcoa$eig), 3) * 100
explainvar3 <- round(fish.pcoa$eig[3] / sum(fish.pcoa$eig), 3) * 100
sum.eig <- sum(explainvar1, explainvar2, explainvar3)

#To do ordination, first define plot parameters
par(mar = c(5,5,1,2) + 0.1)
```

```r
#Then, initiate plot
plot(fish.pcoa$points[ ,1], fish.pcoa$points[,2], ylim = c(-0.2, 0.7),
     xlab = paste("PCoA 1 (", explainvar1, "%)", sep = ""),
     ylab = paste("PCoA 2 (", explainvar2, "%)", 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)

#Add points and labels
points(fish.pcoa$points[ ,1], fish.pcoa$points[ ,2],
  pch = 19, cex = 3, bg = "gray", col = "gray")

text(fish.pcoa$points[ ,1], fish.pcoa$points[ ,2],
     labels = row.names(fish.pcoa$points))

#In order to show were particular species would show up on ordination (rather than just site #s), first

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

#Calculate and add species scores
fish.pcoa <- add.spec.scores(fish.pcoa, fishREL, method = "pcoa.scores")
text(fish.pcoa$cproj[ ,1], fish.pcoa$cproj[ ,2],
  labels = rownames(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.

```
#To determine the correlation coefficient of each species, and specify a correlation coefficient cut of

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

#Permutation test for species abundances across axes
fit <- envfit(fish.pcoa, fishREL, per = 999)
fit
```

```
##
## ***VECTORS
##
##          Dim1     Dim2     r2 Pr(>r)
## Cogo -0.83884 -0.54438 0.2982  0.008 **
## Satr -0.99904  0.04371 0.4326  0.005 **
## Phph -0.94110 -0.33813 0.7814  0.001 ***
## Neba -0.91413 -0.40543 0.6234  0.001 ***
## Thth -0.87692 -0.48063 0.2634  0.021 *
## Teso -0.44704 -0.89452 0.1700  0.081 .
## Chna  0.99707 -0.07644 0.4612  0.003 **
```

```
## Chto  0.42032 -0.90738 0.2579  0.017 *
## Lele  0.33041 -0.94384 0.0495  0.512
## Lece  0.06856  0.99765 0.3399  0.007 **
## Baba  0.54118 -0.84091 0.6752  0.001 ***
## Spbi  0.57341 -0.81927 0.4138  0.003 **
## Gogo  0.97507  0.22188 0.3753  0.006 **
## Eslu  0.72044 -0.69352 0.1673  0.083 .
## Pefl  0.43762 -0.89916 0.3048  0.009 **
## Rham  0.72476 -0.68901 0.8301  0.001 ***
## Legi  0.93461 -0.35568 0.7016  0.001 ***
## Scer  0.98569  0.16858 0.3533  0.008 **
## Cyca  0.68181 -0.73153 0.7743  0.001 ***
## Titi  0.64378 -0.76521 0.4586  0.001 ***
## Abbr  0.77254 -0.63497 0.7128  0.001 ***
## Icme  0.75626 -0.65427 0.5270  0.002 **
## Acce  0.88799  0.45986 0.6294  0.001 ***
## Ruru  0.48379  0.87518 0.5177  0.001 ***
## Blbj  0.95802 -0.28671 0.6766  0.001 ***
## Alal  0.28755  0.95777 0.8592  0.001 ***
## Anan  0.74277 -0.66954 0.7894  0.001 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Permutation: free
## Number of permutations: 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***: Though there are exceptions, it seems that fish community composition is related to location. For example, sites 23, 24, and 25 are clumped nearest 'Alal,' and and sites 1, 2, 3, and 4 are clumped around 'Satr.'
> ***Answer 7b***: Abundance of Alburnus alburnus is negatively correlated with concentration of E. coli in the water.

SYNTHESIS

Using the jelly bean data from class (i.e., JellyBeans.Source.txt and JellyBeans.txt):

1) Compare the average pairwise similarity among subsamples in group A to the average pairswise similarity among subsamples in group B. Use a t-test to determine whether compositional similarity was affected by the "vicariance" event. Finally, compare the compositional similarity of jelly beans in group A and group B to the source community?

```
Source <- read.table("JellyBeans.Source.txt")
Samples <- read.csv("JellyBeanSampleData.csv")

#Load manipulated data files
SamplesB2 <- read.csv("JellyBeanSampleDataB2.csv")
SamplesA2 <- read.csv("JellyBeanSampleDataA2.csv")
```

```r
#Compare all of the sites in B
beta.w(SamplesB2, 1, 2, pairwise = TRUE)
```

```
## [1] 0.182
```

```r
beta.w(SamplesB2, 1, 3, pairwise = TRUE)
```

```
## [1] 0.19
```

```r
beta.w(SamplesB2, 1, 4, pairwise = TRUE)
```

```
## [1] 0.163
```

```r
beta.w(SamplesB2, 2, 3, pairwise = TRUE)
```

```
## [1] 0.2
```

```r
beta.w(SamplesB2, 2, 4, pairwise = TRUE)
```

```
## [1] 0.268
```

```r
beta.w(SamplesB2, 3, 4, pairwise = TRUE)
```

```
## [1] 0.333
```

```r
#Make a vector of B pairwise beta.w values
PairwiseSampleB <- c(0.182, 0.19, 0.163, 0.2, 0.268, 0.333)

#Compare all of the sites in A
beta.w(SamplesA2, 1, 2, pairwise = TRUE)
```

```
## [1] 0.111
```

```r
beta.w(SamplesA2, 1, 3, pairwise = TRUE)
```

```
## [1] 0.182
```

```r
beta.w(SamplesA2, 1, 4, pairwise = TRUE)
```

```
## [1] 0.106
```

```r
beta.w(SamplesA2, 1, 5, pairwise = TRUE)
```

```
## [1] 0.13
```

```
beta.w(SamplesA2, 2, 3, pairwise = TRUE)
```

```
## [1] 0.171
```

```
beta.w(SamplesA2, 2, 4, pairwise = TRUE)
```

```
## [1] 0.136
```

```
beta.w(SamplesA2, 2, 5, pairwise = TRUE)
```

```
## [1] 0.163
```

```
beta.w(SamplesA2, 3, 4, pairwise = TRUE)
```

```
## [1] 0.209
```

```
beta.w(SamplesA2, 3, 5, pairwise = TRUE)
```

```
## [1] 0.19
```

```
beta.w(SamplesA2, 4, 5, pairwise = TRUE)
```

```
## [1] 0.111
```

```
PairwiseSamplesA <- c(0.111, 0.182, 0.106, 0.13, 0.171, 0.136, 0.163, 0.209, 0.19, 0.111)

#Should've used vegdist function?

#Compared the two samples using a t-test
t.test(PairwiseSampleB, PairwiseSamplesA)
```

```
##
##  Welch Two Sample t-test
##
## data:  PairwiseSampleB and PairwiseSamplesA
## t = 2.4803, df = 6.9895, p-value = 0.04224
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.003325572 0.140207761
## sample estimates:
## mean of x mean of y
## 0.2226667 0.1509000
```

```
#P value of .04, so there appears to be a significant difference between site A and B.

#I'm not sure how to compare with the source community.
```

2) Create a cluster diagram or ordination using the jelly bean data. Are there any visual trends that would suggest a difference in composition between group A and group B?

   The separate PCoA plots that I generated for site A and site B were similiar in that they both didn't have any points that were clumped together. I know that I still need to improve my understanding of PCoA, and am aware that I could've gone about doing this comparison wrong.

```
#For A

sampleA.db <- vegdist(SamplesA2, method = "bray")
sampleA.pcoa <- cmdscale(sampleA.db, eig = TRUE, k = 3)
sampleA.pcoa
```

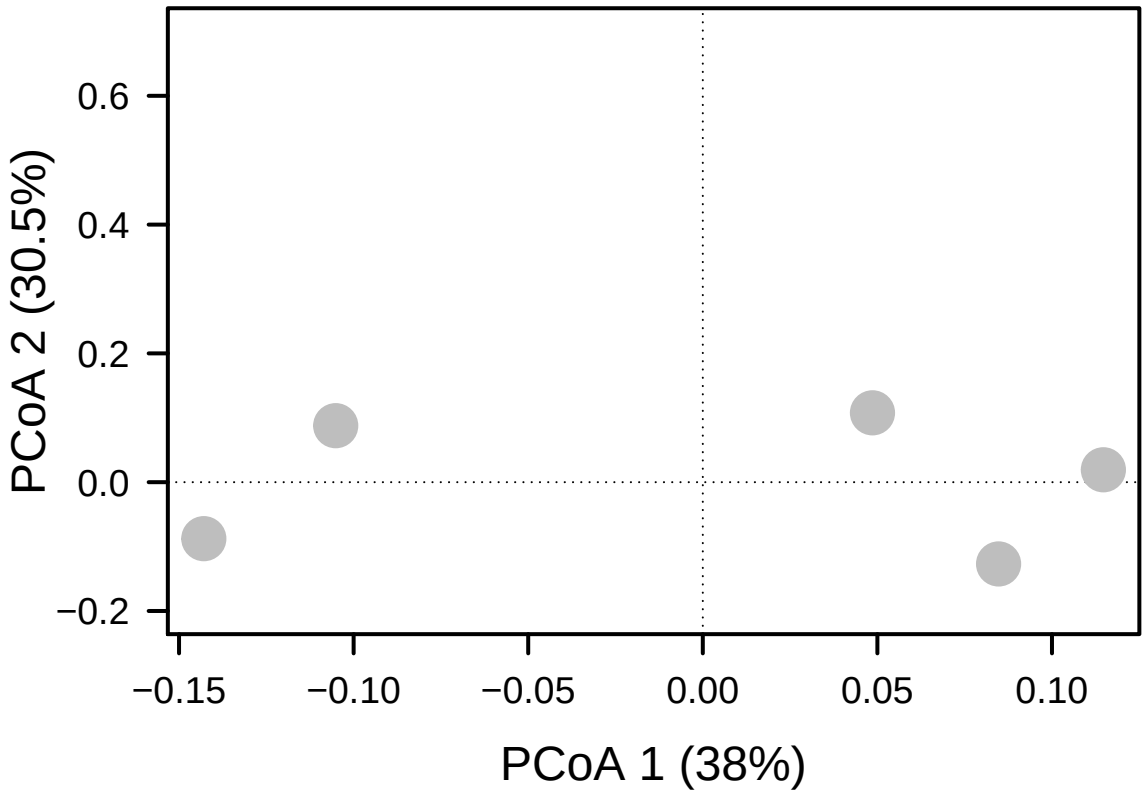
```
## $points
##             [,1]        [,2]        [,3]
## [1,]  0.11472303  0.01926287 -0.07035910
## [2,]  0.08470896 -0.12698233  0.09167583
## [3,] -0.14290128 -0.08772547 -0.08864036
## [4,]  0.04859651  0.10770716 -0.03436836
## [5,] -0.10512721  0.08773778  0.10169199
##
## $eig
## [1] 5.417111e-02 4.349008e-02 3.273442e-02 1.223826e-02 9.020562e-17
##
## $x
## NULL
##
## $ac
## [1] 0
##
## $GOF
## [1] 0.9141981 0.9141981
```

```
xplainvar1 <- round(sampleA.pcoa$eig[1] / sum(sampleA.pcoa$eig), 3) * 100
xplainvar2 <- round(sampleA.pcoa$eig[2] / sum(sampleA.pcoa$eig), 3) * 100
xplainvar3 <- round(sampleA.pcoa$eig[3] / sum(sampleA.pcoa$eig), 3) * 100
sum.eigA <- sum(explainvar1, explainvar2, explainvar3)

sum.eigA
```

```
## [1] 81.3
```

```
#To do ordination, first define plot parameters
par(mar = c(5,5,1,2) + 0.1)

#Then, initiate plot
plot(sampleA.pcoa$points[ ,1], sampleA.pcoa$points[,2], ylim = c(-0.2, 0.7),
     xlab = paste("PCoA 1 (", xplainvar1, "%)", sep = ""),
     ylab = paste("PCoA 2 (", xplainvar2, "%)", sep = ""),
     pch = 16, cex = 2.0, type = "n", cex.lab = 1.5, cex.axis = 1.2, axes = FALSE)

#Add axes
```

```r
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)

#Add points and labels
points(sampleA.pcoa$points[ ,1], sampleA.pcoa$points[ ,2],
  pch = 19, cex = 3, bg = "gray", col = "gray")

text(sampleA.pcoa$points[ ,1], sampleA.pcoa$points[ ,2],
     labels = row.names(sampleA.pcoa$points))
```




```r
#For B
sampleB.db <- vegdist(SamplesB2, method = "bray")
sampleB.pcoa <- cmdscale(sampleB.db, eig = TRUE, k = 3)
sampleB.pcoa
```

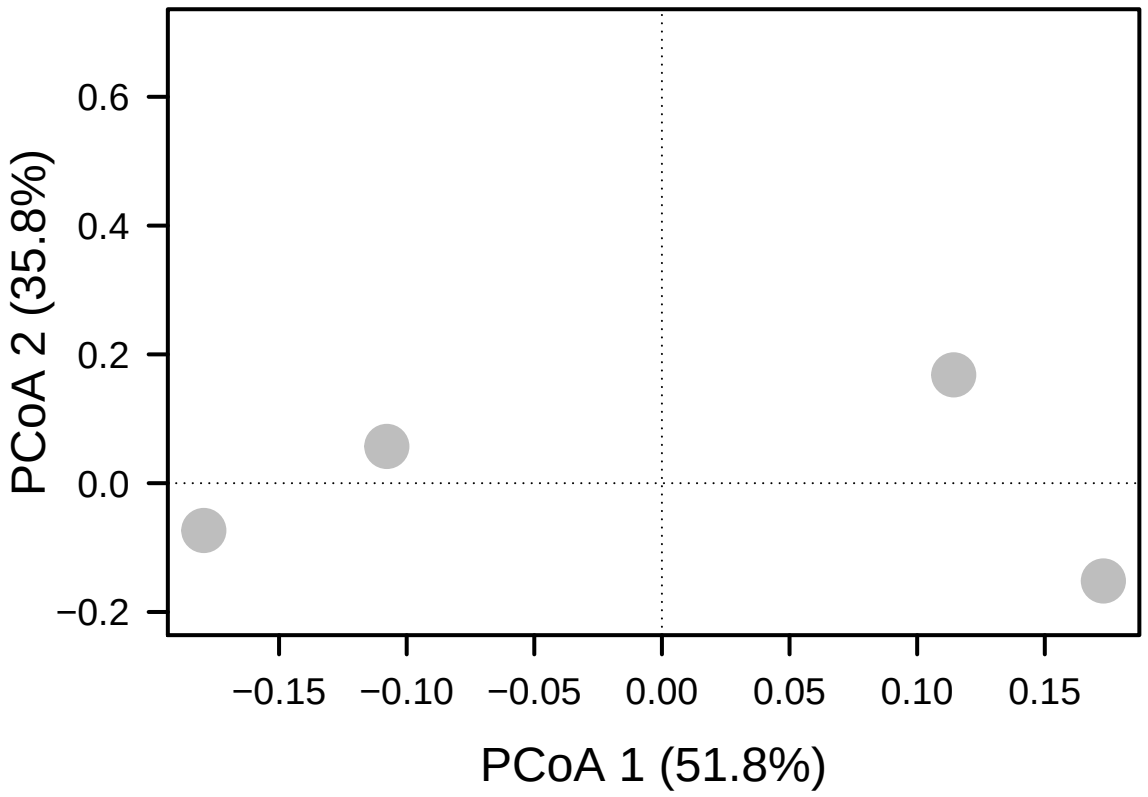
```
## $points
##            [,1]        [,2]        [,3]
## [1,] -0.1077832  0.05707839  0.10773785
## [2,] -0.1794738 -0.07345535 -0.07726931
## [3,]  0.1729458 -0.15184337  0.02093308
## [4,]  0.1143111  0.16822033 -0.05140162
##
## $eig
```

```
## [1]  8.680534e-02  6.000812e-02  2.065831e-02 -1.248783e-17
##
## $x
## NULL
##
## $ac
## [1] 0
##
## $GOF
## [1] 1 1
```

```
Explainvar1 <- round(sampleB.pcoa$eig[1] / sum(sampleB.pcoa$eig), 3) * 100
Explainvar2 <- round(sampleB.pcoa$eig[2] / sum(sampleB.pcoa$eig), 3) * 100
Explainvar3 <- round(sampleB.pcoa$eig[3] / sum(sampleB.pcoa$eig), 3) * 100
sum.eigB <- sum(explainvar1, explainvar2, explainvar3)

sum.eigB
```

```
## [1] 81.3
```

```
#To do ordination, first define plot parameters
par(mar = c(5,5,1,2) + 0.1)

#Then, initiate plot
plot(sampleB.pcoa$points[ ,1], sampleB.pcoa$points[,2], ylim = c(-0.2, 0.7),
     xlab = paste("PCoA 1 (", Explainvar1, "%)", sep = ""),
     ylab = paste("PCoA 2 (", Explainvar2, "%)", 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)

#Add points and labels
points(sampleB.pcoa$points[ ,1], sampleB.pcoa$points[ ,2],
  pch = 19, cex = 3, bg = "gray", col = "gray")

text(sampleB.pcoa$points[ ,1], sampleB.pcoa$points[ ,2],
     labels = row.names(sampleB.pcoa$points))
```


0.6
0.4
0.2
0.0
−0.2
PCoA 2 (35.8%)
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15
PCoA 1 (51.8%)