# 9.Phylogenetic Diversity - Communities

Erica Nadolski; Z620: Quantitative Biodiversity, Indiana University

28 February, 2023

## OVERVIEW

Complementing taxonomic measures of $\alpha$- and $\beta$-diversity with evolutionary information yields insight into a broad range of biodiversity issues including conservation, biogeography, and community assembly. In this worksheet, you will be introduced to some commonly used methods in phylogenetic community ecology.

After completing this assignment you will know how to:

1. incorporate an evolutionary perspective into your understanding of community ecology
2. quantify and interpret phylogenetic $\alpha$- and $\beta$-diversity
3. evaluate the contribution of phylogeny to spatial patterns of biodiversity

### Directions:

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 '9.PhyloCom' folder. 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 *9.PhyloCom_Worksheet.Rmd* and the PDF output of `Knitr` (*9.PhyloCom_Worksheet.pdf*). The completed exercise is due on **Wednesday, March 1st, 2023 before 12:00 PM (noon)**.

## 1) 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 `/9.PhyloCom` folder,
4. load all of the required R packages (be sure to install if needed), and
5. load the required R source file.

```
package.list <- c("picante",'ape', 'seqinr', "vegan","fossil","reshape",
                  "devtools","BiocManager","ineq","labdsv",
                  "matrixStats","pROC",'phylobase', 'adephylo', 'geiger',
                  'stats', 'RColorBrewer', 'caper', 'phylolm', 'pmc',
                  'ggplot2', 'tidyr', 'dplyr', 'phangorn', 'pander',
                  'phylogram', 'dendextend')
for (package in package.list) {
```

```
  if (!require(package, character.only=TRUE, quietly=TRUE)) {
    install.packages(package)
    library(package, character.only=TRUE)
  }
}
```

```
## This is vegan 2.6-4
```

```
##
## Attaching package: 'seqinr'
```

```
## The following object is masked from 'package:nlme':
##
##     gls
```

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

```
## The following objects are masked from 'package:ape':
##
##     as.alignment, consensus
```

```
##
## Attaching package: 'shapefiles'
```

```
## The following objects are masked from 'package:foreign':
##
##     read.dbf, write.dbf
```

```
##
## Attaching package: 'devtools'
```

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

```
##
## Attaching package: 'BiocManager'
```

```
## The following object is masked from 'package:devtools':
##
##     install
```

```
## This is mgcv 1.8-41. For overview type 'help("mgcv-package")'.
```

```
## Registered S3 method overwritten by 'labdsv':
##   method       from
##   summary.dist ade4
```

```
## This is labdsv 2.0-1
## convert existing ordinations with as.dsvord()

##
## Attaching package: 'labdsv'

## The following object is masked from 'package:stats':
##
##     density

##
## Attaching package: 'matrixStats'

## The following object is masked from 'package:seqinr':
##
##     count

## Type 'citation("pROC")' for a citation.

##
## Attaching package: 'pROC'

## The following objects are masked from 'package:stats':
##
##     cov, smooth, var

##
## Attaching package: 'phylobase'

## The following object is masked from 'package:ape':
##
##     edges

##
## Attaching package: 'tidyr'

## The following objects are masked from 'package:reshape':
##
##     expand, smiths

##
## Attaching package: 'dplyr'

## The following object is masked from 'package:MASS':
##
##     select

## The following object is masked from 'package:matrixStats':
##
##     count
```

```
## The following object is masked from 'package:reshape':
##
##     rename

## The following object is masked from 'package:seqinr':
##
##     count

## The following object is masked from 'package:nlme':
##
##     collapse

## The following object is masked from 'package:ape':
##
##     where

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

##
## Attaching package: 'phangorn'

## The following object is masked from 'package:pROC':
##
##     coords

## The following objects are masked from 'package:vegan':
##
##     diversity, treedist

##
## Attaching package: 'phylogram'

## The following object is masked from 'package:phylobase':
##
##     prune

## Registered S3 methods overwritten by 'dendextend':
##   method              from
##   as.dendrogram.phylo phylogram
##   rev.hclust          vegan

##
## --------------------
## Welcome to dendextend version 1.16.0
## Type citation('dendextend') for how to cite the package.
```

```
##
## Type browseVignettes(package = 'dendextend') for the package vignette.
## The github page is: https://github.com/talgalili/dendextend/
##
## Suggestions and bug-reports can be submitted at: https://github.com/talgalili/dendextend/issues
## You may ask questions at stackoverflow, use the r and dendextend tags:
##   https://stackoverflow.com/questions/tagged/dendextend
##
##  To suppress this message use:  suppressPackageStartupMessages(library(dendextend))
## ----------------------

##
## Attaching package: 'dendextend'

## The following object is masked from 'package:phylogram':
##
##     prune

## The following object is masked from 'package:geiger':
##
##     is.phylo

## The following objects are masked from 'package:phylobase':
##
##     labels<-, prune

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

## The following objects are masked from 'package:ape':
##
##     ladderize, rotate

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

```
source("/Users/ericanadolski/GitHub/QB2023_Nadolski/2.Worksheets/9.PhyloCom/bin/MothurTools.R")
```

## 2) DESCRIPTION OF DATA

**need to discuss data set from spatial ecology!**

We sampled >50 forested ponds in Brown County State Park, Yellowood State Park, and Hoosier National Forest in southern Indiana. In addition to measuring a suite of geographic and environmental variables, we characterized the diversity of bacteria in the ponds using molecular-based approaches. Specifically, we amplified the 16S rRNA gene (i.e., the DNA sequence) and 16S rRNA transcripts (i.e., the RNA transcript of the gene) of bacteria. We used a program called `mothur` to quality-trim our data set and assign sequences to operational taxonomic units (OTUs), which resulted in a site-by-OTU matrix.
In this module we will focus on taxa that were present (i.e., DNA), but there will be a few steps where we need to parse out the transcript (i.e., RNA) samples. See the handout for a further description of this week's dataset.

## 3) LOAD THE DATA

In the R code chunk below, do the following:
1. load the environmental data for the Brown County ponds (*20130801_PondDataMod.csv*),
2. load the site-by-species matrix using the `read.otu()` function,
3. subset the data to include only DNA-based identifications of bacteria,
4. rename the sites by removing extra characters,
5. remove unnecessary OTUs in the site-by-species, and
6. load the taxonomic data using the `read.tax()` function from the source-code file.

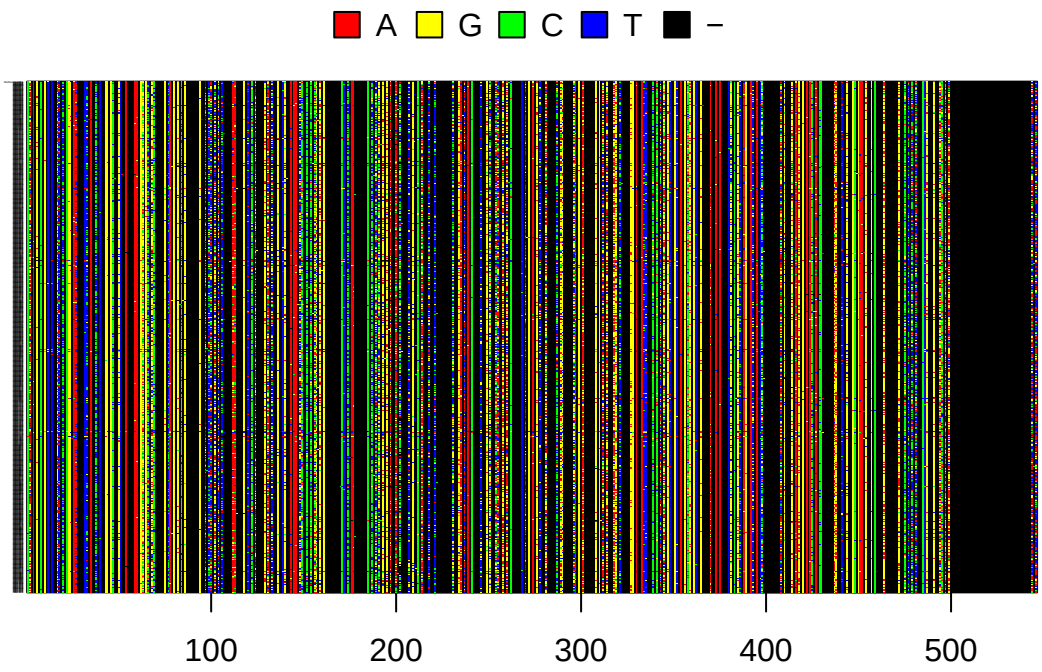
```
env <- read.table("/Users/ericanadolski/GitHub/QB2023_Nadolski/2.Worksheets/9.PhyloCom/data/20130801_Po
env <- na.omit(env)

# site by species
comm <- read.otu(shared = "/Users/ericanadolski/GitHub/QB2023_Nadolski/2.Worksheets/9.PhyloCom/data/INP

# select DNA using grep
comm <- comm[grep("*-DNA", rownames(comm)), ]

# perform replacement of all matches with gsub
rownames(comm) <- gsub("\\-DNA", "", rownames(comm))
rownames(comm) <- gsub("_", "", rownames(comm))

# remove sites not in env dataset
comm <- comm[rownames(comm) %in% env$Sample_ID, ]

# remove zero-abundance OTUs from data set
comm <- comm[ , colSums(comm) > 0]

# import taxonomy
tax <- read.tax(taxonomy = "/Users/ericanadolski/GitHub/QB2023_Nadolski/2.Worksheets/9.PhyloCom/data/IN
```

```
## Warning in type.convert.default(as.character(x)): 'as.is' should be specified
## by the caller; using TRUE

## Warning in type.convert.default(as.character(x)): 'as.is' should be specified
## by the caller; using TRUE

## Warning in type.convert.default(as.character(x)): 'as.is' should be specified
## by the caller; using TRUE

## Warning in type.convert.default(as.character(x)): 'as.is' should be specified
## by the caller; using TRUE

## Warning in type.convert.default(as.character(x)): 'as.is' should be specified
## by the caller; using TRUE

## Warning in type.convert.default(as.character(x)): 'as.is' should be specified
## by the caller; using TRUE
```

Next, in the R code chunk below, do the following:
1. load the FASTA alignment for the bacterial operational taxonomic units (OTUs), 2. rename the OTUs by removing everything before the tab (\t) and after the bar (|),

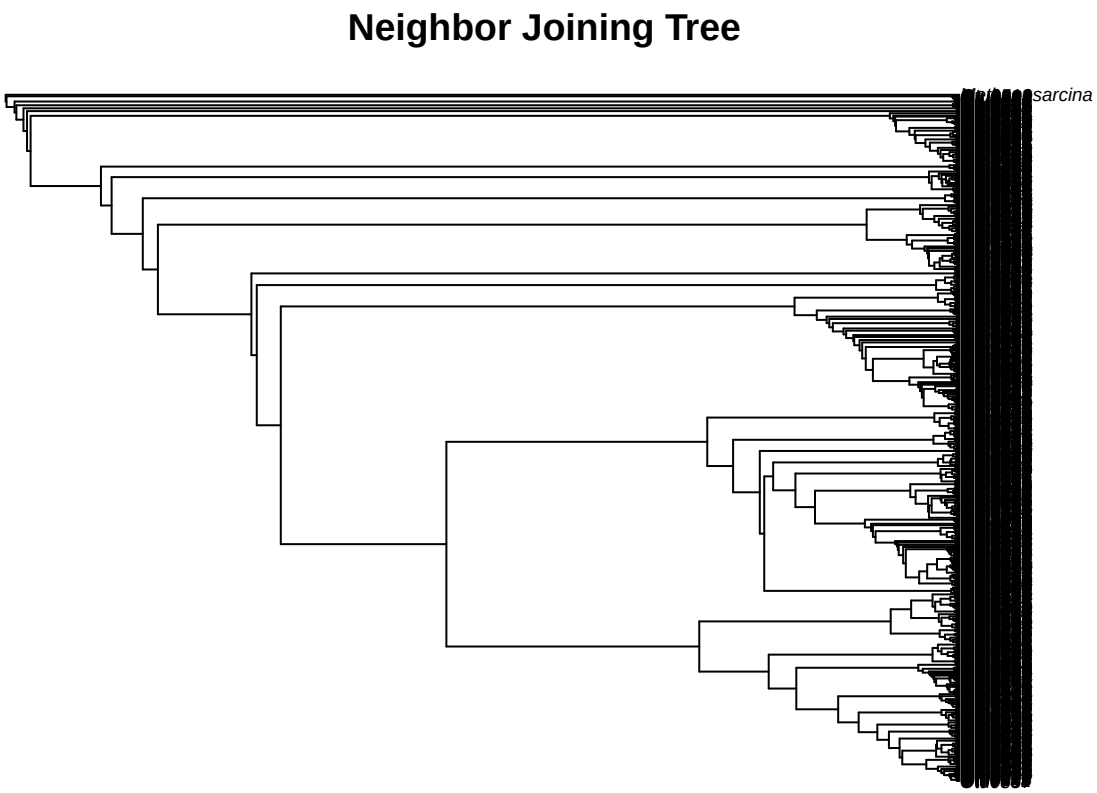
3. import the *Methanosarcina* outgroup FASTA file,
4. convert both FASTA files into the DNAbin format and combine using `rbind()`, 5. visualize the sequence alignment,
6. using the alignment (with outgroup), pick a DNA substitution model, and create a phylogenetic distance matrix,
7. using the distance matrix above, make a neighbor joining tree,
8. remove any tips (OTUs) that are not in the community data set,
9. plot the rooted tree.

```
# import alignment file
ponds.cons <- read.alignment(file = "/Users/ericanadolski/GitHub/QB2023_Nadolski/2.Worksheets/9.PhyloCo

# rename OTUs in fasta
ponds.cons$nam <- gsub("\\|.*$", "", gsub("^.*?\t", "", ponds.cons$nam))

# import outgroup
outgroup <- read.alignment(file="/Users/ericanadolski/GitHub/QB2023_Nadolski/2.Worksheets/9.PhyloCom/da

#convert alignment files to DNAbin and combine
DNAbin <- rbind(as.DNAbin(outgroup),as.DNAbin(ponds.cons))

#visualize alignment
image.DNAbin(DNAbin, show.labels=T, cex.lab=0.05, las=1)
```




```
# make distance matrix {ape}
seq.dist.jc <- dist.dna(DNAbin, model="JC", pairwise.deletion=FALSE)

# make neighbor joining tree file {ape}
phy.all <- bionj(seq.dist.jc)


# drop tips of zero occurrence OTUs {ape}
```

```
phy <- drop.tip(phy.all, phy.all$tip.label[!phy.all$tip.label %in% c(colnames(comm), "Methanosarcina")]
# ID outgroup
outgroup <- match("Methanosarcina", phy$tip.label)
# root tree
phy <- root(phy,"Methanosarcina",resolve.root = TRUE)
#plot rooted tree {ape}
par(mar = c(1, 1, 2, 1) + 0.1)
plot.phylo(phy, main = "Neighbor Joining Tree", "phylogram",
      show.tip.label =TRUE, use.edge.length = FALSE,
      direction = "right", cex = 0.6, label.offset = 1)
```




## 4) PHYLOGENETIC ALPHA DIVERSITY

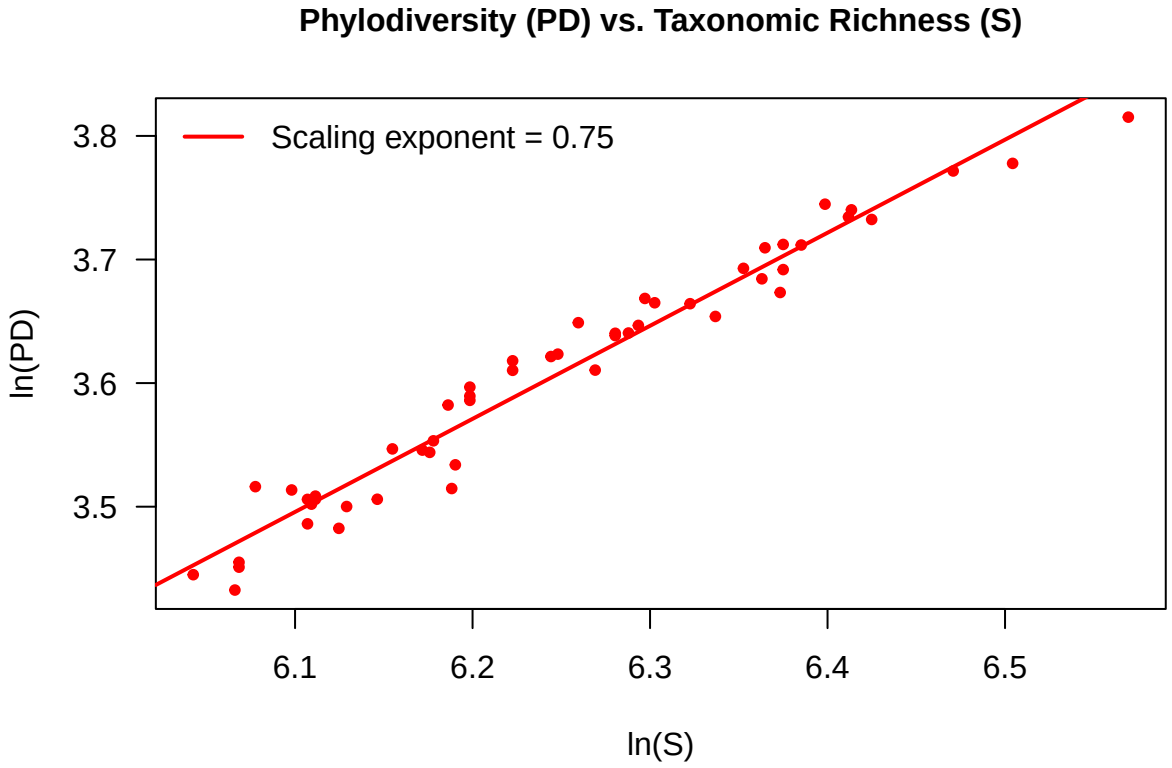
### A. Faith's Phylogenetic Diversity (PD)

In the R code chunk below, do the following:
1. calculate Faith's D using the `pd()` function.

```
# calculate PD and S {picante}
pd <- pd(comm,phy,include.root=FALSE)
```

In the R code chunk below, do the following:
1. plot species richness (S) versus phylogenetic diversity (PD),
2. add the trend line, and
3. calculate the scaling exponent.

```r
# test of power law relationship
fit <- lm(log(pd$PD) ~ log(pd$S))
exponent <- round(coefficients(fit)[2], 2)

# bi plot of S and PD
par(mar=c(5,5,4,1)+0.1)

plot(log(pd$S), log(pd$PD),
     pch=20,col="red",las=1, xlab="ln(S)", ylab="ln(PD)", cex.main=1,
     main="Phylodiversity (PD) vs. Taxonomic Richness (S)");
abline(fit, col="red",lw=2);
legend("topleft", legend=paste("Scaling exponent = ", exponent, sep=""),
       bty="n", lw=2, col="red")
```




***Question 1***: Answer the following questions about the PD-S pattern.
a. Based on how PD is calculated, why should this metric be related to taxonomic richness? b. Describe the relationship between taxonomic richness and phylodiversity. c. When would you expect these two estimates of diversity to deviate from one another? d. Interpret the significance of the scaling PD-S scaling exponent.

***Answer 1a***: PD is calculated by summing the branch lengths for each species in the phylogenetic tree, so it will increase with increasing richness (+1 richness = +1 tip to the tree). ***Answer 1b***: The graph shows that natural log of PD scales to the natural log of richness with an exponent of 0.75, meaning that for an increase in ln(s) of 1, ln(PD) will increase 0.75. ***Answer 1c***: One situation in which they would differ would be if a sample had very high richness, but all species are extremely recently diverged (high S, low PD). The reverse would also cause them to differ, if a sample had low richness (only a few species), but all species are extremely distantly related. ***Answer 1d***: The scaling exponent means that all else equal, measures of PD will increase at a slower rate than the corresponding measures of richness.

### i. Randomizations and Null Models

In the R code chunk below, do the following:
1. estimate the standardized effect size of PD using the `richness` randomization method.

```
ses.pd <- ses.pd(comm[1:2,], phy, null.model="richness",runs=25,
                 include.root=FALSE)
ses.pd
```

```
##       ntaxa   pd.obs pd.rand.mean pd.rand.sd pd.obs.rank   pd.obs.z  pd.obs.p
## BC001   668 43.71912     43.79703  0.6910385          14 -0.1127371 0.5384615
## BC002   587 40.94334     40.09115  0.7849885          22  1.0856043 0.8461538
##       runs
## BC001   25
## BC002   25
```

```
ses.pd.freq <- ses.pd(comm[1:2,], phy, null.model="frequency",runs=25,
                 include.root=FALSE)
ses.pd.freq
```

```
##       ntaxa   pd.obs pd.rand.mean pd.rand.sd pd.obs.rank  pd.obs.z   pd.obs.p
## BC001   668 43.71912     42.43159  0.5259435          26  2.448040 1.00000000
## BC002   587 40.94334     42.22456  0.5191435           1 -2.467951 0.03846154
##       runs
## BC001   25
## BC002   25
```

```
ses.pd.pool <- ses.pd(comm[1:2,], phy, null.model="sample.pool",runs=25,
                 include.root=FALSE)
ses.pd.pool
```

```
##       ntaxa   pd.obs pd.rand.mean pd.rand.sd pd.obs.rank   pd.obs.z  pd.obs.p
## BC001   668 43.71912     44.13720  0.9218173           7 -0.4535353 0.2692308
## BC002   587 40.94334     39.76267  0.7758621          25  1.5217497 0.9615385
##       runs
## BC001   25
## BC002   25
```

***Question 2***: Using `help()` and the table above, run the `ses.pd()` function using two other null models and answer the following questions:

a. What are the null and alternative hypotheses you are testing via randomization when calculating `ses.pd`?
b. How did your choice of null model influence your observed ses.pd values? Explain why this choice affected or did not affect the output.

***Answer 2a***: The null hypothesis is that the phylogenetic diversity of species in the sample matches an expectation based on random brownian motion. The alternative hypotheses are that the phylogenetic diversity in our sample is higher or lower than expected by the null model.
***Answer 2b***: Honestly I couldn't figure out how to interpret the output columns of ses.pd, sorry!

### B. Phylogenetic Dispersion Within a Sample

Another way to assess phylogenetic $\alpha$-diversity is to look at dispersion within a sample.

#### i. Phylogenetic Resemblance Matrix

In the R code chunk below, do the following:
1. calculate the phylogenetic resemblance matrix for taxa in the Indiana ponds data set.

```
# phylogenetic distance matrix
phydist <- cophenetic.phylo(phy)
```

#### ii. Net Relatedness Index (NRI)

In the R code chunk below, do the following:
1. Calculate the NRI for each site in the Indiana ponds data set.

```
# estimate standardized effect size for NRI, mean phylo distance
ses.mpd <- ses.mpd(comm, phydist, null.model="taxa.labels",
                   abundance.weighted=FALSE, runs=25)
# calculate NRI
NRI <- as.matrix(-1 * ((ses.mpd[,2] - ses.mpd[,3]) / ses.mpd[,4]))
rownames(NRI) <- row.names(ses.mpd)
colnames(NRI) <- "NRI"

## abundance weighted
# estimate standardized effect size for NRI, mean phylo distance
ses.mpd.ab <- ses.mpd(comm, phydist, null.model="taxa.labels",
                   abundance.weighted=TRUE, runs=25)
# calculate NRI
NRI.ab <- as.matrix(-1 * ((ses.mpd.ab[,2] - ses.mpd.ab[,3]) / ses.mpd.ab[,4]))
rownames(NRI.ab) <- row.names(ses.mpd.ab)
colnames(NRI.ab) <- "NRI"
```

#### iii. Nearest Taxon Index (NTI)

In the R code chunk below, do the following: 1. Calculate the NTI for each site in the Indiana ponds data set.

```
# estimate standardized effect size for NTI, mean nearest neighbor distance
ses.mntd <- ses.mntd(comm, phydist, null.model = "taxa.labels",
                     abundance.weighted = FALSE, runs=25)

# calculate NTI
NTI <- as.matrix(-1 * ((ses.mntd[,2] - ses.mntd[,3]) / ses.mntd[,4]))
rownames(NTI) <- row.names(ses.mntd)
colnames(NTI) <- "NTI"

## abundance weighted
# estimate standardized effect size for NRI, mean phylo distance
ses.mntd.ab <- ses.mpd(comm, phydist, null.model="taxa.labels",
                   abundance.weighted=TRUE, runs=25)
# calculate NRI
NTI.ab <- as.matrix(-1 * ((ses.mntd.ab[,2] - ses.mntd.ab[,3]) / ses.mntd.ab[,4]))
rownames(NTI.ab) <- row.names(ses.mntd.ab)
colnames(NTI.ab) <- "NTI"
```

***Question 3***:

a. In your own words describe what you are doing when you calculate the NRI.
b. In your own words describe what you are doing when you calculate the NTI.
c. Interpret the NRI and NTI values you observed for this dataset.
d. In the NRI and NTI examples above, the arguments "abundance.weighted = FALSE" means that the indices were calculated using presence-absence data. Modify and rerun the code so that NRI and NTI are calculated using abundance data. How does this affect the interpretation of NRI and NTI?

***Answer 3a***: The NRI calculation is a way to test for phylogenetic clustering or overdispersion, by calculating the average pairwise branch length separating the tips of the tree (MPD), and then by calculating the difference between observed MPD and randomized mean MPD divided by the standard deviation of the randomized MPD values. The index value tells you if the pairwise branch lengths in your tree are more dispersed or more clustered than expected by chance under a null model. ***Answer 3b***: The NTI calculation is another way to test for phylogenetic clustering or overdispersion, by calculating the average pairwise branch length separating each tip of the tree from its most closely related neighbor (MNND), and then by calculating the difference between observed MNND and randomized mean MNND divided by the standard deviation of the randomized MNND values. The index value tells you if the branch lengths between each tip of your tree and its closest neighbor are more dispersed or more clustered than expected by chance under a null model. ***Answer 3c***:
***Answer 3d***:

## 5) PHYLOGENETIC BETA DIVERSITY

### A. Phylogenetically Based Community Resemblance Matrix

In the R code chunk below, do the following:
1. calculate the phylogenetically based community resemblance matrix using Mean Pair Distance, and
2. calculate the phylogenetically based community resemblance matrix using UniFrac distance.

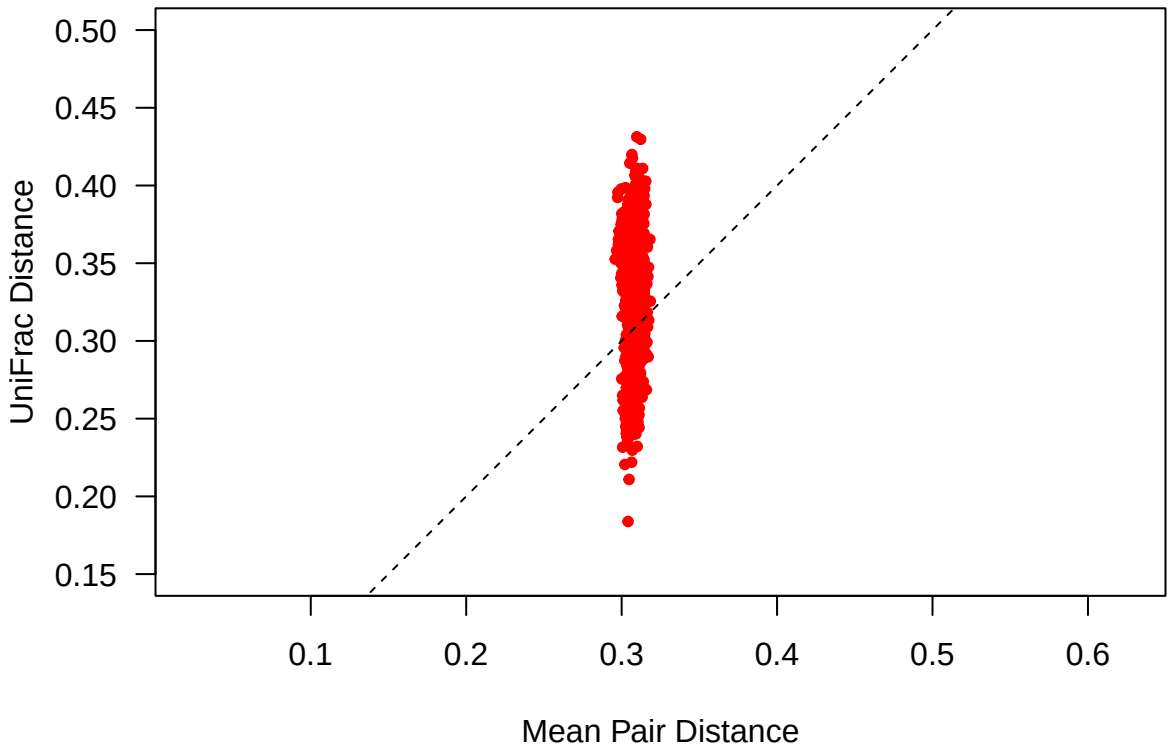
```
# mean pairwise distance
dist.mp <- comdist(comm, phydist)
```

```
## [1] "Dropping taxa from the distance matrix because they are not present in the community data:"
## [1] "Methanosarcina"
```

```
# unifrac distance
dist.uf <- unifrac(comm, phy)
```

In the R code chunk below, do the following:
1. plot Mean Pair Distance versus UniFrac distance and compare.

```
par(mar=c(5,5,2,1)+0.1)
plot(dist.mp,dist.uf,
     pch=20, col="red", las=1, asp=1, xlim=c(0.15,0.5), ylim=c(0.15,0.5), xlab="Mean Pair Distance", yl
abline(b = 1, a = 0, lty=2);
text(o=0.5,0.47, "1:1")
```




***Question 4***:

a. In your own words describe Mean Pair Distance, UniFrac distance, and the difference between them.
b. Using the plot above, describe the relationship between Mean Pair Distance and UniFrac distance. Note: we are calculating unweighted phylogenetic distances (similar to incidence based measures). That means that we are not taking into account the abundance of each taxon in each site.
c. Why might MPD show less variation than UniFrac?

***Answer 4a***: Both methods are used to calculate community resemblance matrices that take phylogenetic information into account. MP distance is based on the mean phylogenetic distance between pairwise taxa. UF distance is calculated for a pair of taxa by dividing the sum of their UNshared branch lengths by the total sum of branch lengths in the tree. ***Answer 4b***: UF distance varies in a range from 0.2 to 0.45, while MP hardly varies from a mean value of 0.3 ***Answer 4c***: MP distance only accounts for a direct path through the tree between two tips, while UF distance accounts for the two tips by taking the proportion of the total branches in the tree that do not form a path between the two tips, so it incorporates more variation relative to the total tree size.

### B. Visualizing Phylogenetic Beta-Diversity

Now that we have our phylogenetically based community resemblance matrix, we can visualize phylogenetic diversity among samples using the same techniques that we used in the $\beta$-diversity module from earlier in the course.

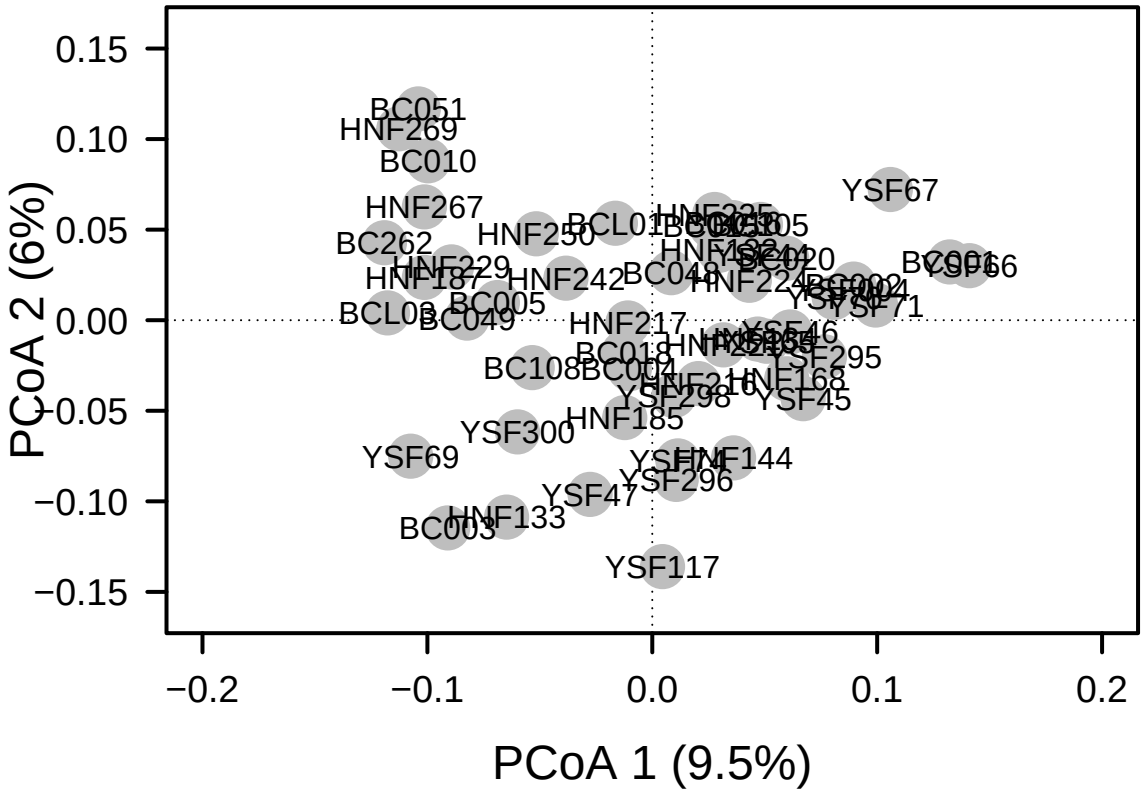
In the R code chunk below, do the following:
1. perform a PCoA based on the UniFrac distances, and
2. calculate the explained variation for the first three PCoA axes.

```
pond.pcoa <- cmdscale(dist.uf, eig=T, k=3)

var1 <- round(pond.pcoa$eig[1] / sum(pond.pcoa$eig), 3) * 100
var2 <- round(pond.pcoa$eig[2] / sum(pond.pcoa$eig), 3) * 100
var3 <- round(pond.pcoa$eig[3] / sum(pond.pcoa$eig), 3) * 100
total.sum.eig <- sum(var1, var2, var3)
```

Now that we have calculated our PCoA, we can plot the results.

In the R code chunk below, do the following:
1. plot the PCoA results using either the R base package or the `ggplot` package,
2. include the appropriate axes,
3. add and label the points, and
4. customize the plot.

```
par(mar=c(5,5,1,2)+0.1)
plot(pond.pcoa$points[ ,1], pond.pcoa$points[ ,2],
     xlim= c(-0.2,0.2), ylim = c(-.16,0.16),
     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);
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(pond.pcoa$points[,1], pond.pcoa$points[,2],
       pch = 19, cex = 3, bg = "gray", col = "gray");
text(pond.pcoa$points[,1], pond.pcoa$points[,2], labels=rownames(pond.pcoa$points))
```




In the following R code chunk: 1. perform another PCoA on taxonomic data using an appropriate measure of dissimilarity, and 2. calculate the explained variation on the first three PCoA axes.

```
# Bray Curtis resemblance matrix
comm.dist <- vegdist(comm, method="bray")

#pcoa
tax.pcoa <- cmdscale(comm.dist, eig=T, k=3)

v1 <- round(tax.pcoa$eig[1] / sum(tax.pcoa$eig), 3) * 100
v2 <- round(tax.pcoa$eig[2] / sum(tax.pcoa$eig), 3) * 100
v3 <- round(tax.pcoa$eig[3] / sum(tax.pcoa$eig), 3) * 100
sum.eig <- sum(v1, v2, v3)
```

***Question 5***: Using a combination of visualization tools and percent variation explained, how does the phylogenetically based ordination compare or contrast with the taxonomic ordination? What does this tell you about the importance of phylogenetic information in this system?

> ***Answer 5***: USing only taxonomic info, 49% of the variation in the data is explained by the first three PCs. When phylogenetic information is included, the variation explained by the first three PCs drops to 20.9%, meaning that the phylogenetic relationships are significantly important for understanding the variation in communities among the ponds because that alone explains ~20-30% of community similarity.

**C. Hypothesis Testing**

**i. Categorical Approach**

In the R code chunk below, do the following:
1. test the hypothesis that watershed has an effect on the phylogenetic diversity of bacterial communities.

```
# define environmental category
watershed <- env$Location

# run permanova with adonis
phylo.adonis<- adonis2(dist.uf ~ watershed, permutations = 999)

# compare to taxonomy permanova
tax.adonis <- adonis2(vegdist(decostand(comm, method = "log"), method="bray") ~ watershed, permutations
```

### ii. Continuous Approach

In the R code chunk below, do the following: 1. from the environmental data matrix, subset the variables related to physical and chemical properties of the ponds, and
2. calculate environmental distance between ponds based on the Euclidean distance between sites in the environmental data matrix (after transforming and centering using `scale()`).

```
# define environmental variables
envs <- env[,5:19]

# remove redundant variables
envs <- envs[, -which(names(envs) %in% c("TDS", "Salinity", "Cal_Volume"))]

# create distance matrix for env. variables
env.dist <- vegdist(scale(envs), method="euclid")
```

In the R code chunk below, do the following:
1. conduct a Mantel test to evaluate whether or not UniFrac distance is correlated with environmental variation.

```
mantel(dist.uf, env.dist)
```

```
##
## Mantel statistic based on Pearson's product-moment correlation
##
## Call:
## mantel(xdis = dist.uf, ydis = env.dist)
##
## Mantel statistic r: 0.1604
##       Significance: 0.063
##
## Upper quantiles of permutations (null model):
##   90%   95% 97.5%   99%
## 0.134 0.173 0.209 0.236
## Permutation: free
## Number of permutations: 999
```

Last, conduct a distance-based Redundancy Analysis (dbRDA).

In the R code chunk below, do the following:
1. conduct a dbRDA to test the hypothesis that environmental variation effects the phylogenetic diversity of bacterial communities,
2. use a permutation test to determine significance, and 3. plot the dbRDA results

```r
# Constrained Ordination
ponds.dbrda <- dbrda(dist.uf ~ ., data = as.data.frame(scale(envs)))

# permutation tests to evaluate significance
anova(ponds.dbrda, by="axis")
```

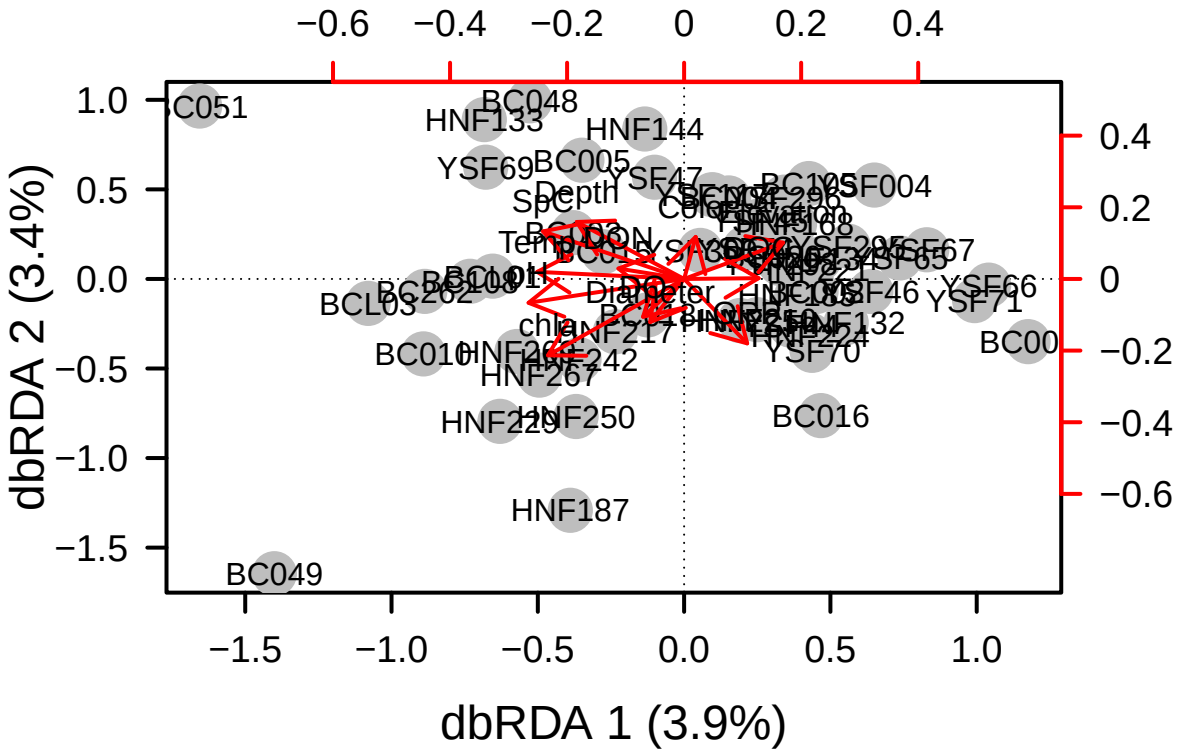
```
## Permutation test for dbrda under reduced model
## Forward tests for axes
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = dist.uf ~ Elevation + Diameter + Depth + ORP + Temp + SpC + DO + pH + Color +
##          Df SumOfSqs      F Pr(>F)
## dbRDA1    1  0.10566 2.0152  0.442
## dbRDA2    1  0.09258 1.7658  0.602
## dbRDA3    1  0.07555 1.4409  0.972
## dbRDA4    1  0.06677 1.2735  0.997
## dbRDA5    1  0.05666 1.0807  1.000
## dbRDA6    1  0.05293 1.0095  1.000
## dbRDA7    1  0.04750 0.9059  1.000
## dbRDA8    1  0.03941 0.7517  1.000
## dbRDA9    1  0.03775 0.7201  1.000
## dbRDA10   1  0.03280 0.6256  1.000
## dbRDA11   1  0.02876 0.5485  1.000
## dbRDA12   1  0.02501 0.4770  0.999
## Residual 39  2.04482
```

```r
envfit(ponds.dbrda, envs, perm=999)
```

```
##
## ***VECTORS
##
##             dbRDA1   dbRDA2     r2 Pr(>r)
## Elevation  0.77670  0.62986 0.0959  0.100 .
## Diameter  -0.27972 -0.96008 0.0541  0.236
## Depth     -0.63137  0.77548 0.1756  0.010 **
## ORP        0.41879 -0.90808 0.1437  0.030 *
## Temp      -0.98250  0.18628 0.1523  0.019 *
## SpC       -0.77101  0.63682 0.2087  0.006 **
## DO        -0.39318 -0.91946 0.0464  0.314
## pH        -0.96210 -0.27270 0.1756  0.010 **
## Color      0.06353  0.99798 0.0464  0.326
## chla      -0.60392 -0.79704 0.2626  0.007 **
## DOC        0.99847 -0.05526 0.0382  0.366
## DON       -0.91633  0.40042 0.0339  0.435
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Permutation: free
## Number of permutations: 999
```

```r
#calculate explained variation
exvar1 <- round(ponds.dbrda$CCA$eig[1] / sum(c(ponds.dbrda$CCA$eig, ponds.dbrda$CA$eig)), 3) *100
```

```
exvar2 <- round(ponds.dbrda$CCA$eig[2] / sum(c(ponds.dbrda$CCA$eig, ponds.dbrda$CA$eig)), 3) *100

# create vectors to plot
vectors <- scores(ponds.dbrda,display="bp")

# now plot selected model, define plot parameters
par(mar=c(5,5,4,4)+0.1)
plot(scores(ponds.dbrda, display="wa"),
     xlab=paste("dbRDA 1 (", exvar1,"%)", sep=""),
     ylab=paste("dbRDA 2 (", exvar2, "%)", 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(scores(ponds.dbrda, display="wa"), pch=19,cex=3,bg="gray",col="gray");
text(scores(ponds.dbrda, display="wa"), labels=row.names(scores(ponds.dbrda, display="wa")));
arrows(0,0,vectors[,1], vectors[,2], lwd=2, lty=1, length=0.2, col="red");
text(vectors[,1], vectors[,2], pos=3,labels=row.names(vectors));
axis(side=3, lwd.ticks=2, cex.axis=1.2,las=1, col="red", lwd=2.2,
     at=pretty(range(vectors[,1]))*2, labels=pretty(range(vectors[,1])));
axis(side=4, lwd.ticks=2, cex.axis=1.2,las=1, col="red", lwd=2.2,
     at=pretty(range(vectors[,2]))*2, labels=pretty(range(vectors[,2])))
```




***Question 6***: Based on the multivariate procedures conducted above, describe the phylogenetic patterns of $\beta$-diversity for bacterial communities in the Indiana ponds.

***Answer 6***: It appears that the BC and HNF sites cluster more closely together into one cluster, and the YSF sites cluster more tightly into another cluster. According to the Mantel test,

16% of the variation in pond community similarity varies in correspondence with pond environmental similarity. The dbRDA also shows that a variety of environmental variables covary with community composition along the first two dbRDA axes.

# 6) SPATIAL PHYLOGENETIC COMMUNITY ECOLOGY

### A. Phylogenetic Distance-Decay (PDD)

A distance decay (DD) relationship reflects the spatial autocorrelation of community similarity. That is, communities located near one another should be more similar to one another in taxonomic composition than distant communities. (This is analagous to the isolation by distance (IBD) pattern that is commonly found when examining genetic similarity of a populations as a function of space.) Historically, the two most common explanations for the taxonomic DD are that it reflects spatially autocorrelated environmental variables and the influence of dispersal limitation. However, if phylogenetic diversity is also spatially autocorrelated, then evolutionary history may also explain some of the taxonomic DD pattern. Here, we will construct the phylogenetic distance-decay (PDD) relationship

First, calculate distances for geographic data, taxonomic data, and phylogenetic data among all unique pair-wise combinations of ponds.

In the R code chunk below, do the following:
1. calculate the geographic distances among ponds,
2. calculate the taxonomic similarity among ponds,
3. calculate the phylogenetic similarity among ponds, and
4. create a dataframe that includes all of the above information.

```
# geographic distances among ponds
long.lat <- as.matrix(cbind(env$long, env$lat))
coord.dist <- earth.dist(long.lat, dist=TRUE)

# taxonomic similarity among ponds (bray curtis distance)
bray.curtis.dist <- 1 - vegdist(comm)

# phylogenetic similarity among ponds (unifrac distance)
unifrac.dist <- 1 - dist.uf

# transform all distances into pairwise long format with melt function from {reshape}
unifrac.dist.mlt <- melt(as.matrix(unifrac.dist))[melt(upper.tri(as.matrix(unifrac.dist)))$value,]
```

```
## Warning in type.convert.default(X[[i]], ...): 'as.is' should be specified by
## the caller; using TRUE

## Warning in type.convert.default(X[[i]], ...): 'as.is' should be specified by
## the caller; using TRUE
```

```
bray.curtis.dist.mlt <- melt(as.matrix(bray.curtis.dist))[melt(upper.tri(as.matrix(bray.curtis.dist)))$
```

```
## Warning in type.convert.default(X[[i]], ...): 'as.is' should be specified by
## the caller; using TRUE

## Warning in type.convert.default(X[[i]], ...): 'as.is' should be specified by
## the caller; using TRUE
```

```
coord.dist.mlt <- melt(as.matrix(coord.dist))[melt(upper.tri(as.matrix(coord.dist)))$value,]
```

```
## Warning in type.convert.default(X[[i]], ...): 'as.is' should be specified by
## the caller; using TRUE

## Warning in type.convert.default(X[[i]], ...): 'as.is' should be specified by
## the caller; using TRUE
```

```
env.dist.mlt <- melt(as.matrix(env.dist))[melt(upper.tri(as.matrix(env.dist)))$value,]
```

```
## Warning in type.convert.default(X[[i]], ...): 'as.is' should be specified by
## the caller; using TRUE

## Warning in type.convert.default(X[[i]], ...): 'as.is' should be specified by
## the caller; using TRUE
```

```
# create a dataframe from the lists of distances
df <- data.frame(coord.dist.mlt, bray.curtis.dist.mlt[,3], unifrac.dist.mlt[,3],env.dist.mlt[,3])
names(df)[3:6] <- c("geo.dist","bray.curtis","unifrac","env.dist")
```

Now, let's plot the DD relationships:
In the R code chunk below, do the following:
1. plot the taxonomic distance decay relationship,
2. plot the phylogenetic distance decay relationship, and
3. add trend lines to each.

```
# initial plot parameters
par(mfrow=c(2,1), mar = c(1,5,2,1) +0.1, oma=c(2,0,0,0))

# plot for taxonomic DD
plot(df$geo.dist, df$bray.curtis, xlab = "", xaxt="n", las = 1, ylim=c(0.1,0.9),
     ylab = "Bray Curtis Similarity", main="Distance Decay", col="steelblue")

# regression for taxonomic DD
DD.reg.bc <- lm(df$bray.curtis ~ df$geo.dist)
summary(DD.reg.bc)
```

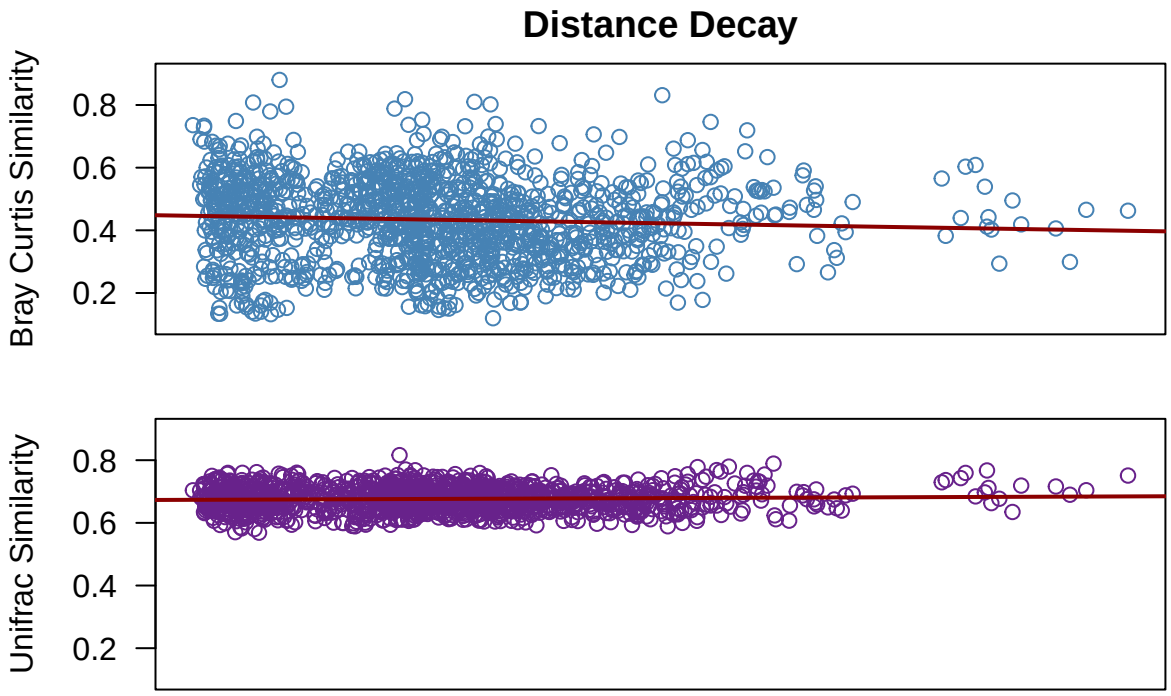
```
##
## Call:
## lm(formula = df$bray.curtis ~ df$geo.dist)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.31151 -0.08843  0.00315  0.09121  0.43817
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.4463453  0.0066883  66.735   <2e-16 ***
## df$geo.dist -0.0013051  0.0005864  -2.226   0.0262 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.1303 on 1324 degrees of freedom
## Multiple R-squared:  0.003728,   Adjusted R-squared:  0.002975
## F-statistic: 4.954 on 1 and 1324 DF,  p-value: 0.0262
```

```
abline(DD.reg.bc, col="red4", lwd=2)

# new plot parameters
par(mar=c(2,5,1,1)+0.1)

# plot phylogenetic DD
plot(df$geo.dist,df$unifrac,xlab = "", xaxt="n", las = 1, ylim=c(0.1,0.9),
     ylab = "Unifrac Similarity", col="darkorchid4")

# regression phylogenetic DD
DD.reg.uni <- lm(df$unifrac ~ df$geo.dist)
summary(DD.reg.uni)
```

```
##
## Call:
## lm(formula = df$unifrac ~ df$geo.dist)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.105629 -0.027107 -0.000077  0.026761  0.140215
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6735186  0.0019206 350.677   <2e-16 ***
## df$geo.dist 0.0002976  0.0001684   1.767   0.0774 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03741 on 1324 degrees of freedom
## Multiple R-squared:  0.002354,   Adjusted R-squared:  0.0016
## F-statistic: 3.124 on 1 and 1324 DF,  p-value: 0.07738
```

```
abline(DD.reg.uni, col="red4", lwd=2)
mtext("Geographic Distance (km)", side=1, adj=0.55, line=0.5, outer=TRUE)
```




In the R code chunk below, test if the trend lines in the above distance decay relationships are different from one another.

```
source("/Users/ericanadolski/GitHub/QB2023_Nadolski/2.Worksheets/9.PhyloCom/bin/diffslope.R")
diffslope(df$geo.dist,df$unifrac,df$geo.dist,df$bray.curtis)
```

```
## $call
## diffslope(x1 = df$geo.dist, y1 = df$unifrac, x2 = df$geo.dist,
##     y2 = df$bray.curtis)
##
## $slope.diff
## [1] 0.001602746
##
## $signif
## [1] 0.002
##
## $permutations
## [1] 1000
##
## $perms
##    [1] -8.022177e-04 -5.958786e-04  2.613778e-06  9.772173e-05  3.369831e-04
##    [6]  4.960021e-04  8.631268e-04  1.284867e-04 -4.699856e-04 -7.220666e-04
##   [11]  1.622965e-03 -5.613657e-04  1.236276e-04  3.036189e-04 -2.357792e-04
##   [16] -8.548197e-04  4.358538e-04  1.016781e-04  4.682616e-04 -6.720537e-05
##   [21] -7.322312e-05 -8.140755e-04  9.920568e-04 -4.759859e-04  9.559999e-06
##   [26]  8.455279e-05 -1.122040e-03  6.024788e-04 -1.947166e-05 -1.294890e-04
##   [31] -7.140890e-04 -4.057906e-04  7.022482e-04  1.357551e-03 -4.667859e-04
```

```
##   [36]  5.721533e-04  4.842749e-04 -4.006966e-04 -7.182303e-04  2.040047e-04
##   [41]  2.325997e-04  2.592066e-04 -2.120034e-04  4.061734e-04  5.245019e-04
##   [46]  1.305214e-03 -7.055719e-05  2.454276e-04 -1.115938e-03  3.376899e-04
##   [51]  8.194092e-04 -6.096002e-05 -1.746611e-04  1.924624e-05  5.273269e-05
##   [56]  4.293253e-04  4.789584e-04  3.988627e-05  1.971060e-04 -6.551980e-04
##   [61] -3.760242e-04 -2.729293e-04  1.069496e-04  1.942535e-04  2.091886e-04
##   [66]  4.210459e-04 -5.506305e-04 -5.091375e-04 -4.717748e-04  7.483975e-04
##   [71]  3.040455e-05 -2.155078e-04 -3.224023e-04 -7.997917e-04  2.229969e-04
##   [76]  3.453810e-04  8.060214e-04  4.880419e-04  3.034062e-04 -1.194356e-03
##   [81]  7.226064e-04 -2.732365e-04  1.844488e-05  4.116758e-05  1.045247e-03
##   [86]  8.133953e-04 -3.301039e-04 -5.916045e-04 -5.740109e-04 -3.522840e-04
##   [91] -1.023836e-03  5.151478e-04  2.195000e-04  7.823075e-04 -6.663121e-04
##   [96]  8.434653e-04 -1.751879e-03 -2.438249e-04  9.667329e-04  2.123702e-04
##  [101] -2.113093e-04  2.268103e-04  4.760590e-04 -3.380146e-04  2.850065e-04
##  [106]  2.217920e-04  7.623951e-04  6.121287e-04 -6.496366e-04 -4.686661e-04
##  [111] -5.706762e-04  2.678232e-04  6.258839e-04 -2.071352e-06 -2.929560e-04
##  [116] -2.735787e-04  3.229900e-04 -4.159001e-04  4.427870e-04 -4.084079e-05
##  [121] -7.656464e-04  1.767226e-04  1.136363e-03  1.940205e-04  3.301001e-04
##  [126] -2.479866e-05 -1.179209e-03  3.931271e-04 -2.291281e-04  6.337893e-04
##  [131] -6.404727e-04  4.952708e-04  1.466563e-03  9.429465e-04  4.321359e-04
##  [136] -5.500245e-04 -4.407054e-04  7.916784e-04  1.087079e-03  3.569538e-05
##  [141] -2.207991e-04  4.910832e-05 -5.525847e-04 -5.483707e-05  5.225098e-04
##  [146]  3.818567e-04  4.125569e-04 -8.149598e-04  9.033619e-04  2.813455e-04
##  [151] -3.835031e-05  6.229194e-04 -2.064322e-04  9.262075e-05 -5.593228e-04
##  [156]  3.971147e-04  7.305309e-04 -3.058786e-04  2.485806e-04  8.646009e-05
##  [161] -5.993252e-05 -4.571833e-04 -3.656916e-04  2.182298e-04 -3.823631e-04
##  [166]  5.450662e-04 -6.316528e-04  2.164802e-04  1.060223e-03 -1.326982e-05
##  [171] -5.384696e-04 -6.599852e-04  9.228757e-04  1.721244e-04 -6.279103e-04
##  [176]  3.967306e-04 -1.052084e-03  3.195391e-04  9.395303e-05 -1.401529e-03
##  [181] -4.809099e-04  4.507079e-04  5.070896e-04  9.189186e-04 -9.526533e-04
##  [186]  7.692727e-05 -2.007481e-04 -1.182724e-03  5.768596e-04 -1.437364e-03
##  [191] -1.191466e-03  1.843957e-04 -1.263304e-04 -7.450927e-04 -6.869057e-04
##  [196] -2.988185e-04  1.034332e-03  3.329312e-04 -1.083516e-03 -2.504291e-04
##  [201]  8.021723e-04 -1.160550e-04 -2.270713e-04  2.948139e-04 -6.372907e-04
##  [206]  3.658198e-05 -1.793819e-04  9.797786e-04 -1.299319e-04  3.504140e-04
##  [211] -1.007733e-03 -6.810471e-04 -6.080307e-04 -6.490305e-04  2.780645e-04
##  [216]  8.727007e-04 -2.651129e-04  7.466573e-04  9.685848e-04 -1.090674e-04
##  [221] -4.505192e-04  2.157162e-04  3.492724e-04  1.676778e-05 -3.949361e-04
##  [226] -1.687000e-04 -4.595090e-04  7.631806e-04  4.485687e-04 -6.801843e-04
##  [231] -2.799717e-04 -2.723528e-04  6.001728e-04 -2.144126e-04 -5.700076e-04
##  [236] -1.398391e-03 -9.945745e-04 -1.143175e-04 -6.482454e-04 -3.364682e-04
##  [241] -5.388588e-04 -8.596326e-04  7.663426e-04  7.523228e-05 -1.277775e-03
##  [246]  1.823607e-04 -4.124414e-04  7.965795e-04 -5.721467e-04  7.630841e-04
##  [251] -9.410094e-04 -1.056018e-03 -7.513317e-04 -2.082863e-04 -2.408400e-04
##  [256] -1.455647e-03  1.413997e-03 -1.447997e-04  6.214849e-05 -8.086630e-04
##  [261]  9.928886e-05  1.303520e-04 -9.599874e-04 -8.424292e-04 -8.516271e-04
##  [266] -1.490847e-04  2.877655e-04  4.056585e-04  2.427468e-04 -3.871847e-04
##  [271]  2.934520e-04 -7.885242e-05  1.226636e-04  4.795278e-04  5.505781e-04
##  [276] -9.285182e-04 -3.424758e-04 -6.831660e-04 -3.374069e-04  9.169522e-04
##  [281]  6.495430e-04  5.790079e-04 -1.661307e-04  2.916642e-04  8.101448e-04
##  [286]  2.300138e-04  1.741613e-04  5.720525e-04  1.220509e-03  8.221652e-04
##  [291]  2.425429e-04  2.995873e-04 -3.890561e-04 -5.593931e-05 -7.187231e-04
##  [296] -4.835349e-04 -2.452402e-03 -1.769627e-04 -1.225203e-03 -4.696642e-04
##  [301] -4.868794e-04  1.014851e-03  1.407493e-04 -1.251473e-03 -3.968285e-04
```

```
##  [306] -1.222638e-03 -6.890675e-04  3.685369e-04  7.593160e-04 -7.082333e-04
##  [311]  1.098064e-04  1.333165e-04 -8.659763e-04  8.963342e-04 -1.385891e-04
##  [316] -8.168156e-04  1.782801e-04 -3.171298e-04 -4.246148e-04 -2.595974e-04
##  [321]  4.765883e-04 -8.474809e-04  1.568658e-04  3.464815e-04 -7.908777e-04
##  [326] -4.190187e-04  4.600201e-04  6.681208e-04  3.802485e-04  5.865635e-04
##  [331]  5.285831e-04  5.125942e-04 -7.800376e-06  5.819496e-04 -6.710740e-05
##  [336] -4.279010e-04  4.255314e-04  4.830149e-04 -1.215047e-03  3.766749e-04
##  [341]  1.274214e-03  4.738176e-04 -3.027622e-04 -1.483631e-04  3.722112e-05
##  [346] -1.430612e-04  4.824860e-04  4.757585e-04  9.269817e-05  1.196578e-04
##  [351]  4.484969e-04  3.890208e-04 -4.526208e-04  6.900455e-04 -3.931357e-04
##  [356] -2.440654e-04  1.158780e-06 -6.130328e-04 -5.343676e-06  9.739321e-04
##  [361]  1.413673e-04  2.916672e-05 -2.215960e-04 -2.810887e-04  8.428149e-04
##  [366]  1.397156e-04  5.577903e-04 -9.473279e-04  6.969586e-04 -3.413641e-04
##  [371]  1.143546e-03 -3.491996e-04 -3.074742e-04 -8.419141e-04 -1.619133e-04
##  [376]  1.691893e-04  4.631618e-05 -3.625679e-04  3.126685e-04  3.869427e-04
##  [381] -2.892022e-04 -1.025530e-04 -6.114736e-04 -1.397344e-04  7.790931e-04
##  [386] -6.873188e-05 -5.350366e-05 -1.269784e-05 -9.177109e-04  1.396151e-04
##  [391] -6.683552e-04  3.436331e-04 -1.277260e-04 -3.393237e-04  1.036229e-03
##  [396] -8.355128e-04  4.348325e-04  1.812326e-04  1.868956e-04 -1.123954e-04
##  [401] -1.004084e-03  5.591112e-04 -1.082150e-04  3.005174e-04  1.030505e-03
##  [406] -1.430009e-04  9.546198e-04 -2.506510e-04  6.570310e-04  1.148877e-03
##  [411]  9.525352e-05 -3.051803e-04 -6.229115e-05 -1.255141e-04 -3.296540e-04
##  [416]  3.068859e-05  4.818578e-04 -6.568785e-04  1.036155e-03 -4.805207e-04
##  [421] -5.324151e-04  3.653222e-04 -2.083297e-04  1.977623e-04 -8.147090e-04
##  [426]  6.302131e-04  5.905607e-04  4.276554e-04 -5.863364e-04 -4.062949e-04
##  [431] -1.843972e-04 -1.762428e-04  7.959982e-04 -2.383437e-04  2.362889e-04
##  [436] -2.684650e-04 -1.156296e-04  6.767942e-04 -1.797689e-04 -2.987480e-04
##  [441] -2.988637e-05 -1.438136e-04  4.374655e-04 -6.764163e-04  1.163234e-04
##  [446]  1.436193e-03 -2.936363e-04 -1.295959e-03  2.814532e-04  1.771168e-04
##  [451]  1.400032e-04 -5.520519e-04  8.795493e-04 -8.418196e-04  9.498275e-05
##  [456]  3.764694e-04  1.221627e-04  1.067528e-03 -7.728058e-04 -6.060715e-04
##  [461]  1.914138e-05 -4.750262e-04 -3.535924e-04 -1.044238e-03 -2.239116e-04
##  [466] -5.161078e-04  2.868097e-04 -1.153761e-03  3.651157e-04 -2.430721e-05
##  [471]  5.418471e-04  3.401271e-04 -8.163063e-04 -8.692082e-04  6.011884e-04
##  [476] -1.006473e-03  1.189875e-03 -2.578514e-04 -6.146276e-04 -4.829504e-04
##  [481]  1.121462e-03  3.764938e-04 -8.178931e-04 -2.705134e-04 -6.730741e-04
##  [486]  2.082929e-04  3.069056e-04 -6.642336e-05  5.094854e-04  2.276096e-04
##  [491] -7.404161e-04  2.604830e-04  7.994025e-05  4.099184e-04  7.419833e-04
##  [496] -2.112622e-04  4.594631e-04  6.805958e-04  3.230712e-04 -6.297009e-05
##  [501] -9.181083e-04  7.249982e-04 -2.240574e-04  2.798265e-05 -9.633530e-04
##  [506]  8.058615e-05  2.489954e-04  3.924750e-04  1.258873e-03 -2.669431e-04
##  [511]  2.703942e-04 -7.509102e-04  7.265837e-04  4.288282e-04  5.383994e-04
##  [516]  1.757386e-04  8.061643e-05 -1.119570e-05  3.173902e-04  3.272810e-05
##  [521] -2.211863e-04  7.182347e-05 -2.237029e-04  1.329221e-03 -6.017837e-04
##  [526] -3.930844e-04  4.553677e-04  2.557191e-04  3.501937e-04 -4.796920e-04
##  [531] -7.773552e-04 -5.168371e-04 -1.652032e-04  1.279298e-03  1.670710e-05
##  [536] -3.229338e-04  2.856582e-05  7.616177e-04 -9.658706e-05  3.125242e-04
##  [541]  8.811681e-05 -9.888400e-05  8.518220e-04 -3.641677e-04  3.220928e-04
##  [546] -1.149665e-03  4.452416e-04 -9.394721e-05 -4.348414e-05 -3.244133e-04
##  [551]  1.960923e-04  4.776483e-04 -1.386863e-04 -5.200333e-04 -1.541710e-04
##  [556] -2.152339e-04  3.763041e-04  6.789522e-04 -1.605697e-03 -3.455072e-04
##  [561] -4.237441e-05  5.072335e-05 -3.949481e-04  8.421106e-04 -3.698190e-05
##  [566]  7.840111e-04 -7.215440e-04 -6.072596e-04  1.016266e-05 -2.997138e-04
##  [571] -4.801678e-04 -1.740237e-03 -7.978633e-04 -3.795893e-04 -3.234139e-04
```

```
##  [576]  3.213892e-04 -1.236694e-04 -8.705239e-04  6.700967e-04 -7.864529e-04
##  [581]  6.904830e-05 -5.451553e-04 -5.384110e-04 -5.434502e-05 -7.256025e-04
##  [586] -7.253930e-04 -1.520079e-04  5.561778e-04 -1.752930e-04 -1.018351e-05
##  [591]  6.130072e-04 -4.773925e-04 -1.397938e-04  7.353284e-04 -3.931579e-05
##  [596]  7.581494e-05 -1.182620e-03 -4.074474e-04  8.970397e-04  4.980860e-04
##  [601] -1.254942e-04 -2.356648e-04  7.998461e-04 -1.128628e-04  1.643791e-04
##  [606]  5.320670e-04  5.200146e-06 -6.807378e-04 -5.910554e-04 -1.932624e-04
##  [611] -7.069284e-04  7.014686e-04  1.242125e-04 -3.508406e-04 -6.152955e-04
##  [616] -6.529522e-05  1.870940e-06 -2.319368e-04 -5.453534e-04 -5.238445e-04
##  [621] -8.539114e-04  1.675396e-04  1.533991e-05  7.969237e-04 -1.625555e-03
##  [626]  3.126261e-04  6.907771e-05  1.905011e-04 -5.419109e-06 -5.685347e-04
##  [631]  3.629671e-04 -4.185058e-04  6.790795e-04  3.536995e-04  8.304201e-04
##  [636] -2.358242e-04 -5.442384e-04  1.111039e-03  1.296487e-03  2.620519e-04
##  [641] -3.316241e-04  9.576454e-04  1.882452e-04  8.521526e-05  4.805700e-05
##  [646]  2.025433e-04  3.052338e-04  2.896131e-04 -1.113881e-03 -1.019713e-03
##  [651]  6.375801e-04 -7.155520e-05 -5.795081e-04 -3.395020e-05  3.904342e-04
##  [656] -7.283221e-04 -4.997012e-04 -7.008604e-04  2.039044e-04  2.057674e-04
##  [661] -3.553330e-04  7.377837e-04 -7.286857e-04  2.434070e-04  4.380210e-04
##  [666]  5.935062e-04  8.921915e-05 -1.116930e-03  2.041615e-04 -1.833089e-04
##  [671]  3.077972e-05  4.397495e-04 -7.642220e-04 -7.618967e-04  1.895027e-04
##  [676]  3.791916e-04 -1.084114e-03 -9.046731e-04  3.546428e-04  8.176519e-05
##  [681] -1.983012e-05  4.721959e-04  4.222762e-04 -1.453227e-04  3.510101e-04
##  [686] -9.920831e-04  4.643887e-04 -1.212417e-03 -8.273791e-04 -7.353580e-04
##  [691]  9.529234e-05  6.464276e-04 -9.390583e-04 -1.179615e-03 -6.019479e-04
##  [696] -2.954365e-04  5.300970e-04 -9.063519e-05  4.516664e-04  4.740736e-05
##  [701]  3.610590e-04 -7.131905e-04  1.121802e-03  4.957737e-04 -6.209490e-04
##  [706]  4.096052e-04  3.246294e-04 -8.716483e-04  2.035138e-04 -4.285232e-04
##  [711]  1.402623e-03 -4.047613e-04 -8.265183e-04 -4.466438e-04 -1.776562e-04
##  [716]  2.810161e-04  1.137935e-05  2.398994e-04 -9.124213e-04 -5.312109e-04
##  [721] -3.261380e-04 -1.170863e-03 -1.221938e-04 -1.154829e-04  6.538197e-04
##  [726] -7.600606e-04 -1.085695e-04  9.783744e-05  1.105382e-04  1.722749e-04
##  [731] -6.134267e-04  1.183110e-04 -6.340483e-04  4.961143e-04 -2.212375e-05
##  [736]  5.084457e-05  5.800314e-04 -4.649932e-04 -1.198826e-03  6.230271e-04
##  [741]  6.764190e-05 -2.265105e-04  7.720214e-04 -4.253936e-04 -1.528719e-04
##  [746] -7.440000e-04 -4.315408e-04 -4.501165e-04  1.943221e-04 -2.413057e-04
##  [751]  1.095558e-03  1.058960e-04  4.519950e-04  4.869791e-04  1.812338e-04
##  [756]  6.645235e-04 -1.821065e-04  2.430877e-04 -6.213226e-04 -1.643476e-04
##  [761]  8.479536e-04  8.649014e-04  4.054296e-05 -3.464583e-04  4.544931e-04
##  [766]  7.856443e-04  8.429579e-04 -2.563487e-04 -8.011270e-04 -5.473560e-04
##  [771] -9.904715e-04  1.206862e-03  5.766807e-04  1.021200e-03  3.943975e-05
##  [776]  2.757286e-04  6.273161e-04 -2.644439e-04  3.873472e-04  9.566022e-04
##  [781] -4.087448e-04 -8.978131e-04 -4.578817e-04  6.085980e-04 -2.918555e-04
##  [786]  2.770873e-05 -1.880870e-04  1.288535e-04 -3.047479e-04 -1.080827e-04
##  [791] -1.679970e-03  3.779494e-04 -4.492346e-04  1.310247e-04  4.870103e-04
##  [796]  9.621843e-04  1.046764e-03  1.874977e-04 -6.591202e-04 -4.895257e-04
##  [801]  1.562749e-04  5.898864e-05 -3.859355e-04  6.245572e-04 -4.897834e-05
##  [806]  8.988108e-04  2.707963e-04 -1.141722e-03  6.099588e-04  4.818479e-04
##  [811] -9.637004e-04  3.989902e-04  1.846944e-04 -1.583538e-03  6.453283e-04
##  [816] -1.568265e-04  7.922574e-04  1.160793e-04 -6.470636e-04  1.947233e-04
##  [821]  3.702927e-04 -5.258937e-04 -1.616230e-04  3.048046e-04  1.298509e-03
##  [826] -4.050255e-07  9.334059e-04 -3.722434e-04 -8.569209e-05 -1.197230e-04
##  [831]  6.219926e-04  5.480964e-04  1.357885e-04  2.639901e-05 -2.125321e-04
##  [836] -1.724071e-04  9.356546e-04 -3.904467e-04  4.054402e-04 -3.920198e-05
##  [841] -8.177270e-04 -4.018310e-04  4.457252e-04 -7.026363e-05 -1.364653e-04
```

```
##  [846]  5.566489e-04  1.088445e-04  2.283943e-04 -8.155838e-04  1.800653e-04
##  [851]  5.834780e-04 -5.837912e-04  1.143638e-03 -2.330532e-04  2.988735e-04
##  [856]  2.191135e-04  5.722923e-04 -7.877802e-04  5.783003e-04  2.181883e-04
##  [861]  1.297833e-03  2.632800e-04 -4.985147e-05 -6.597573e-04 -1.114585e-03
##  [866] -1.636453e-04 -7.452212e-04  4.056322e-04 -6.075852e-04 -5.844164e-04
##  [871]  2.896236e-04 -3.110815e-04  1.454925e-04  8.148482e-05 -8.125319e-04
##  [876] -4.106531e-04 -4.739663e-04  6.917968e-04 -1.182534e-03 -1.173791e-03
##  [881]  3.678285e-04 -8.774525e-05 -9.578932e-04  9.212382e-04  4.478526e-04
##  [886]  5.455630e-04 -3.093397e-04 -9.208157e-06  6.565043e-04 -2.723062e-04
##  [891] -7.685900e-05 -2.393488e-04 -9.665626e-05  3.582906e-04  7.700566e-04
##  [896]  3.403841e-04 -4.652634e-04 -3.059598e-04 -1.018004e-03  9.225648e-04
##  [901]  4.159738e-04  5.139847e-04  1.313264e-04 -1.258016e-04 -3.504158e-04
##  [906]  5.613881e-04 -5.345992e-04 -5.477740e-04  6.049072e-04 -8.067146e-04
##  [911]  8.536684e-04 -2.696529e-04 -1.595072e-04 -3.130322e-04 -6.392321e-04
##  [916]  3.147474e-04  4.989085e-04  3.832273e-04 -1.258594e-05  1.431508e-04
##  [921]  5.568388e-04  1.697895e-03  5.100898e-04  4.983133e-04 -1.331031e-04
##  [926] -1.858448e-04  1.314600e-03  3.576874e-04  4.914619e-04 -6.742978e-04
##  [931]  2.658504e-04 -2.277271e-04 -6.197838e-04  3.445965e-04 -3.644981e-04
##  [936] -2.284042e-05  4.858532e-04 -7.399795e-04 -6.385015e-04  4.186625e-04
##  [941]  7.513191e-04 -2.142236e-04  9.925788e-05 -9.305254e-05  2.640978e-05
##  [946]  5.026179e-04  8.336417e-04  2.284485e-04  1.003920e-03  1.219520e-04
##  [951] -1.060170e-03  3.603330e-04 -7.774080e-04 -6.915881e-05  1.609385e-05
##  [956] -1.458594e-05 -3.479340e-04  5.974003e-04  1.214354e-03 -3.117195e-04
##  [961] -8.611058e-04 -1.565341e-04 -1.548090e-03 -7.920623e-04 -7.035596e-04
##  [966]  6.348721e-04  1.185359e-04 -1.614405e-04 -3.708062e-04 -6.539352e-04
##  [971]  4.738712e-04 -2.961043e-04  5.271393e-04 -7.625975e-04 -1.116437e-03
##  [976]  6.188964e-04  4.886473e-04  5.228078e-04 -5.023593e-04 -6.265282e-04
##  [981]  5.319406e-04 -1.308496e-03  5.107584e-05  6.246126e-05 -1.126323e-04
##  [986]  3.786694e-04 -5.311331e-04  5.745940e-04  1.028966e-03 -6.791092e-04
##  [991] -2.646824e-04  6.099896e-04 -1.119553e-03 -9.381475e-04 -5.364834e-04
##  [996]  1.105820e-03  4.480160e-05 -1.641621e-04  2.812408e-04 -7.395974e-04
##
## attr(,"class")
## [1] "dsl"
```

***Question 7***: Interpret the slopes from the taxonomic and phylogenetic DD relationships. If there are differences, hypothesize why this might be.

> ***Answer 7***: The distance decay slopes indicate how community similarity changes with increasing geographic distance, this relationship reflects "spatial autocorrelation" of community composition due to geographic similarity. The slope for taxonomic DD is slightly negative, indicating that distant communities are less similar than communities close to each other. The phylogenetic DD slope appears to be close to zero, which I believe means that when phylogenetic information is taken into account, this information explains spatial autocorrelation, and then geographic distance doesn't explain more of the community similarity on top of what phylogeny explains.

## SYNTHESIS

Ignoring technical or methodological constraints, discuss how phylogenetic information could be useful in your own research. Specifically, what kinds of phylogenetic data would you need? How could you use it to answer important questions in your field? In your response, feel free to consider not only phylogenetic approaches related to phylogenetic community ecology, but also those we discussed last week in the PhyloTraits module, or any other concepts that we have not covered in this course.

***Synthesis Answer***: Phylogenetic information is hugely important for work in evo devo and genomics of development. In my work, I am interested in assessing how the developmental programs that generate sexual dimorphisms evolve between species, and are employed differentially in different sexually dimrophic body regions. To answer the first type of question, we can look at evolutionary differences between transcriptomes, and to get any meaningful inference about how these transcriptomes have changes due to selection for or against some sexually dimorphic trait, we HAVE to take into account phylogenetic information and in a way, test for how much divergence we would expect between species based only on random changes since their split from their most recent common ancestor.