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

Emmi Mueller; 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.pdf**).

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/emmim/GitHub/QB2019_Mueller/2.Worksheets/8.BetaDiversity"
```

```
setwd("~/GitHub/QB2019_Mueller/2.Worksheets/8.BetaDiversity")
packages <- c("vegan","ade4","viridis","gplots","BiodiversityR","indicspecies", "gtools")
for(package in packages){
   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)
```

```
##
## Attaching package: 'gtools'
```

```
## The following object is masked from 'package:permute':
##
##     permute
```

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, max.level = 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:
```

```
head(doubs$env)
```

```
##   dfs alt   slo flo pH har pho nit amm oxy bdo
## 1   3 934 6.176  84 79  45   1  20   0 122  27
## 2  22 932 3.434 100 80  40   2  20  10 103  19
## 3 102 914 3.638 180 83  52   5  22   5 105  35
## 4 185 854 3.497 253 80  72  10  21   0 110  13
## 5 215 849 3.178 264 81  84  38  52  20  80  62
## 6 324 846 3.497 286 79  60  20  15   0 102  53
```

```
xy <- as.data.frame(doubs["xy"])
fish <- as.data.frame(doubs["fish"])
```

```r
species <- as.data.frame(doubs["species"])
env <- as.data.frame(doubs["env"])
```

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 four lists in the doubs dataset: env, fish, xy, species. ***Answer 1b***: There are 27 species in the doubs dataset. ***Answer 1c***: There are 30 species in the doubs dataset.

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: Species richness increases throughout the river but a few individual sites downstream have lower species richness than earlier sites. ***Answer 2b***: Brown Trout abundance decreases from upstream to downstream. ***Answer 2c***: While richness is an important metric, it does not show how individual species vary with distance from the upstream sites.

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.

```r
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(fish)
```

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

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

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

```
beta.w(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: Based on the calculation of Whittaker's β-Diversity as $\beta_w = \gamma/\alpha$, we know that $beta_w x \alpha$ is equal to γ so the turnover between sites multiplied by the average local diversity is equal to the regional diversity. ***Answer 3b***: The fish assemblage at site 1 is more similar to the assemblage at site 2 than at site 10 as is shown by the higher Whittaker's (β)-Diversity between 1 and 10 (0.714) than between 1 and 2 (0.5). ***Answer 3c***: If we defined beta attitively ($\beta = \gamma - \alpha$), it would be the difference between the total diversity and the average diversity as opposed to a comparison of how many more times diverse the regional pool is than the average site diversity.

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 metrics give the same weight to rare species as to abundance species whereas abundance-based metrics weight abundance species more heavily than rare species.

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 <- fish[-8,]
fish.dj <- vegdist(fish, method = "jaccard", binary = TRUE)
fish.db <- vegdist(fish, method = "bray")
fish.ds <- vegdist(fish, method = "bray", binary = TRUE)
fish.db.full <- vegdist(fish, method = "bray", upper = TRUE, diag = 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 resemblance matrix represents dissimilarity because when site 1 is compared to site 1, a β diversity of 0 is calculated. As site 1 and site 1 are completely the same, 0 represents total similarity and 1 represents total dissimilarity. ***Answer 5b***: As Bray-Curtis distances are based on abundances and Sørensen distances are based on incidence data, Sørensen distances show that the sites are more similar than the Bray-Curtis distances. This makes sense as communities with the same species but in different abundances will be seen as completely similar by the Sørensen calculation but not by the Bray-Curtis calculation.

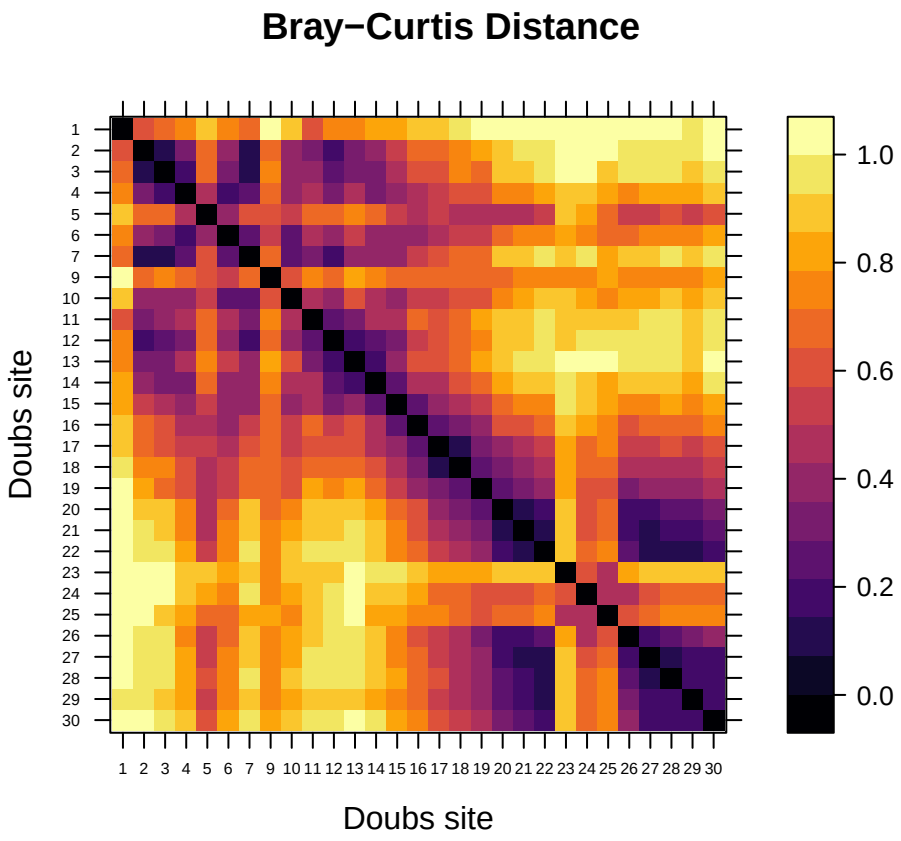
4) VISUALIZING BETA-DIVERSITY

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.full, "Labels"))
levelplot(as.matrix(fish.db.full)[, order], aspect = "iso", col.regions = inferno, xlab = "Doubs site",
```

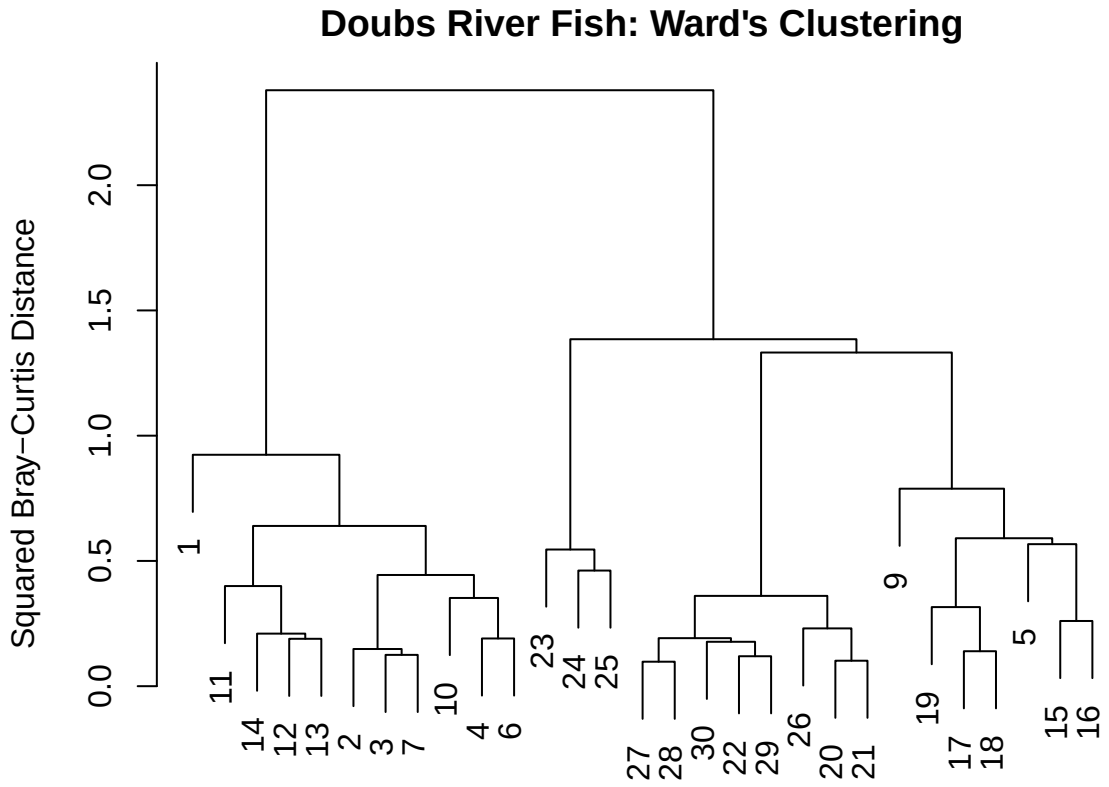


B. Cluster Analysis

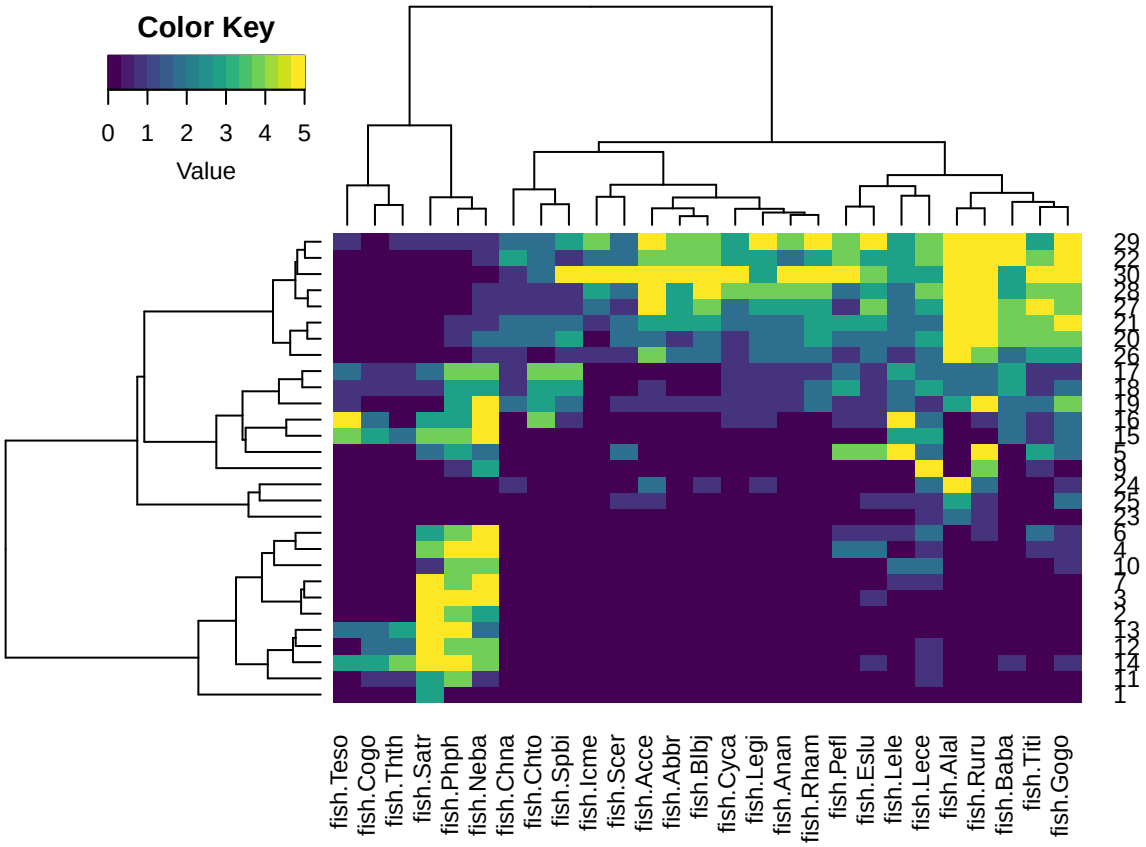
In the R code chunk below, do the following:

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

```r
fish.ward <- hclust(fish.db.full, 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")
```




```r
gplots::heatmap.2(as.matrix(fish), distfun = function(x) vegdist(x, method = "bray"), hclustfun = functi
```



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***: Sites in the Doubs river that are closer together tend to have similar environments leading to similar compositions of fish. There are other sites along the river where the environment tends to vary quickly through the river, maybe due to physical aspects of the river that prevent dispersal of species of fish between sites that are close together.

C. Ordination

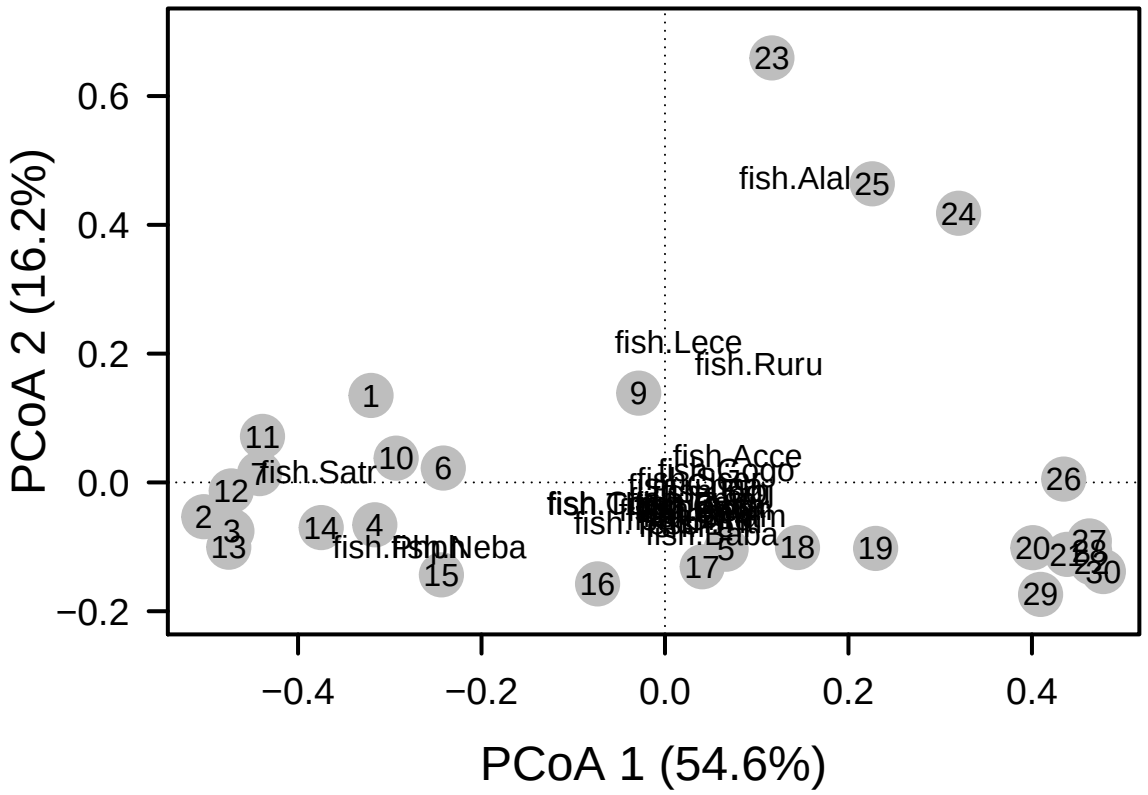
Principal Coordinates Analysis (PCoA)

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.full, eig = TRUE, k= 3)
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

par(mar = c(5,5,1,2) + 0.1)

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)

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 = "gray")

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

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

fish.pcoa <- add.spec.scores(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.

```r
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,]
imp.spp
```

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

```r
fit <- envfit(fish.pcoa, fishREL, perm = 999)
fit
```

```
##
## ***VECTORS
##
##               Dim1     Dim2     r2 Pr(>r)
## fish.Cogo -0.83884 -0.54438 0.2982  0.016 *
## fish.Satr -0.99904  0.04371 0.4326  0.003 **
## fish.Phph -0.94110 -0.33813 0.7814  0.001 ***
## fish.Neba -0.91413 -0.40543 0.6234  0.001 ***
## fish.Thth -0.87692 -0.48063 0.2634  0.028 *
## fish.Teso -0.44704 -0.89452 0.1700  0.081 .
## fish.Chna  0.99707 -0.07644 0.4612  0.001 ***
## fish.Chto  0.42032 -0.90738 0.2579  0.020 *
## fish.Lele  0.33041 -0.94384 0.0495  0.503
## fish.Lece  0.06856  0.99765 0.3399  0.009 **
## fish.Baba  0.54118 -0.84091 0.6752  0.001 ***
## fish.Spbi  0.57341 -0.81927 0.4138  0.001 ***
## fish.Gogo  0.97507  0.22188 0.3753  0.002 **
## fish.Eslu  0.72044 -0.69352 0.1673  0.099 .
## fish.Pefl  0.43762 -0.89916 0.3048  0.013 *
## fish.Rham  0.72476 -0.68901 0.8301  0.001 ***
## fish.Legi  0.93461 -0.35568 0.7016  0.001 ***
## fish.Scer  0.98569  0.16858 0.3533  0.008 **
## fish.Cyca  0.68181 -0.73153 0.7743  0.001 ***
## fish.Titi  0.64378 -0.76521 0.4586  0.001 ***
## fish.Abbr  0.77254 -0.63497 0.7128  0.001 ***
## fish.Icme  0.75626 -0.65427 0.5270  0.002 **
## fish.Acce  0.88799  0.45986 0.6294  0.001 ***
## fish.Ruru  0.48379  0.87518 0.5177  0.002 **
## fish.Blbj  0.95802 -0.28671 0.6766  0.001 ***
## fish.Alal  0.28755  0.95777 0.8592  0.001 ***
## fish.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: The sites in the Doubs River that are closer together are more similar in fish community composition than those farther away. ***Answer 7b***: Based on the PCoA, I expect that the bleak, roach, stone loach, minnow, and brown trout are potential indicators of river quality.

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?

```
jelly <- as.data.frame(read.table("./JellyBeans.Source.txt", header = TRUE, sep = "\t", row.names = "St
jelly <- as.data.frame(t(jelly))
jelly <- as.data.frame(jelly["Count",])

jellyAB <- read.table("./JellyBeans_site.txt", header = TRUE, sep = "\t", row.names = 1)
jellybeangroups <- as.data.frame(jellyAB["Group"])
jellyAB[,"GreenTrans"] <- jellyAB[,"GreenTrans"]+jellyAB[,"GreenTrans2"]
jellyAB[,"Rainbow"] <- jellyAB[,"WhiteSolid"] + jellyAB[,"Rainbow"]
jellyAB[,c("Group", "GreenTrans2", "WhiteSolid")] <- list(NULL)

betaA <- 0
betaAcount <- 0
betaA.df <- as.data.frame(NA)
betaB <- 0
betaBcount <- 0
betaB.df <- as.data.frame(NA)
end <- nrow(jellyAB)
for(row in row.names(jellyAB)){
  n <- which(rownames(jellyAB) == row)
  rows <- row.names(jellyAB[n:end,])
  for(compare in rows){
    if(jellybeangroups[row,"Group"] == jellybeangroups[compare,"Group"]){
      if(as.character(jellybeangroups[row,"Group"]) == "A" & (row != compare)){
        betaA.df <- rbind(betaA.df, beta.w(jellyAB, row, compare, pairwise = TRUE))
        betaA <- betaA + beta.w(jellyAB, row, compare, pairwise = TRUE)
        betaAcount <- betaAcount + 1
      }
      else if (as.character(jellybeangroups[row,"Group"]) == "B" & (row != compare)){
        betaB.df <- rbind(betaB.df, beta.w(jellyAB, row, compare, pairwise = TRUE))
        betaB <- betaB + beta.w(jellyAB, row, compare, pairwise = TRUE)
        betaBcount <- betaBcount + 1
      }
    }
  }
}
avgbetaA <- betaA/betaAcount
avgbetaB <- betaB/betaBcount
cat("Average pairwise similarity for group A: ", avgbetaA, "\n")
```

```
## Average pairwise similarity for group A:  0.1509
```

```
cat("Average pairwise similarity for group B: ", avgbetaB)
```

```
## Average pairwise similarity for group B:  0.2226667
```

```
betaA.df <- betaA.df[-1,]
betaB.df <- betaB.df[-1,]

t.test(betaA.df, betaB.df)
```

```
##
##  Welch Two Sample t-test
##
## data:  betaA.df and betaB.df
## 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:
##  -0.140207761 -0.003325572
## sample estimates:
## mean of x mean of y
## 0.1509000 0.2226667
```

Synthesis 1:The t-test between the pairwise beta diversity of Group A and Group B shows that Group B has a significantly higher pairwise dissimilarity than Group A when looking exclusively at the incidence data of both groups to calculate beta diversity.

```
x <- rep(0, ncol(jellyAB))
names <- row.names(jellyAB)
jellyABcombine <- rbind(x, x, jellyAB)
row.names(jellyABcombine) <- c("GroupA", "GroupB", names)
for(column in colnames(jellyABcombine)){
  for (row in rownames(jellyABcombine)){
    if (row == "GroupA" || row == "GroupB"){}
    else if (jellybeangroups[row,] == "A"){
          jellyABcombine["GroupA", column] <- jellyABcombine["GroupA", column] + jellyABcombine[row, co
    }
    else if (jellybeangroups[row,] == "B"){
          jellyABcombine["GroupB", column] <- jellyABcombine["GroupB", column] + jellyABcombine[row, co
    }
  }
}
jelly <- t(read.table("./JellyBeans.Source.reduced.txt", header = TRUE, row.names = "Student.ID"))

jellyABcombine <- smartbind(jellyABcombine, jelly)
row.names(jellyABcombine) <- c("GroupA", "GroupB", names, "Source")
jellyABcombine <- jellyABcombine[c("GroupA", "GroupB", "Source"),]

vegdist(jellyABcombine, method = "bray", upper = TRUE, diag = TRUE)
```

```
##           GroupA    GroupB    Source
## GroupA 0.0000000 0.2507463 0.4315322
## GroupB 0.2507463 0.0000000 0.5434956
## Source 0.4315322 0.5434956 0.0000000
```

Synthesis 1: The pairwise distance between Group B and Source is higher than between Group A and Source. Group A and Group B are more similar to each other than either of them is to

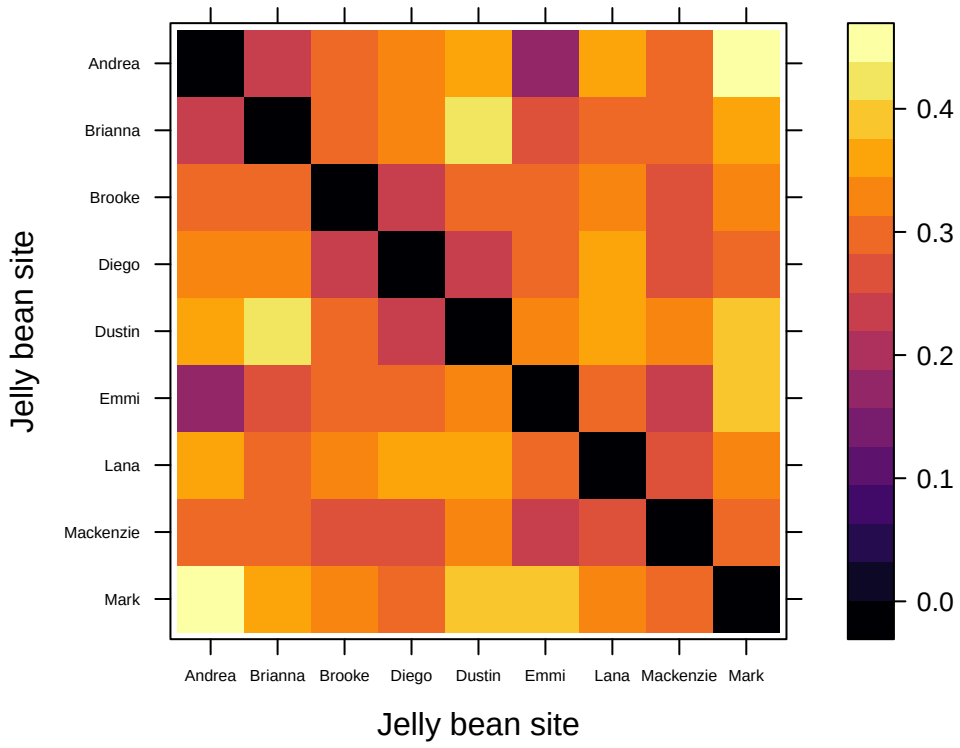
Source but this may be due to not having been rarified before calculating the beta diversity.

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?

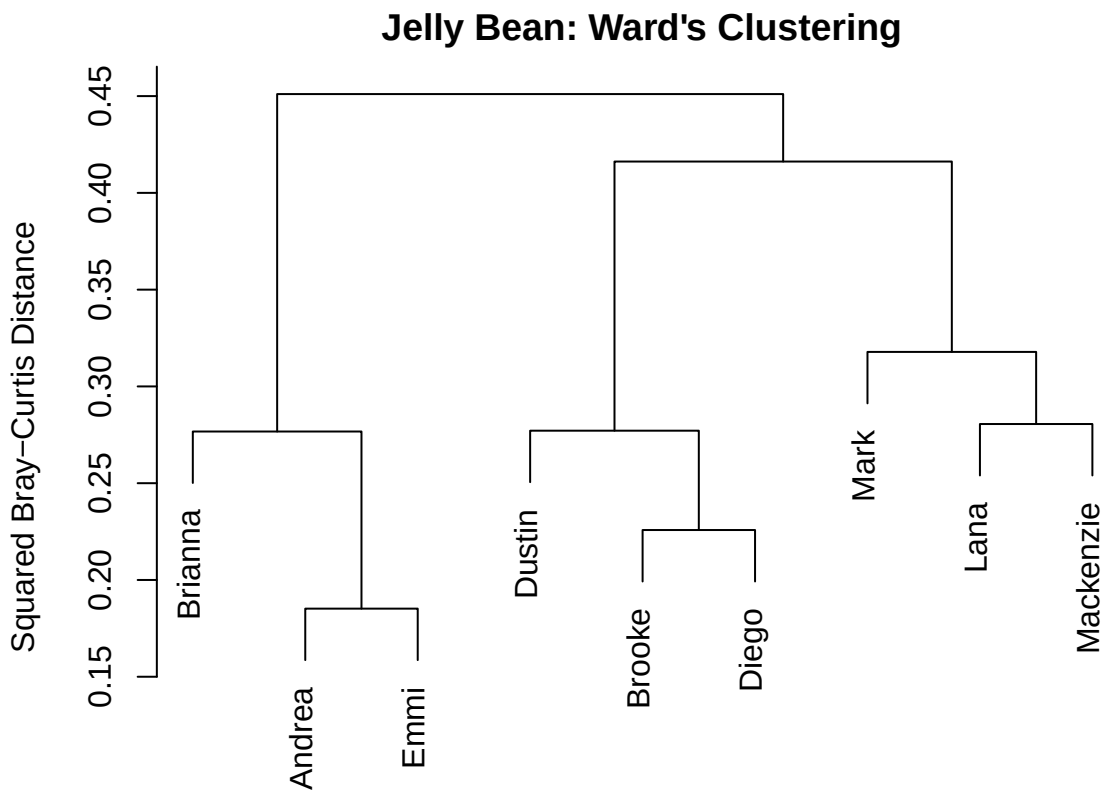
```
jellyAB.db <- vegdist(jellyAB, method = "bray", upper = TRUE, diag = TRUE)
print(jellyAB.db)
```

```
##              Andrea   Brianna    Brooke     Diego    Dustin      Emmi
## Andrea    0.0000000 0.2387097 0.2903226 0.3250000 0.3750000 0.1851852
## Brianna   0.2387097 0.0000000 0.2933333 0.3161290 0.4193548 0.2738854
## Brooke    0.2903226 0.2933333 0.0000000 0.2258065 0.2903226 0.2866242
## Diego     0.3250000 0.3161290 0.2258065 0.0000000 0.2375000 0.3086420
## Dustin    0.3750000 0.4193548 0.2903226 0.2375000 0.0000000 0.3333333
## Emmi      0.1851852 0.2738854 0.2866242 0.3086420 0.3333333 0.0000000
## Lana      0.3469388 0.2957746 0.3380282 0.3605442 0.3741497 0.2885906
## Mackenzie 0.2894737 0.2925170 0.2653061 0.2763158 0.3157895 0.2337662
## Mark      0.4388489 0.3731343 0.3432836 0.2949640 0.3812950 0.3758865
##                Lana Mackenzie      Mark
## Andrea    0.3469388 0.2894737 0.4388489
## Brianna   0.2957746 0.2925170 0.3731343
## Brooke    0.3380282 0.2653061 0.3432836
## Diego     0.3605442 0.2763158 0.2949640
## Dustin    0.3741497 0.3157895 0.3812950
## Emmi      0.2885906 0.2337662 0.3758865
## Lana      0.0000000 0.2805755 0.3333333
## Mackenzie 0.2805755 0.0000000 0.2824427
## Mark      0.3333333 0.2824427 0.0000000
```

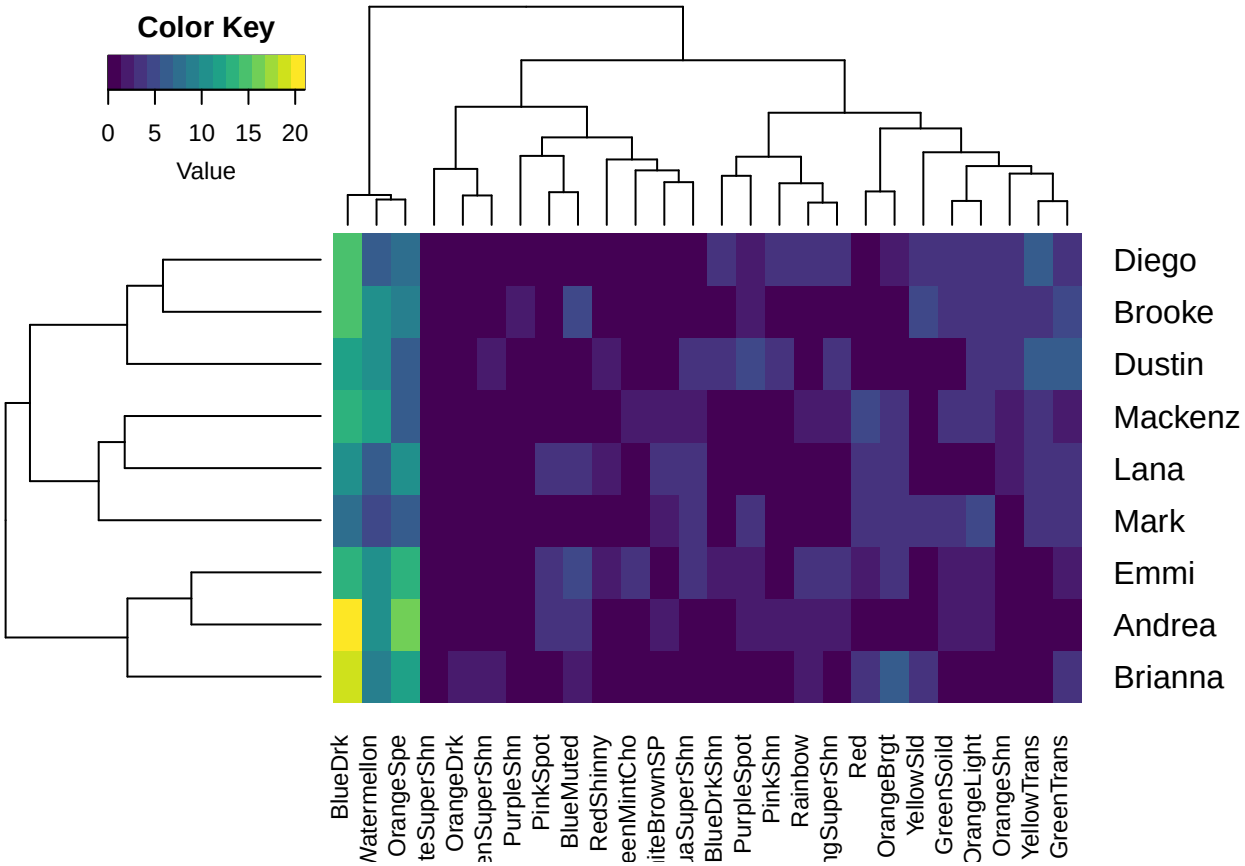
```
order <- rev(attr(jellyAB.db, "Labels"))
levelplot(as.matrix(jellyAB.db)[, order], aspect = "iso", col.regions = inferno, xlab = "Jelly bean sit
```




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

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




```
gplots::heatmap.2(as.matrix(jellyAB), distfun = function(x) vegdist(x, method = "bray"), hclustfun = fu
```



>***Synthesis 2:***While a few of the Group A sites group together (Andrea, Emmi, Brianna), the other two sites are found in clusters with Group B sites meaning the vicariance event is not detectable by Ward's clustering of community composition diversity.