Diversity Project Abundances

Joshua

4/1/2021

```
rm(list=ls())
package.list <- c('vegan', 'data.table', 'reshape2', 'ggplot2')
for (package in package.list){
  if (!require(package, character.only = TRUE, quietly = TRUE)) {
    install.packages(package)
    library(package,character.only = TRUE)
  }
}
```

```
## Warning: package 'vegan' was built under R version 4.0.4
```

```
## Warning: package 'permute' was built under R version 4.0.4
```

```
## This is vegan 2.5-7
```

```
## Warning: package 'data.table' was built under R version 4.0.5
```

```
## Warning: package 'reshape2' was built under R version 4.0.5
```

```
##
## Attaching package: 'reshape2'
```

```
## The following objects are masked from 'package:data.table':
##
##     dcast, melt
```

Importing site-species data into R

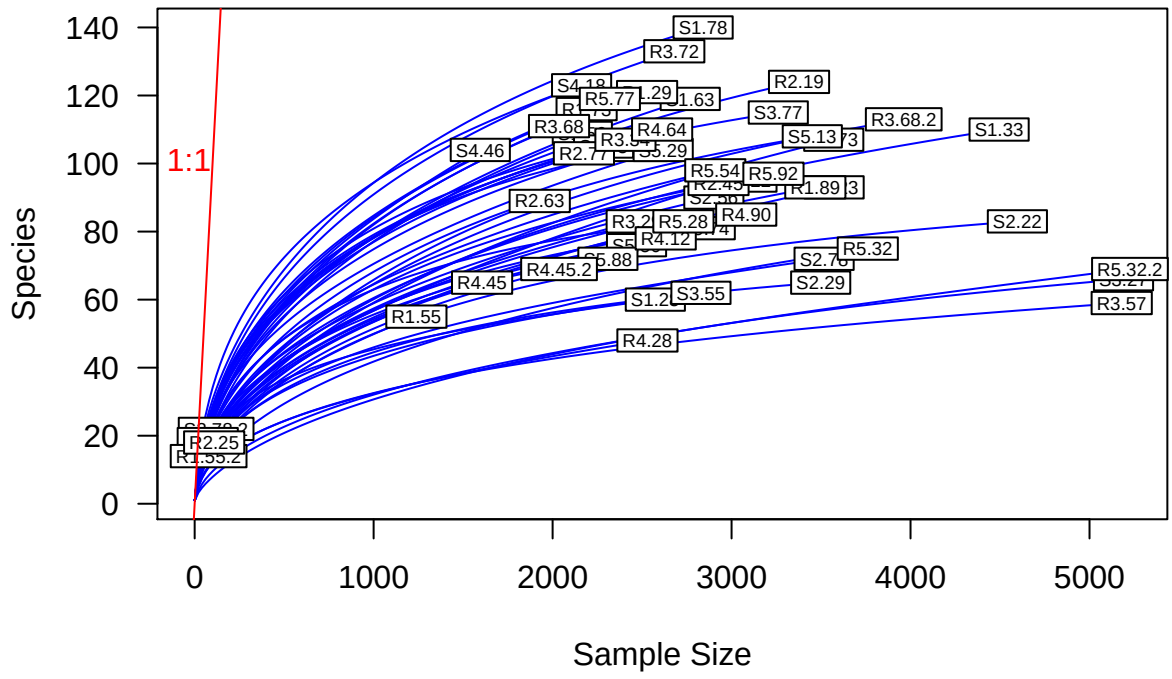
```
site_species <- read.csv("alpine_ridge_data/OTU_table.csv", header = TRUE)
site_species.t <- t(site_species)
```

#Rarefaction

```
#Visualization of rarecurves for all samples
richness <- rowSums((site_species.t > 0) * 1)
print(richness)
```

```
##   S1.28   S1.33   S1.56   S1.63   S1.78   S2.22   S2.29   S2.56   S2.73   S2.78
##      60     110     109     119     140      83      65      90     105      72
## S2.78.2   S3.23   S3.27   S3.55   S3.73   S3.77   S4.18   S4.22   S4.46   S4.68
##      22      93      66      62     107     115     123      95     104     104
##   S4.73   S5.13   S5.29   S5.56   S5.74   S5.88   R1.14   R1.29   R1.55 R1.55.2
##     105     108     104      76      81      72      19     121      55      14
##   R1.73   R1.89   R2.19   R2.25   R2.45   R2.63   R2.77   R3.29   R3.34   R3.57
##     116      93     124      18      94      89     103      83     107      59
##   R3.68 R3.68.2   R3.72   R4.12   R4.28   R4.45 R4.45.2   R4.64   R4.90   R5.28
##     111     113     133      78      48      65      69     110      85      83
##   R5.32 R5.32.2   R5.54   R5.77   R5.92
##      75      69      98     119      97
```

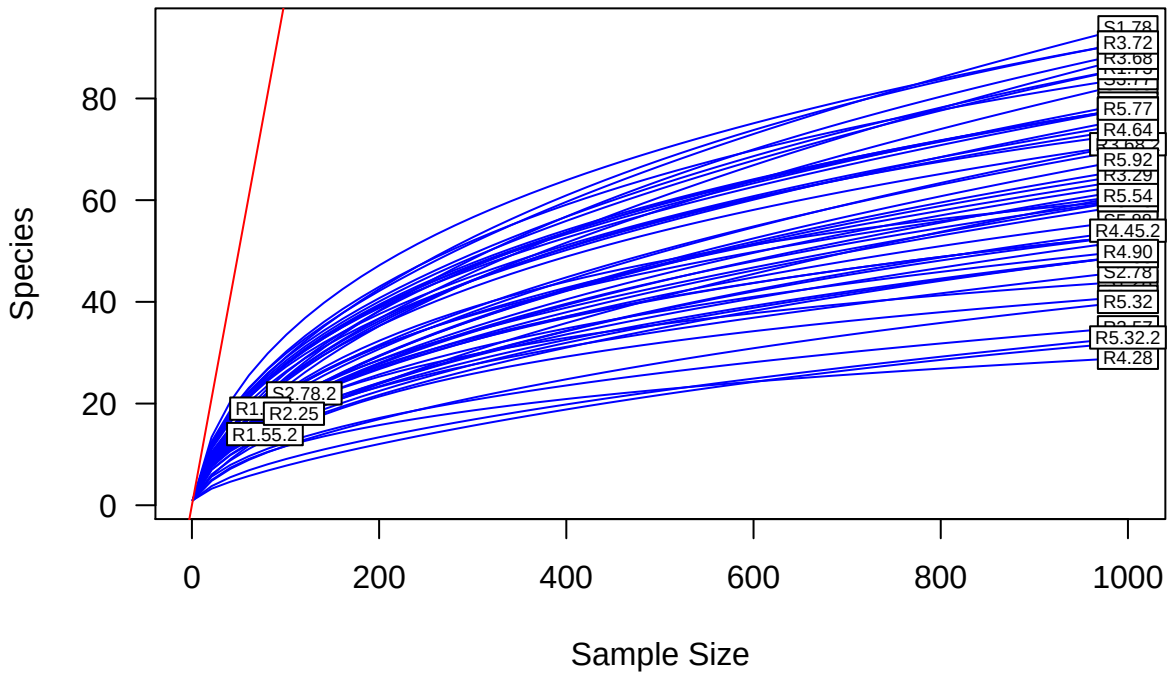
```
minimum.r <- min(rowSums(site_species.t))
rarefy <- rarefy(x = site_species.t, sample = minimum.r, se = TRUE)

rarecurve(x = site_species.t, step = 20, col = "blue", cex = .6, las = 1)
abline(0, 1, col = 'red')
text(200, 100, "1:1", pos = 2, col = 'red')
```




```
#Rarefaction of samples
site_species.r <- rrarefy(site_species.t, 1000)
```

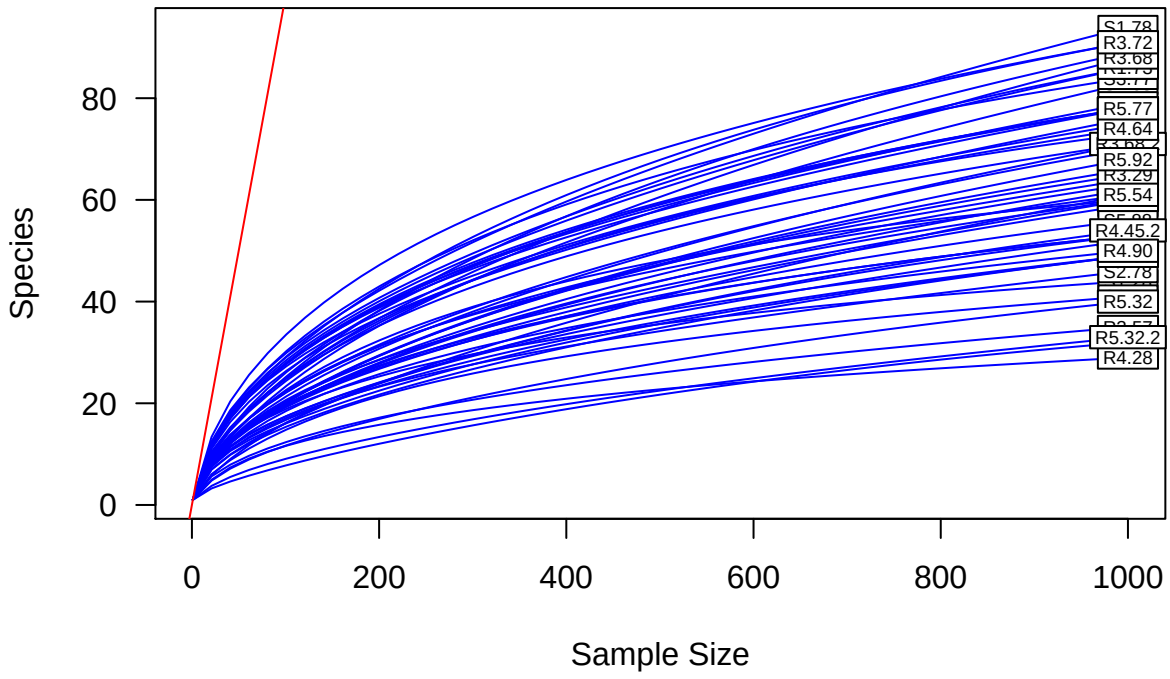
```
## Warning in rrarefy(site_species.t, 1000): some row sums < 'sample' and are not
## rarefied
```

```
richness <- rowSums((site_species.r > 0) * 1)
minimum.r <- min(rowSums(site_species.r))
rarefy <- rarefy(x = site_species.r, sample = minimum.r, se = TRUE)

rarecurve(x = site_species.r, step = 20, col = "blue", cex = .6, las = 1)
abline(0, 1, col = 'red')
text(200, 100, "1:1", pos = 2, col = 'red')
```




```
#Remove samples containing less than 1000 reads (R1.14, R1.55.2, R2.25, S2.78.2)

df.site_species.r <- as.data.frame(site_species.r)
rarefied_site_species <- data.frame()

for (i in 1:nrow(df.site_species.r)){
  if (rowSums(df.site_species.r[i,]) >= 1000){
    rarefied_site_species <- rbind(rarefied_site_species, df.site_species.r[i,])
  }
}

#Visualizing
richness <- rowSums((rarefied_site_species > 0) * 1)
minimum.r <- min(rowSums(rarefied_site_species))
rarefy <- rarefy(x = rarefied_site_species, sample = minimum.r, se = TRUE)

rarecurve(x = rarefied_site_species, step = 20, col = "blue", cex = .6, las = 1)
```

```
abline(0, 1, col = 'red')
text(200, 100, "1:1", pos = 2, col = 'red')
```



#Removing samples to match environmental data downstream

```
rarefied_site_species <- rarefied_site_species[-c(38,43,48),]
```

Importing Environmental Data

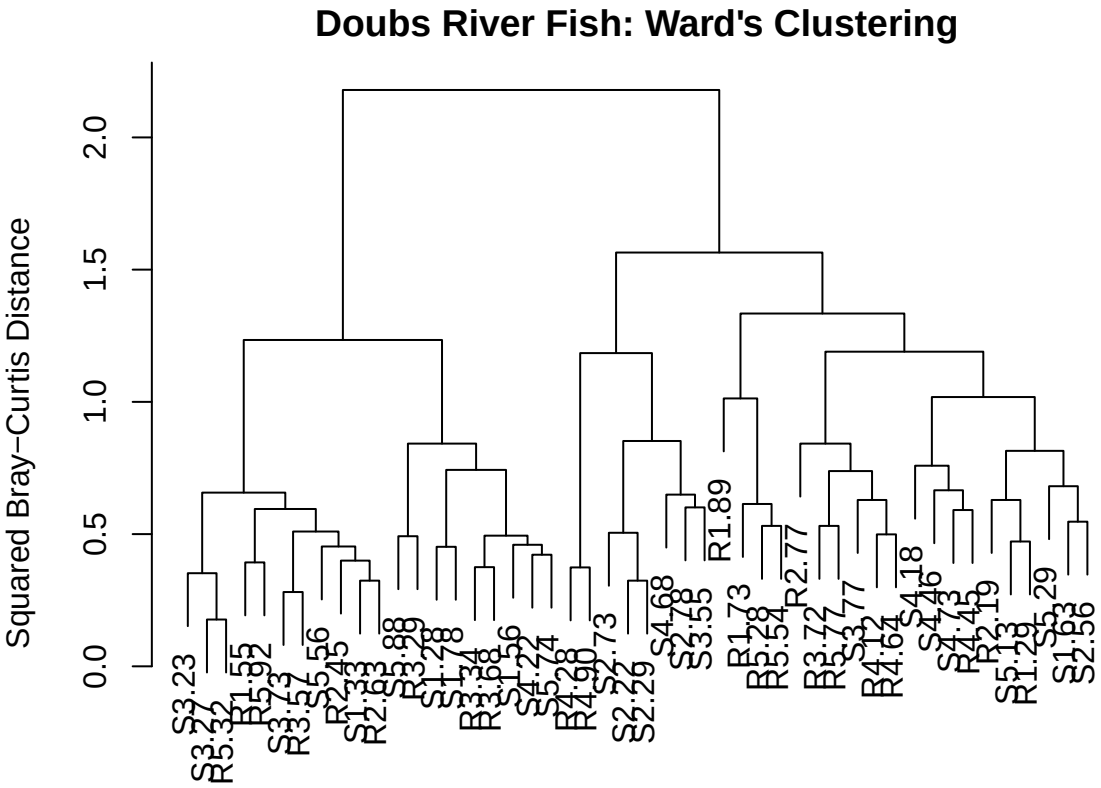
```
env <- read.csv("alpine_ridge_data/variables.txt", header = TRUE, sep = "\t")
env <- env[-c(26,31),]
```

Calculating Bray-Curtis Beta-Diversity

```
fungalBC <- vegdist(rarefied_site_species, method = "bray")
```

Cluster Analysis of Fungal Communities

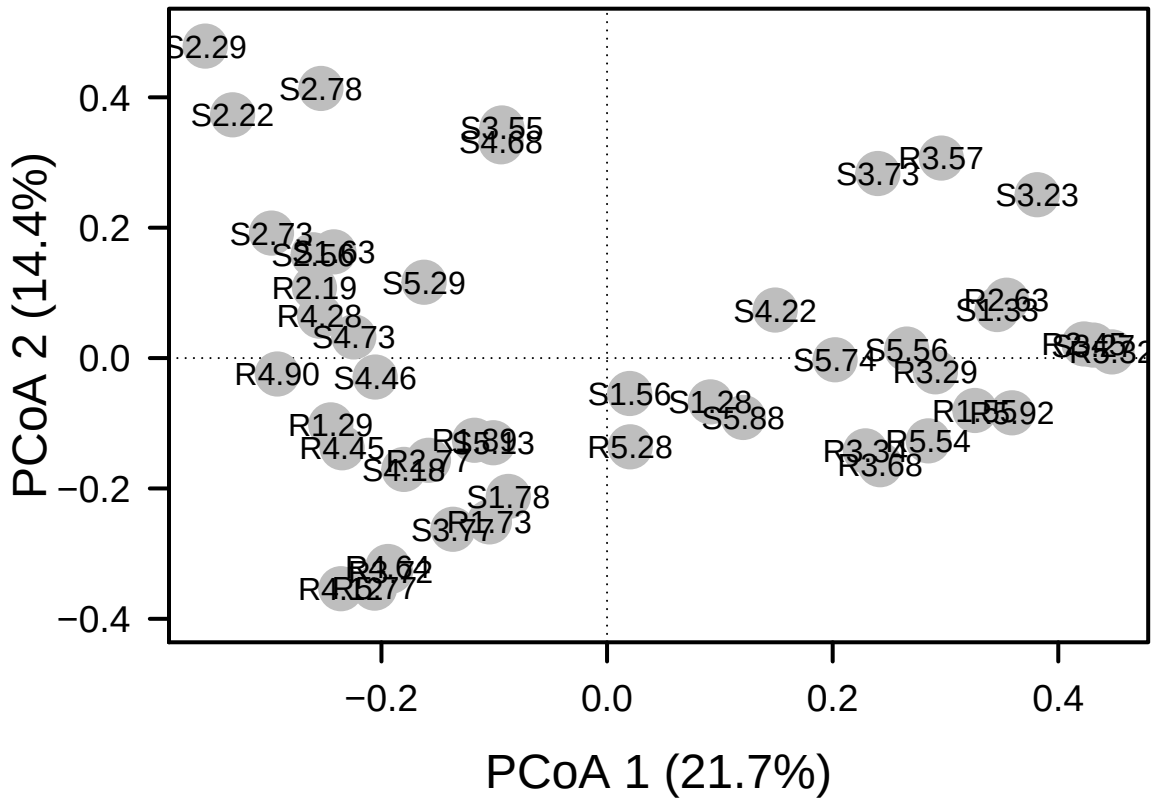
```
#Performing Cluster Analysis
fungal.ward <- hclust(fungalBC, method = "ward.D2")

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



PCoA of Fungal Communities

```
fungal.pcoa <- cmdscale(fungalBC, eig = TRUE, k = 3)

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

#Define Plot Parameters
par(mar = c(5,5,1,2), .1)
```

```
#Initiate Plot
plot(fungal.pcoa$points[,1], fungal.pcoa$points[,2], ylim = c(-.4, .5),
     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 axis
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 & Labels
points(fungal.pcoa$points[,1], fungal.pcoa$points[,2],
       pch = 19, cex = 3, bg = "gray", col = "gray")
text(fungal.pcoa$points[,1], fungal.pcoa$points[,2],
     labels = row.names(fungal.pcoa$points))
```



How much variance is explained by site location (Bray-Curtis)

```
site <- c(rep("S1", 5), rep("S2", 5), rep("S3", 5), rep("S4", 5), rep("S5", 5), rep("R1", 4), rep("R2",
```

```
adonis(rarefied_site_species ~ env$V + site, permutations = 999)
```

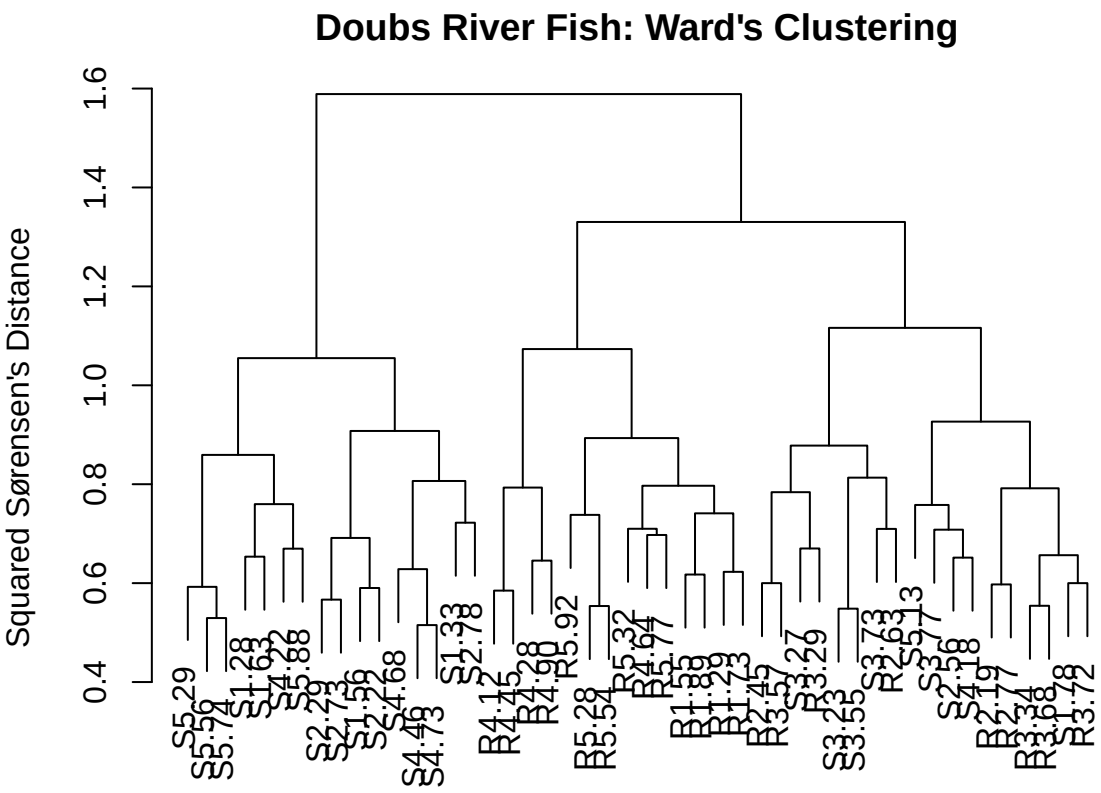
```
##
## Call:
## adonis(formula = rarefied_site_species ~ env$V + site, permutations = 999)
##
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## env$V      1    0.9460 0.94597  3.9732 0.06750  0.001 ***
## site       8    4.0212 0.50265  2.1112 0.28693  0.001 ***
## Residuals 38    9.0474 0.23809         0.64557
## Total     47   14.0146                1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Calculating Sorensen Beta-Diversity

```
fungalS <- vegdist(rarefied_site_species, method = "bray", binary = "TRUE")
```

Cluster Analysis of Fungal Communities

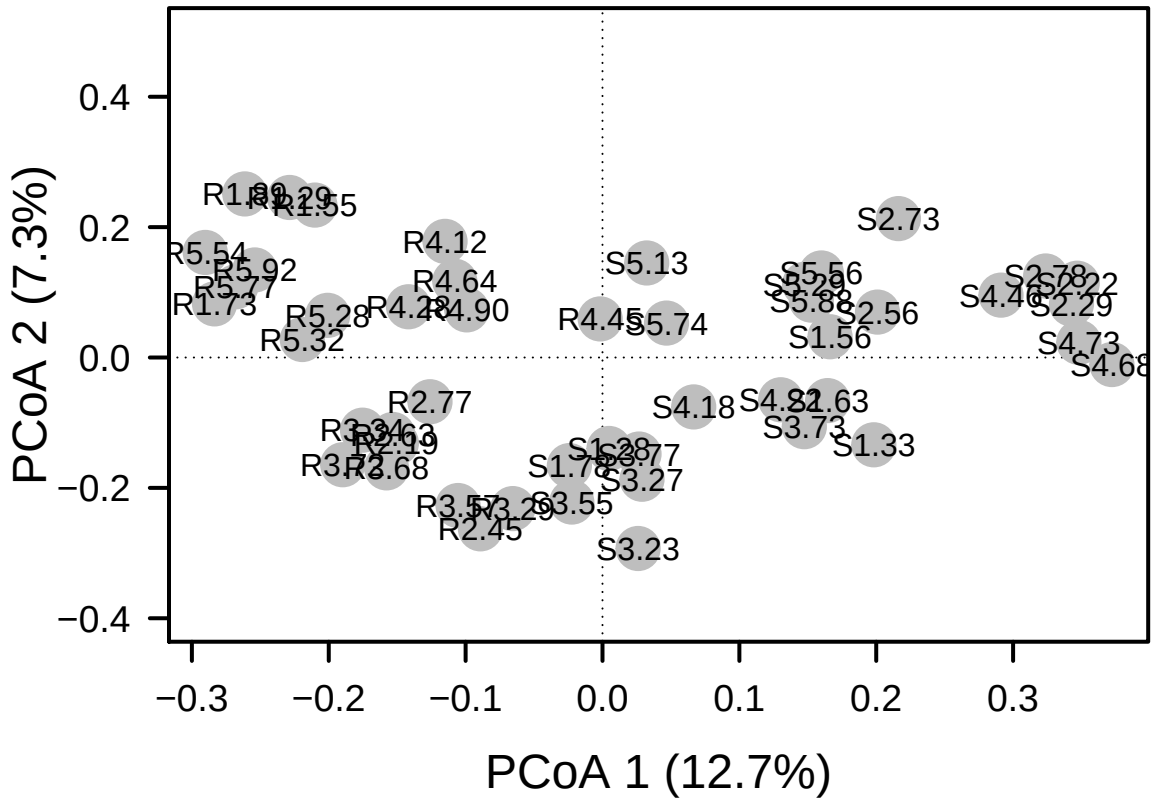
```
#Performing Cluster Analysis
fungal.wardS <- hclust(fungalS, method = "ward.D2")

#Plotting Cluster
par(mar = c(1,5,2,2) + .1)
plot(fungal.wardS, main = "Doubs River Fish: Ward's Clustering",
     ylab = "Squared Sørensen's Distance")
```



PCoA of Fungal Communities

```
fungal.S.pcoa <- cmdscale(fungalS, eig = TRUE, k = 3)

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

#Define Plot Parameters
par(mar = c(5,5,1,2), .1)

#Initiate Plot
plot(fungal.S.pcoa$points[,1], fungal.S.pcoa$points[,2], ylim = c(-.4, .5),
     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 axis
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 & Labels
points(fungal.S.pcoa$points[,1], fungal.S.pcoa$points[,2],
       pch = 19, cex = 3, bg = "gray", col = "gray")
text(fungal.S.pcoa$points[,1], fungal.S.pcoa$points[,2],
     labels = row.names(fungal.S.pcoa$points))
```



How much variance is explained by site location (Sorensen)

```
site <- c(rep("S1", 5), rep("S2", 5), rep("S3", 5), rep("S4", 5), rep("S5", 5), rep("R1", 4), rep("R2",

adonis(rarefied_site_species ~ env$V + site, binary = TRUE, permutations = 999)
```

```
##
## Call:
## adonis(formula = rarefied_site_species ~ env$V + site, permutations = 999,      binary = TRUE)
##
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
```

```
##           Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## env$V      1    1.3835 1.38353  5.8981 0.09704  0.001 ***
## site       8    3.9594 0.49493  2.1099 0.27772  0.001 ***
## Residuals 38    8.9137 0.23457         0.62523
## Total     47   14.2567                1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

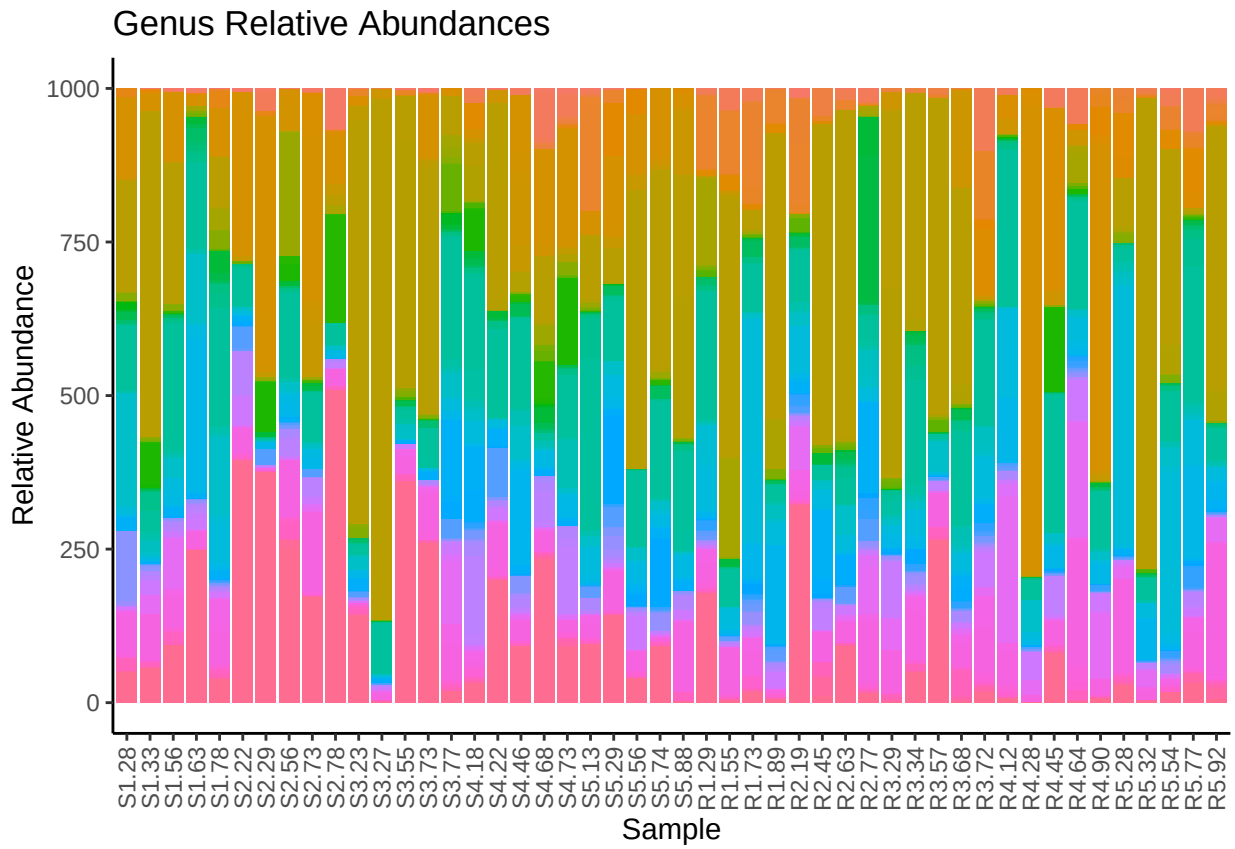
Relative Abundance Visualization

```
genus <- read.csv("alpine_ridge_data/genus_table.csv", header = TRUE)
#convert rownames into column

genus.1 <- as.data.frame(t(rarefied_site_species))
setDT(genus.1, keep.rownames = TRUE)[]
```

```
##        rn S1.28 S1.33 S1.56 S1.63 S1.78 S2.22 S2.29 S2.56 S2.73 S2.78 S3.23
##   1:   V1     0     0     0     1     0     0     0     0     0     0     0
##   2:   V2     0     0     0     0     0     0     0     0     0     0     0
##   3:   V3     0     0     0     0     1     0     0     0     0     0     0
##   4:   V4     0     0     0     1     0     0     0     4     0     0     0
##   5:   V5     0     0     0     0     0     0     0     0     0     0     0
##  ---
## 803: V803     0     1     0     0     1     0     3     1     0     0     0
## 804: V804     0     0     0     0     0     0     0     0     4     0     0
## 805: V805     0     0     0     0     0     0     0     0     0     0     0
## 806: V806     0     0     0     0     0     0     0     0     0     0     0
## 807: V807     0     0     0     0     0     0     0     0     0     0     0
##      S3.27 S3.55 S3.73 S3.77 S4.18 S4.22 S4.46 S4.68 S4.73 S5.13 S5.29 S5.56
##   1:     0     0     0     0     0     0     0     0     0     0     0     0
##   2:     0     0     0     0     6     0    29     0     0     0    14    21
##   3:     0     0     0     0     0     0     0     0     0     0     0     0
##   4:     0     0     0     0     0     0     0     0     0     0     0     0
##   5:     0     0     0     0     0     0     1     0     0     0     0     0
##  ---
## 803:     0     0     0     1     6     2     0     3     0     0     0     0
## 804:     0     1     0     2     0     0     0     0     0     0     0     0
## 805:     0     0     0     0     0     0     0     0     0     0     0     0
## 806:     0     0     0     0     0     0     0     0     0     0     0     0
## 807:     0     0     0     0     0     0     0     0     0     0     0     0
##      S5.74 S5.88 R1.29 R1.55 R1.73 R1.89 R2.19 R2.45 R2.63 R2.77 R3.29 R3.34
##   1:     0     0     0     0     0     0     0     0     0     0     0     0
##   2:     7     1     0     0     0     0     0     0     0     0     0     0
##   3:     0     0     0     0     0     0     0     0     0     0     0     0
##   4:     0     0     0     0     0     0     0     0     0     0     0     0
##   5:     0     0     0     0     0     0     0     0     0     1     0     0
##  ---
## 803:     0     0     0     0     0     0     0     3     0     1     0     1
## 804:     0     0     7     1     6     0     0     0     0     0     0     0
## 805:     0     0     0     0     0     0     0     1     0     0     0     0
## 806:     0     0     0     0     0     0     0     0     0     0     0     0
## 807:     0     0     0     0     0     0     0     0     0     0     0     0
```

```
##      R3.57 R3.68 R3.72 R4.12 R4.28 R4.45 R4.64 R4.90 R5.28 R5.32 R5.54 R5.77
##   1:     0     0     0     0     0     0     0     0     0     0     0     0
##   2:     0     0     0     0     0     0     0     0     0     0     0     0
##   3:     0     0     0     0     0     0     0     0     0     0     0     0
##   4:     0     0     0     0     0     0     0     0     0     0     0     0
##   5:     0     0     0     0     0     0     0     0     0     0     0     0
##  ---
## 803:     0     0     3     0     0     0     1     0     0     2     0     0
## 804:     0     0     0     0     1     0     0     6    11     0    16    30
## 805:     0     0     0     0     0     0     0     0     0     0     0     0
## 806:     0     0     0     0     0     0     0     0     2     0     0     1
## 807:     0     0     0     0     0     0     0     0     0     0     0     0
##      R5.92
##   1:     0
##   2:     0
##   3:     0
##   4:     0
##   5:     0
##  ---
## 803:     0
## 804:     1
## 805:     0
## 806:     0
## 807:     0
```

```
# Converting to Long Format
genus_long <- melt(genus.1, id.vars = "rn", variable.name = "Sample")

# Creating Graph of data
genus_graph <- ggplot(data = genus_long, mapping = aes(x = Sample,  y = value, fill = rn))
genus_graph <- genus_graph + geom_bar(stat="identity")
genus_graph <- genus_graph + labs(y = "Relative Abundance", x = "Sample", title = "Genus Relative Abunda
genus_graph <- genus_graph + theme(legend.position = "None")
genus_graph <- genus_graph + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
genus_graph
```



Importing environmental data and testing for significance

```
env <- read.csv("alpine_ridge_data/variables.txt", header = TRUE, sep = "\t")
env <- env[-c(26,31),]

site <- c(rep("S1", 5), rep("S2", 5), rep("S3", 5), rep("S4", 5), rep("S5", 5), rep("R1", 4), rep("R2",
adonis(rarefied_site_species ~ env$V + env$P + env$N + env$C + site, method = "bray", permutations = 99
```

```
##
## Call:
## adonis(formula = rarefied_site_species ~ env$V + env$P + env$N +      env$C + site, permutations = 9
##
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## env$V      1    0.9460 0.94597  3.9025 0.06750  0.001 ***
## env$P      1    0.2870 0.28705  1.1842 0.02048  0.257
## env$N      1    0.2937 0.29372  1.2117 0.02096  0.250
## env$C      1    0.3065 0.30650  1.2644 0.02187  0.188
## site       8    3.6973 0.46216  1.9066 0.26382  0.001 ***
```

```
## Residuals 35    8.4841 0.24240         0.60538
## Total     47   14.0146                 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Constructing Constrained Ordination

```
env.chem <- as.matrix(env[,c(2:4)])


S.dbrda <- dbrda(fungalS ~ ., as.data.frame(env.chem))
#ordiplot(S.dbrda)

S.dbrda0 <- dbrda(fungalS ~ 1, as.data.frame(env.chem))
S.dbrda1 <- dbrda(fungalS ~ ., as.data.frame(env.chem))

S.dbrda <- ordiR2step(S.dbrda0, S.dbrda1, perm.max = 999)
```

```
## Step: R2.adj= 0
## Call: fungalS ~ 1
##
##                 R2.adjusted
## <All variables>  0.03575746
## + N              0.02805636
## + C              0.02285932
## + P              0.00376483
## <none>           0.00000000
##
##     Df    AIC      F Pr(>F)
## + N  1 128.14 2.3567  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.02805636
## Call: fungalS ~ N
##
##                 R2.adjusted
## <All variables>  0.03575746
## + C              0.03520912
## <none>           0.02805636
## + P              0.02735468
##
##     Df    AIC     F Pr(>F)
## + C  1 128.73 1.341  0.066 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
permutest(S.dbrda, permutations = 999)
```

```
##
```

```
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fungalS ~ N, data = as.data.frame(env.chem))
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     1  0.6948 2.3567  0.001 ***
## Residual 46 13.5619
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
envfit(S.dbrda, env.chem, permutations = 999)
```

```
##
## ***VECTORS
##
##     dbRDA1     MDS1     r2 Pr(>r)
## P  0.87525 -0.48367 0.0913  0.119
## N -0.77808  0.62817 0.8279  0.001 ***
## C -0.75222  0.65892 0.8016  0.001 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Permutation: free
## Number of permutations: 999
```

```
#Calculating explained variation on axes
S.explainvar1 <- round(S.dbrda$CCA$eig[1]/
                           sum(c(S.dbrda$CCA$eig, S.dbrda$CA$eig)),
                         3 ) * 100
S.explainvar2 <- round(S.dbrda$CCA$eig[2]/
                           sum(c(S.dbrda$CCA$eig, S.dbrda$CA$eig)),
                         3 ) * 100

#Plotting constrained ordination results
par(mar = c(5,5,4,4) + .1)

plot(scores(S.dbrda, display = "wa"), xlim = c(-2, 2.1), ylim = c(-2.3, 2.0),
     xlab = paste("dbRDA 1 (", S.explainvar1, "%)", sep = ""),
     ylab = paste("dbRDA 2 (", S.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(scores(S.dbrda, display = "wa"),
       pch = 19, cex = 3, bg = "gray", col = "gray")
text(scores(S.dbrda, display = "wa"),
     labels = row.names(scores(S.dbrda, display = "wa")))
```

```r
#Plotting vectors for enfluence of environmental factors
vectors <- scores(S.dbrda, display = "bp")
arrows(0, 0, vectors[,1], vectors[,2],
       lwd = 2, lty = 1, length = .2, col = "red")
text(vectors[,1], vectors[,2], pos = 3,
     labels = row.names(vectors), col = "red")
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])))
```