# Dimensions of Biodiversity - Aim 1, Persistence

*Jay T. Lennon and Stuart E. Jones*

*21 April, 2015*

## 1) SETUP

### A. Retrieve and Set Your Working Directory

```
rm(list = ls())
getwd()
setwd("~/GitHub/Dimensions/Aim1/DeathCurves/Phylo")
```

### B. Load Packages

```
require("muscle")
require("seqinr")
require("ape")
require("phylobase")
require("adephylo")
require("geiger")
require("picante")
require("stats")
require("RColorBrewer")
require("caper")
```

## 2) Read in FASTA file and take a look at lenghts of each sequence

```
fasta <- read.fasta(file = "./persistence.fasta", seqtype = "DNA")
summary(fasta)
```

```
##                Length Class       Mode
## KBS0701        1511   SeqFastadna character
## KBS0702        1517   SeqFastadna character
## KBS0703        1191   SeqFastadna character
## KBS0704        1479   SeqFastadna character
## KBS0705        1203   SeqFastadna character
## KBS0706        1035   SeqFastadna character
## KBS0707         791   SeqFastadna character
## KBS0710        1035   SeqFastadna character
## KBS0711        1519   SeqFastadna character
## KBS0712         758   SeqFastadna character
## KBS0713        1395   SeqFastadna character
## KBS0714        1514   SeqFastadna character
## KBS0715        1513   SeqFastadna character
## KBS0721        1010   SeqFastadna character
```

```
## KBS0722         1512   SeqFastadna character
## KBS0724         1508   SeqFastadna character
## KBS0725B        1482   SeqFastadna character
## KBS0727B        1482   SeqFastadna character
## KBS0801         1524   SeqFastadna character
## KBS0802         1228   SeqFastadna character
## KBS0812         1429   SeqFastadna character
## ATCC13985       1530   SeqFastadna character
## ATCC43928       1487   SeqFastadna character
## KBS0816         1419   SeqFastadna character
## Methanosarcina 1426   SeqFastadna character
```

## 3) Use mothur alingment based on Silva reference to make a tree

### A. Performing mothur alignment

i) Copy FASTA file from AFS to Mason Open terminal and type the following commands:

```
ssh karst.uits.iu.edu
cd /afs/iu.edu/home/l/e/lennon/Lennon_Shared/Long-term_Dormancy/Sequences
kinit
aklog
cd /afs/iu.edu/home/l/e/lennon/Lennon_Shared/Long-term_Dormancy/Sequences
cp persistence.fasta /N/dc2/projects/Lennon_Sequences/Persistence
```

ii) Perform mothur alingnment

```
ssh lennonj@mason.indiana.edu
cd /N/dc2/projects/Lennon_Sequences/Persistence
module load gcc
module load mothur/1.31.2
mothur
align.seqs(fasta=persistence.fasta, reference=silva.bacteria.fasta, flip=T, processors=4)
align.seqs(fasta=persistence.fasta, reference=silva.bacteria.rdp.tax, flip=T, processors=4)
summary.seqs(fasta=persistence.align)
screen.seqs(fasta=persistence.align, minlength=758)
filter.seqs(fasta=persistence.good.align, vertical=T, trump=.)
quit()
```

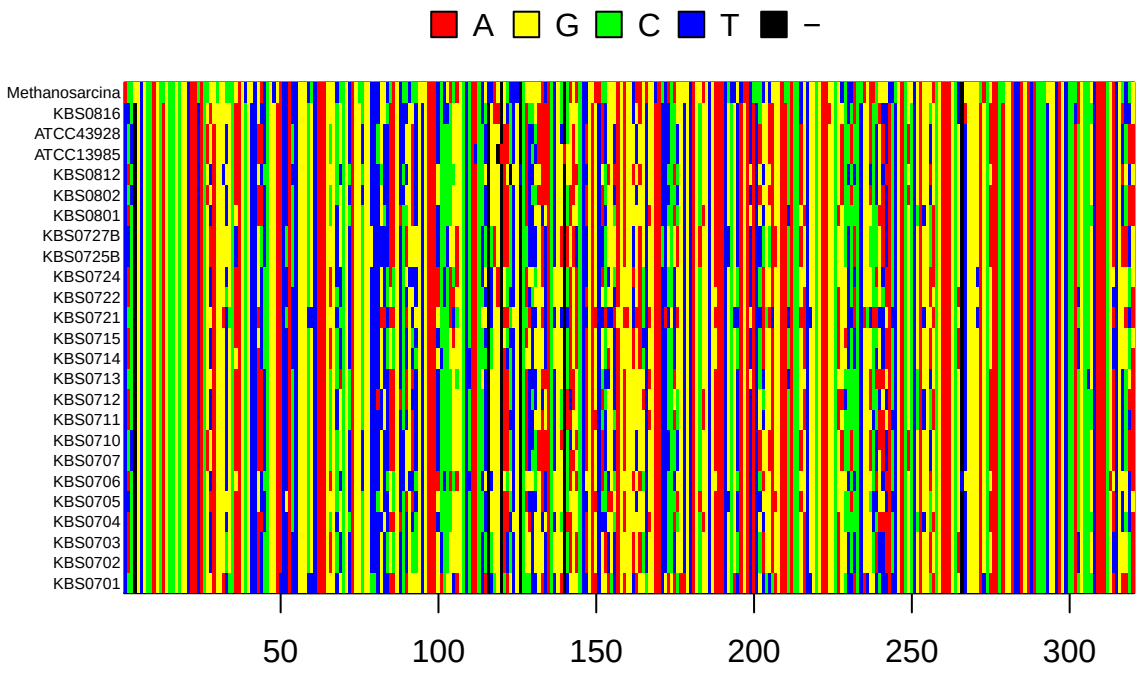
Open new terminal and cd to Github project folder.
Move the file using the following commands:

```
scp lennonj@karst.uits.iu.edu:/N/dc2/projects/Lennon_Sequences/Persistence/persistence.good.filter.
```

### B. Visualize alignments

```
# Read mothur alignment file {seqinr}
read.aln.M <- read.alignment(file = "./persistence.good.filter.silva.afa", format = "fasta")

# # Read arb full alignment file {seqinr}
# read.aln.M <- read.alignment(file = "./persistence.arb.none.afa", format = "fasta")
```

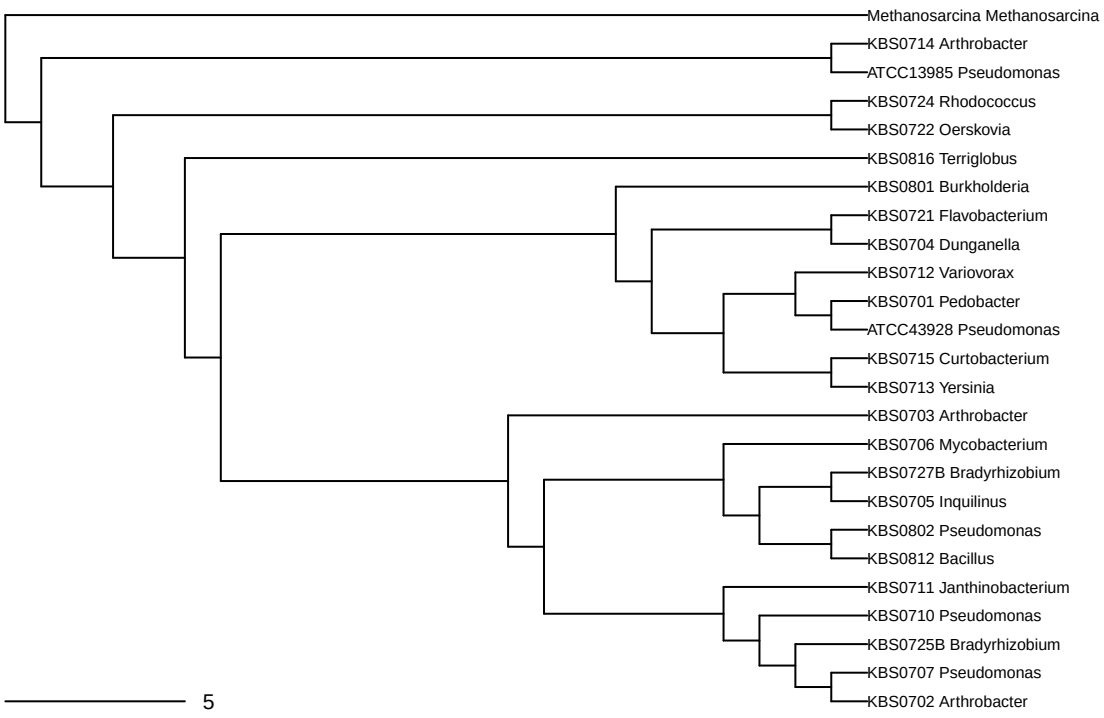
```
#
# # Read arb alignment file with gap columns removed {seqinr}
# read.aln.M <- read.alignment(file = "./persistence.arb.vert.afa", format = "fasta")
#
# # Read mega alignment file with gap columns removed {seqinr}
# read.aln.M <- read.alignment(file = "./persistence.mega.short.afa", format = "fasta")
#
# # Read RDP alignment file {seqinr}
# read.aln.M <- read.alignment(file = "./persistence.rdp.afa", format = "fasta")

# Convert Alignment File to DNAbin Object {ape}
p.DNAbin.M <- as.DNAbin(read.aln.M)

# Identify Base Pair Region of 16S rRNA Gene to Visuzlize (adjust range)
window.M <- p.DNAbin.M[, 1:320]

# Command to Visusalize Sequence Alignment {ape}
image.DNAbin(window.M, cex.lab = 0.50)
```



### 4) Make neibhor-joining tree

## Read Alignment File {seqinr}

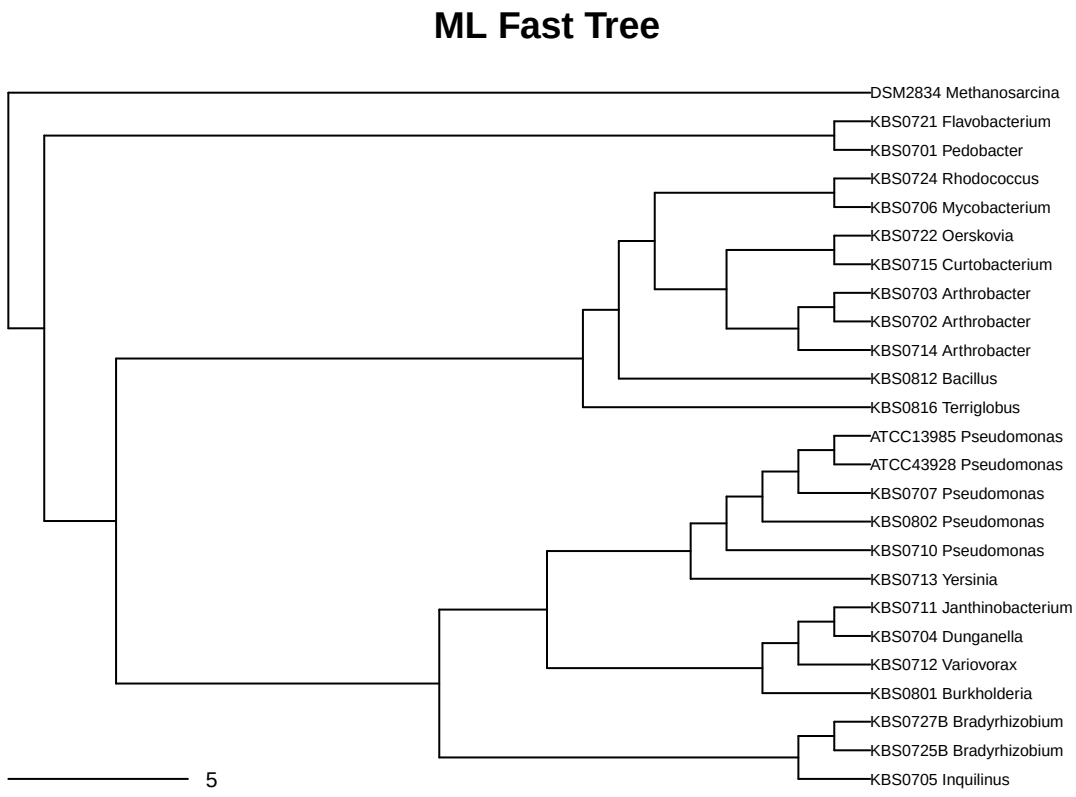
```
# Create Distance Matrix with "raw" Model {ape}
seq.dist.raw.M <- dist.dna(p.DNAbin.M, model = "K80", pairwise.deletion = FALSE)

# Neighbor Joining Algorithm to Construct Tree, a 'phylo' Object {ape}
nj.tree <- bionjs(seq.dist.raw.M)

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

# Root the Tree {ape}
nj.rooted <- root(nj.tree, outgroup.M, resolve.root = TRUE)

# Load phylo taxonomy data
tax <- read.table("persistence.phylo.txt", sep = "\t", header = TRUE)
rownames(tax) <- tax$Code
tax.2 <- tax[nj.tree$tip.label, ]
tax.name <- paste(tax.2$Code, tax.2$Genus)
tax.name[25] <- "DSM2834 Methanosarcina"

tax.name<-paste(tax$Code,tax$Genus)
nj.tree$tip.label <- match(nj.tree$tip.label,tax.name)

# Plot the Rooted Tree{ape}
par(mar = c(1,1,2,1) + 0.1)
plot.phylo(nj.rooted, main = "Neigbor Joining Tree from mothur Alignment",
           "phylogram", use.edge.length = FALSE, direction = "right",
           cex = 0.6, label.offset = 1, show.tip.label = FALSE, x.lim = 30)

tiplabels(tax.name, adj = c(0,0.5), cex = 0.5, frame = "none",
          pch = NULL, thermo = NULL, pie = NULL, piecol = NULL,
          col = NULL, bg = NULL)

add.scale.bar(cex = 0.7)
```




## 5) Read in maximum likelihood tree

```
# Read tree
ml.tree <- read.tree("persistence.arb.non.fasttree.tre")

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

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

# Load phylo taxonomy data
tax <- read.table("persistence.phylo.txt", sep = "\t", header = TRUE)
rownames(tax) <- tax$Code
tax.2 <- tax[ml.tree$tip.label, ]
tax.name <- paste(tax.2$Code, tax.2$Genus)
tax.name[15] <- "DSM2834 Methanosarcina"

# Plot the Rooted Tree{ape}
par(mar = c(1,1,2,1) + 0.1)
plot.phylo(ml.rooted, main = "ML Fast Tree",
           "phylogram", use.edge.length = FALSE, direction = "right",
           cex = 0.6, label.offset = 1, show.tip.label = FALSE, x.lim = 30)
```

```
tiplabels(tax.name, adj = c(0,0.5), cex = 0.5, frame = "none",
          pch = NULL, thermo = NULL, pie = NULL, piecol = NULL,
          col = NULL, bg = NULL)

add.scale.bar(cex = 0.7)
```

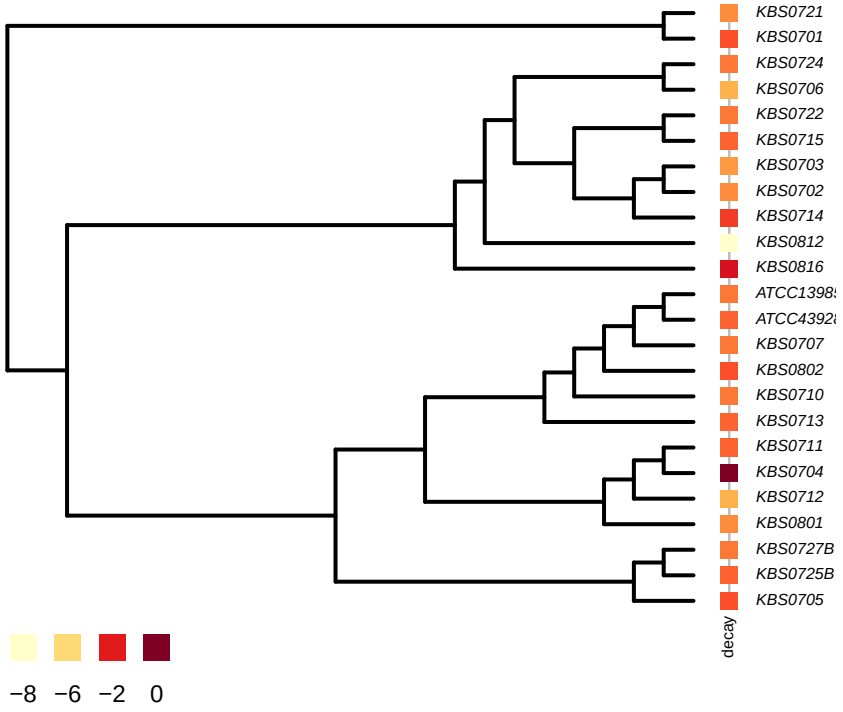



## 6) Map traits onto tree

```
# Keep Rooted but Drop Outgroup Branch
ml.rooted <- root(ml.tree, outgroup, resolve.root = TRUE)
ml.rooted <- drop.tip(ml.rooted, "Methanosarcina")

# Define Color Palette
mypalette <- colorRampPalette(brewer.pal(9, "YlOrRd"))

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

decay <- as.matrix(log10(tax[1:24,7]))
rownames(decay) <- tax[1:24,1]
colnames(decay) <- c("Decay")
x.decay <- phylo4d(ml.rooted, decay, check.node.labels = "drop")

table.phylo4d(x.decay, treetype = "phylo", symbol = "colors", show.node = TRUE,
```

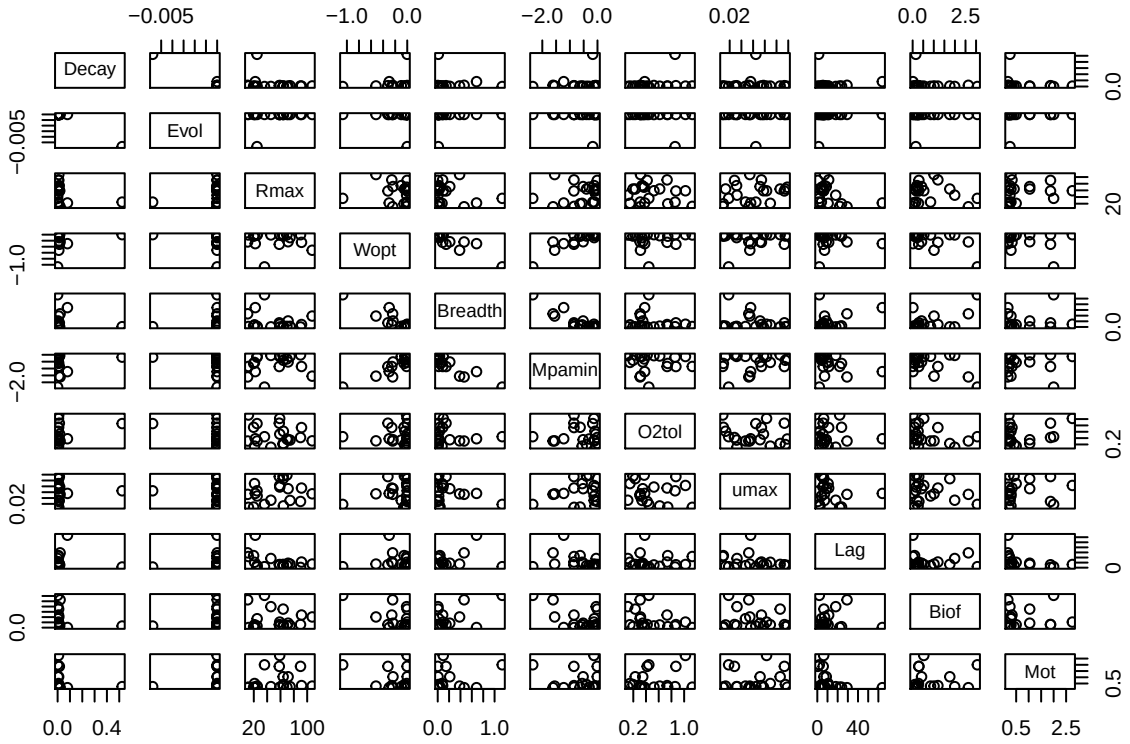
```
            cex.label = 0.5, scale = FALSE, use.edge.length = FALSE,
            edge.color = "black", edge.width = 2, box = FALSE,
            col=mypalette(25), pch = 15, cex.symbol = 1.25, var.label=("       decay"),
            ratio.tree = 0.90, cex.legend = 1.5, center = FALSE)
```

```
## Warning: There may not be enough room left to plot data; you may consider
## reducing ratio.tree or cex.label.
```

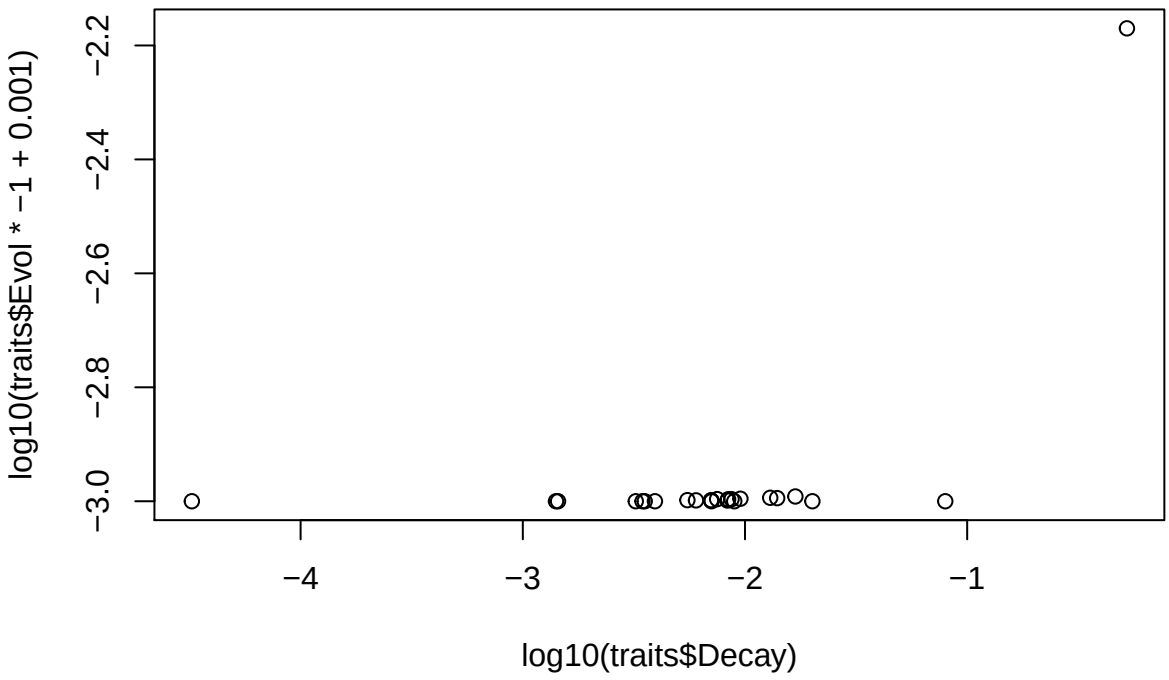



```
## 6) Look at some trait correlations
```

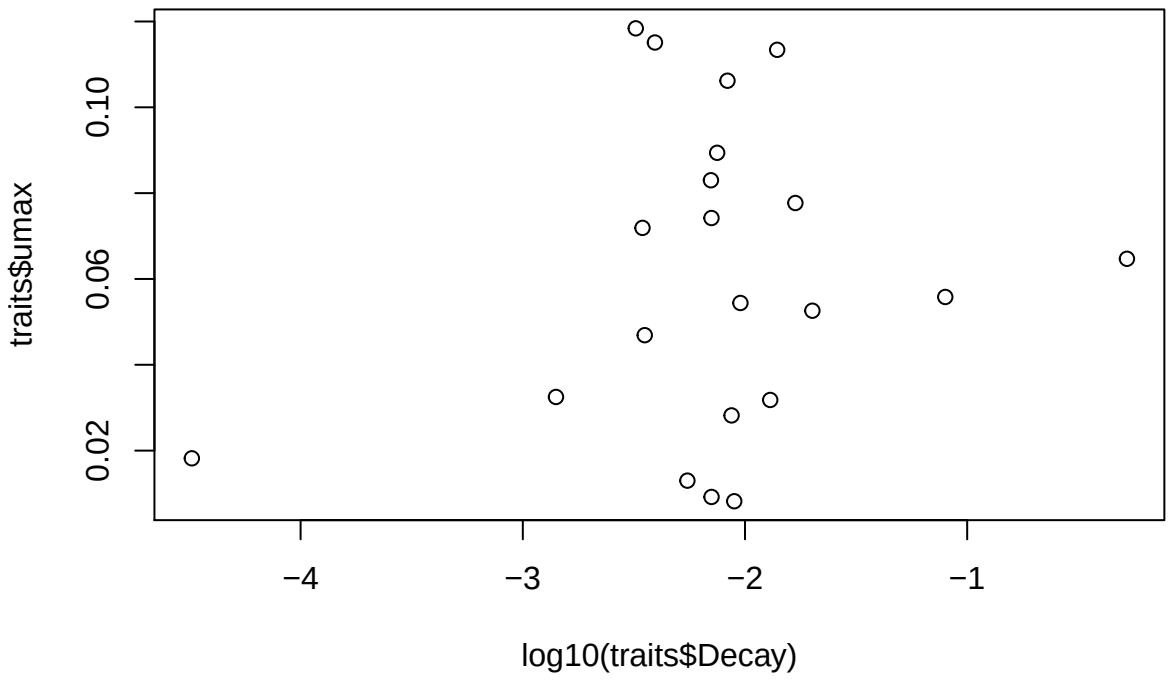
```
traits <- tax[,7:17]
pairs(traits)
```



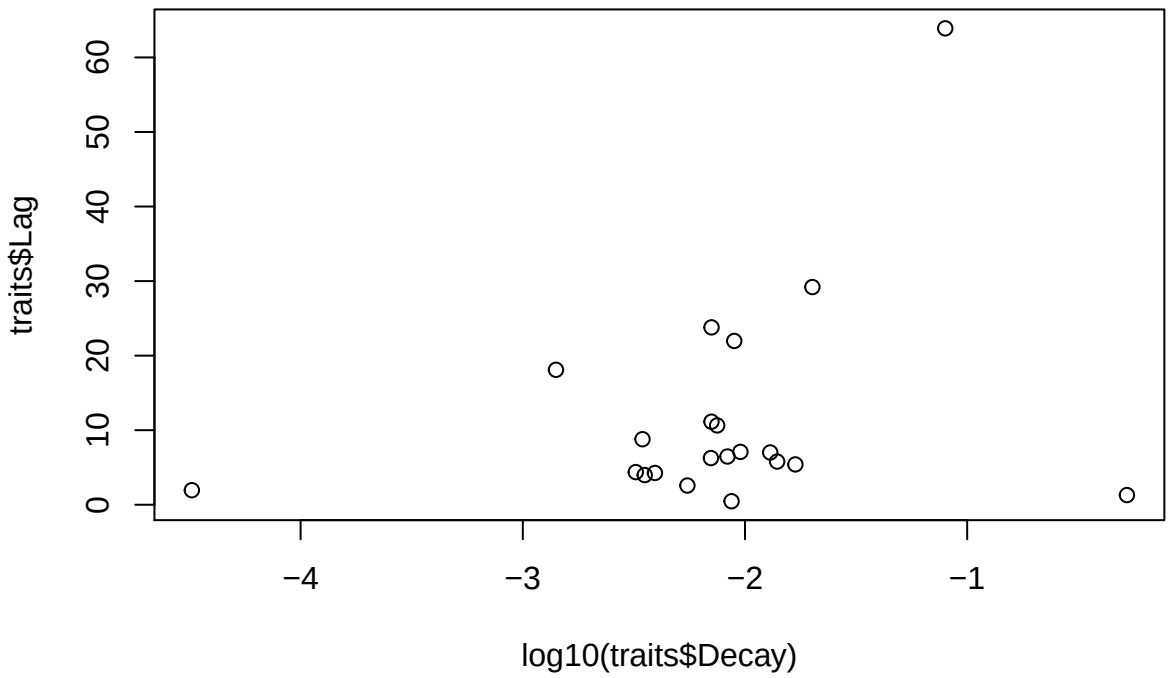

```
fold.decay <- max(traits$Decay, na.rm = TRUE)/min(traits$Decay, na.rm = TRUE)
plot(log10(traits$Decay),log10(traits$Evol*-1 + 0.001))
```




```
plot(log10(traits$Decay),traits$umax)
```




```
plot(log10(traits$Decay),traits$Lag)
```


traits$Lag
0
10
20
30
40
50
60
−4
−3
−2
−1
log10(traits$Decay)