The geographical ecology of pond bacteria

March 3, 2015

```
## Loading required package: sp
## Loading required package: geoR
## Loading required package: MASS
## ------------------------------------------------------------
##  Analysis of geostatistical data
##  For an Introduction to geoR go to http://www.leg.ufpr.br/geoR
##  geoR version 1.7-4.1 (built on 2012-06-29) is now loaded
## ------------------------------------------------------------
##
## Loading required package: rgdal
## rgdal: version: 0.9-1, (SVN revision 518)
## Geospatial Data Abstraction Library extensions to R successfully loaded
## Loaded GDAL runtime: GDAL 1.11.1, released 2014/09/24
## Path to GDAL shared files: /usr/local/Cellar/gdal/1.11.1_3/share/gdal
## Loaded PROJ.4 runtime: Rel. 4.8.0, 6 March 2012, [PJ_VERSION: 480]
## Path to PROJ.4 shared files: (autodetected)
## Loading required package: raster
##
## Attaching package: 'raster'
##
## The following objects are masked from 'package:MASS':
##
##     area, select
##
## Loading required package: maptools
## Checking rgeos availability: TRUE
## Loading required package: picante
## Loading required package: ape
##
## Attaching package: 'ape'
##
## The following objects are masked from 'package:raster':
##
##     rotate, zoom
##
## Loading required package: vegan
## Loading required package: permute
## Loading required package: lattice
## This is vegan 2.2-1
## Loading required package: nlme
##
## Attaching package: 'nlme'
##
## The following object is masked from 'package:raster':
##
##     getData
##
## Loading required package: seqinr
## Loading required package: ade4
```

```
##
## Attaching package: 'ade4'
##
## The following object is masked from 'package:vegan':
##
##     cca
##
##
## Attaching package: 'seqinr'
##
## The following object is masked from 'package:nlme':
##
##     gls
##
## The following object is masked from 'package:permute':
##
##     getType
##
## The following objects are masked from 'package:ape':
##
##     as.alignment, consensus
##
## Loading required package: fossil
## Loading required package: maps
## Loading required package: shapefiles
## Loading required package: foreign
##
## Attaching package: 'shapefiles'
##
## The following objects are masked from 'package:foreign':
##
##     read.dbf, write.dbf
##
## Loading required package: simba
## This is simba 0.3-5
##
## Attaching package: 'simba'
##
## The following object is masked from 'package:picante':
##
##     mpd
##
## The following object is masked from 'package:stats':
##
##     mad
##
## Loading required package: reshape
```

Overview

Here, we will explore our primary geographical patterns of interest: the taxa-area relationship (TAR), the phylogenetic diversity-area relationship, and the distance-decay relationship in taxonomic and phylogenetic community similarity.

Study area

We analyzed environmental and bacterial community data from a survey of shallow ponds found east of Bloomington, IN. These ponds were constructed in the 1940s as wildlife refuge ponds, and are scattered throughout Brown County State Park, Yellowood State Forest, and Hoosier National Forest. In the summer of 2013, we visited approximately 50 of these ponds and recorded their geographic locations. We sampled aspects of water chemistry, physical properties, and bacterial community composition.

Figure 1. Spatially explicit data on environmental and geographic features.

```
# Load Environmental and Geographical Data
env <- read.table("DATA/CurrentData/20130801_PondDataMod.csv", sep = ",", header = TRUE)
lats <- as.numeric(env[, 3]) # latitudes (north and south)
lons <- as.numeric(env[, 4]) # longitudes (east and west)
```

Environmental data

We measured 19 environmental and geographic variables. These included elevation (m), geographical coordinates (lat-long; data: WGS84), temperature (C), Diameter(m), Depth(m), redox potential (ORP), specific conductivity or SpC (uS/cm), dissolved Oxygen (mg/L), total dissolved solids (g/L), salinity (p.s.u.=ppm), color - measured at absorbance = 660; an estimate of carbon in the water sample, chlorophyll a (ug/ml), dissolved organic carbon (mg/L), dissolved organic nitrogen (mg/L), and total phosphorus (ug/L).

Microbial community data

In addition to measuring a suite of geographic and environmental variables, we characterized the diversity of bacteria in the ponds using molecular-based approaches. Specifically, we amplified the 16S rRNA gene (i.e., "DNA") and 16S rRNA transcripts (i.e., "RNA") of bacteria using barcoded primers on the Illumina MiSeq platform. We then used a `mothur` pipeline to quality-trim our data set and assign sequences to operational taxonomic units (OTU).

Constructing the phylogeny...

For each pond, we used the observed taxonomic richness (S), total number of gene reads (N), and number of gene reads per OTU (Ni) to estimate Shannon's diversity index (H), and Simpson's evenness (D/S). We should estimate a handful of diversity and evenness metrics, as well conduct richness estimation for each site (Chao1, ACE, rarefaction, jacknife). These will provide basic diversity-related variables to explore with respect to geograpy and environmental conditions.

```
# Load Site-by-Species Matrix
comm <- read.otu(shared = "./DATA/CurrentData/INPonds.final.rdp.shared", cutoff = "1")

# Select DNA Data: Use the `grep()` Command and Rename with `gsub()`

# The active portion, based on cDNA
active.comm <- comm[grep("*-cDNA", rownames(comm)), ]
rownames(active.comm) <- gsub("\\-cDNA", "", rownames(active.comm))
rownames(active.comm) <- gsub("\\_", "", rownames(active.comm))

# The community without respect to active or not, 16S rRNA gene sequences
all.comm <- comm[grep("*-DNA", rownames(comm)), ]
rownames(all.comm) <- gsub("\\-DNA", "", rownames(all.comm))
rownames(all.comm) <- gsub("\\_", "", rownames(all.comm))
```

```r
# Remove Sites Not in the Environmental Data Set
active.comm <- active.comm[rownames(active.comm)  %in% env$Sample_ID, ]
all.comm <- all.comm[rownames(all.comm)  %in% env$Sample_ID, ]

# Remove Zero-Occurrence Taxa
active.comm <- active.comm[ , colSums(active.comm) > 0]
all.comm <- all.comm[ , colSums(all.comm) > 0]

# Import Taxonomy Data Using `read.tax()` from Source Code
tax <- read.tax(taxonomy = "./DATA/CurrentData/INPonds.final.rdp.1.cons.taxonomy")

#############
# Now, let's load and process the phylogenetic data
#############

# Import the Alignment File {seqinr}
ponds.cons <- read.alignment(file = "./DATA/CurrentData/INPonds.final.rdp.1.rep.fasta", format = "fasta

# Rename OTUs in the FASTA File
ponds.cons$nam <- gsub("\\|.*$", "", gsub("^.*?\t", "", ponds.cons$nam))

# Import the Outgroup FASTA File {seqinr}
outgroup <- read.alignment(file = "./DATA/CurrentData/methanosarcina.fasta", format = "fasta")

# Convert Alignment File to DNAbin Object {ape}
DNAbin <- rbind(as.DNAbin(outgroup), as.DNAbin(ponds.cons))

# Visusalize Sequence Alignment {ape}
#image.DNAbin(DNAbin, show.labels=T, cex.lab = 0.05, las = 1)

# Create Distance Matrix with the Jukes Cantor "JC" Model {ape}
seq.dist.jc <- dist.dna(DNAbin, model = "JC", pairwise.deletion = FALSE)

# Use Neighbor Joining Algorithm to Construct a Full Tree (DNA and RNA sequences) {ape}
phy.all <- bionj(seq.dist.jc)
phy <- phy.all

############# Keep all for overall analysis?
# Drop Tips of Zero-Occurrence OTUs (Removes Taxa Only Found via RNA Sequencing) {ape}
#phy <- drop.tip(phy.all, phy.all$tip.label[!phy.all$tip.label %in% c(colnames(comm), "Methanosarcina")
############# Keep all for overall analysis?

# Identify Outgroup Sequence
outgroup <- match("Methanosarcina", phy$tip.label)

# Root the Tree {ape}
phy <- root(phy, outgroup, resolve.root = TRUE)

#### A function to generate observed richness
S.obs <- function(x = ""){ rowSums(x > 0) * 1}
####

#### For each site:
```

```r
# N equals numbers of reads
env$active.N <- as.vector(rowSums(active.comm))
env$all.N <- as.vector(rowSums(all.comm))

# S equals the number of non-zero abundances
env$active.S <- S.obs(active.comm)
env$all.S <- S.obs(all.comm)

# Diversity is Shannon's
env$active.H <- as.vector(diversity(active.comm, index = "shannon"))
env$all.H <- as.vector(diversity(all.comm, index = "shannon"))

# Evenness is Simpsons; divide Simpson's Diversity by S
env$active.De <- as.vector(diversity(active.comm, index = "invsimpson")/env$active.S)
env$all.De <- as.vector(diversity(all.comm, index = "invsimpson")/env$all.S)
####

# Create a Phylogenetic Distance Matrix {picante}
phydist <- cophenetic.phylo(phy)

# Mean Pairwise Distance
active.dist.mp <- comdist(active.comm, phydist)
all.dist.mp <- comdist(all.comm, phydist)

# UniFrac Distance (Note: This Takes a Few Minutes; Be Patient)
active.dist.uf <- unifrac(active.comm, phy)
all.dist.uf <- unifrac(all.comm, phy)
```

Primary geographic patterns

We examined four primary geographic patterns, i.e., Distanct-decay (DD), Taxa-area relationship (TAR), Phylogenetic diversity-area relationship (PDAR), and the specific spatial abundance distribution (SSAD). While the DD and TAR have been more or less frequently studied in microbial ecology and microbial biogeography, the PDAR and SSAD have been mainly, in not entirely examined only in studies of macroscopic plants and animals.

1.) Distance Decay, taxonomic and phylogenetic

Tobler's first law of geography states that "Everything is related to everything else, but near things are more related than distant things" (Tobler 1970). This law is a formulation of the concept of spatial autocorrelation. In short, spatial autocorrelation is the degree to which spatial variables are either clustered in space (positive autocorrelation) or over-dispersed (negative autocorrelation).

The distance-decay relationship is a primary biogeographic pattern of spatial autocorrelation, and captures the rate of decreasing similarity with increasing distance. This pattern addresses whether communities close to one another are more similar than communities that are farther away. The distance-decay pattern can also be used to address whether near environments have greater similarity than far ones. We looked at decay in both taxonomic level compositional similarity via bray-curtis (should also do for Sorensens) and phylogenetic distance via unifrac distance.

2.) Species- or taxa- area relationship (SAR)

The species-area relationshp describes the rate at which species are discovred with increasing area. The SAR one of ecology's oldest and most intensively studied patterns. Arrhenius (1921) first described the general form of the *species-area relationship (SAR)* as a power-law: $S = cA^z$ where S is species richnness and A is area. Arrhenius's formula predicts a rate of increase in richness that is approximately linear in log-log space. That is, $log(S) = c + zlog(A)$, where z is the scaling exponent.

```
# A function to generate the species-area relationship
SAR <- function(OTUs){
  Alist <- c()
  Slist <- c()

  num.ponds <- c(1,2,3,4,6,8,12,16,24,32,42,51)
  for (i in num.ponds) {
    areas <- c() # hold iterated area values
    Ss <- c() # hold iterated S values

    for(j in 1:200){
      pond.sample <- sample(51, replace = FALSE, size = i)
      area <- 0
      cum.abs <- vector(mode = "numeric", length = length(comm))

      for (k in pond.sample) { # Loop through each randomly drawn pond
        area <- area + pond.areas[k] # aggregating area, (incrementing)
        cum.abs <- cum.abs + as.numeric(comm[k, ])

      } # End random pond samples loop
      S <- length(cum.abs[cum.abs > 0])
      areas <- c(areas, area)
      Ss <- c(Ss, S)
    }
    Alist <- rbind(Alist, mean(areas))
    Slist <- rbind(Slist, mean(Ss))
    print(c(mean(areas), mean(Ss)))
  }
  return(cbind(log10(Alist), log10(Slist)))
}
```

3.) Phylogenetic diversity-area relationship

Helmus and Ives (2012) developed methods to study how phylogenetic diversity changes with increasing area. This phylogenetic diversity-area relationship (PDAR) is analogous to the species-area relationship (SAR) that we learned about in the Geographical Ecology module. Note, however, that while the SAR is a cumulative relationship (S = cA^z) and so, cannot be negative. In contrast, the PDAR can increase or decrease with area.

In their study of phylogenetic diversity-area relationships (PDARs), Helmus and Ives, increased area by aggregating plots in two ways. First, they aggregated plots with respect to whether the plots were adjacent; what they called 'spatial'. Second, they aggregated plots at random ('non-spatial'). While both spatial and non-spatial sampling methods capture the effect of increasing area, sampling with respect to location (i.e., spatial) also captures the effect of increasing distance or spatial autocorrelation.

Helmus and Ives (2012) used the phylogenetic species variability (PSV) metric to quantify phylogenetic diversity. PSV quantifies how phylogenetic relatedness decreases the variance of a hypothetical neutral trait

shared by all species in a community. We constructed PDARs using the random aggregation approach of Helmus and Ives (2012).

We examined the relationship between phylogenetic diversity and area using both Spearman's correlation coefficient (S) and Pearson's correlation coefficient (P). It is informative to use both because while S is computed on ranks and depicts monotonic relationships (the degree to which the relationship is continually increasing or decreasing), P is computed on the observed values and therefore depicts linear relationships.

```
# A function to generate the non-spatial PDAR
PDAR <- function(com, tree){

  # Create Objects to Hold Areas and Diversity
  areas <- c()
  diversity <- c()

  # Create Vector Increasing Number of Plots by 2x
  num.plots <- c(2,3,4,6,8,12,16,24,32,42,51)

  for (i in num.plots){
    # Create vectors to hold areas and diversity form iterations, used for means
    areas.iter <- c()
    diversity.iter <- c()

    # Iterate 20 Times Per Sample Size
    for (j in 1:20){
      # Sample w/o replacement
      pond.sample <- sample(51, replace = FALSE, size = i)

      # Create Variable and Vector to Hold Accumulating Area and Taxa
      area <- 0
      sites <- c()

      for (k in pond.sample) {       # Loop Through Each Randomly Drawn Pond
        area <- area + pond.areas[k] # Aggregating Area (Roughly Doubling)
        sites <- rbind(sites, com[k, ])    # And Sites
        }

      # Concatenate the rea to areas.iter
      areas.iter <- c(areas.iter, area)
      # Calculate PSV or Other Phylogenetic Alpha-Diversity Metric
      psv.vals <- psv(sites, tree, compute.var = FALSE)
      psv <- psv.vals$PSVs[1]
      diversity.iter <- c(diversity.iter, as.numeric(psv))
      }

    diversity <- c(diversity, mean(diversity.iter)) # Let Diversity be the Mean PSV
    areas <- c(areas, mean(areas.iter))             # Let areas be the Average Area
    print(c(i, mean(diversity.iter), mean(areas.iter))) # Print As We Go
    }
  # Return Vectors of Areas (x) and Diversity (y)
  return(cbind(areas, diversity))
  }
```

RESULTS: Phylogenetic and taxonomic distance-Decay

```
#plot.new()
#par(mfrow=c(2, 2))

# Geographic Distances (Kilometers) Among Ponds
long.lat <- as.matrix(cbind(env$long, env$lat))
coord.dist <- earth.dist(long.lat, dist = TRUE)

# Taxonomic Distances Among Ponds (Bray-Curits)
active.bray.curtis.dist <- 1 - vegdist(active.comm)
all.bray.curtis.dist <- 1 - vegdist(all.comm)

# Phylogenetic Distances Among Ponds
active.unifrac.dist <- 1 - active.dist.uf
all.unifrac.dist <- 1 - all.dist.uf

# Transform All Distances Into List Format:
active.unifrac.dist.ls <- liste(active.unifrac.dist, entry = "unifrac")
all.unifrac.dist.ls <- liste(all.unifrac.dist, entry = "unifrac")

active.bray.curtis.dist.ls <- liste(active.bray.curtis.dist, entry = "bray.curtis")
all.bray.curtis.dist.ls <- liste(all.bray.curtis.dist, entry = "bray.curtis")
coord.dist.ls <- liste(coord.dist, entry = "geo.dist")

# Create a Data Frame from the Lists of Distances
df <- data.frame(coord.dist.ls,
          active.bray.curtis.dist.ls[, 3],
          all.bray.curtis.dist.ls[, 3],
          active.unifrac.dist.ls[, 3],
          all.unifrac.dist.ls[, 3])
names(df)[4:7] <- c("active.bray.curtis", "all.bray.curtis",
                              "active.unifrac", "all.unifrac")
attach(df)

# Now, let's plot the DD relationships:

# Set Initial Plot Parameters
par(mfrow=c(2, 2))#, mar = c(1, 5, 2, 1) + 0.1, oma = c(2, 0, 0, 0))

# Make Plot for Taxonomic DD
plot(coord.dist, active.bray.curtis, xlab = "", xaxt = "n", las = 1, ylim = c(0.1, 0.9),
     ylab="Bray-Curtis Similarity",
     main = "Distance Decay, Active taxa", col = "SteelBlue")

# Regression for Taxonomic DD
DD.reg.bc <- lm(active.bray.curtis ~ geo.dist)
summary(DD.reg.bc)
```

```
##
## Call:
## lm(formula = active.bray.curtis ~ geo.dist)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.23114 -0.07733 -0.00265  0.06916  0.39188
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.4747669  0.0057171   83.04   <2e-16 ***
## geo.dist    -0.0065592  0.0005152  -12.73   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1025 on 1273 degrees of freedom
## Multiple R-squared:  0.1129, Adjusted R-squared:  0.1122
## F-statistic: 162.1 on 1 and 1273 DF,  p-value: < 2.2e-16
```

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

# Make Plot for Taxonomic DD
plot(coord.dist, all.bray.curtis, xlab = "", xaxt = "n", las = 1, ylim = c(0.1, 0.9),
     ylab="Bray-Curtis Similarity",
     main = "Distance Decay, All taxa", col = "SteelBlue")

# Regression for Taxonomic DD
DD.reg.bc <- lm(all.bray.curtis ~ geo.dist)
summary(DD.reg.bc)
```

```
##
## Call:
## lm(formula = all.bray.curtis ~ geo.dist)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.32739 -0.08562  0.00755  0.08627  0.42761
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.4739804  0.0071731  66.077  < 2e-16 ***
## geo.dist    -0.0038858  0.0006464  -6.011  2.4e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1286 on 1273 degrees of freedom
## Multiple R-squared:  0.0276, Adjusted R-squared:  0.02684
## F-statistic: 36.13 on 1 and 1273 DF,  p-value: 2.404e-09
```

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

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

# Make Plot for Phylogenetic DD
plot(coord.dist, active.unifrac, xlab = "", las = 1, ylim = c(0.1, 0.9),
```

```
     ylab = "Unifrac Similarity", col = "SteelBlue",
     main = "Distance Decay, Active taxa")

# Regression for Phylogenetic DD
DD.reg.uni <- lm(active.unifrac ~ coord.dist)
summary(DD.reg.uni)
```

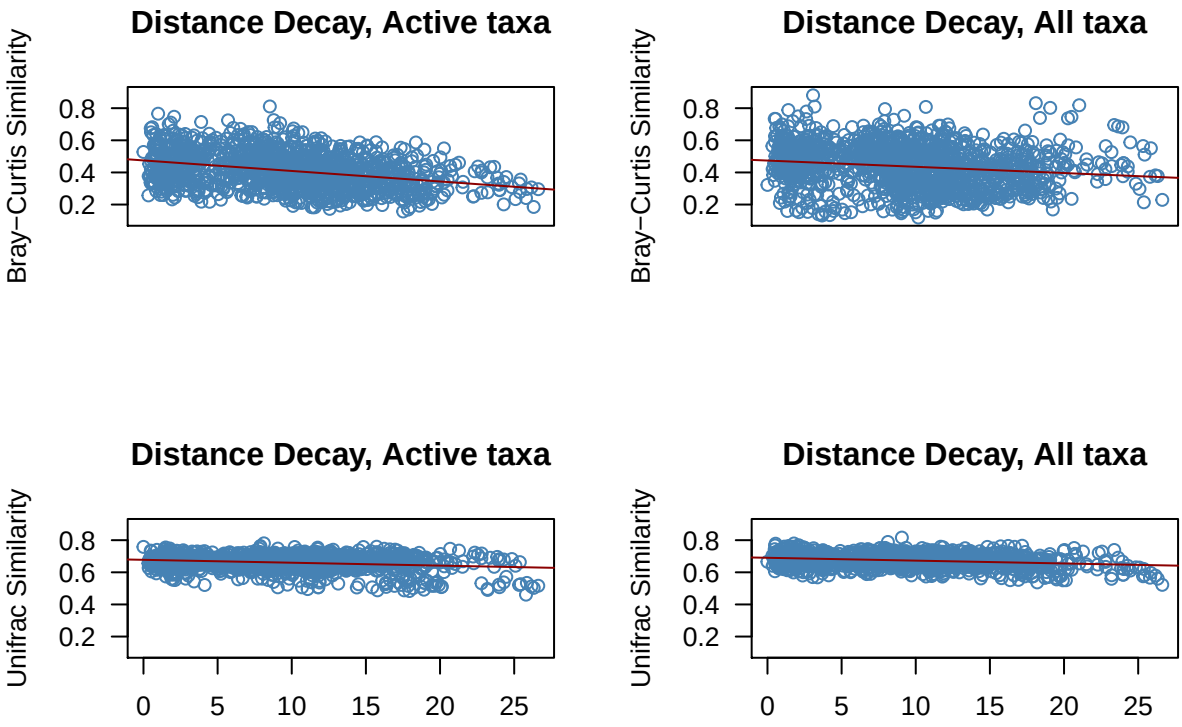
```
##
## Call:
## lm(formula = active.unifrac ~ coord.dist)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.170478 -0.021114  0.004084  0.031163  0.117319
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.677860   0.002563 264.502  < 2e-16 ***
## coord.dist  -0.001803   0.000231  -7.804 1.24e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04596 on 1273 degrees of freedom
## Multiple R-squared:  0.04566,    Adjusted R-squared:  0.04491
## F-statistic: 60.91 on 1 and 1273 DF,  p-value: 1.239e-14
```

```
abline(DD.reg.uni, col = "red4")

# Make Plot for Phylogenetic DD
plot(coord.dist, all.unifrac, xlab = "", las = 1, ylim = c(0.1, 0.9),
     ylab = "Unifrac Similarity", col = "SteelBlue", main = "Distance Decay, All taxa")

# Regression for Phylogenetic DD
DD.reg.uni <- lm(all.unifrac ~ coord.dist)
summary(DD.reg.uni)
```

```
##
## Call:
## lm(formula = all.unifrac ~ coord.dist)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.121366 -0.025420  0.002015  0.027773  0.141728
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.6904710  0.0021841 316.141   <2e-16 ***
## coord.dist  -0.0017706  0.0001968  -8.996   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03917 on 1273 degrees of freedom
```

```
## Multiple R-squared:  0.05977,    Adjusted R-squared:  0.05903
## F-statistic: 80.93 on 1 and 1273 DF,  p-value: < 2.2e-16
```

```
abline(DD.reg.uni, col = "red4")

# Add X-Axis Label to Plot
mtext("Geographic Distance, km", side = 1, adj = 0.55,
      line = 0.5, outer = TRUE)
```




```
diffslope(geo.dist, active.unifrac, geo.dist, all.unifrac)
diffslope(geo.dist, active.bray.curtis, geo.dist, all.bray.curtis)
```

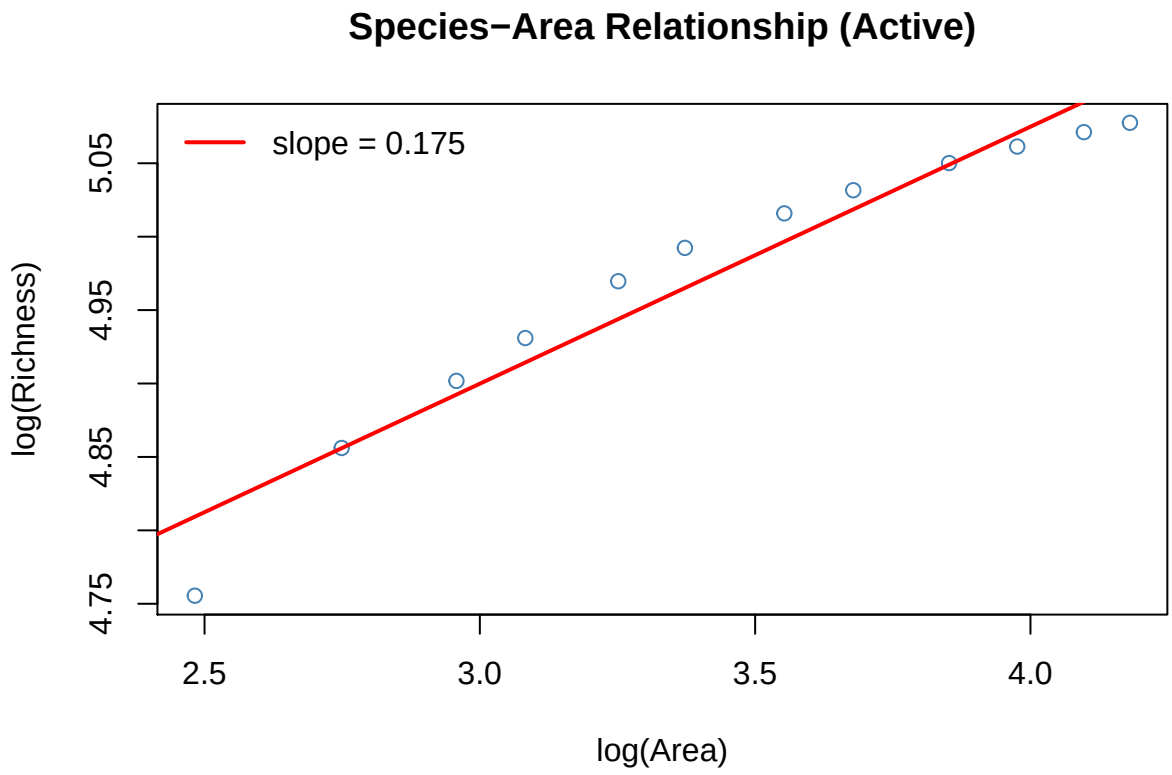
RESULTS: Taxa-area relationship

```
#plot.new()
#par(mfrow=c(2, 1))
pond.areas <- as.vector(pi * (env$Diameter/2)^2) # Find areas of all 51 ponds
coms <- as.list(active.comm, all.comm)

sar <- SAR(active.com)
```

```
## [1]   303.6563 56951.4350
```

```
## [1]   561.1703 71796.2450
## [1]   906.6696 79759.3550
## [1]  1209.325 85304.265
## [1]  1783.856 93249.860
## [1]  2357.336 98259.150
## [1]   3572.519 103722.700
## [1]   4766.478 107556.225
## [1]   7115.861 112235.555
## [1]   9461.905 115180.900
## [1]  12502.65 117816.06
## [1]  15159.69 119554.00
```

```
sar <- as.data.frame(sar)
plot(sar, xlab = "log(Area)", ylab = "log(Richness)",
  main = "Species-Area Relationship (Active)", col = "SteelBlue")

OLS <- lm(sar$V2 ~ sar$V1)
abline(OLS, col = "red", lw = 2)
slope <- round(coefficients(OLS)[2], 3)
  legend("topleft", legend = paste("slope =", slope),
         bty = "n", lw = 2, col = "red")
```

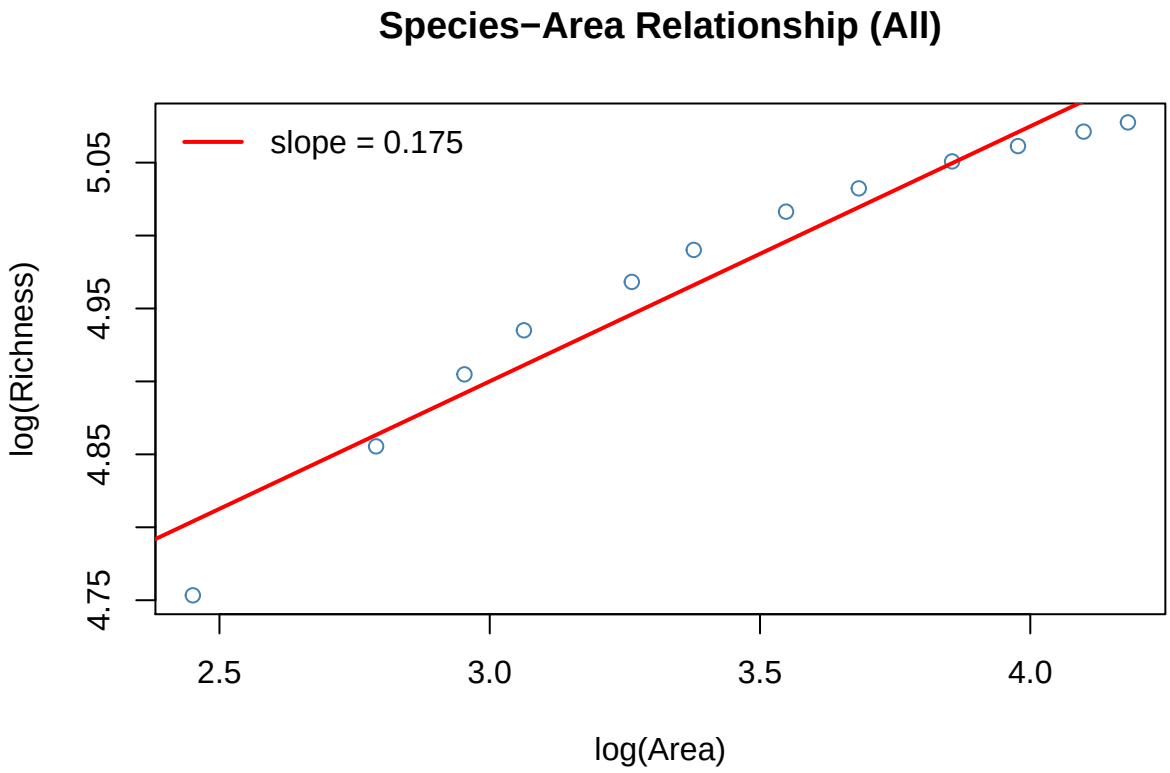



```
sar <- SAR(all.com)
```

```
## [1]   282.31 56672.89
```

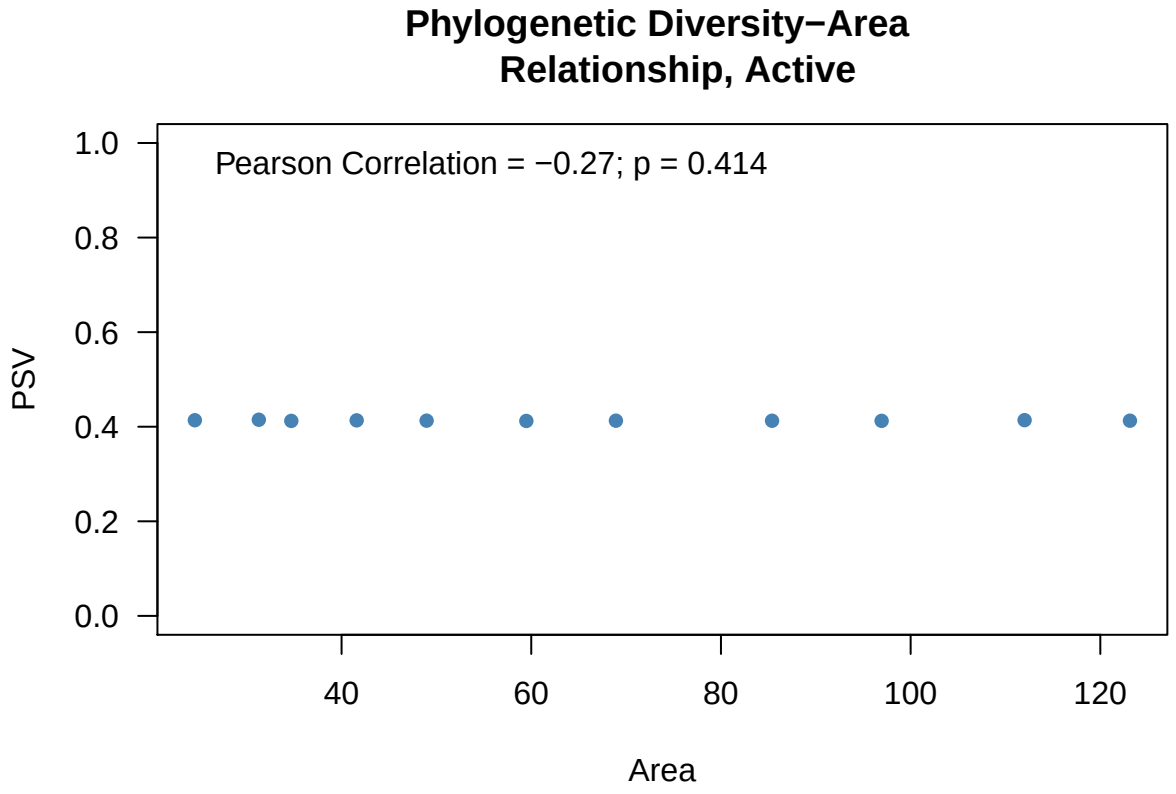
```
## [1]   616.2475 71692.2850
## [1]   897.801 80323.225
## [1]  1156.731 86107.695
## [1]  1831.62 92944.76
## [1]  2384.593 97759.125
## [1]   3533.203 103844.175
## [1]   4816.858 107732.505
## [1]   7168.642 112417.485
## [1]   9488.621 115167.340
## [1]  12551.07 117827.36
## [1]  15159.69 119554.00
```

```r
sar <- as.data.frame(sar)
plot(sar, xlab = "log(Area)", ylab = "log(Richness)",
  main = "Species-Area Relationship (All)", col = "SteelBlue")

OLS <- lm(sar$V2 ~ sar$V1)
abline(OLS, col = "red", lw = 2)
slope <- round(coefficients(OLS)[2], 3)
  legend("topleft", legend = paste("slope =", slope),
         bty = "n", lw = 2, col = "red")
```



RESULTS: Phylogenetic diversity-area relationship

```
#plot.new()
#par(mfrow=c(1, 1))

# Compute the PDAR
pdar <- PDAR(active.comm, phy)
pdar <- as.data.frame(pdar)
pdar$areas <- sqrt(pdar$areas)

# Calculate Perason's Correlation Coefficient
Pearson <- cor.test(pdar$areas, pdar$diversity, method = "pearson")
P <- round(Pearson$estimate, 2)
Pp <- round(Pearson$p.value, 3)

# Plot the PDAR
#par(mar = c(5, 5, 4, 2) + 0.1)
plot(pdar[, 1], pdar[, 2], xlab = "Area", ylab = "PSV", ylim = c(0, 1),
   main = "Phylogenetic Diversity-Area
   Relationship, Active",
   col = "Steelblue", pch = 16, las = 1)

legend("topleft", legend=paste("Pearson Correlation = ", P, "; p = ", Pp, sep = ""),
              bty = "n", col = "red")
```



```r
# Compute the PDAR
pdar <- PDAR(all.comm, phy)
pdar <- as.data.frame(pdar)
pdar$areas <- sqrt(pdar$areas)

# Calculate Perason's Correlation Coefficient
Pearson <- cor.test(pdar$areas, pdar$diversity, method = "pearson")
P <- round(Pearson$estimate, 2)
Pp <- round(Pearson$p.value, 3)

# Plot the PDAR
#par(mar = c(5, 5, 4, 2) + 0.1)
plot(pdar[, 1], pdar[, 2], xlab = "Area", ylab = "PSV", ylim = c(0, 1),
   main = "Phylogenetic Diversity-Area
   Relationship, All",
   col = "Steelblue", pch = 16, las = 1)

legend("topleft", legend=paste("Pearson Correlation = ", P, "; p = ", Pp, sep = ""),
              bty = "n", col = "red")
```