Resource Heterogeneity Structures Microbial Communities

Mario E. Muscarella

19 April, 2016

Introduction

Much is already know about how spatial gradients in resource availability contribution to the structure and fucntion of microbial communities. However, we are begining to appreciate the molecular diversity within the resource pool. Resources can be just as diverse as consumers and differ in their quality and availability to consumers. As such, resource diverse represents a mechanisms to understand spatial distirubitons of consumers. In this study, we explore how both the concentration and diversity of resources contribute to the structure and function of aquatic microbial communities.

We explore three hypotheses:

1. Is there a relationship between resource concentration and diversity?

- Dilution Hypothesis: Arrieta et al. 2015 Science

2. Do the concentrations and diversity of resources explain differences between communities?

- Resource Heterogeneity Hypothesis

3. Do concentration and diversity explain different aspects of diversity

- Grinnellian (environmental habitat) vs. Eltonian (food habitat) Niche Hypothesis
- Resource Substitution Hypothesis

Initial Setup

```
rm(list=ls())
getwd()
setwd("~/GitHub/ResourceHeterogeneity/analyses")

# Import Tools and Standard Functions
source("../bin/MothurTools.R")
se <- function(x, ...){sd(x, na.rm = TRUE)/sqrt(length(na.omit(x)))}
CV <- function(x, ...){(sd(x, na.rm = TRUE)/mean(x, na.rm = TRUE))*100}

# Save Standard Plot Settings
opar <- par(no.readonly = TRUE)  # Saves plot defaults

# Load Required Packages
require("xtable");require("png");require("grid");require("vegan")
require("picante");require("phyloseq");require("car")
require("colorspace");require("bioDist");require("gplots")
require("igraph")
```

Study System

We sampled 10 lakes in the Huron Mountains of Michigan. The Huron Mountains are located in the Superior Bedrock Uplands region of the Michigan Upper Peninsula (Schaetzl et al 2013).

The region is classified as Superior Bedrock Uplands The forests around the lakes are old-growth hemlock-nothern hardwoods (Kerry Woods) The lakes are part of the Pine River Watershed which drains into Lake Superior (see: www.co.marquette.mi.us/departments/plannings/docs/watershed.pdf)

Water Body Physical Information

```
lake.data <- read.csv("../data/lake_data2.txt", row.names=1)
colnames(lake.data) <- c("lat", "long", "area", "pH", "DO1", "DO2", "Temp1", "Temp2")
lake.data <- lake.data[sort(row.names(lake.data)), ]
```

Table 1: Lake Physical Properties

```
addtorow <- list()
addtorow$pos <- list(0)
addtorow$command <- c("Lake & Area (Km$^2$)  & pH \\\\\n")
lake.tab <- xtable(lake.data[,c(3,4)], digits = c(0,1,3))
align(lake.tab) <- "ccc"
print(lake.tab, add.to.row = addtorow, include.colnames = FALSE,
      type= "latex", file="../tables/Table1.tex",
      hline.after = c(-1, -1, 0, nrow(lake.tab)))
print(lake.tab, add.to.row = addtorow, include.colnames = FALSE,
      comment = FALSE, hline.after = c(-1, -1, 0, nrow(lake.tab)))
```

Lake	Area (Km^2)	pH
Ann	0.3	7.860
Canyon	0.0	7.020
Howe	0.7	7.780
Ives	1.9	8.100
Lily	0.0	5.510
Mountain	3.4	8.310
Pony	0.0	5.390
Rush	1.3	8.140
SecondPine	0.7	8.090
UpperPine	0.2	7.790

Supplemental Figure 1: System Map

This study system is know to have differences in the concentrations of growth limiting nutrients. Specifically, the concentrations of dissolved organic carbon and phosphorus differ between the lakes. We also have data for total nitrogen, but the values are odd and may not be reasonable for interpretation. The reason the nitrogen values are odd is because many were at the lower detection level of the instrument.

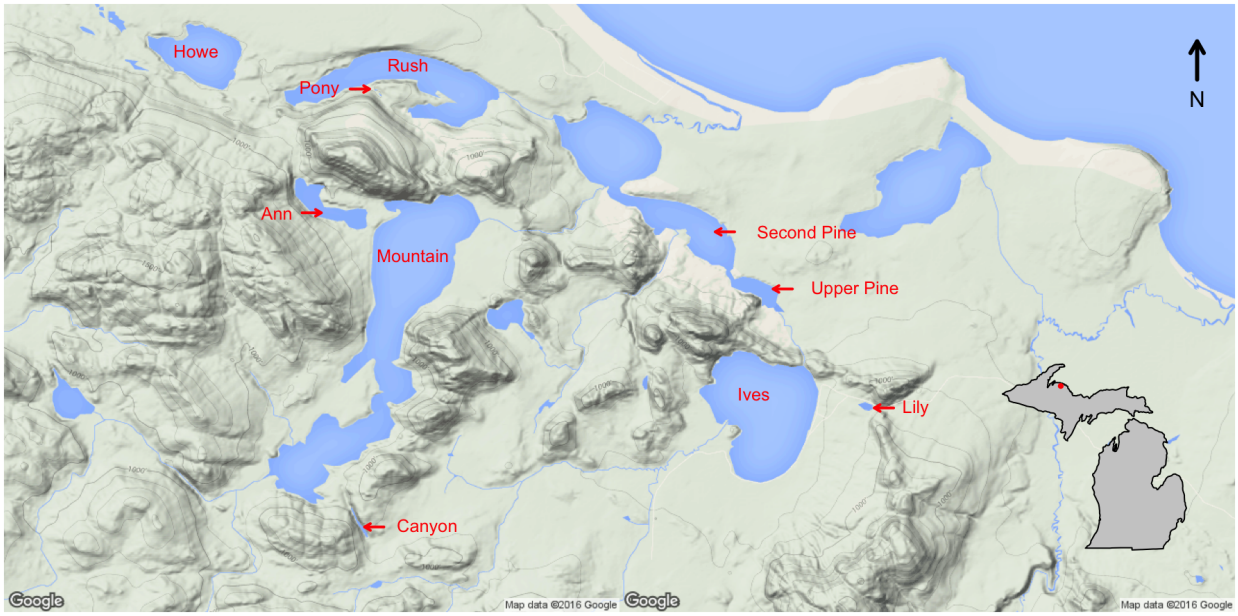


Figure 1: Study System Map

Lake Nutrient Concentrations

```
# DOC
DOC2011 <- read.delim("../data/2011DOC_data.txt", header=T)
DOC2012 <- read.delim("../data/2012DOC_data.txt", header=T)
DOC <- rbind(DOC2011, DOC2012)
DOC <- DOC[grep("MEM*", DOC$Sample), ]
colnames(DOC) <- c("sample", "conc", "LCL", "UCL", "se")
DOCkey <- read.delim("../data/DOC_KEY_epi.txt", header=T)
DOC$code <- DOC$sample
DOC <- DOC[which(DOC$code %in% DOCkey$Sample.Name), ]
DOC$sample <- DOCkey$Site[match(DOCkey$Sample.Name, DOC$code)]
DOC$year <- substr(DOC$code, 4, 7)
DOC$conc <- pmax(DOC$conc, 0)
DOC2 <- data.frame("sample" = DOC$sample, "year" = DOC$year,
                   "conc" =  DOC$conc)[order(DOC$sample, DOC$year), ]
DOC$sample[grep("Pony", DOC$sample)] <- "Pony"
DOC <- droplevels(DOC)

# Total Nitrogen
TN <- read.delim("../data/HMWF_TN.txt")
colnames(TN) <- c("sample", "year", "conc")
TN2 <- data.frame("sample" = TN$sample, "year" = TN$year,
                  "conc" = TN$conc)[order(TN$sample, TN$year), ]
TN <- droplevels(TN)

# Total Phosphorus
TP2011 <- read.delim("../data/2011TP_data.txt")
TP2012 <- read.delim("../data/2012TP_data.txt")
```

```
TP2011$year <- rep("2011", dim(TP2011)[1])
TP2012$year <- rep("2012", dim(TP2012)[1])
TP <- rbind(TP2011, TP2012)
TP <- TP[grep("*iltered", TP$Sample), ]
colnames(TP) <- c("sample", "conc", "LCL", "UCL", "se", "year")
TP$code <- TP$sample
TDP <- TP[grep("*Filtered", TP$sample), ]
TP <- TP[grep("*Unfiltered", TP$sample), ]
TP$sample <- gsub(" Unfiltered", "", TP$sample)
TDP$sample <- gsub(" Filtered", "", TDP$sample)
TP[6, ] <- TDP[6, ]
TP <- TP[-c(which(TP$sample == "CanyonHypo" | TP$sample == "CanyonChemo")), ]
TP$sample <- gsub("CanyonEpi", "Canyon", TP$sample)
TP$sample <- as.factor(TP$sample)
TP$conc <- pmax(TP$conc, 0)
TP2 <- data.frame("sample" = TP$sample, "year" = TP$year,
                  "conc" = TP$conc)[order(TP$sample, TP$year), ]
TP$sample[grep("Pony", TP$sample)] <- "Pony"
TP <- droplevels(TP)
```

Organize Data Table

```
DOC2 <- aggregate(conc ~ sample + year, DOC2, mean)
TN2 <- aggregate(conc ~ sample + year, TN2, mean)
TP2 <- aggregate(conc ~ sample + year, TP2, mean)

nuts <- data.frame("sample" = DOC2$sample, "year" = DOC2$year,
              "DOC" = DOC2$conc, "TP" = TP2$conc)

nuts <- data.frame(nuts[-which(nuts$sample == "Pony.N"), ])
nuts$sample[grep("Pony", nuts$sample)] <- "Pony"
nuts <- droplevels(nuts)
nuts$TN <- TN2$conc
```

Statistical Tests of Nutrients

```
a.tn <- Anova(lm(TN ~ as.factor(sample) + as.factor(year), data = nuts))
a.doc <- Anova(lm(DOC ~ as.factor(sample) + as.factor(year), data = nuts))
a.tp <- Anova(lm(TP ~ as.factor(sample) + as.factor(year), data = nuts))
```

Table 2: Lake Nutrients

```
nuts2 <- data.frame(matrix(NA, 10, 6))
row.names(nuts2) <- levels(nuts$sample)
colnames(nuts2) <- c("DOC11", "DOC12", "TP11", "TP12", "TN11", "TN12")
for (i in row.names(nuts2)){
  nuts2[i, 1] <- round(nuts[nuts$sample == i & nuts$year == "2011", 3], 2)
```

```
  nuts2[i, 2] <- round(nuts[nuts$sample == i & nuts$year == "2012", 3], 2)
  nuts2[i, 3] <- round(nuts[nuts$sample == i & nuts$year == "2011", 4], 2)
  nuts2[i, 4] <- round(nuts[nuts$sample == i & nuts$year == "2012", 4], 2)
  nuts2[i, 5] <- round(nuts[nuts$sample == i & nuts$year == "2011", 5], 2)
  nuts2[i, 6] <- round(nuts[nuts$sample == i & nuts$year == "2012", 5], 2)
}

addtorow <- list()
addtorow$pos <- list(0, 0, 0)
addtorow$command <- c(" & \\multicolumn{2}{c}{DOC} &
                      \\multicolumn{2}{c}{TP} &
                      \\multicolumn{2}{c}{TN}\\\\\\n",
                      "Lake & \\multicolumn{2}{c}{(mg C L$^{-1}$)} &
                      \\multicolumn{2}{c}{($\\mu$g P L$^{-1}$)} &
                      \\multicolumn{2}{c}{(mg N L$^{-1}$)}\\\\\\n",
                      " & 2011 & 2012 & 2011 & 2012 & 2011 & 2012 \\\\\\n")
nut.tab <- xtable(nuts2)
align(nut.tab) <- "crrrrrr"
print(nut.tab, add.to.row = addtorow, include.colnames = FALSE,
      type= "latex", file="../tables/Table2.tex",
      hline.after = c(-1, -1, 0, nrow(nut.tab)))
print(nut.tab, add.to.row = addtorow, include.colnames = FALSE,
      comment = FALSE, hline.after = c(-1, -1, 0, nrow(nut.tab)))
```

Lake	DOC (mg C L^{-1})		TP (μg P L^{-1})		TN (mg N L^{-1})	
	2011	2012	2011	2012	2011	2012
Ann	6.15	5.97	3.98	7.27	0.42	0.43
Canyon	7.62	7.23	2.45	2.64	0.44	0.38
Howe	6.88	7.04	1.86	5.21	0.56	0.57
Ives	9.54	6.91	1.35	9.15	0.42	0.38
Lily	13.36	14.35	4.74	11.55	0.82	0.93
Mountain	5.41	5.27	2.11	4.87	0.34	0.34
Pony	31.65	28.99	1.52	49.95	1.56	1.86
Rush	4.44	4.22	3.55	3.84	0.30	0.41
SecondPine	7.20	6.26	10.76	12.92	0.43	0.44
UpperPine	7.99	7.84	2.96	11.21	0.59	0.57

Patterns of Resource Diversity

A major distinction of this project is the addition of organic matter diversity

Import Raw Data

```
# Define Inputs
# Resource = raw site-by-resource matrix
resource.pos <- "../data/SpecAbundAvePos.csv"
resource.neg <- "../data/SpecAbundAveNeg.csv"
```

```
design.in <- "../data/design.txt"

# Import Design
design <- read.delim(design.in, header=T, row.names=1)

# Import Resources
res.in <- read.csv(resource.neg, header=T, row.names=1)
rownames(res.in) <- c("Ann", "blank", "CanyonChemo", "Canyon", "CanyonHypo",
                      "CanyonI", "CanyonII", "CanyonIII", "CanyonIV", "Howe",
                      "Ives", "Jordan", "Lily", "Mountain", "Pony", "Rush",
                      "SecondPine", "UpperPine")

blank <- unlist(res.in["blank", ])
res.hmwf <- res.in[-c(which(rownames(res.in) %in% c("blank", "CanyonChemo",
                           "CanyonHypo", "CanyonI", "CanyonII",
                           "CanyonIII", "CanyonIV", "Jordan"))), ]
```

Remove Major Peaks from Blanks

```
# summary(blank)
# blank[which(blank > 2 * sd(blank))]
# res.hmwf <- res.hmwf[, -c(which(blank > sd(blank)))]

# What other peaks should be removed
for (i in 1:dim(res.hmwf)[1]){
  res.hmwf[i, ] <- res.hmwf[i, ] - blank * 1.1
}

res.hmwf[res.hmwf < 50] <- 0
res.hmwf <- res.hmwf[,colSums(res.hmwf) > 0]
```

Data Transformations

```
# Remove OTUs with less than two occurences across all sites
res <- res.hmwf

# Sequencing Coverage
coverage <- rowSums(res)

# # Remove Low Coverage Samples (This code removes two sites: Site 5DNA, Site 6cDNA)
# lows <- which(coverage < 10000)
# OTUs <- OTUs[-which(coverage < 10000), ]
# design <- design[-which(coverage < 10000), ]

# Make Relative Abundence Matrices
resREL <- res
for(i in 1:dim(res)[1]){
  resREL[i,] <- res[i,]/sum(res[i,])
}
```

```r
# Log Transform Relative Resource Abundance
resREL.log <- decostand(resREL, method="log")
```

```
## Warning: non-integer data: divided by smallest positive value
```

Calculate Alpha Diversity

```r
# Observed Richness
S.res <- rowSums((res > 0) * 1)

# Simpson's Evenness
SimpE <- function(x = ""){
  x <- as.data.frame(x)
  D <- vegan::diversity(x, "inv")
  S <- sum((x > 0) * 1)
  E <- (D)/S
  return(E)
}

res.simpsE <- round(apply(res, 1, SimpE), 3)

# Shannon's Diversity
res.shan2 <- round(vegan::diversity(res, index = "shannon"), 2)

res.div <- as.data.frame(cbind(S.res, res.simpsE, res.shan2))

# Summary Stats
range(res.div$S.res)
```

```
## [1] 529 569
```

```r
range(res.div$res.shan2)
```

```
## [1] 4.89 5.56
```

```r
range(res.div$res.simpsE)
```

```
## [1] 0.053 0.152
```

```r
CV(res.div$S.res)
```

```
## [1] 2.118764
```

```r
CV(res.div$res.shan2)
```

```
## [1] 3.629068
```

```
CV(res.div$res.simpsE)
```

```
## [1] 32.91581
```

Calculate and Visualize Resource Beta Diversity

```
# Calculate Bray-Curtis
hmwf.bray.res <- vegdist(res, method = "bray")

dis.mean <- mean(hmwf.bray.res)

# Principal Coordinates Analysis
pcoa.res <- cmdscale(hmwf.bray.res, eig = TRUE, k = 3)
explainvar1.res <- round(pcoa.res$eig[1] / sum(pcoa.res$eig), 3) * 100
explainvar2.res <- round(pcoa.res$eig[2] / sum(pcoa.res$eig), 3) * 100
explainvar3.res <- round(pcoa.res$eig[3] / sum(pcoa.res$eig), 3) * 100
sum.eig.res <- sum(explainvar1.res, explainvar2.res, explainvar3.res)
```

Figure 1: Organic Matter Ordination Figure

```
design2 <- design[design$Molecule == "DNA" & design$Year == "2012", ]
# Custom palette
palette(rainbow_hcl(10, c = 80, l = 60))
lake.col <- rep(NA, length(unique(design2$Lake)))
names(lake.col) <- unique(design2$Lake)
lake.col <- as.numeric(factor(design2$Lake))

png(filename="../figures/Figure1.png",
    width = 900, height = 900, res = 96*2)
par(opar)

# Define Plot Parameters
layout(matrix(1))
par(mar = c(5, 5, 0, 0) + 0.5)

plot(pcoa.res$points[ ,1], pcoa.res$points[ ,2],
     ylim = c(-0.25, 0.3), xlim = c(-0.25, 0.3),
     xlab = paste("PCoA 1 (", explainvar1.res, "%)", sep = ""),
     ylab = paste("PCoA 2 (", explainvar2.res, "%)", sep = ""),
     #xlab = "", ylab = "",
     xaxt = "n", yaxt = "n",
     pch = 17, cex = 2.0, type = "n", cex.lab = 1.5, cex.axis = 1,
     axes = FALSE)

  # Add Axes
  axis(side = 1, labels = T, lwd.ticks = 2, cex.axis = 1, las = 1)
  axis(side = 2, labels = T, lwd.ticks = 2, cex.axis = 1, las = 1)
  axis(side = 3, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=-0.02)
  axis(side = 4, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=-0.02)
```

```
  axis(side = 1, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
  axis(side = 2, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
  axis(side = 3, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
  axis(side = 4, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
  abline(h = 0, v = 0, lty = 3)
  box(lwd = 2)

# Add Points & Labels
points(pcoa.res$points[ ,1], pcoa.res$points[ ,2], pch = 15,
       cex = 4, bg = "gray", col = lake.col)
text(pcoa.res$points[ ,1], pcoa.res$points[ ,2], labels = row.names(pcoa.res$points))

dev.off() # this writes plot to folder
graphics.off() # shuts down open devices
par(opar)
```

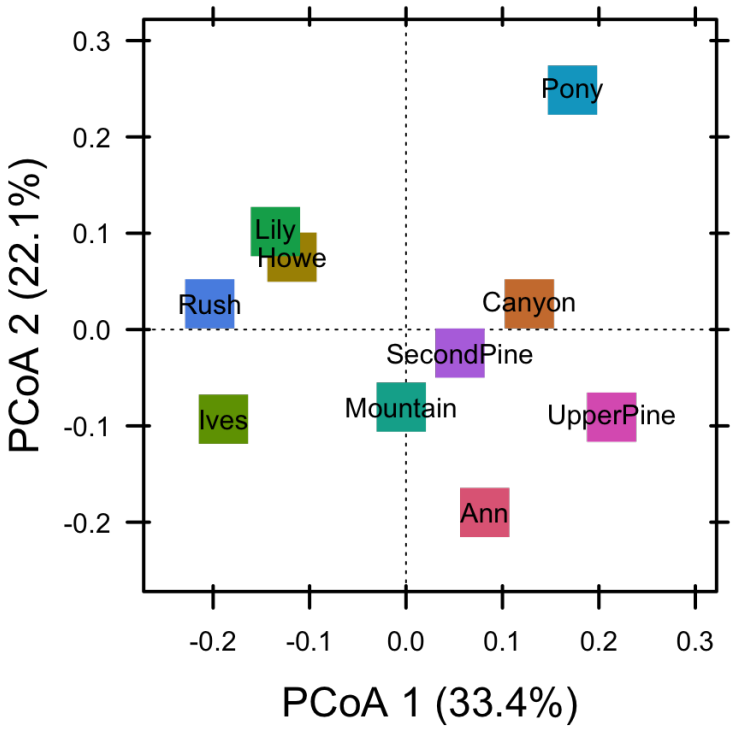


Figure 2: PCoA Plot Resources

Patterns of Bacterial Diversity

The major difference that we are interested in is bacterial diversity across the sites. We know that the lakes have different microbiomes. These microbiomes can be influenced by physical, chemical and biological interactions within each lake.

Import Raw Data

```
# Define Inputs
# Design = general design file for experiment
# shared = OTU table from mothur with sequence similarity clustering
# Taxonomy = Taxonomic information for each OTU
design.in <- "../data/design.txt"
shared <- "../data/HMWF.bac.final.shared"
taxon  <- "../data/HMWF.bac.final.0.03.taxonomy"

# Import Design
design <- read.delim(design.in, header=T, row.names=1)

# Import Shared Files
OTUs.in <- read.otu(shared = shared, cutoff = "0.03")    # 97% Similarity

# Import Taxonomy
OTU.tax <- read.tax(taxonomy = taxon, format = "rdp")
```

Data Transformations

```
# Reorder Site
OTUs.hmwf <- OTUs.in[rownames(design), ]

# Remove OTUs with less than two occurences across all sites
# OTUs <- OTUs.hmwf[, which(colSums(OTUs.hmwf) >= 2)]
OTUs <- OTUs.hmwf[, colSums((OTUs.hmwf > 0) * 1)  >= 2 | colSums(OTUs.hmwf >= 10)]

# Sequencing Coverage
coverage <- rowSums(OTUs)

# Good's Coverage
goods <- function(x = ""){
  1 - (sum(x == 1) / rowSums(x))
}
goods.c <- goods(OTUs)

# Make Presence Absence Matrix
OTUsPA <- (OTUs > 0) * 1

# Make Relative Abundence Matrices
OTUsREL <- OTUs
for(i in 1:dim(OTUs)[1]){
  OTUsREL[i,] <- OTUs[i,]/sum(OTUs[i,])
}

# Log Transform Relative Abundances
OTUsREL.log <- decostand(OTUs, method="log")
```

Calculate Alpha Diversity

```
# Observed Richness
S.obs <- rowSums((OTUs > 0) * 1)

# Simpson's Evenness
SimpE <- function(x = ""){
  x <- as.data.frame(x)
  D <- vegan::diversity(x, "inv")
  S <- sum((x > 0) * 1)
  E <- (D)/S
  return(E)
}

simpsE <- round(apply(OTUs, 1, SimpE), 3)

# Shannon's Diversity
H <- function(x = ""){
  x <- x[x>0]
  H = 0
  for (n_i in x){
    p = n_i / sum(x)
    H = H - p*log(p)
  }
  return(H)
}

shan <- round(apply(OTUs, 1, H), 2)
shan2 <- vegan::diversity(OTUs, index = "shannon")

# Rarefied Richness
S.rar <- round(rarefy(OTUs, min(rowSums(OTUs))), 0)

design <- droplevels(design)

alpha.div <- cbind(design, S.obs, simpsE, shan, S.rar)
alpha.div <- alpha.div[order(alpha.div$Lake, alpha.div$Year, alpha.div$Molecule), ]
```

Alpha Diversity Statistics

```
div.mod1 <- aov(S.obs ~ Lake, data = alpha.div)
# TukeyHSD(div.mod1)
div.mod2 <- aov(simpsE ~ Lake, data = alpha.div)
# TukeyHSD(div.mod2)
div.mod3 <- aov(S.rar ~ Lake + Molecule, data = alpha.div)
# TukeyHSD(div.mod3)
div.mod3 <- lm(S.rar ~ Lake + Molecule, data = alpha.div)
# Anova(div.mod3)
div.mod4 <- aov(S.rar ~ Lake + as.factor(Year), data = alpha.div)
# TukeyHSD(div.mod4)
div.mod5 <- aov(S.rar ~ Lake, data = alpha.div[alpha.div$Lake != "Pony" &
```

```
                                                    alpha.div$Lake != "Lily", ])
div.mod5 <- lm(S.rar ~ Lake + Molecule, data = alpha.div[alpha.div$Lake != "Pony" &
                                                    alpha.div$Lake != "Lily", ])
Anova(div.mod5)
```

```
## Anova Table (Type II tests)
##
## Response: S.rar
##           Sum Sq Df F value Pr(>F)
## Lake      127554  7  1.6305 0.1769
## Molecule   26623  1  2.3821 0.1364
## Residuals 257049 23
```

Lake Phylogenetic Diversity

```
# Import Tree with *ape*
hmwf.tree <- read.tree("../fasttree/HMWF.bac.0.03.gg.tree")
# tips <- paste(rep("Otu", length(hmwf.tree$tip.label)),
#               formatC(seq(1:length(hmwf.tree$tip.label)),
#                       width = 6, format = "d", flag = "0"), sep = "")
# hmwf.tree$tip.label <- tips
hmwf.tree <- root(hmwf.tree, "Otu011336")
prunedTree <- prune.sample(OTUs,hmwf.tree)

OTUs.rar <- rrarefy(OTUs, sample = min(rowSums(OTUs)))

faiths <- pd(OTUs, prunedTree, include.root = FALSE)
# ses.pd(OTUs, prunedTree, include.root = FALSE, null.model = "independentswap", runs = 9, iterations =
PSR <- psr(OTUs, prunedTree, compute.var = F)
PSR <- psr(OTUs.rar, prunedTree, compute.var = F)
PSD <- psd(OTUs, prunedTree, compute.var = F)
# specaccum.psr(OTUs, prunedTree, permutations = 100, method = "random")
# plot(faiths$PD ~ faiths$SR)
# plot(PSR$PSR ~ PSR$SR)
```

Phylogenetic Diversity Statistics

```
faiths2 <- cbind(faiths, alpha.div)
phy.mod1 <- aov(PD ~ Lake, data = faiths2)
summary(phy.mod1)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## Lake         9  84569    9397   7.525 1.13e-05 ***
## Residuals   30  37464    1249
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# TukeyHSD(phy.mod1)
phy.mod2 <- aov(PD ~ Lake + Molecule, data = faiths2[faiths2$Lake != "Pony" &
                                                      faiths2$Lake != "Lily", ])
summary(phy.mod2)
```

```
##             Df Sum Sq Mean Sq F value  Pr(>F)
## Lake         7   7926    1132   1.723 0.15302
## Molecule     1   5552    5552   8.449 0.00795 **
## Residuals   23  15114     657
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
phy.mod3 <- aov(PD ~ Lake + Molecule, data = faiths2)
summary(phy.mod3)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## Lake         9  84569    9397    9.32 1.74e-06 ***
## Molecule     1   8227    8227    8.16  0.00784 **
## Residuals   29  29237    1008
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
phy.mod4 <- lm(PD ~ Lake + Molecule, data = faiths2)

DNA.PD <- mean(faiths2$PD[faiths2$Molecule == "DNA"])
RNA.PD <- mean(faiths2$PD[faiths2$Molecule == "RNA"])

phy.mod4 <- lm(PD ~ SR + Lake + Molecule, data = faiths2)
Anova(phy.mod4)
```

```
## Anova Table (Type II tests)
##
## Response: PD
##            Sum Sq Df  F value    Pr(>F)
## SR        28409.8  1 961.6673 < 2.2e-16 ***
## Lake       2437.1  9   9.1661 2.592e-06 ***
## Molecule   2885.6  1  97.6779 1.247e-10 ***
## Residuals   827.2 28
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 3: Bacterial Diversity

```r
div.raw <- cbind(alpha.div, faiths)
div <- data.frame(matrix(NA, ncol = 9, nrow = 10))
colnames(div) <- c("Lake", "SR_T_11", "SR_T_12", "PR_T_11", "PR_T_12",
                   "SR_A_11", "SR_A_12", "PR_A_11", "PR_A_12")
row.names(div) <- div.raw$Lake[div.raw$Molecule == "DNA" & div.raw$Year == 2011]
div[, 1] <- div.raw$S.rar[div.raw$Molecule == "DNA" & div.raw$Year == 2011]
```

```
div[, 2] <- div.raw$S.rar[div.raw$Molecule == "DNA" & div.raw$Year == 2012]
div[, 3] <- div.raw$PD[div.raw$Molecule == "DNA" & div.raw$Year == 2011]
div[, 4] <- div.raw$PD[div.raw$Molecule == "DNA" & div.raw$Year == 2012]
div[, 5] <- div.raw$S.rar[div.raw$Molecule == "RNA" & div.raw$Year == 2011]
div[, 6] <- div.raw$S.rar[div.raw$Molecule == "RNA" & div.raw$Year == 2012]
div[, 7] <- div.raw$PD[div.raw$Molecule == "RNA" & div.raw$Year == 2011]
div[, 8] <- div.raw$PD[div.raw$Molecule == "RNA" & div.raw$Year == 2012]

div[, 1:8] <- round(div[, 1:8], 0)

addtorow <- list()
addtorow$pos <- list(0, 0, 0)
addtorow$command <- c(" & \\multicolumn{4}{c}{Total} &
                      \\multicolumn{4}{c}{Active} \\\\\n",
                      "Lake & \\multicolumn{2}{c}{S\\textsubscript{spec}} &
                      \\multicolumn{2}{c}{S\\textsubscript{phy}} &
                      \\multicolumn{2}{c}{S\\textsubscript{spec}} &
                      \\multicolumn{2}{c}{S\\textsubscript{phy}} \\\\\n",
                      " & 2011 & 2012 & 2011 & 2012 &
                      2011 & 2012 & 2011 & 2012 \\\\\n")
div.tab <- xtable(div, auto = TRUE)
align(div.tab) <- c("c ","r ", "r ", "r ", "r "," r ", "r ", "r ", "r ", "r ")
print(div.tab, add.to.row = addtorow, include.colnames = FALSE,
      type= "latex", file="../tables/Table3.tex",
      hline.after = c(-1, -1, 0, nrow(div.tab)))
print(div.tab, add.to.row = addtorow, include.colnames = FALSE,
      comment = FALSE, hline.after = c(-1, -1, 0, nrow(div.tab)))
```

Lake	Total				Active			
	S_{spec}		S_{phy}		S_{spec}		S_{phy}	
	2011	2012	2011	2012	2011	2012	2011	2012
Ann	765	678	203	184	860	705	165	153
Canyon	709	727	186	195	907	801	179	161
Howe	736	627	196	163	720	649	115	127
Ives	544	521	144	134	562	589	107	127
Lily	777	1076	227	284	703	1254	135	266
Mountain	707	690	185	179	757	653	163	139
Pony	1337	1956	305	313	1154	1863	248	328
Rush	776	575	203	149	822	560	178	119
SecondPine	717	530	181	129	905	584	192	121
UpperPine	745	543	212	148	891	548	196	129

Paired T-Tests

```
div.2011 <- div.raw[div.raw$Year == "2011", ]
div.2012 <- div.raw[div.raw$Year == "2012", ]
t.test(div.2011$S.rar, div.2012$S.rar, paired = T)
```

```
##
##  Paired t-test
```

```
##
## data:  div.2011$S.rar and div.2012$S.rar
## t = -0.025738, df = 19, p-value = 0.9797
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -144.0598  140.5598
## sample estimates:
## mean of the differences
##                   -1.75
```

```
div.D <- div.raw[div.raw$Molecule == "DNA", ]
div.R <- div.raw[div.raw$Molecule == "RNA", ]
t.test(div.D$S.rar, div.R$S.rar, paired = T)
```

```
##
##  Paired t-test
##
## data:  div.D$S.rar and div.R$S.rar
## t = -1.7397, df = 19, p-value = 0.09808
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -82.725922   7.625922
## sample estimates:
## mean of the differences
##                  -37.55
```

```
mod1 <- lm(div.raw$S.rar ~ as.factor(div.raw$Lake) +
             as.factor(div.raw$Year) + as.factor(div.raw$Molecule))
Anova(mod1)
```

```
## Anova Table (Type II tests)
##
## Response: div.raw$S.rar
##                               Sum Sq Df F value    Pr(>F)
## as.factor(div.raw$Lake)     3019518  9 10.0934 1.008e-06 ***
## as.factor(div.raw$Year)          31  1  0.0009    0.9760
## as.factor(div.raw$Molecule)   14100  1  0.4242    0.5202
## Residuals                    930712 28
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Calculate and Visualize Beta Diversity

```
beta.w <- function(site1 = "", site2 = ""){
  site1 = subset(site1, select = site1 > 0)              # Removes absences
  site2 = subset(site2, select = site2 > 0)              # Removes absences
  gamma = union(colnames(site1), colnames(site2))        # Gamma species pool
  s     = length(gamma)                                  # Gamma richness
  a.bar = mean(c(specnumber(site1), specnumber(site2)))  # Mean sample richness
  b.w   = round(s/a.bar - 1, 3)
```

```
  return(b.w)
}

# Calculate Bray-Curtis
hmwf.bray.PA <- vegdist(OTUsPA, method = "bray")
hmwf.bray.REL <- vegdist(OTUsREL, method = "bray")
hmwf.bray.Log <- vegdist(OTUsREL.log, method = "bray")

# Import UniFrac Distances
hmwf.uni.in <- read.table("../fasttree/HMWF.bac.0.03.gg.tree1.weighted.phylip.dist", skip=1, row.names =
hmwf.uni <- as.dist(hmwf.uni.in)

OTU.tab <- otu_table(OTUsREL, taxa_are_rows = F)
OTU.tab.l <- otu_table(OTUsREL.log, taxa_are_rows = F)
PHY.tree <- phy_tree(hmwf.tree)
phylo.seq <- phyloseq(OTU.tab, PHY.tree)
phylo.seq.l <- phyloseq(OTU.tab.l, PHY.tree)
uni.frac.u <- UniFrac(phylo.seq, weighted = FALSE, normalized = TRUE)
uni.frac.w <- UniFrac(phylo.seq, weighted = TRUE, normalized = TRUE)
uni.frac.wl <- UniFrac(phylo.seq.l, weighted = TRUE, normalized = TRUE)
```

Principal Coordinates Analysis

```
# Bray Curtis PA
pcoa.pa <- cmdscale(hmwf.bray.PA, eig = TRUE, k = 3)
explainvar1.pa <- round(pcoa.pa$eig[1] / sum(pcoa.pa$eig), 3) * 100
explainvar2.pa <- round(pcoa.pa$eig[2] / sum(pcoa.pa$eig), 3) * 100
explainvar3.pa <- round(pcoa.pa$eig[3] / sum(pcoa.pa$eig), 3) * 100
sum.eig.pa <- sum(explainvar1.pa, explainvar2.pa, explainvar3.pa)

# Bray Curtis REL
pcoa.rel <- cmdscale(hmwf.bray.REL, eig = TRUE, k = 3)
explainvar1.rel <- round(pcoa.rel$eig[1] / sum(pcoa.rel$eig), 3) * 100
explainvar2.rel <- round(pcoa.rel$eig[2] / sum(pcoa.rel$eig), 3) * 100
explainvar3.rel <- round(pcoa.rel$eig[3] / sum(pcoa.rel$eig), 3) * 100
sum.eig.rel <- sum(explainvar1.rel, explainvar2.rel, explainvar3.rel)

# Bray Curtis REL Log
pcoa.log <- cmdscale(hmwf.bray.Log, eig = TRUE, k = 3)
explainvar1.log <- round(pcoa.log$eig[1] / sum(pcoa.log$eig), 3) * 100
explainvar2.log <- round(pcoa.log$eig[2] / sum(pcoa.log$eig), 3) * 100
explainvar3.log <- round(pcoa.log$eig[3] / sum(pcoa.log$eig), 3) * 100
sum.eig.log <- sum(explainvar1.log, explainvar2.log, explainvar3.log)

# UniFrac Unweighted
pcoa.ufu <- cmdscale(uni.frac.u, eig = TRUE, k = 3)
explainvar1.ufu <- round(pcoa.ufu$eig[1] / sum(pcoa.ufu$eig), 3) * 100
explainvar2.ufu <- round(pcoa.ufu$eig[2] / sum(pcoa.ufu$eig), 3) * 100
explainvar3.ufu <- round(pcoa.ufu$eig[3] / sum(pcoa.ufu$eig), 3) * 100
sum.eig.ufu <- sum(explainvar1.ufu, explainvar2.ufu, explainvar3.ufu)
```

```r
# UniFrac Weighted
pcoa.ufw <- cmdscale(uni.frac.w, eig = TRUE, k = 3)
explainvar1.ufw <- round(pcoa.ufw$eig[1] / sum(pcoa.ufw$eig), 3) * 100
explainvar2.ufw <- round(pcoa.ufw$eig[2] / sum(pcoa.ufw$eig), 3) * 100
explainvar3.ufw <- round(pcoa.ufw$eig[3] / sum(pcoa.ufw$eig), 3) * 100
sum.eig.ufwl <- sum(explainvar1.ufw, explainvar2.ufw, explainvar3.ufw)

# UniFrac Weighted Log
pcoa.ufwl <- cmdscale(uni.frac.wl, eig = TRUE, k = 3)
explainvar1.ufwl <- round(pcoa.ufwl$eig[1] / sum(pcoa.ufwl$eig), 3) * 100
explainvar2.ufwl <- round(pcoa.ufwl$eig[2] / sum(pcoa.ufwl$eig), 3) * 100
explainvar3.ufwl <- round(pcoa.ufwl$eig[3] / sum(pcoa.ufwl$eig), 3) * 100
sum.eig.ufwl <- sum(explainvar1.ufwl, explainvar2.ufwl, explainvar3.ufwl)

dis.meanB <- mean(hmwf.bray.REL)
dis.meanU <- mean(uni.frac.w)
```

Figure 2: Bacterial Community Composition Ordination Figures (Rel Abundance)

```r
design$Lake_Mol <- paste(design$Lake, design$Molecule, sep = "_")

# after specifying custom palette
palette(rainbow_hcl(10, c = 80, l = 60)) #[c(5, 2, 6, 3, 1)])
lake.col <- rep(NA, length(unique(design$Lake)))
names(lake.col) <- unique(design$Lake)
lake.col <- as.numeric(factor(design$Lake))
design$lake.col <- NA
for (i in 1:dim(design)[1]){
  design$lake.col[i] <- which(levels(design$Lake) == design$Lake[i])
}

png(filename="../figures/Figure2.png",
    width = 1800, height = 900, res = 96*2)
par(opar)
# Define Plot Parameters
layout(matrix(c(1, 1, 2, 2, 3), ncol = 5, byrow = T))
par(mar = c(4, 5, 1, 1) + 0.5, oma = c(1, 1, 1, 1))

# Define Plot Symbols
lake.pch <- rep(NA, length(design$Molecule))
for (i in 1:length(design$Molecule)){
if (design$Molecule[i] == "DNA"){
    lake.pch[i] <- 16
  }else{
    lake.pch[i] <- 17
  }}

pcoa.plots <- list(pcoa.pa, pcoa.ufu,
                   pcoa.rel, pcoa.ufw,
                   pcoa.log, pcoa.ufwl)
```

```
explainvar1 <- c(explainvar1.pa, explainvar1.rel, explainvar1.log, explainvar1.ufu,
                 explainvar1.ufw, explainvar1.ufwl)
explainvar2 <- c(explainvar2.pa, explainvar2.rel, explainvar2.log, explainvar2.ufu,
                 explainvar2.ufw, explainvar2.ufwl)

xlabel <- c(F, F, T, T, T, T)
ylabel <- c(T, F, T, F, T, F)

# Initiate Plot 1
plot(pcoa.rel$points[ ,1], pcoa.rel$points[ ,2],
     ylim = c(-0.4, 0.4), xlim = c(-0.25, 0.55),
     xlab = paste("PCoA 1 (", explainvar1.rel, "%)", sep = ""),
     ylab = paste("PCoA 2 (", explainvar2.rel, "%)", sep = ""),
     #xlab = "", ylab = "", xaxt = "n", yaxt = "n",
     pch = lake.pch, cex = 2.0, type = "n", cex.lab = 1.5, cex.axis = 1,
     axes = FALSE)

# Add Axes
axis(side = 1, labels = T, lwd.ticks = 2, cex.axis = 1, las = 1)
axis(side = 2, labels = T, lwd.ticks = 2, cex.axis = 1, las = 1)
axis(side = 3, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=-0.02)
axis(side = 4, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=-0.02)
axis(side = 1, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
axis(side = 2, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
axis(side = 3, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
axis(side = 4, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
abline(h = 0, v = 0, lty = 3)
box(lwd = 2)

# Add Points & Labels
points(pcoa.rel$points[ ,1], pcoa.rel$points[ ,2], pch = lake.pch,
       cex = 4, bg = "gray", col = lake.col)

# Add Molecule Hulls
ordihull(cbind(pcoa.rel$points[ ,1], pcoa.rel$points[ ,2]),
         design$Molecule, lwd=2, draw = "polygon", col="gray80", border = "black",
         label=TRUE, cex=1, bty = 'n')

# Initiate Plot 2
plot(pcoa.ufw$points[ ,1], pcoa.ufw$points[ ,2],
     ylim = c(-0.4, 0.4), xlim = c(-0.25, 0.55),
     xlab = paste("PCoA 1 (", explainvar1.ufw, "%)", sep = ""),
     ylab = paste("PCoA 2 (", explainvar2.ufw, "%)", sep = ""),
     #xlab = "", ylab = "", xaxt = "n", yaxt = "n",
     pch = lake.pch, cex = 2.0, type = "n", cex.lab = 1.5, cex.axis = 1,
     axes = FALSE)

# Add Axes
axis(side = 1, labels = T, lwd.ticks = 2, cex.axis = 1, las = 1)
axis(side = 2, labels = T, lwd.ticks = 2, cex.axis = 1, las = 1)
axis(side = 3, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=-0.02)
axis(side = 4, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=-0.02)
axis(side = 1, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
```

```
axis(side = 2, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
axis(side = 3, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
axis(side = 4, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
abline(h = 0, v = 0, lty = 3)
box(lwd = 2)

# Add Points & Labels
points(pcoa.ufw$points[ ,1], pcoa.ufw$points[ ,2], pch = lake.pch,
       cex = 4, bg = "gray", col = lake.col)

# Add Molecule Hulls
ordihull(cbind(pcoa.ufw$points[ ,1], pcoa.ufw$points[ ,2]),
         design$Molecule, lwd=2, draw = "polygon", col="gray80", border = "black",
         label=TRUE, cex=1, bty = 'n')

# Add Legends as Plot 3
par(mar = c(4, 0, 1, 0) + 0.5)
plot.new()
legend("bottomleft", c("DNA", "RNA"), pch = c(16, 17),
       col = "gray", bty = "n", pt.cex = 1.5, cex = 1.5,
       y.intersp = 1.25, inset = c(0, 0, 0, 0.1))
legend("topleft", levels(design$Lake), ncol = 1, pch = 16, col = 1:10,
       bty = "n", pt.cex = 1.5, cex = 1.5, y.intersp = 1.25,
       inset = c(0, 0, 0, 0.1))



dev.off() # this writes plot to folder
graphics.off() # shuts down open devices
par(opar)
```

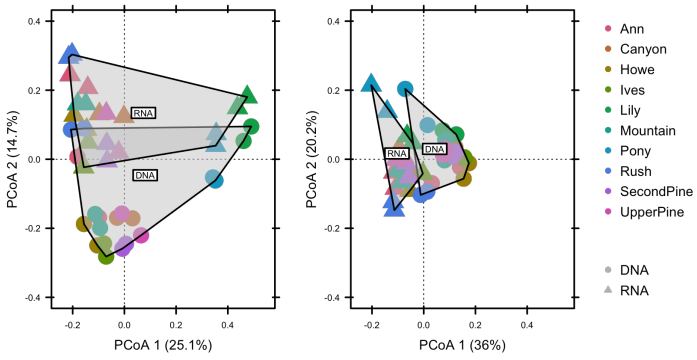


Figure 3: PCoA Plots

PERMANOVA

```
beta.dis <- betadisper(vegdist(OTUsREL, "bray"), design$Lake)
permutest(beta.dis)
```

```
##
## Permutation test for homogeneity of multivariate dispersions
## Permutation: free
## Number of permutations: 999
##
## Response: Distances
##           Df   Sum Sq   Mean Sq      F N.Perm Pr(>F)
## Groups     9 0.030134 0.0033482 1.7859    999  0.135
## Residuals 30 0.056244 0.0018748
```

```
per.bray.pa <- adonis(OTUsPA ~ design$Lake + design$Molecule, method = "bray")
per.bray.rel <- adonis(OTUsREL ~ design$Lake + design$Molecule, method = "bray")
per.bray.l <- adonis(OTUsREL.log ~ design$Lake + design$Molecule, method = "bray")
per.uni.pa <- adonis(uni.frac.u ~ design$Lake + design$Molecule)
per.uni.rel <- adonis(uni.frac.w ~ design$Lake + design$Molecule)
per.uni.l <- adonis(uni.frac.wl ~ design$Lake + design$Molecule)
```

PERMANOVA Table

```
per.models <- data.frame(matrix(NA, ncol = 4, nrow = 6))
colnames(per.models) <- c("Lake R2", "Lake P", "Molecule R2", "Molecule P")
row.names(per.models) <- c("Bray-Curtis -- PA", "Bray-Curtis -- REL", "Bray-Curtis -- Log",
                    "UniFrac -- PA", "UniFrac -- REL", "UniFrac -- Log")
per.models[1, ] <- c(per.bray.pa$aov.tab$R2[1], per.bray.pa$aov.tab$`Pr(>F)`[1],
                     per.bray.pa$aov.tab$R2[2], per.bray.pa$aov.tab$`Pr(>F)`[2])
per.models[2, ] <- c(per.bray.rel$aov.tab$R2[1], per.bray.rel$aov.tab$`Pr(>F)`[1],
                     per.bray.rel$aov.tab$R2[2], per.bray.rel$aov.tab$`Pr(>F)`[2])
per.models[3, ] <- c(per.bray.l$aov.tab$R2[1], per.bray.l$aov.tab$`Pr(>F)`[1],
                     per.bray.l$aov.tab$R2[2], per.bray.l$aov.tab$`Pr(>F)`[2])
per.models[4, ] <- c(per.uni.pa$aov.tab$R2[1], per.uni.pa$aov.tab$`Pr(>F)`[1],
                     per.uni.pa$aov.tab$R2[2], per.uni.pa$aov.tab$`Pr(>F)`[2])
per.models[5, ] <- c(per.uni.rel$aov.tab$R2[1], per.uni.rel$aov.tab$`Pr(>F)`[1],
                     per.uni.rel$aov.tab$R2[2], per.uni.rel$aov.tab$`Pr(>F)`[2])
per.models[6, ] <- c(per.uni.l$aov.tab$R2[1], per.uni.l$aov.tab$`Pr(>F)`[1],
                     per.uni.l$aov.tab$R2[2], per.uni.l$aov.tab$`Pr(>F)`[2])

per.models[, c(1,3)] <- round(per.models[, c(1,3)], 2)

addtorow <- list()
addtorow$pos <- list(0, 0)
addtorow$command <- c("Model & \\multicolumn{2}{c}{Lake} & \\multicolumn{2}{c}{Molecule} \\\\\n",
                      " & \\multicolumn{1}{c}{R\\textsuperscript{2}} &
                          \\multicolumn{1}{c}{\\emph{P}} &
                          \\multicolumn{1}{c}{R\\textsuperscript{2}} &
                          \\multicolumn{1}{c}{\\emph{P}} \\\\\n")
per.tab <- xtable(per.models, auto = TRUE)
align(per.tab) <- c("l ","r ", "r ", "r ", "r ")
print(per.tab, add.to.row = addtorow, include.colnames = FALSE,
      type= "latex", file="../tables/Table5.tex",
      hline.after = c(-1, -1, 0, nrow(per.tab)))
print(per.tab, add.to.row = addtorow, include.colnames = FALSE,
      comment = FALSE, hline.after = c(-1, -1, 0, nrow(per.tab)))
```

Model	Lake		Molecule	
	R^2	P	R^2	P
Bray-Curtis – PA	0.51	0.001	0.04	0.004
Bray-Curtis – REL	0.64	0.001	0.11	0.001
Bray-Curtis – Log	0.60	0.001	0.03	0.004
UniFrac – PA	0.45	0.001	0.05	0.001
UniFrac – REL	0.53	0.001	0.28	0.001
UniFrac – Log	0.58	0.001	0.11	0.001

Resource Diversity (Evennes) and Nutrient Concentrations

```
nuts2012 <- nuts[nuts$year == "2012", ]
evenmod <- glm(res.div$res.simpsE ~ nuts2012$DOC+nuts2012$TN+nuts2012$TP)
# Not using this because variation in resource alpha doesn't appear to be meaningful
```

Nutrient Concentrations Explain BCC and DOM Composition

Distance Based Redundancy Analysis

```
# Define Environmental Matrix
env.chem <- as.data.frame(scale(apply(nuts[,3:5], 2, log)))
row.names(env.chem) <- paste(nuts$sample, nuts$year, sep = "-")
env.chem <- as.data.frame(env.chem[sort(row.names(env.chem)), ])

env.chem2012 <- as.data.frame(env.chem[grep("2012", row.names(env.chem)), ])
designDNA <- design[design$Molecule == "DNA", ]
designRNA <- design[design$Molecule == "RNA", ]
designRES <- design[design$Molecule == "DNA" & design$Year == "2012", ]

# Define DNA Community
OTUsREL.D <- OTUsREL[design$Molecule == "DNA" , ]

# Define RNA Community
OTUsREL.R <- OTUsREL[design$Molecule == "RNA" , ]


# Calculate Bray-Curtis Distances for Bacteria
Bray.REL.D <- vegdist(OTUsREL.D, "bray")
Bray.REL.R <- vegdist(OTUsREL.R, "bray")

# Calculate UniFrac Distances for Bacteria
OTU.tab.D <- otu_table(OTUsREL.D, taxa_are_rows = F)
OTU.tab.R <- otu_table(OTUsREL.R, taxa_are_rows = F)
PHY.tree <- phy_tree(hmwf.tree)
phylo.seq.D <- phyloseq(OTU.tab.D, PHY.tree)
phylo.seq.R <- phyloseq(OTU.tab.R, PHY.tree)
Uni.REL.D <- UniFrac(phylo.seq.D, weighted = TRUE, normalized = TRUE)
Uni.REL.R <- UniFrac(phylo.seq.R, weighted = TRUE, normalized = TRUE)
```

```
# Bray-Curtis Distances for Organic Matter
Bray.OM <- vegdist(res, "bray")

# Conduct dbRDA
hmwf.bray.D.rda <- capscale(Bray.REL.D ~ env.chem$DOC + env.chem$TP + env.chem$TN,
                            comm = OTUsREL.D, add = T)
hmwf.bray.R.rda <- capscale(Bray.REL.R ~ env.chem$DOC + env.chem$TP + env.chem$TN,
                            comm = OTUsREL.R, add = T)
hmwf.uni.D.rda <- capscale(Uni.REL.R ~ env.chem$DOC + env.chem$TP + env.chem$TN,
                           comm = OTUsREL.D, add = T)
hmwf.uni.R.rda <- capscale(Uni.REL.R ~ env.chem$DOC + env.chem$TP + env.chem$TN,
                            comm = OTUsREL.R, add = T)
hmwf.res.rda <- capscale(Bray.OM ~ env.chem2012$DOC + env.chem2012$TP + env.chem2012$TN,
                         comm = res, add = T)

anova(hmwf.bray.D.rda)
RsquareAdj(hmwf.bray.D.rda)
anova(hmwf.bray.R.rda)
RsquareAdj(hmwf.bray.R.rda)
anova(hmwf.uni.D.rda)
RsquareAdj(hmwf.uni.D.rda)
anova(hmwf.uni.R.rda)
RsquareAdj(hmwf.uni.R.rda)
anova(hmwf.res.rda)
RsquareAdj(hmwf.res.rda)

rda.D.explainvar1 <- round(hmwf.bray.D.rda$CCA$eig[1] / sum(hmwf.bray.D.rda$CCA$eig), 3) * 100
rda.D.explainvar2 <- round(hmwf.bray.D.rda$CCA$eig[2] / sum(hmwf.bray.D.rda$CCA$eig), 3) * 100
rda.R.explainvar1 <- round(hmwf.bray.R.rda$CCA$eig[1] / sum(hmwf.bray.R.rda$CCA$eig), 3) * 100
rda.R.explainvar2 <- round(hmwf.bray.R.rda$CCA$eig[2] / sum(hmwf.bray.R.rda$CCA$eig), 3) * 100
rda.UD.explainvar1 <- round(hmwf.uni.D.rda$CCA$eig[1] / sum(hmwf.uni.D.rda$CCA$eig), 3) * 100
rda.UD.explainvar2 <- round(hmwf.uni.D.rda$CCA$eig[2] / sum(hmwf.uni.D.rda$CCA$eig), 3) * 100
rda.UR.explainvar1 <- round(hmwf.uni.R.rda$CCA$eig[1] / sum(hmwf.uni.R.rda$CCA$eig), 3) * 100
rda.UR.explainvar2 <- round(hmwf.uni.R.rda$CCA$eig[2] / sum(hmwf.uni.R.rda$CCA$eig), 3) * 100
rda.Res.explainvar1 <- round(hmwf.res.rda$CCA$eig[1] / sum(hmwf.res.rda$CCA$eig), 3) * 100
rda.Res.explainvar2 <- round(hmwf.res.rda$CCA$eig[2] / sum(hmwf.res.rda$CCA$eig), 3) * 100

# Remove Lily and Pony
env.chem.L <- env.chem[nuts$sample != "Pony" & nuts$sample != "Lily", ]
OTUsREL.D.L <- OTUsREL[design$Molecule == "DNA" & design$Lake != "Pony" & design$Lake != "Lily", ]
OTUsREL.R.L <- OTUsREL[design$Molecule == "RNA" & design$Lake != "Pony" & design$Lake != "Lily", ]
Bray.REL.D.L <- vegdist(OTUsREL.D.L, "bray")
Bray.REL.R.L <- vegdist(OTUsREL.R.L, "bray")
hmwf.bray.D.L.rda <- capscale(Bray.REL.D.L ~ env.chem.L$DOC + env.chem.L$TP + env.chem.L$TN,
                            comm = OTUsREL.D.L, add = T)
hmwf.bray.R.L.rda <- capscale(Bray.REL.R.L ~ env.chem.L$DOC + env.chem.L$TP + env.chem.L$TN,
                            comm = OTUsREL.R.L, add = T)
anova(hmwf.bray.D.L.rda)
RsquareAdj(hmwf.bray.D.L.rda)
anova(hmwf.bray.R.L.rda)
RsquareAdj(hmwf.bray.R.L.rda)

# Partition Each
```

```
anova(hmwf.bray.D.rda, by = "terms")
anova(hmwf.bray.R.rda, by = "terms")
anova(hmwf.uni.D.rda, by = "terms")
anova(hmwf.uni.R.rda, by = "terms")


hmwf.bray.D.L.rda <- capscale(Bray.REL.D.L ~ env.chem.L$DOC + env.chem.L$TN,
                              comm = OTUsREL.D.L, add = T)
hmwf.bray.R.L.rda <- capscale(Bray.REL.R.L ~ env.chem.L$DOC + env.chem.L$TN,
                              comm = OTUsREL.R.L, add = T)
anova(hmwf.bray.D.L.rda)
RsquareAdj(hmwf.bray.D.L.rda)
anova(hmwf.bray.R.L.rda)
RsquareAdj(hmwf.bray.R.L.rda)

hmwf.bray.D.L.rda <- capscale(Bray.REL.D.L ~ env.chem.L$TP,
                              comm = OTUsREL.D.L, add = T)
hmwf.bray.R.L.rda <- capscale(Bray.REL.R.L ~ env.chem.L$TP,
                              comm = OTUsREL.R.L, add = T)
anova(hmwf.bray.D.L.rda)
RsquareAdj(hmwf.bray.D.L.rda)
anova(hmwf.bray.R.L.rda)
RsquareAdj(hmwf.bray.R.L.rda)
```

Figure 3: dbRDA

```
png(filename="../figures/Figure3.png",
    width = 1800, height = 1800, res = 96*2)
par(opar)
# Define Plot Parameters
layout(matrix(c(1:4), ncol = 2, byrow = T))
par(mar = c(5, 5, 3, 2) + 0.1, oma = c(1, 1, 1, 1))

# Initiate Plot 1
plot(scores(hmwf.bray.D.rda, display = "wa"), xlim = c(-1, 2), ylim = c(-1.5, 2),
     xlab = paste("dbRDA 1 (", rda.D.explainvar1, "%)", sep = ""),
     ylab = paste("dbRDA 2 (", rda.D.explainvar2, "%)", sep = ""),
     pch = 16, cex = 2.0, type = "n", cex.lab = 1.5, cex.axis = 1.2, axes = FALSE)

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

# Add Points & Labels
points(scores(hmwf.bray.D.rda, display = "wa"),
       pch = 16, cex = 2, bg = "gray", col = designDNA$lake.col)

# Add Environmental Vectors
vectors <- scores(hmwf.bray.D.rda, display = "bp")
```

```
row.names(vectors) <- c("DOC", "TP", "TN")
arrows(0, 0, vectors[,1] , vectors[, 2],
       lwd = 2, lty = 1, length = 0.2, col = "red")
text(vectors[,1], vectors[, 2], pos = c(3, 1, 1),
     labels = row.names(vectors), col = "red", cex = 0.8)
axis(side = 3, lwd.ticks=2, cex.axis=  0.8, las = 1, col = "red", lwd = 2.2,
     at = c(-0.5, 0, 0.5, 1, 1.5),
     labels = c(-0.5, 0, 0.5, 1, 1.5), col.axis = "red")
axis(side = 4, lwd.ticks=2, cex.axis = 0.8, las = 1, col = "red", lwd = 2.2,
     at = c(-1, 0, 1),
     labels = c(-1, 0, 1), col.axis = "red")

# Initiate Plot 2
plot(scores(hmwf.bray.R.rda, display = "wa"), xlim = c(-1, 2), ylim = c(-1.5, 2),
     xlab = paste("dbRDA 1 (", rda.R.explainvar1, "%)", sep = ""),
     ylab = paste("dbRDA 2 (", rda.R.explainvar2, "%)", sep = ""),
     pch = 17, cex = 2.0, type = "n", cex.lab = 1.5, cex.axis = 1.2, axes = FALSE)

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

# Add Points & Labels
points(scores(hmwf.bray.R.rda, display = "wa"),
       pch = 17, cex = 2, bg = "gray", col = designRNA$lake.col)

# Add Environmental Vectors
vectors <- scores(hmwf.bray.R.rda, display = "bp")
row.names(vectors) <- c("DOC", "TP", "TN")
arrows(0, 0, vectors[,1] , vectors[, 2],
       lwd = 2, lty = 1, length = 0.2, col = "red")
text(vectors[,1], vectors[, 2], pos = c(3, 1, 1),
     labels = row.names(vectors), col = "red", cex = 0.8)
axis(side = 3, lwd.ticks=2, cex.axis=  0.8, las = 1, col = "red", lwd = 2.2,
     at = c(-0.5, 0, 0.5, 1, 1.5),
     labels = c(-0.5, 0, 0.5, 1, 1.5), col.axis = "red")
axis(side = 4, lwd.ticks=2, cex.axis = 0.8, las = 1, col = "red", lwd = 2.2,
     at = c(-1, 0, 1),
     labels = c(-1, 0, 1), col.axis = "red")

# Initiate Plot 3
plot(scores(hmwf.res.rda, display = "wa"), xlim = c(-1, 2), ylim = c(-1.5, 2),
     xlab = paste("dbRDA 1 (", rda.Res.explainvar1, "%)", sep = ""),
     ylab = paste("dbRDA 2 (", rda.Res.explainvar2, "%)", sep = ""),
     pch = 15, cex = 2.0, type = "n", cex.lab = 1.5, cex.axis = 1.2, axes = FALSE)

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

```
# Add Points & Labels
points(scores(hmwf.res.rda, display = "wa"),
       pch = 15, cex = 2, bg = "gray", col = designRES$lake.col)

# Add Environmental Vectors
vectors <- scores(hmwf.res.rda, display = "bp")
row.names(vectors) <- c("DOC", "TP", "TN")
arrows(0, 0, vectors[,1] , vectors[, 2],
       lwd = 2, lty = 1, length = 0.2, col = "red")
text(vectors[,1], vectors[, 2], pos = c(3, 1, 1),
     labels = row.names(vectors), col = "red", cex = 0.8)
axis(side = 3, lwd.ticks=2, cex.axis=  0.8, las = 1, col = "red", lwd = 2.2,
     at = c(-0.5, 0, 0.5, 1, 1.5),
     labels = c(-0.5, 0, 0.5, 1, 1.5), col.axis = "red")
axis(side = 4, lwd.ticks=2, cex.axis = 0.8, las = 1, col = "red", lwd = 2.2,
     at = c(-1, 0, 1),
     labels = c(-1, 0, 1), col.axis = "red")

plot.new()
par(mar = c(0, 0, 0, 0) + 0.5)
legend("topleft", legend = designRES$Lake, pch = 16, col = designRES$lake.col,
       bty = "n", ncol = 2, cex = 1.5, pt.cex = 1.5, y.intersp = 1.5, inset = c(-0.1, 0, 0, 0))
legend("bottomleft", legend = c("DNA", "RNA", "OM"), pch = c(16, 17, 15),
       col = "gray", bty = "n", cex = 1.5, pt.cex = 1.5, y.intersp = 1.5, inset = c(-0.1, 0, 0, 0.1))

dev.off() # this writes plot to folder
graphics.off() # shuts down open devices
par(opar)
```

Influence of Organic Matter on Community Composition

```
# PCoA of OM
res.mat <- hmwf.bray.res
PcoA.RES <- cmdscale(res.mat, eig = TRUE, k = 4)
var.exp <- sum(round(PcoA.RES$eig[1] / sum(PcoA.RES$eig), 3) * 100,
               round(PcoA.RES$eig[2] / sum(PcoA.RES$eig), 3) * 100,
               round(PcoA.RES$eig[3] / sum(PcoA.RES$eig), 3) * 100,
               round(PcoA.RES$eig[4] / sum(PcoA.RES$eig), 3) * 100)

# 2012 Env Data
env.chem2012 <- as.data.frame(env.chem[grep("2012", row.names(env.chem)), ])

# 2012 Communities Only
OTUsREL2012.D <- OTUsREL[design$Molecule == "DNA" & design$Year == "2012", ]
OTUsREL2012.R <- OTUsREL[design$Molecule == "RNA" & design$Year == "2012", ]

# Calculate Bray-Curtis Distances for Bacteria
Bray.REL2012.D <- vegdist(OTUsREL2012.D, "bray")
Bray.REL2012.R <- vegdist(OTUsREL2012.R, "bray")
```

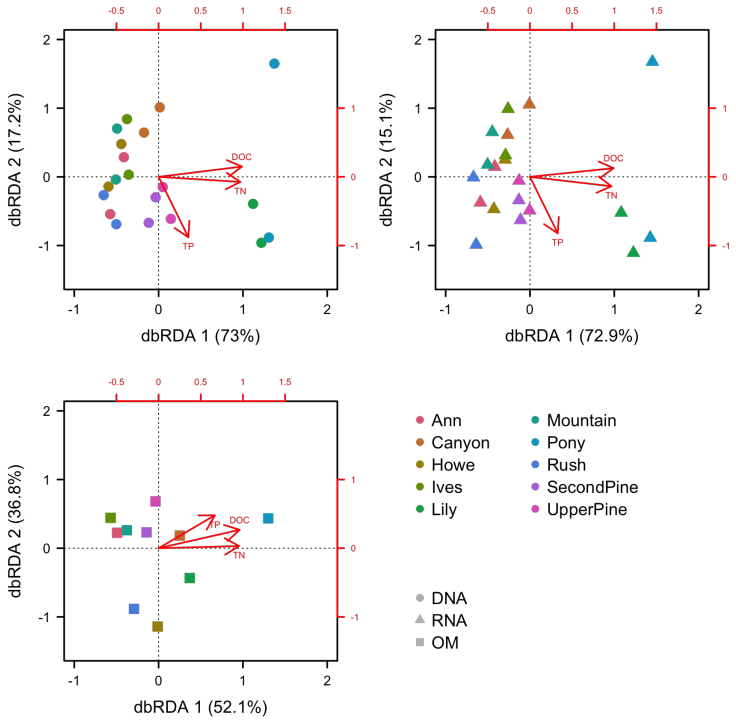


Figure 4: PCoA Supplement

```
# Calculate UniFrac Distances for Bacteria
OTU.tab2012.D <- otu_table(OTUsREL2012.D, taxa_are_rows = F)
OTU.tab2012.R <- otu_table(OTUsREL2012.R, taxa_are_rows = F)
PHY.tree <- phy_tree(hmwf.tree)
phylo.seq2012.D <- phyloseq(OTU.tab2012.D, PHY.tree)
phylo.seq2012.R <- phyloseq(OTU.tab2012.R, PHY.tree)
Uni.REL2012.D <- UniFrac(phylo.seq2012.D, weighted = TRUE, normalized = TRUE)
Uni.REL2012.R <- UniFrac(phylo.seq2012.R, weighted = TRUE, normalized = TRUE)

# Conduct dbRDA
hmwf.bray.D.res.rda <- capscale(Bray.REL2012.D ~ PcoA.RES$points,
                          comm = OTUsREL2012.D, add = T)
hmwf.bray.R.res.rda <- capscale(Bray.REL2012.R ~ PcoA.RES$points,
                          comm = OTUsREL2012.R, add = T)
hmwf.uni.D.res.rda <- capscale(Uni.REL2012.R ~ PcoA.RES$points,
                         comm = OTUsREL2012.D, add = T)
hmwf.uni.R.res.rda <- capscale(Uni.REL2012.R ~ PcoA.RES$points,
                          comm = OTUsREL2012.R, add = T)

# Tests
anova(hmwf.bray.D.res.rda)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
```

```
## Model: capscale(formula = Bray.REL2012.D ~ PcoA.RES$points, comm = OTUsREL2012.D, add = T)
##          Df SumOfSqs      F Pr(>F)
## Model     4  0.88630 1.7945  0.019 *
## Residual  5  0.61738
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RsquareAdj(hmwf.bray.D.res.rda)
```

```
## $r.squared
## [1] 0.5894219
##
## $adj.r.squared
## [1] 0.2609595
```

```
anova(hmwf.bray.R.res.rda)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = Bray.REL2012.R ~ PcoA.RES$points, comm = OTUsREL2012.R, add = T)
##          Df SumOfSqs      F Pr(>F)
## Model     4  1.04816 1.6888  0.016 *
## Residual  5  0.77583
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RsquareAdj(hmwf.bray.R.res.rda)
```

```
## $r.squared
## [1] 0.5746515
##
## $adj.r.squared
## [1] 0.2343727
```

```
anova(hmwf.uni.D.res.rda)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = Uni.REL2012.R ~ PcoA.RES$points, comm = OTUsREL2012.D, add = T)
##          Df SumOfSqs      F Pr(>F)
## Model     4 0.115802 1.5682  0.062 .
## Residual  5 0.092302
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
RsquareAdj(hmwf.uni.D.res.rda)
```

```
## $r.squared
## [1] 0.5564607
##
## $adj.r.squared
## [1] 0.2016292
```

```r
anova(hmwf.uni.R.res.rda)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = Uni.REL2012.R ~ PcoA.RES$points, comm = OTUsREL2012.R, add = T)
##          Df SumOfSqs      F Pr(>F)
## Model     4 0.115802 1.5682  0.037 *
## Residual  5 0.092302
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
RsquareAdj(hmwf.uni.R.res.rda)
```

```
## $r.squared
## [1] 0.5564607
##
## $adj.r.squared
## [1] 0.2016292
```

```r
# Full Model With Nutrients
PCoA.RES <- cmdscale(res.mat, eig = TRUE, k = 1)
env.chem2012 <- as.matrix(env.chem2012)
env.red <- princomp(env.chem2012)$scores[,1]
hmwf.bray.D.res.rda2 <- capscale(Bray.REL2012.D ~ PCoA.RES$points + env.red,
                            comm = OTUsREL2012.D, add = T)
hmwf.bray.R.res.rda2 <- capscale(Bray.REL2012.R ~ PCoA.RES$points + env.red,
                            comm = OTUsREL2012.R, add = T)
hmwf.uni.D.res.rda2 <- capscale(Uni.REL2012.R ~ PCoA.RES$points + env.red,
                           comm = OTUsREL2012.D, add = T)
hmwf.uni.R.res.rda2 <- capscale(Uni.REL2012.R ~ PCoA.RES$points + env.red,
                            comm = OTUsREL2012.R, add = T)

# Tests
anova(hmwf.bray.D.res.rda2)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = Bray.REL2012.D ~ PCoA.RES$points + env.red, comm = OTUsREL2012.D, add = T)
##          Df SumOfSqs      F Pr(>F)
```

```
## Model     2  0.59059 2.2638  0.012 *
## Residual  7  0.91310
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RsquareAdj(hmwf.bray.D.res.rda2)
```

```
## $r.squared
## [1] 0.3927599
##
## $adj.r.squared
## [1] 0.2192627
```

```
anova(hmwf.bray.R.res.rda2)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = Bray.REL2012.R ~ PCoA.RES$points + env.red, comm = OTUsREL2012.R, add = T)
##          Df SumOfSqs      F Pr(>F)
## Model     2  0.69916 2.1755  0.006 **
## Residual  7  1.12483
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RsquareAdj(hmwf.bray.R.res.rda2)
```

```
## $r.squared
## [1] 0.3833126
##
## $adj.r.squared
## [1] 0.2071162
```

```
anova(hmwf.uni.D.res.rda2)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = Uni.REL2012.R ~ PCoA.RES$points + env.red, comm = OTUsREL2012.D, add = T)
##          Df SumOfSqs      F Pr(>F)
## Model     2 0.077697 2.0853  0.008 **
## Residual  7 0.130406
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RsquareAdj(hmwf.uni.D.res.rda2)
```

```
## $r.squared
## [1] 0.3733581
##
## $adj.r.squared
## [1] 0.1943175
```

```
anova(hmwf.uni.R.res.rda2)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = Uni.REL2012.R ~ PCoA.RES$points + env.red, comm = OTUsREL2012.R, add = T)
##          Df SumOfSqs      F Pr(>F)
## Model     2 0.077697 2.0853  0.008 **
## Residual  7 0.130406
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RsquareAdj(hmwf.uni.R.res.rda2)
```

```
## $r.squared
## [1] 0.3733581
##
## $adj.r.squared
## [1] 0.1943175
```

Co-Occurance Analysis

To attempt to understand the relationship between organic matter and community composition, we decided to use co-occurance analaysis to do pair-wise comparisions.

Co-Orrurance Analysis Setup

```
# Subset OTUs for most dominant
OTUsREL.dom <- OTUsREL[,which(colSums(as.matrix(OTUsREL)) > 0.05)]
OTUsREL.dom2012 <- OTUsREL.dom[grep("2012_RNA", rownames(OTUsREL.dom)),]
OTUsREL.dom2012 <- OTUsREL.dom2012[,order(colSums(as.matrix(OTUsREL.dom2012)),
                                          decreasing = T)[1:100]]

# Subset Resources for most dominant
resREL.dom <- resREL[,which(colSums(as.matrix(resREL)) > 0.022)]
resREL.dom <- resREL.dom[,order(colSums(as.matrix(resREL.dom)), decreasing = T)[1:100]]

# Mege OTUs and Resources
ConRes1 <- cbind(as.matrix(OTUsREL.dom2012), as.matrix(resREL.dom))
ConRes2 <- cbind(as.matrix(resREL.dom), as.matrix(OTUsREL.dom2012))
ConRes <- rbind(ConRes1, ConRes2)
```

```
# Calculate 1 - Spearman Correlation Coefficients: Spearman Distance
spear.bac <- spearman.dist(t(OTUsREL.dom2012), abs = FALSE)
spear.res <- spearman.dist(t(as.matrix(resREL.dom)), abs = FALSE)
spear.ConRes <- spearman.dist(t(as.matrix(ConRes)), abs = FALSE, diag = T, upper = T)

spear.bac2 <- spear.bac - 1
spear.bac3 <- spear.bac2
spear.bac3[which(spear.bac2 < 0.5 & spear.bac2 > -0.5)] <- 0

spear.res2 <- spear.res - 1
spear.res3 <- spear.res2
spear.res3[which(spear.res2 < 0.5 & spear.res2 > -0.5)] <- 0

spear.ConRes2 <- spear.ConRes - 1
spear.ConRes3 <- spear.ConRes2
spear.ConRes3[which(spear.ConRes2 < 0.5 & spear.ConRes2 > -0.5)] <- 0
spear.ConRes4 <- as.matrix(spear.ConRes3)[1:100, 101:200]
```

```
## Warning in as.matrix.dist(spear.ConRes3): number of items to replace is not
## a multiple of replacement length
```

```
# Custome Color Palette
jet.colors <- colorRampPalette(c("#00007F", "blue", "#007FFF", "cyan",
                                 "#7FFF7F", "yellow", "#FF7F00", "red",
                                 "#7F0000"))
jet.colors.W <- colorRampPalette(c("#00007F", "blue", "#007FFF", "cyan",
                                 "white", "white", "white", "white",
                                 "yellow", "#FF7F00", "red", "#7F0000"))
```

Co-Occurence Plots

```
png(filename="../figures/Figure4A.png",
    width = 1800, height = 1800, res = 96*2)
par(opar)
heatmap.2(as.matrix(spear.bac3), col = jet.colors.W, distfun = dist,
          dendrogram = "both", na.rm = F, na.color = "white", trace = "none",
          density.info = "none", key.xlab = "Correlation", key.title = "",
          cexRow = 0.4, cexCol = 0.4, main = "Species-Species", lhei = c(1.5, 8))
dev.off() # this writes plot to folder
graphics.off() # shuts down open devices
par(opar)
```

```
png(filename="../figures/Figure4B.png",
    width = 1800, height = 1800, res = 96*2)
par(opar)
heatmap.2(as.matrix(spear.res3), col = jet.colors.W, distfun = dist,
          dendrogram = "both", na.rm = F, na.color = "white", trace = "none",
          density.info = "none", key.xlab = "Correlation", key.title = "",
          cexRow = 0.4, cexCol = 0.4, main = "Resource-Resource", lhei = c(1.5, 8))
dev.off() # this writes plot to folder
```

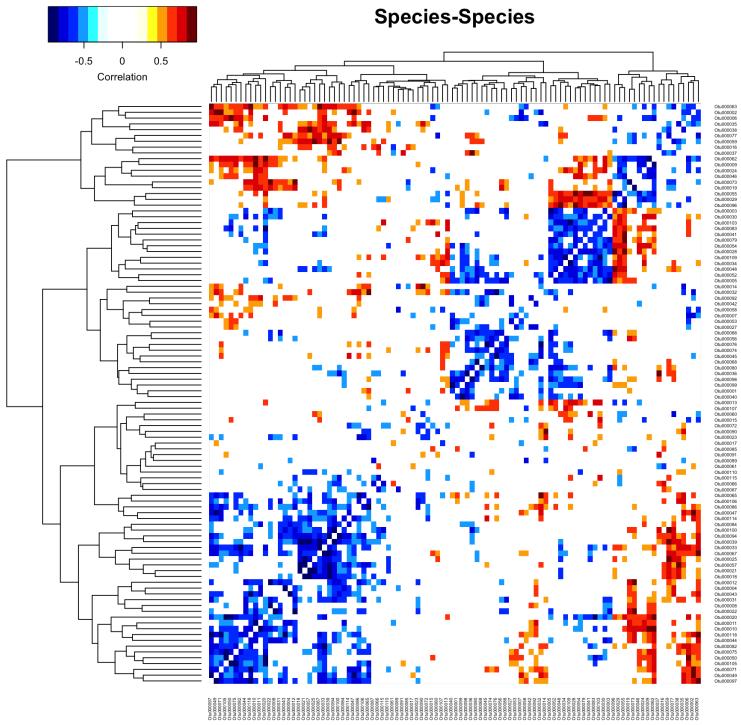


Figure 5: PCoA Supplement

```
graphics.off() # shuts down open devices
par(opar)
```

```
png(filename="../figures/Figure4C.png",
    width = 1800, height = 1800, res = 96*2)
par(opar)
heatmap.2(as.matrix(spear.ConRes4), col = jet.colors.W, distfun = dist,
          dendrogram = "both", na.rm = F, na.color = "white", trace = "none",
          density.info = "none", key.xlab = "Correlation", key.title = "",
          cexCol = 0.4, cexRow = 0.4, main = "Species-Resource", lhei = c(2, 8))
dev.off() # this writes plot to folder
graphics.off() # shuts down open devices
par(opar)
```

Co-Occurence Summary

```
proC <- length(which(spear.bac3 > 0))
antC <- length(which(spear.bac3 < 0))

(proC + antC) / length(spear.bac3)
```

```
## [1] 0.240404
```

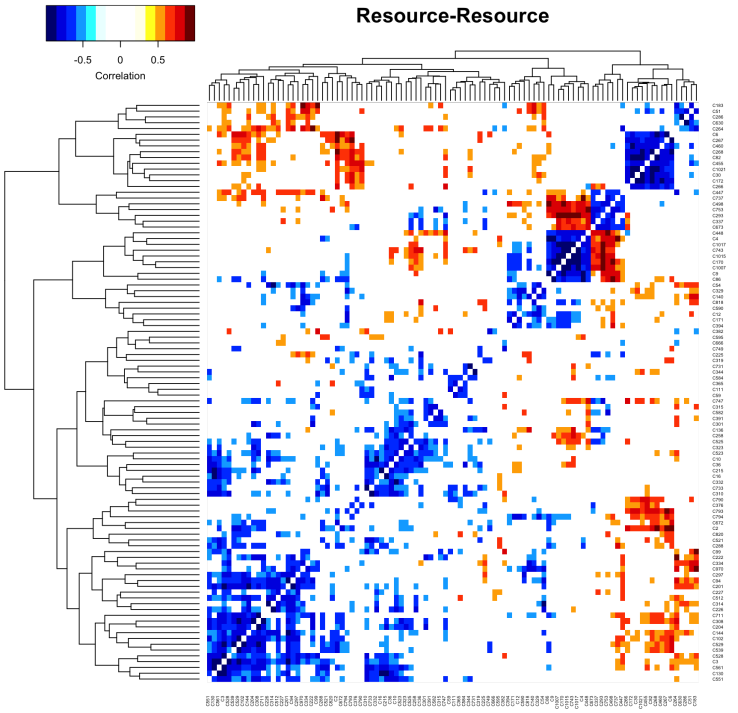


Figure 6: PCoA Supplement

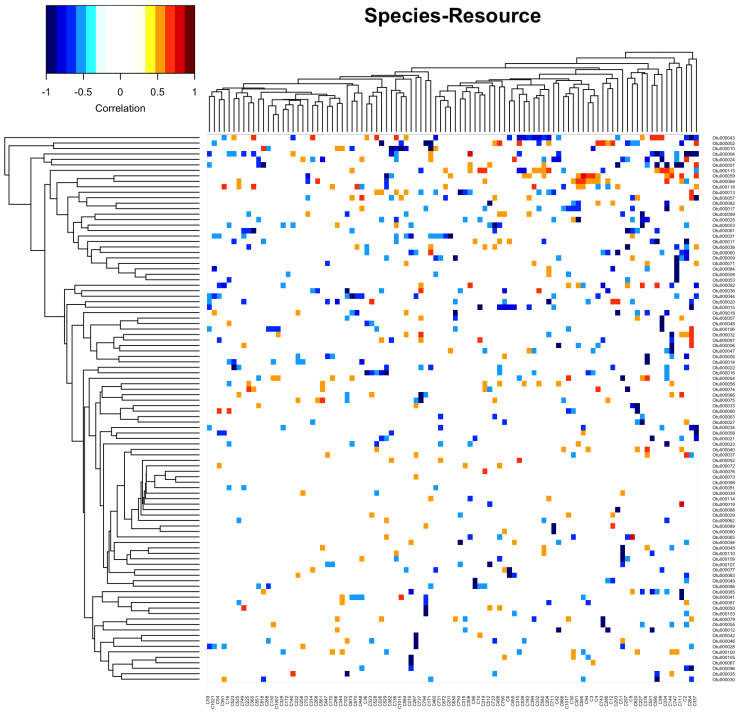


Figure 7: PCoA Supplement

```
proC <- colSums(as.matrix(spear.bac3) > 0)
conC <- colSums(as.matrix(spear.bac3) < 0)

proR <- length(which(spear.res3 > 0))
antR <- length(which(spear.res3 < 0))

prosR <- colSums(as.matrix(spear.res3) > 0)
consR <- colSums(as.matrix(spear.res3) < 0)
(proR + antR) / length(spear.res3)
```

```
## [1] 0.2385859
```

```
proCR <- sum(spear.ConRes4 > 0)
conCR <- sum(spear.ConRes4 < 0)
(proCR + conCR) / sum(spear.ConRes4 != "NA")
```

```
## [1] 0.0747
```

```
specC <- rowSums(as.matrix(spear.ConRes4) < 0)
genC <- rowSums(as.matrix(spear.ConRes4) > 0)
```

Graph Analysis

```
png(filename="../figures/Figure5.png",
    width = 1800, height = 900, res = 96*2)
par(opar)

# Reorganize Data
str(spear.ConRes4)
bac <- rep(rownames(spear.ConRes4), 100)
res <- rep(colnames(spear.ConRes4), each = 100)
int <- as.numeric(spear.ConRes4[, 1])
for (i in 2:100){
  int = append(int, as.numeric(spear.ConRes4[, i]))
}
bac2 <- bac
bac2 <- gsub("Otu000", "O", bac2)

# Connectedness

Mode <- function(x) {
  ux <- unique(x)
  ux[which.max(tabulate(match(x, ux)))]
}


conn <- data.frame(bac2, res, int)
connC <- conn[conn$int < -0.5, ]
dim(connC)[1]
length(connC$bac2)
```

```r
length(unique(connC$bac2))
length(unique(connC$res))
Cbac <- unique(connC$bac2)
Cbac_c <- rep(NA, length(Cbac))
for (i in 1:length(Cbac)){
  Cbac_c[i] <- sum(connC$bac2 == Cbac[i])
}
mean(Cbac_c)
sum(Cbac_c > 3)
Cres <- unique(connC$res)
Cres_c <- rep(NA, length(Cres))
for (i in 1:length(Cres)){
  Cres_c[i] <- sum(connC$res == Cres[i])
}
mean(Cres_c)
sum(Cres_c > 3)
connP <- conn[conn$int > 0.5, ]
Pbac <- unique(connP$bac2)
Pbac_c <- rep(NA, length(Pbac))
for (i in 1:length(Pbac)){
  Pbac_c[i] <- sum(connP$bac2 == Pbac[i])
}
mean(Pbac_c)
Pres <- unique(connP$res)
Pres_c <- rep(NA, length(Pres))
for (i in 1:length(Pbac_c)){
  Pres_c[i] <- sum(connP$res == Pres[i])
}
mean(Pres_c)

graph.list.full <- data.frame(bac2, res, int)
graph.list <- graph.list.full[graph.list.full$int > 0.5, ]
hmwf.network <- graph.data.frame(graph.list, directed=F)
hmwf.network <- simplify(hmwf.network)
bets <- betweenness(hmwf.network)
mean(bets[grep("O", names(bets))])
mean(bets[grep("C", names(bets))])

graph.list <- graph.list.full[graph.list.full$int < -0.5, ]
hmwf.network <- graph.data.frame(graph.list, directed=F)
hmwf.network <- simplify(hmwf.network)
bets <- betweenness(hmwf.network)
mean(bets[grep("O", names(bets))])
mean(bets[grep("C", names(bets))])


# Plot Settings
layout(matrix(1:2, ncol = 2))
par(mar = c(1, 1, 2, 1))

# Positive Network
graph.list.full <- data.frame(bac2, res, int)
graph.list <- graph.list.full[graph.list.full$int > 0.65, ]
```

```
hmwf.network <- graph.data.frame(graph.list, directed=F)
hmwf.network <- simplify(hmwf.network)
bets <- betweenness(hmwf.network)
mean(bets[grep("O", names(bets))])
mean(bets[grep("C", names(bets))])
V(hmwf.network)$color <- V(hmwf.network)$name
V(hmwf.network)$color[grepl("O",V(hmwf.network)$color)] <- "cornflowerblue"
V(hmwf.network)$color[grepl("C",V(hmwf.network)$color)] <- "wheat3"
plot(hmwf.network, layout=layout.fruchterman.reingold,
     main='Production Network', vertex.label.dist=0,
     vertex.label.color='black', vertex.label.cex=0.5)

# Negative Network
graph.list <- graph.list.full[graph.list.full$int < -0.999, ]
hmwf.network <- graph.data.frame(graph.list, directed=F)
bets <- betweenness(hmwf.network)
mean(bets[grep("O", names(bets))])
mean(bets[grep("C", names(bets))])
V(hmwf.network)$color <- V(hmwf.network)$name
V(hmwf.network)$color[grepl("O",V(hmwf.network)$color)] <- "cornflowerblue"
V(hmwf.network)$color[grepl("C",V(hmwf.network)$color)] <- "wheat3"
plot(hmwf.network, layout=layout.fruchterman.reingold,
     main='Consumption Network', vertex.label.dist=0,
     vertex.label.color='black', vertex.label.cex=0.5)
dev.off() # this writes plot to folder
graphics.off() # shuts down open devices
par(opar)
```

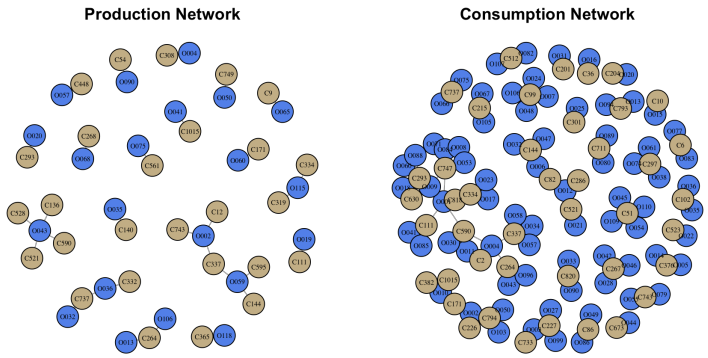


Figure 8: Networks

Species Distributions and Generalism

```
lake.yr <- paste(design$Lake[design$Molecule == "DNA"],
                 design$Year[design$Molecule == "DNA"], sep = "")

total <- OTUs[design$Molecule == "DNA", ]
row.names(total) <- lake.yr
```

```
totalPA <- (total > 0) * 1

active.rna <- OTUs[design$Molecule == "RNA", ]
row.names(active.rna) <- lake.yr
activePA <- (active.rna > 0) * 1
active <- total * (activePA)
row.names(active) <- lake.yr

activePA.2 <- (active.rna > (total * 0.25)) * 1
active.2 <- total * (activePA.2)
row.names(active.2) <- lake.yr

per <- c(0.001, 0.005, 0.01, 0.02, 0.03, 0.04, 0.05, 0.075, 0.10)
per.act <- matrix(NA, dim(total)[1], length(per))
per.actT <- matrix(NA, dim(total)[1], length(per))
colnames(per.act) <- per
colnames(per.actT) <- per

for (i in 1:length(per)){
  activePA.temp <- (active.rna > (total * per[i])) * 1
  active.temp <- total * activePA.temp
  per.act[,i] <- rowSums(active.temp) / rowSums(total)
  per.actT[,i] <- rowSums(activePA.temp) / rowSums(totalPA)
}

mean.per.act <- colMeans(per.act)
se.per.act <- apply(per.act, 2, se)

mean.per.actT <- colMeans(per.actT)
se.per.actT <- apply(per.actT, 2, se)
```

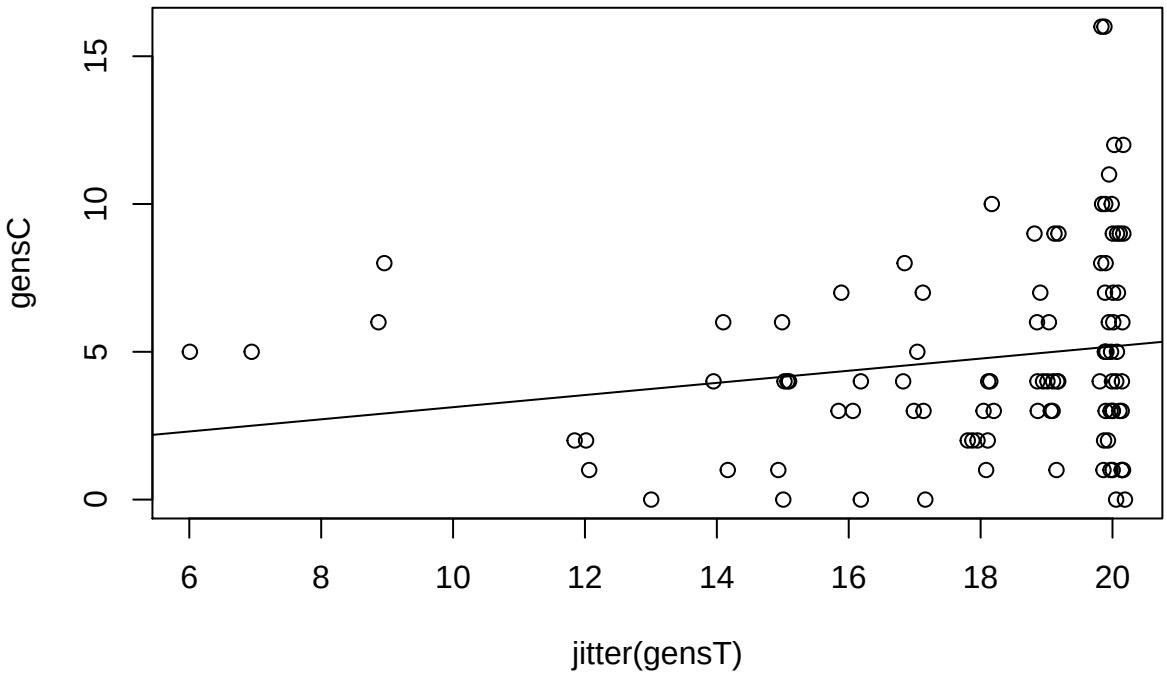
Supplemental Figure for Active Percent Cutoff

```
png(filename="../figures/Supp2.png",
    width = 900, height = 900, res = 96*2)
par(opar)
par(mar = c(4.5,5,1,1) + 0.1)
plot(mean.per.act ~ per, las = 1, type = "n",
     ylim = c(0.875, 0.99),
     xaxt = "n", yaxt = "n", xlab = "", ylab = "")

arrows(x0 = per, y0 = mean.per.act, y1 = mean.per.act + se.per.act, angle= 90, length = 0.1, lwd = 2)
arrows(x0 = per, y0 = mean.per.act, y1 = mean.per.act - se.per.act, angle= 90, length = 0.1, lwd = 2)
points(per, mean.per.act, pch = 15, col = "firebrick", cex = 1.25)

axis(side=1, lwd.ticks = 2, tck=-0.02, labels = T, cex.axis = 1)
axis(side=1, lwd.ticks = 2, tck=0.01, labels = F, cex.axis = 1)
axis(side=2, lwd.ticks = 2, labels = T, cex.axis = 1, las = 1,
     at = c(0.89, 0.92, 0.95, 0.98))
axis(side=2, lwd.ticks = 2, tck=0.01, labels = F, cex.axis = 1, at = c(0.89, 0.92, 0.95, 0.98))
axis(side=3, lwd.ticks = 2, tck=-0.02, labels = F, cex.axis = 1)
```

```
axis(side=3, lwd.ticks = 2, tck=0.01, labels = F, cex.axis = 1)
axis(side=4, lwd.ticks = 2, tck=-0.02, labels = F, cex.axis = 1, at = c(0.89, 0.92, 0.95, 0.98))
axis(side=4, lwd.ticks = 2, tck=0.01, labels = F, cex.axis = 1, at = c(0.89, 0.92, 0.95, 0.98))
mtext(side = 2, "Percent Active", line = 3, cex = 1.5)
mtext(side = 1, "Cutoff", line = 2.5, cex = 1.5)
box(lwd = 2)
dev.off() # this writes plot to folder
graphics.off() # shuts down open devices
par(opar)
```

Generalist Taxa

```
# Generalists by Site
gens <- data.frame(matrix(NA, 21, 3))
colnames(gens) <- c("sites", "taxaA", "taxaT")
gens$sites <- c(1:21)

for (i in 1:21){
  gens$taxaA[i] <- sum(colSums(activePA) == i)
  gens$taxaT[i] <- sum(colSums(totalPA) == i)
}

# Generalists by Consumption
consum <- as.matrix(spear.ConRes4) < -0.5
gensC <- rowSums(abs(consum))
chem.gen <- names(which(genC > 0))

activePA2 <- activePA[, colnames(activePA) %in% names(gensC)]
totalPA2 <- totalPA[, colnames(totalPA) %in% names(gensC)]
gensA <- colSums(activePA2)
gensT <- colSums(totalPA2)
gen.mod <- lm(gensC ~ gensT)
plot(gensC ~ jitter(gensT))
abline(gen.mod)
```




```
spatial.gen <- names(which(colSums(activePA) >= 18))
length(spatial.gen)
```

```
## [1] 129
```

```
length(intersect(chem.gen, spatial.gen))
```

```
## [1] 58
```

```
length(intersect(chem.gen, spatial.gen)) / length(spatial.gen)
```

```
## [1] 0.4496124
```

Generalists Plot

```
png(filename="../figures/Supp3.png",
    width = 900, height = 900, res = 96*2)
par(opar)
# Define Plot Parameters
par(mar = c(5, 6, 1, 1) + 0.1)
plot(log10(gens$taxaT) ~ gens$sites,
     yaxt = "n", xaxt = "n", ylab = "", xlab = "", las = 1,
     pch = 16, col = "cornflowerblue", xlim = c(0, 22), ylim = c(1, 3.8))
points(log10(gens$taxaA) ~ gens$sites, pch = 17, col = "darkolivegreen3")

mtext(side = 1, "Number of Sites", line = 3, cex = 1.5)
mtext(side = 2, "Number of Taxa", line = 3.5, cex = 1.5)
axis(side = 1, lwd = 2, labels = T)
```

```
axis(side = 2, lwd = 2, at = c(1, 2, 3, 3.7), labels = c(10, 100, 1000, 5000), las = 1)
axis(side = 3, lwd =2, tck = -0.02, labels = F)
axis(side = 4, lwd =2, tck = -0.02, labels = F, at = c(1, 2, 3, 3.7))
axis(side = 1, lwd =2, tck = 0.02, labels = F)
axis(side = 2, lwd =2, tck = 0.02, labels = F, at = c(1, 2, 3, 3.7))
axis(side = 3, lwd =2, tck = 0.02, labels = F)
axis(side = 4, lwd =2, tck = 0.02, labels = F, at = c(1, 2, 3, 3.7))
legend("topright", pch = c(16, 17), c("Total", "Active"), bty = "n",
       col = c("cornflowerblue", "darkolivegreen3"))
box(lwd = 2)
dev.off() # this writes plot to folder
graphics.off() # shuts doum open devices
par(opar)
```

```
inactive <- total * (1 - activePA)
inactivePA <- (inactive > 0) * 1


# Inactive Taxa
inactivePA <- totalPA - activePA
inactive <- total * inactivePA
```

Species Abundance Distribution for Active and Inactive

```
png(filename="../figures/Supp4.png",
    width = 1200, height = 1000, res = 96*2)
par(opar)

par(mar = c(4, 5, 1, 1) + 0.1)
plot(sort(log10(colSums(active)), decreasing = T), col = "darkolivegreen3",
     las = 1, xaxt="n", xlab = "", yaxt="n", ylab = "", xlim = c(0, 7000))
points(sort(log10(colSums(inactive)), decreasing = T), col = "cornflowerblue")
axis(side=1, lwd.ticks = 2, tck=-0.02, labels = T, cex.axis = 1, at = c(0, 1500, 3000, 4500, 6000))
axis(side=1, lwd.ticks = 2, tck=0.01, labels = F, cex.axis = 1, at = c(0, 1500, 3000, 4500, 6000))
axis(side=2, lwd.ticks = 2, labels = c("10", "1000", "100000"), cex.axis = 1, las = 1,
     at = c(1, 3, 5))
axis(side=2, lwd.ticks = 2, tck=0.01, labels = F, cex.axis = 1, at = c(1, 3, 5))
axis(side=3, lwd.ticks = 2, tck=-0.02, labels = F, cex.axis = 1, at = c(0, 1500, 3000, 4500, 6000))
axis(side=3, lwd.ticks = 2, tck=0.01, labels = F, cex.axis = 1, at = c(0, 1500, 3000, 4500, 6000))
axis(side=4, lwd.ticks = 2, tck=-0.02, labels = F, cex.axis = 1, at = c(1, 3, 5))
axis(side=4, lwd.ticks = 2, tck=0.01, labels = F, cex.axis = 1, at = c(1, 3, 5))
mtext(side = 2, "Abundance", line = 3.5, cex = 1.5)
mtext(side = 1, "Index", line = 2.5, cex = 1)
legend("topright", c("Inactive", "Active"), fill = c("cornflowerblue", "darkolivegreen3"), bty = "n")
box(lwd = 2)
dev.off() # this writes plot to folder
graphics.off() # shuts doum open devices
par(opar)
```

Data Checks

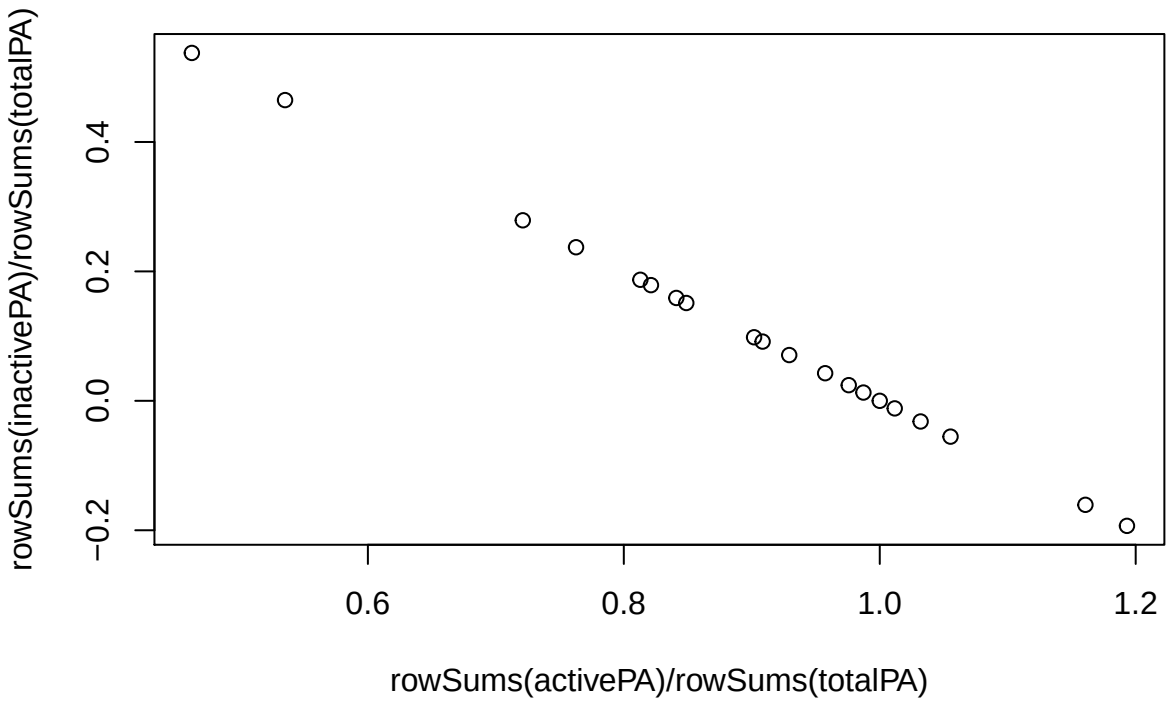
```
pro.act <- rowSums(active)/rowSums(total)

rowSums(inactivePA)/rowSums(totalPA)
```

```
##        Ann2011        Ann2012     Canyon2011     Canyon2012       Howe2011
##     0.17879162     0.09821429    -0.05531609     0.09153318     0.53757225
##       Howe2012       Ives2011       Ives2012       Lily2011       Lily2012
##     0.18707811     0.27895753     0.01277584     0.46472564    -0.03192796
##   Mountain2011   Mountain2012       Pony2011       Pony2012       Rush2011
##     0.04266467     0.23732904     0.15118062    -0.16071429     0.02421925
##       Rush2012 SecondPine2011 SecondPine2012  UpperPine2011  UpperPine2012
##     0.15896885    -0.19319728     0.00000000    -0.01169916     0.07073955
```

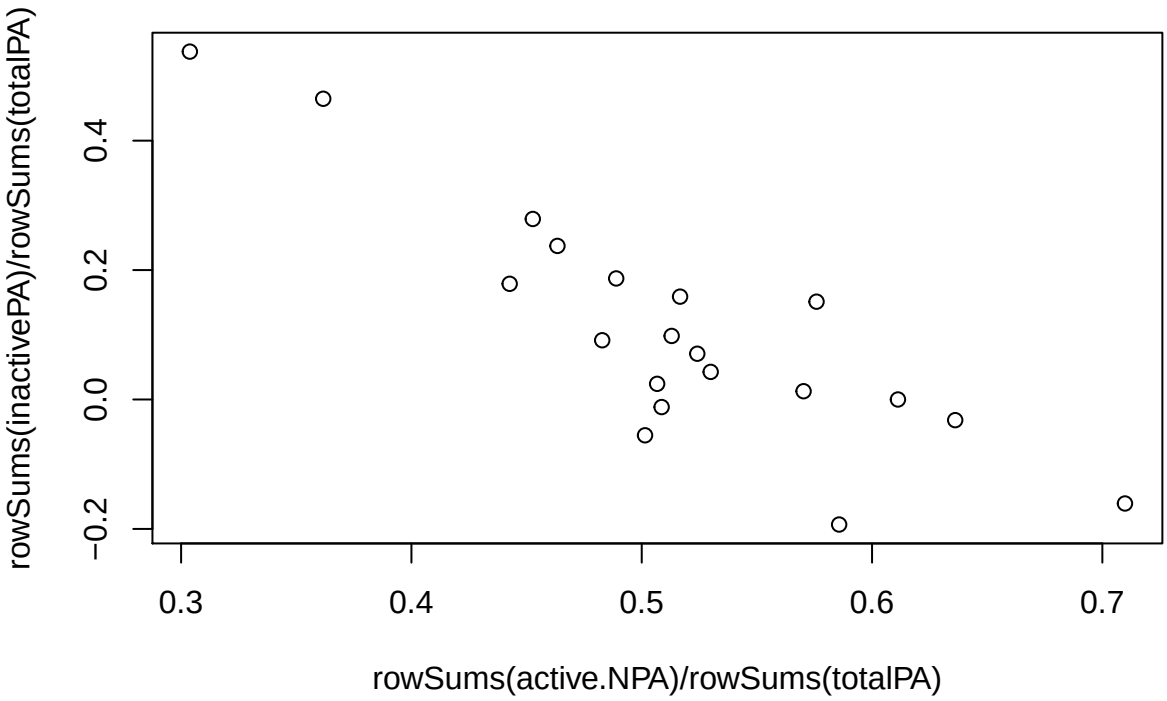
```
rowSums(activePA)/rowSums(totalPA)
```

```
##        Ann2011        Ann2012     Canyon2011     Canyon2012       Howe2011
##      0.8212084      0.9017857      1.0553161      0.9084668      0.4624277
##       Howe2012       Ives2011       Ives2012       Lily2011       Lily2012
##      0.8129219      0.7210425      0.9872242      0.5352744      1.0319280
##   Mountain2011   Mountain2012       Pony2011       Pony2012       Rush2011
##      0.9573353      0.7626710      0.8488194      1.1607143      0.9757808
##       Rush2012 SecondPine2011 SecondPine2012  UpperPine2011  UpperPine2012
##      0.8410311      1.1931973      1.0000000      1.0116992      0.9292605
```

```
plot(rowSums(activePA)/rowSums(totalPA), rowSums(inactivePA)/rowSums(totalPA))
```



```
inactive <- total * inactivePA
active.N <- total * (1-inactivePA)
active.NPA <- (active.N > 0 ) * 1

plot(rowSums(active.NPA)/rowSums(totalPA), rowSums(inactivePA)/rowSums(totalPA))
```



Microbial Function

Phosphorus Contributes to Activity

```
resp <- read.delim("../data/Respiration.txt")
chla <- read.delim("../data/ChlorophyllA.txt")
```

Table S1: Ecosystem and Microbial Processes

```
eco <- data.frame(matrix(NA, 10, 4))
row.names(eco) <- levels(chla$Lake)
colnames(eco) <- c("Chl11", "Chl12", "Resp11", "Resp12")
for (i in row.names(eco)){
  eco[i, 1] <- round(mean(chla[chla$Lake == i & chla$Year == "2011", 3]), 2)
  eco[i, 2] <- round(mean(chla[chla$Lake == i & chla$Year == "2012", 3]), 2)
  eco[i, 3] <- round(mean(resp[resp$sample == i & resp$year == "2011", 4]), 3)
  eco[i, 4] <- round(mean(resp[resp$sample == i & resp$year == "2012", 4]), 3)
}
```

```
addtorow <- list()
addtorow$pos <- list(0, 0, 0)
addtorow$command <- c(" & \\multicolumn{2}{c}{Chl $\\emph{a}$} &
                      \\multicolumn{2}{c}{Resp.} \\\\\n",
                      "Lake & \\multicolumn{2}{c}{($\\mu$g L$^{-1}$)} &
                      \\multicolumn{2}{c}{($\\mu$M Hr$^{-1}$)} \\\\\n",
                      " & 2011 & 2012 & 2011 & 2012 \\\\\n")
eco.tab <- xtable(eco)
align(eco.tab) <- c("c ","r ", "r ", "r ", "r ")
print(eco.tab, add.to.row = addtorow, include.colnames = FALSE,
      type= "latex", file="../tables/TableS1.tex",
      hline.after = c(-1, -1, 0, nrow(eco.tab)))
print(eco.tab, add.to.row = addtorow, include.colnames = FALSE,
      comment = FALSE, hline.after = c(-1, -1, 0, nrow(eco.tab)))
```

Lake	Chl *a* (μg L^{-1})		Resp. (μM Hr^{-1})	
	2011	2012	2011	2012
Ann	1.31	1.25	1.96	1.26
Canyon	3.70	1.63	1.78	1.32
Howe	0.75	1.85	1.48	0.97
Ives	2.03	1.39	1.42	0.80
Lily	5.77	3.55	2.06	0.94
Mountain	1.80	2.14	1.91	1.42
Pony	24.58	16.35	3.05	1.69
Rush	0.65	1.23	1.75	1.22
SecondPine	2.13	3.76	1.46	1.17
UpperPine	2.14	8.55	1.73	1.26

Phosphorus Contributes to Microbial Activity

```
png(filename="../figures/Figure6.png",
    width = 1800, height = 1000, res = 96*2)
par(opar)
layout(matrix(c(1:2), ncol = 2))
par(mar=c(2, 5, 1, 1) + 0.1, oma = c(4, 1, 1, 1))

# TP and Active Taxa
plot(rowSums(active.NPA)/rowSums(totalPA) ~ log10(nuts$TP[c(order(nuts$sample))]),
     xlab = "",
     ylab = "",
     las = 1, cex.lab = 1.5, xlim = c(0, 1.8), ylim = c(0.25, 0.75), xaxt = "n",
     pch = 15, col = "darkorchid4", cex = 1.5)
axis(side = 1, lwd = 2, at = c(0, 0.7, 1.7), labels = c(0, 5, 50))
axis(side = 2, lwd = 2, labels = F)
axis(side = 3, lwd = 2, tck = 0.02, labels = F, at = c(0, 0.7, 1.7))
axis(side = 4, lwd = 2, tck = 0.02, labels = F)
axis(side = 1, lwd = 2, tck = -0.02, labels = F, at = c(0, 0.7, 1.7))
axis(side = 2, lwd = 2, tck = -0.02, labels = F)
axis(side = 3, lwd = 2, tck = -0.02, labels = F, at = c(0, 0.7, 1.7))
```

```
axis(side = 4, lwd = 2, tck = -0.02, labels = F)
mtext("Proportion of Active Taxa", side = 2, line = 3, cex = 1.5)
box(lwd = 2)
pcoa.res <- cmdscale(res.mat, eig = TRUE, k = 2)

phos <- lm(rowSums(active.NPA)/rowSums(totalPA) ~ log10(nuts$TP[c(order(nuts$sample))]))
nitro <- lm(rowSums(active.NPA)/rowSums(totalPA) ~ log10(nuts$TN[c(order(nuts$sample))])) # NS
carb <- lm(rowSums(active.NPA)/rowSums(totalPA) ~ log10(nuts$DOC[c(order(nuts$sample))])) # NS
res1 <- lm(rowSums(active.NPA)/rowSums(totalPA) ~ rep(pcoa.res$points[, 1], each = 2)) # NS
res2 <- lm(rowSums(active.NPA)/rowSums(totalPA) ~ rep(pcoa.res$points[, 2], each = 2)) # NS
mod1 <- summary(phos)
abline(phos, lty = 2, lwd = 4, col = "darkred")
Rsq <- round(mod1$r.squared, 2)
Pv <- round(mod1$coefficients[2,4], 3)
text(1.5, 0.35, bquote(italic(R)^2 == .(format(Rsq, digits = 3))))
text(1.5, 0.3, bquote(italic(p) == .(format(Pv, digits = 3))))

# TP and Respiration
bac.resp <- c(eco$Resp11, eco$Resp12)
outl <- nuts$TP[17]
nuts$TP[17] <- NA

phos <- lm(log(bac.resp) ~ log10(nuts$TP))
nitro <- lm(log(bac.resp) ~ log10(nuts$TN)) # NS
carb <- lm(log(bac.resp) ~ log10(nuts$DOC)) # NS
res1 <- lm(log(bac.resp) ~ rep(pcoa.res$points[, 1], 2)) # NS
res2 <- lm(log(bac.resp) ~ rep(pcoa.res$points[, 2], 2)) # NS

plot(log(bac.resp) ~ log10(nuts$TP),
     xlab = "",
     ylab = "",
     las = 1, cex.lab = 1.5, xlim = c(0, 1.8), ylim = c(-0.3, 1.2), xaxt = "n", yaxt = "n",
     pch = 15, col = "darkorchid4", cex = 1.5)
points(x = log10(outl), y = log10(bac.resp)[17], pch = 15, col = "darkorchid4", cex = 1.5)
axis(side = 1, lwd = 2, at = c(0, 0.7, 1.7), labels = c(0, 5, 50))
axis(side = 2, lwd = 2, at = c(0, 0.77, 1.1), labels = c(0, 6, 12), las = 1)
axis(side = 3, lwd = 2, tck = 0.02, labels = F, at = c(0, 0.7, 1.7))
axis(side = 4, lwd = 2, at = c(0, 0.77, 1.1), tck = 0.02, labels = F)
axis(side = 1, lwd = 2, tck = -0.02, labels = F, at = c(0, 0.7, 1.7))
axis(side = 2, lwd = 2, at = c(0, 0.77, 1.1), tck = -0.02, labels = F)
axis(side = 3, lwd = 2, tck = -0.02, labels = F, at = c(0, 0.7, 1.7))
axis(side = 4, lwd = 2, at = c(0, 0.77, 1.1), tck = -0.02, labels = F)
mtext(expression(paste("Respiration (", mu , "M C L"^-1, ")")), side = 2.5, line = 2, cex = 1.5)
box(lwd = 2)

mod1 <- summary(phos)
abline(phos, lty = 2, lwd = 4, col = "darkred")
Rsq <- round(mod1$r.squared, 2)
Pv <- round(mod1$coefficients[2,4], 3)
text(1.5, 1, bquote(italic(R)^2 == .(format(Rsq, digits = 3))))
text(1.5, 0.85, bquote(italic(p) == .(format(Pv, digits = 3))))

mtext(expression(paste("Total Phosphorus (", mu , "g L"^-1, ")")), side = 1, outer = T, line = 1.5, cex
```

```
dev.off() # this writes plot to folder
graphics.off() # shuts down open devices
par(opar)
```

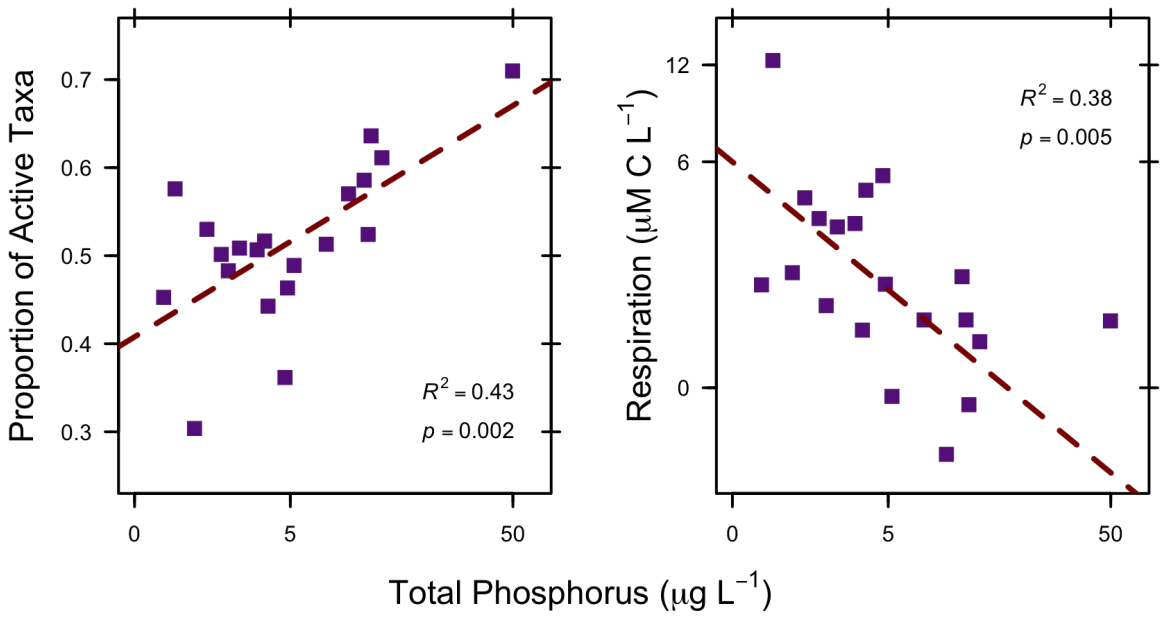


Figure 9: Phosphorus Contributes to Activity

Hypothesis that resource diversity influences consumer diversity

```
alpha.div2012 <- alpha.div[alpha.div$Year == "2012" & alpha.div$Molecule == "RNA", c(1, 4:6)]
rownames(alpha.div2012) <- alpha.div2012[, 1]
alpha.div2012 <- alpha.div2012[, -1]
rownames(res.div) == rownames(alpha.div2012)
```

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

```
cor(res.div, alpha.div2012)
```

```
##                 S.obs     simpsE       shan
## S.res      -0.79777446  0.8464796 0.09607441
## res.simpsE  0.08866852 -0.1099914 0.15291370
## res.shan2   0.14004230 -0.1763717 0.09211756
```

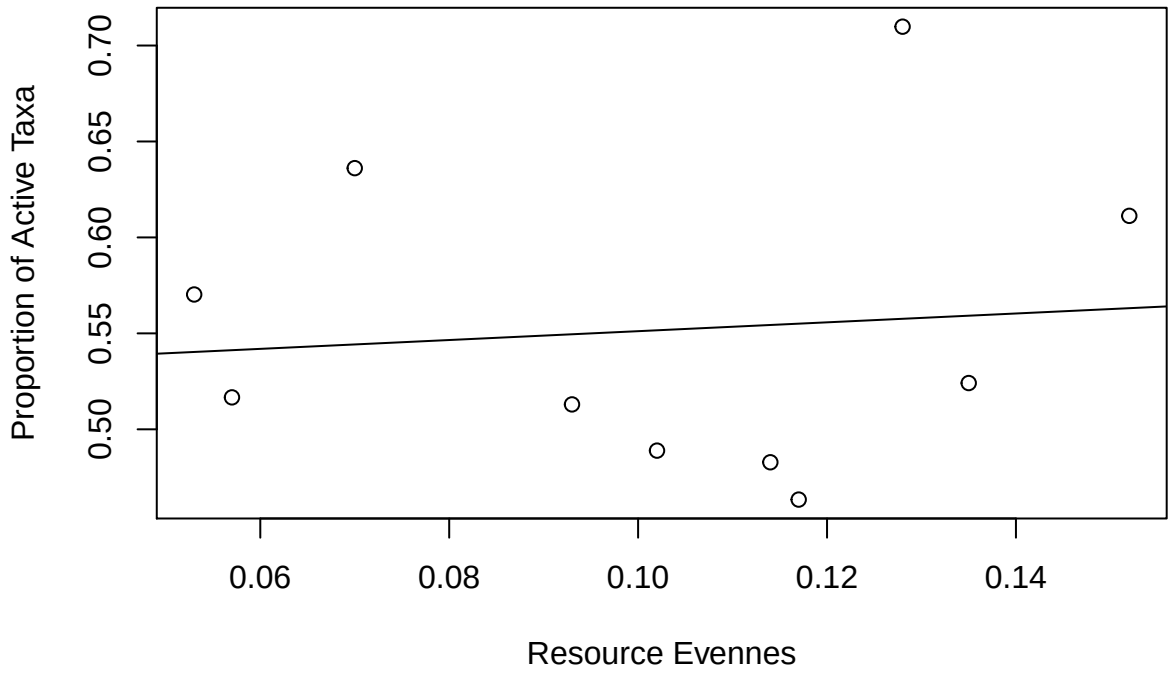
```
rich.mod1 <- lm(alpha.div2012$S.obs ~ res.div$S.res)
rich.mod2 <- lm(alpha.div2012$S.obs ~ res.div$res.simpsE)
summary(rich.mod1)
```

```
##
## Call:
## lm(formula = alpha.div2012$S.obs ~ res.div$S.res)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1110.98  -145.15    50.12   292.89   849.90
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   37226.68    9585.83   3.884  0.00465 **
## res.div$S.res   -64.65      17.27  -3.742  0.00569 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 609.2 on 8 degrees of freedom
## Multiple R-squared:  0.6364, Adjusted R-squared:  0.591
## F-statistic:    14 on 1 and 8 DF,  p-value: 0.005687
```

```
summary(rich.mod2)
```

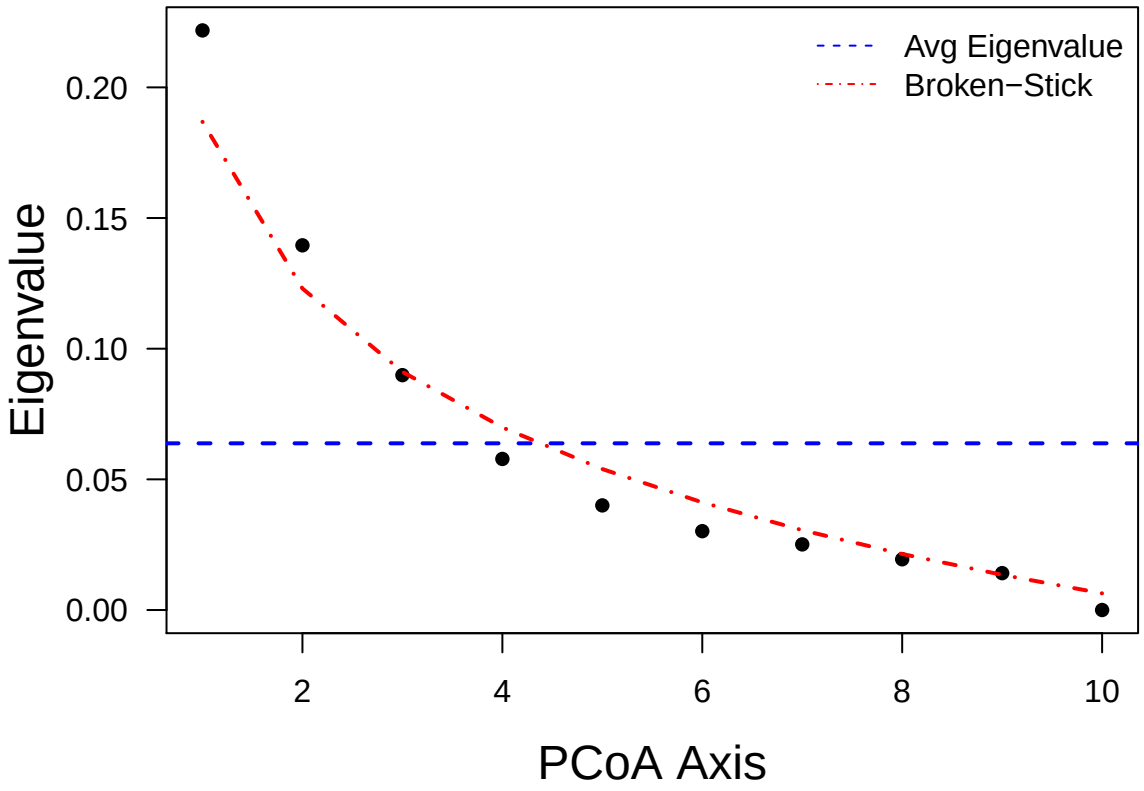
```
##
## Call:
## lm(formula = alpha.div2012$S.obs ~ res.div$res.simpsE)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -632.2 -504.3 -418.8 -206.5 2214.1
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)             1104       1068   1.034    0.331
## res.div$res.simpsE      2513       9982   0.252    0.808
##
## Residual standard error: 1006 on 8 degrees of freedom
## Multiple R-squared:  0.007862,   Adjusted R-squared:  -0.1162
## F-statistic: 0.0634 on 1 and 8 DF,  p-value: 0.8076
```

```
plot((rowSums(active.NPA)/rowSums(totalPA))[seq(2, 20, by = 2)] ~ res.div$res.simpsE,
     xlab = "Resource Evennes", ylab = "Proportion of Active Taxa")
res.ev <- lm((rowSums(active.NPA)/rowSums(totalPA))[seq(2, 20, by = 2)] ~ res.div$res.simpsE)
abline(res.ev)
```



Between site comparisions of resources

```
# Calculate Bray-Curtis
res.db <- vegdist(resREL, method = "bray")

res.pcoa <- cmdscale(res.db, eig = TRUE, k = 3)
explainvar1 <- round(res.pcoa$eig[1] / sum(res.pcoa$eig), 3) * 100
explainvar2 <- round(res.pcoa$eig[2] / sum(res.pcoa$eig), 3) * 100
explainvar3 <- round(res.pcoa$eig[3] / sum(res.pcoa$eig), 3) * 100
sum.eig <- sum(explainvar1, explainvar2, explainvar3)

# Define Plot Parameters
par(mar = c(5, 5, 1, 2) + 0.1)

# Plot Eigenvalues
plot(res.pcoa$eig, xlab = "PCoA Axis", ylab = "Eigenvalue",
     las = 1, cex.lab = 1.5, pch = 16)

# Add Expectation based on Kaiser-Guttman criterion and Broken Stick Model
abline(h = mean(res.pcoa$eig), lty = 2, lwd = 2, col = "blue")
b.stick <- bstick(10, sum(res.pcoa$eig))
lines(1:10, b.stick, type = "l", lty = 4, lwd = 2, col = "red")

# Add Legend
legend("topright", legend = c("Avg Eigenvalue", "Broken-Stick"),
       lty = c(2, 4), bty = "n", col = c("blue", "red"))
```



Resource Explainations of Differences

```
OTUsREL2012.DNA <- OTUsREL[which(design$Year == "2012" & design$Molecule == "DNA"), ]
OTUsREL2012.DNA <- OTUsREL2012.DNA[ , colSums(OTUsREL2012.DNA > 0)]
OTUsREL2012.RNA <- OTUsREL[which(design$Year == "2012" & design$Molecule == "DNA"), ]
OTUsREL2012.RNA <- OTUsREL2012.RNA[ , colSums(OTUsREL2012.RNA > 0)]
active.N2012 <- active.N[grep("2012", rownames(active.N)), ]
rownames(OTUsREL2012.DNA) <- design$Lake[which(design$Year == "2012" & design$Molecule == "DNA")]
rownames(OTUsREL2012.RNA) <- design$Lake[which(design$Year == "2012" & design$Molecule == "DNA")]
rownames(active.N2012) <- design$Lake[which(design$Year == "2012" & design$Molecule == "DNA")]

# DNA
dbrda2012 <- capscale(OTUsREL2012.DNA ~ nuts2012$DOC + nuts2012$TN + nuts2012$TP, add =T, distance = "b
```

```
## Warning in cmdscale(X, k = k, eig = TRUE, add = add): only 8 of the first 9
## eigenvalues are > 0
```

```
anova(dbrda2012)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = OTUsREL2012.DNA ~ nuts2012$DOC + nuts2012$TN + nuts2012$TP, distance = "bra
##          Df SumOfSqs      F Pr(>F)
```

```
## Model     3  0.18760 1.9167  0.087 .
## Residual  6  0.19575
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RsquareAdj(dbrda2012)
```

```
## $r.squared
## [1] 0.4893716
##
## $adj.r.squared
## [1] 0.2340574
```

```
coef(dbrda2012)
```

```
##                      CAP1        CAP2        CAP3
## nuts2012$DOC -0.08665049 -0.20602323 -0.20832737
## nuts2012$TN   3.32155319  2.23604603  2.80303220
## nuts2012$TP  -0.06574191  0.01700405  0.03371554
```

```
# RNA
dbrda2012 <- capscale(OTUsREL2012.RNA ~ nuts2012$DOC + nuts2012$TN + nuts2012$TP, add =T, distance = "b
```

```
## Warning in cmdscale(X, k = k, eig = TRUE, add = add): only 8 of the first 9
## eigenvalues are > 0
```

```
anova(dbrda2012)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = OTUsREL2012.RNA ~ nuts2012$DOC + nuts2012$TN + nuts2012$TP, distance = "bra
##          Df SumOfSqs      F Pr(>F)
## Model     3  0.18760 1.9167  0.098 .
## Residual  6  0.19575
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RsquareAdj(dbrda2012)
```

```
## $r.squared
## [1] 0.4893716
##
## $adj.r.squared
## [1] 0.2340574
```

```
coef(dbrda2012)
```

```
##                      CAP1        CAP2        CAP3
## nuts2012$DOC -0.08665049 -0.20602323 -0.20832737
## nuts2012$TN   3.32155319  2.23604603  2.80303220
## nuts2012$TP  -0.06574191  0.01700405  0.03371554
```

```
# Active
dbrda2012 <- capscale(active.N2012 ~ nuts2012$DOC + nuts2012$TN + nuts2012$TP, add =T, distance = "bray
anova(dbrda2012)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = active.N2012 ~ nuts2012$DOC + nuts2012$TN + nuts2012$TP, distance = "bray"
##          Df SumOfSqs      F Pr(>F)
## Model     3  0.75254 1.9712  0.022 *
## Residual  6  0.76353
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RsquareAdj(dbrda2012)
```

```
## $r.squared
## [1] 0.496375
##
## $adj.r.squared
## [1] 0.2445625
```

```
coef(dbrda2012)
```

```
##                      CAP1         CAP2         CAP3
## nuts2012$DOC  0.08055203 -0.003445359  0.294709469
## nuts2012$TN   0.13440004 -1.707176730 -4.577891568
## nuts2012$TP  -0.02643324  0.070566186 -0.008340943
```

```
# Resoruces
chem.dbrda <- capscale(resREL ~ nuts2012$DOC + nuts2012$TN + nuts2012$TP, add =T, distance = "bray")
anova(chem.dbrda)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = resREL ~ nuts2012$DOC + nuts2012$TN + nuts2012$TP, distance = "bray", add =
##          Df SumOfSqs     F Pr(>F)
## Model     3  0.25058 1.294  0.183
## Residual  6  0.38728
```

```
RsquareAdj(chem.dbrda)
```

```
## $r.squared
## [1] 0.392843
##
## $adj.r.squared
## [1] 0.08926447
```

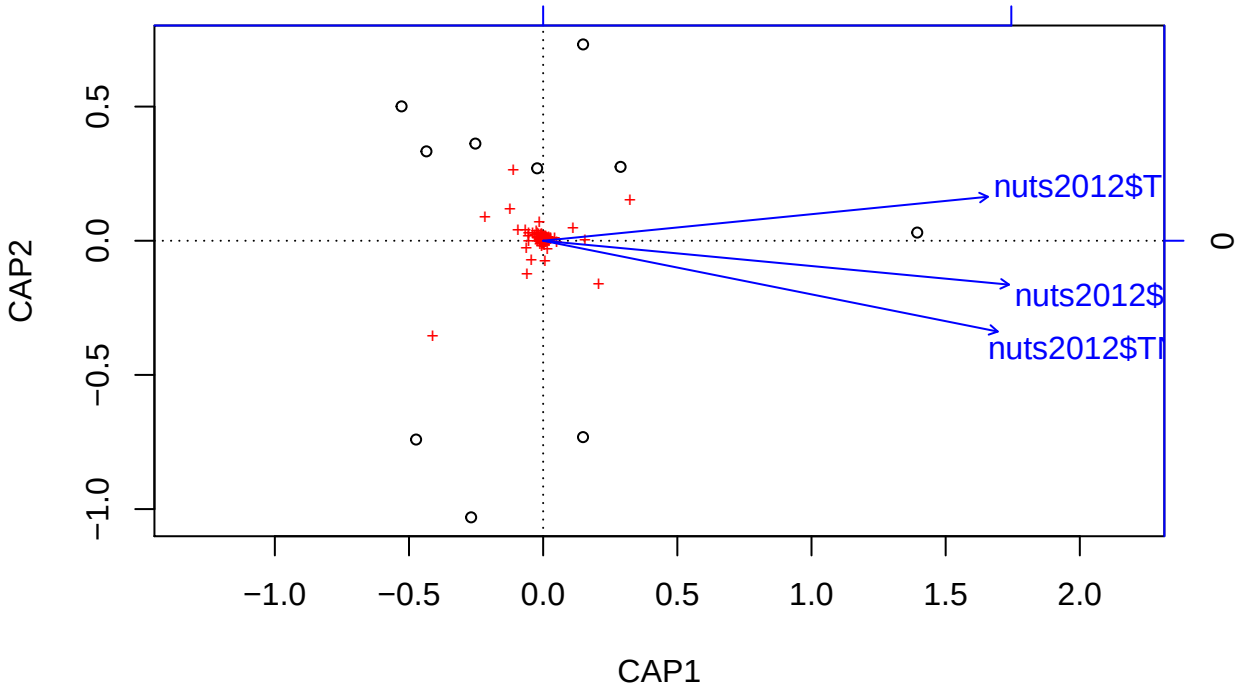
```
coef(chem.dbrda)
```

```
##                       CAP1       CAP2        CAP3
## nuts2012$DOC  0.066586751  0.1474111  0.25921101
## nuts2012$TN  -0.457552304 -4.0261543 -2.73319131
## nuts2012$TP   0.003780914  0.0557068 -0.05128675
```

```
anova.cca(chem.dbrda, step=1000)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = resREL ~ nuts2012$DOC + nuts2012$TN + nuts2012$TP, distance = "bray", add =
##          Df SumOfSqs     F Pr(>F)
## Model     3  0.25058 1.294  0.178
## Residual  6  0.38728
```

```
plot(chem.dbrda)
```




```
lmod <- as.mlm(chem.dbrda)
lmod
```

```
##
## Call:
## lm(formula = x$CCA$wa ~ . - 1, data = as.data.frame(X))
##
## Coefficients:
##                  CAP1       CAP2       CAP3
## `nuts2012$DOC`   0.066587   0.147411   0.259211
## `nuts2012$TN`   -0.457552  -4.026154  -2.733191
## `nuts2012$TP`    0.003781   0.055707  -0.051287
```

```
summary(lmod)
```

```
## Response CAP1 :
##
## Call:
## lm(formula = CAP1 ~ (`nuts2012$DOC` + `nuts2012$TN` + `nuts2012$TP`) -
##     1, data = as.data.frame(X))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.280107 -0.079988 -0.006128  0.083630  0.250485
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## `nuts2012$DOC`  0.066587   0.052926   1.258    0.249
## `nuts2012$TN`  -0.457552   0.846653  -0.540    0.606
## `nuts2012$TP`   0.003781   0.013133   0.288    0.782
##
## Residual standard error: 0.1732 on 7 degrees of freedom
## Multiple R-squared:  0.8264, Adjusted R-squared:  0.752
## F-statistic: 11.11 on 3 and 7 DF,  p-value: 0.004718
##
##
## Response CAP2 :
##
## Call:
## lm(formula = CAP2 ~ (`nuts2012$DOC` + `nuts2012$TN` + `nuts2012$TP`) -
##     1, data = as.data.frame(X))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.18966 -0.16391 -0.01948  0.05003  0.48816
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## `nuts2012$DOC`  0.14741    0.07189   2.051  0.07946 .
## `nuts2012$TN`  -4.02615    1.14998  -3.501  0.00998 **
## `nuts2012$TP`   0.05571    0.01784   3.123  0.01677 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2353 on 7 degrees of freedom
## Multiple R-squared:  0.7207, Adjusted R-squared:  0.601
## F-statistic: 6.022 on 3 and 7 DF,  p-value: 0.0237
```

```
##
##
## Response CAP3 :
##
## Call:
## lm(formula = CAP3 ~ (`nuts2012$DOC` + `nuts2012$TN` + `nuts2012$TP`) -
##     1, data = as.data.frame(X))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3227 -0.2485 -0.0477  0.1862  0.4466
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## `nuts2012$DOC`  0.25921    0.09768   2.654   0.0328 *
## `nuts2012$TN`  -2.73319    1.56266  -1.749   0.1238
## `nuts2012$TP`  -0.05129    0.02424  -2.116   0.0722 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3197 on 7 degrees of freedom
## Multiple R-squared:  0.5829, Adjusted R-squared:  0.4042
## F-statistic: 3.261 on 3 and 7 DF,  p-value: 0.08954
```

```
chem.dbrda <- capscale(resREL ~ log(nuts2012$DOC), add =T, distance = "bray")
anova(chem.dbrda)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = resREL ~ log(nuts2012$DOC), distance = "bray", add = T)
##          Df SumOfSqs      F Pr(>F)
## Model     1  0.12420 1.9343  0.035 *
## Residual  8  0.51366
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RsquareAdj(chem.dbrda)
```

```
## $r.squared
## [1] 0.1947118
##
## $adj.r.squared
## [1] 0.09405081
```

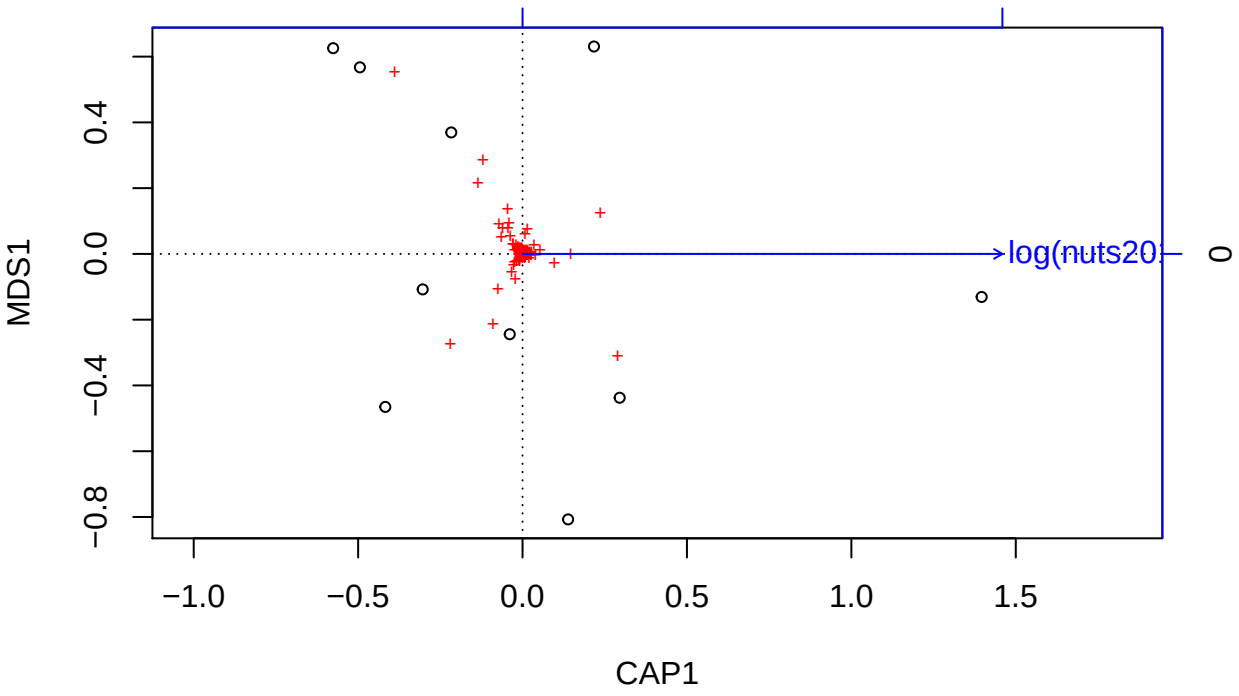
```
coef(chem.dbrda)
```

```
##                         CAP1
## log(nuts2012$DOC) 0.6004698
```

```
anova.cca(chem.dbrda, step=1000)
```

```
## Permutation test for capscale under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: capscale(formula = resREL ~ log(nuts2012$DOC), distance = "bray", add = T)
##          Df SumOfSqs      F Pr(>F)
## Model     1  0.12420 1.9343  0.039 *
## Residual  8  0.51366
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(chem.dbrda)
```




```
lmod <- as.mlm(chem.dbrda)
lmod
```

```
##
## Call:
## lm(formula = x$CCA$wa ~ . - 1, data = as.data.frame(X))
##
## Coefficients:
## `log(nuts2012$DOC)`
##              0.6005
```

```
summary(lmod)
```

```
##
```

```
## Call:
## lm(formula = x$CCA$wa ~ . - 1, data = as.data.frame(X))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.29143 -0.09346  0.05226  0.10989  0.24425
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## `log(nuts2012$DOC)`   0.6005     0.1000   6.004 0.000201 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1666 on 9 degrees of freedom
## Multiple R-squared:  0.8002, Adjusted R-squared:  0.778
## F-statistic: 36.05 on 1 and 9 DF,  p-value: 0.0002015
```

```
require(cocorresp)
```

```
## Loading required package: cocorresp
```

```
#test1 <- coca(OTUsREL2012.RNA, resREL, n.axes = 4)

# Variance Partitioning

nuts3 <- cbind(scale(log(nuts2012$DOC)), scale(log1p(nuts2012$TP)))
# nuts3 <- scale(log(nuts2012$DOC))
rownames(nuts3) <- nuts2012$sample

res.db <- vegdist(resREL, method = "altGower")
res.pcoa <- cmdscale(res.db, eig = TRUE, k = 5)

OTUsREL.log2012 <- OTUsREL.log[design$Year == "2012" & design$Molecule == "RNA", ]
OTUsREL2012 <- OTUsREL[design$Year == "2012" & design$Molecule == "RNA", ]
spe.pcoa <- cmdscale(vegdist(OTUsREL2012, method="bray"), eig = TRUE, k = 5)


# Active
HMWFvarpart <- varpart(spe.pcoa$points[, 1:2], nuts3, res.pcoa$points[, 1:2])
HMWFvarpart <- varpart(OTUsREL.log2012, nuts3, res.pcoa$points)
HMWFvarpart
```

```
##
## Partition of variation in RDA
##
## Call: varpart(Y = OTUsREL.log2012, X = nuts3, res.pcoa$points)
##
## Explanatory tables:
## X1:  nuts3
## X2:  res.pcoa$points
##
## No. of explanatory tables: 2
## Total variation (SS): 112296
```

```
##             Variance: 12477
## No. of observations: 10
##
## Partition table:
##                       Df R.squared Adj.R.squared Testable
## [a+b] = X1             2   0.43386       0.27210     TRUE
## [b+c] = X2             5   0.69171       0.30634     TRUE
## [a+b+c] = X1+X2        7   0.84950       0.32274     TRUE
## Individual fractions
## [a] = X1|X2            2                 0.01640     TRUE
## [b]                    0                 0.25570    FALSE
## [c] = X2|X1            5                 0.05064     TRUE
## [d] = Residuals                          0.67726    FALSE
## ---
## Use function 'rda' to test significance of fractions of interest
```

```
png(filename="../figures/Figure7.png",
    width = 1200, height = 1200, res = 96*2)

plot(HMWFvarpart)

dev.off() # this writes plot to folder
```

```
## pdf
##   2
```

```
graphics.off() # shuts down open devices
par(opar)

anova.cca(rda(spe.pcoa$points, nuts3), step=1000) #[a+b]
```

```
## Permutation test for rda under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: rda(X = spe.pcoa$points, Y = nuts3)
##          Df Variance      F Pr(>F)
## Model     2 0.073408 2.8109   0.01 **
## Residual  7 0.091404
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova.cca(rda(spe.pcoa$points, res.pcoa$points), step=1000) #[b+c]
```

```
## Permutation test for rda under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: rda(X = spe.pcoa$points, Y = res.pcoa$points)
##          Df Variance      F Pr(>F)
## Model     5 0.112439 1.7175  0.082 .
## Residual  4 0.052373
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova.cca(rda(spe.pcoa$points, cbind(nuts3, res.pcoa$points)), step=1000) #[a+b+c]
```

```
## Permutation test for rda under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: rda(X = spe.pcoa$points, Y = cbind(nuts3, res.pcoa$points))
##          Df Variance      F Pr(>F)
## Model     7 0.134654 1.2757  0.348
## Residual  2 0.030158
```

```
anova.cca(rda(spe.pcoa$points, res.pcoa$points, nuts3), step=1000) # [a]
```

```
## Permutation test for rda under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: rda(X = spe.pcoa$points, Y = res.pcoa$points, Z = nuts3)
##          Df Variance      F Pr(>F)
## Model     5 0.061246 0.8123   0.67
## Residual  2 0.030158
```

```
anova.cca(rda(spe.pcoa$points, nuts3, res.pcoa$points), step=1000) # [c]
```

```
## Permutation test for rda under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: rda(X = spe.pcoa$points, Y = nuts3, Z = res.pcoa$points)
##          Df Variance      F Pr(>F)
## Model     2 0.022215 0.7366  0.677
## Residual  2 0.030158
```

Microbial Functional Groups

Define RDP microbial groups Test each along with resource differences Who are the generalist taxa (which are active everywhere) Are generalists more abundant when resource concentration is higher?

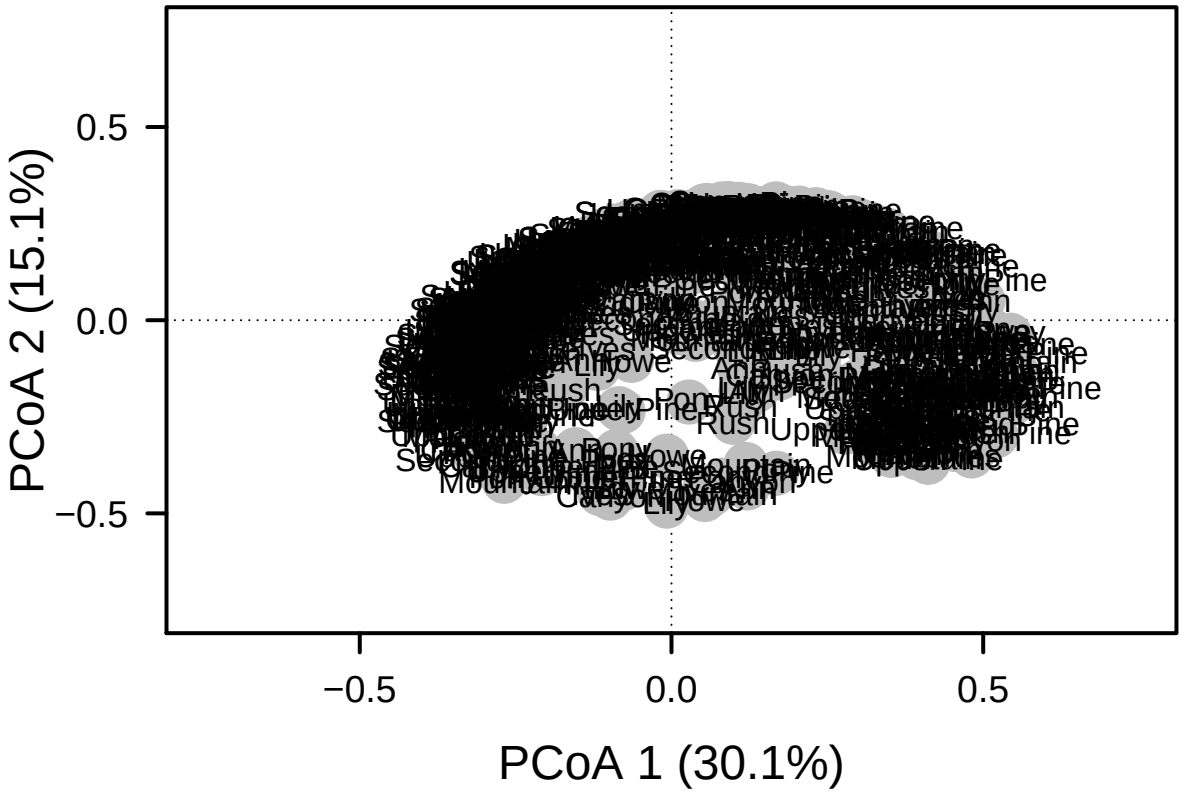
Can we group resources

What are the similar groups of resources: cluster resources based on abundance Can we cluster based on chemical data?

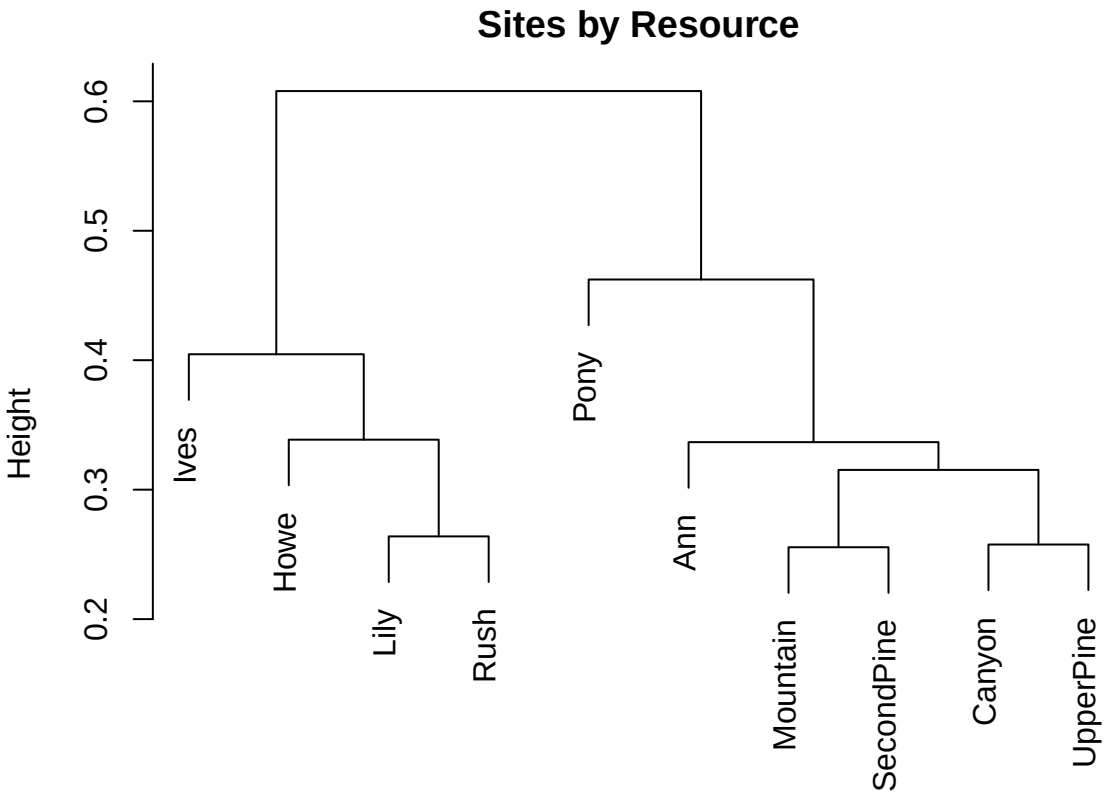
Supplemental

Supplemental Analysis: Cluster Analysis

```
res.dist <- vegdist(t(resREL), "bray")

res2.pcoa <- cmdscale(res.dist, eig = TRUE, k = 3)
explainvar1 <- round(res2.pcoa$eig[1] / sum(res2.pcoa$eig), 3) * 100
explainvar2 <- round(res2.pcoa$eig[2] / sum(res2.pcoa$eig), 3) * 100
explainvar3 <- round(res2.pcoa$eig[3] / sum(res2.pcoa$eig), 3) * 100
sum.eig <- sum(explainvar1, explainvar2, explainvar3)

# Define Plot Parameters
par(mar = c(5, 5, 1, 1) + 0.1)

# Initiate Plot
plot(res2.pcoa$points[ ,1], res2.pcoa$points[ ,2], ylim = c(-0.75, 0.75),
     xlim = c(-0.75, 0.75),
     xlab = paste("PCoA 1 (", explainvar1, "%)", sep = ""),
     ylab = paste("PCoA 2 (", explainvar2, "%)", sep = ""),
     pch = 16, cex = 2.0, type = "n", cex.lab = 1.5, cex.axis = 1.2, axes = FALSE)

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

# Add Points & Labels
points(res2.pcoa$points[ ,1], res2.pcoa$points[ ,2],
       pch = 19, cex = 3, bg = "gray", col = "gray")
text(res2.pcoa$points[ ,1], res2.pcoa$points[ ,2],
     labels = row.names(res.pcoa$points))
```




```
# Distances between molecules given sites

# Calculate distances
res.dist <- vegdist(resREL, method = "bray")
res.dist2 <- vegdist(t(resREL), method = "bray")

res.pcoa <- cmdscale(res.dist2, eig = TRUE, k = 3)




# Perform Cluster Analysis
res.ward <- hclust(res.dist, method = "ward.D2")
# Plot Cluster
par(mar = c(1, 5, 2, 2) + 0.1)
plot(res.ward, main = "Sites by Resource")
```




```
bac.dist <- vegdist(OTUsREL[design$Year == "2012" &
                            design$Molecule == "DNA", ], method="bray")
bac.ward <- hclust(bac.dist, method = "ward.D2")
```

Supplemental Figure 5: Eigenvalue Analysis Plots

```
png(filename="../figures/Supp5.png",
    width = 1600, height = 1100, res = 96*2)
# Define Plot Parameters
layout(as.matrix(cbind(1,2)))
par(mar = c(1, 1.5, 3, 1.5) + 0.1, oma = c(5, 5, 0.5, 0.5))

# Bray Curtis Analysis
# Plot Eigenvalues
plot(pcoa.rel$eig, main = "Bray Curtis",
     #xlab = "PCoA Axis", ylab = "Eigenvalue",
     las = 1, cex.lab = 1.5, pch = 16)

# Add Expectation based on Kaiser-Guttman criterion and Broken Stick Model
abline(h = mean(pcoa.rel$eig), lty = 2, lwd = 2, col = "blue")
b.stick <- bstick(42, sum(pcoa.rel$eig))
lines(1:42, b.stick, type = "l", lty = 4, lwd = 2, col = "red")

# Add Legend
legend("topright", legend = c("Avg Eigenvalue", "Broken-Stick"),
       lty = c(2, 4), bty = "n", col = c("blue", "red"))
```

```r
axis(1, lwd=2, labels = F)
axis(2, lwd=2, labels = F)
#axis(3, lwd=2, tck=-0.01, labels = F)
#axis(4, lwd=2, tck=-0.01, labels = F)
axis(1, lwd=2, tck=0.01, labels = F)
axis(2, lwd=2, tck=0.01, labels = F)
#axis(3, lwd=2, tck=0.01, labels = F)
#axis(4, lwd=2, tck=0.01, labels = F)

box(lwd=2)

# UniFrac Analysis
# Plot Eigenvalues
plot(pcoa.ufw$eig, main = "UniFrac",
     #xlab = "PCoA Axis", ylab = "Eigenvalue",
     las = 1, cex.lab = 1.5, pch = 16)

# Add Expectation based on Kaiser-Guttman criterion and Broken Stick Model
abline(h = mean(pcoa.ufw$eig), lty = 2, lwd = 2, col = "blue")
b.stick <- bstick(42, sum(pcoa.ufw$eig))
lines(1:42, b.stick, type = "l", lty = 4, lwd = 2, col = "red")

# Add Legend
legend("topright", legend = c("Avg Eigenvalue", "Broken-Stick"),
       lty = c(2, 4), bty = "n", col = c("blue", "red"))

axis(1, lwd=2, labels = F)
axis(2, lwd=2, labels = F)
#axis(3, lwd=2, tck=-0.01, labels = F)
#axis(4, lwd=2, tck=-0.01, labels = F)
axis(1, lwd=2, tck=0.01, labels = F)
axis(2, lwd=2, tck=0.01, labels = F)
#axis(3, lwd=2, tck=0.01, labels = F)
#axis(4, lwd=2, tck=0.01, labels = F)

box(lwd=2)
mtext("PCoA Axis", side = 1, outer = T, line = 2.5, cex = 2)
mtext("Eigenvalue", side = 2, outer = T, line = 2.5, cex = 2)

dev.off() # this writes plot to folder
```

```
## pdf
##   2
```

```r
graphics.off() # shuts down open devices
par(opar)
```

Supplemental Figure 6: PCoA w/ All Distances

```
png(filename="../figures/Supp6.png",
    width = 1800, height = 1800, res = 96*2)
par(opar)

# Define Plot Parameters
layout(matrix(c(7, 7, 7, 7,
                1, 1, 2, 2,
                1, 1, 2, 2,
                3, 3, 4, 4,
                3, 3, 4, 4,
                5, 5, 6, 6,
                5, 5, 6, 6), ncol = 4, byrow = T))
par(mar = c(0, 0, 0, 0) + 0.5, oma = c(4, 4, 1, 1))

# Define Plot Symbols
lake.pch <- rep(NA, length(design$Molecule))
for (i in 1:length(design$Molecule)){
if (design$Molecule[i] == "DNA"){
    lake.pch[i] <- 16
  }else{
    lake.pch[i] <- 17
  }}

pcoa.plots <- list(pcoa.pa, pcoa.ufu,
                   pcoa.rel, pcoa.ufw,
                   pcoa.log, pcoa.ufwl)
explainvar1 <- c(explainvar1.pa, explainvar1.rel, explainvar1.log, explainvar1.ufu,
                 explainvar1.ufw, explainvar1.ufwl)
explainvar2 <- c(explainvar2.pa, explainvar2.rel, explainvar2.log, explainvar2.ufu,
                 explainvar2.ufw, explainvar2.ufwl)

xlabel <- c(F, F, F, F, T, T)
ylabel <- c(T, F, T, F, T, F)

for (i in (1:length(pcoa.plots))){
  # Initiate Plot
  plot(pcoa.plots[[i]]$points[ ,1], pcoa.plots[[i]]$points[ ,2],
     ylim = c(-0.4, 0.4), xlim = c(-0.25, 0.55),
     #xlab = paste("PCoA 1 (", explainvar1[i], "%)", sep = ""),
     #ylab = paste("PCoA 2 (", explainvar2[i], "%)", sep = ""),
     xlab = "", ylab = "", xaxt = "n", yaxt = "n",
     pch = lake.pch, cex = 2.0, type = "n", cex.lab = 1.5, cex.axis = 1,
     axes = FALSE)

  # Add Axes
  axis(side = 1, labels = xlabel[i], lwd.ticks = 2, cex.axis = 1, las = 1)
  axis(side = 2, labels = ylabel[i], lwd.ticks = 2, cex.axis = 1, las = 1)
  axis(side = 3, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=-0.02)
  axis(side = 4, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=-0.02)
  axis(side = 1, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
  axis(side = 2, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
  axis(side = 3, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
  axis(side = 4, labels = F, lwd.ticks = 2, cex.axis = 1, las = 1, tck=0.01)
```

```
  abline(h = 0, v = 0, lty = 3)
  box(lwd = 2)

  # Add Percent Explained

  # Add Points & Labels
  points(pcoa.plots[[i]]$points[ ,1], pcoa.plots[[i]]$points[ ,2], pch = lake.pch,
         cex = 2.5, bg = "gray", col = lake.col)

  # Add Molecule Hulls
  ordihull(cbind(pcoa.plots[[i]]$points[ ,1], pcoa.plots[[i]]$points[ ,2]),
           design$Molecule, lwd=2, draw = "polygon", col="gray80", border = "black",
           label=TRUE, cex=1, bty = 'n')

  # Add Y Axis Label
  if (i == 3){
    mtext(side = 2, "PCoA 2", outer = F, line = 2.5)
  }

  # Add Molecule Legend to Plot 6
  if (i == 6){
    legend("topright", c("DNA", "RNA"), pch = c(16, 17),
           col = "gray", bty = "n", pt.cex = 1.25)
  }

}

plot.new()
# par(mar = c(0, 0, 5, 0) + 0.5)
legend("center", levels(design$Lake), ncol = 5, pch = 16, col = 1:10, bty = "n")
mtext(side = 1, "PCoA 1", outer = T, line = 2)

dev.off() # this writes plot to folder
graphics.off() # shuts down open devices
par(opar)
```