

# Host-specificity and core taxa of seagrass leaf microbiome identified across tissue age and geographical regions | *Sanders-Smith, R. & Segovia, B.T.(joint contribution)*, Forbes, C., Hessing-Lewis, M., Morien, E., Lemay, M.A., O'Connor, M. I., Parfrey, L.W.

Bianca Trevizan Segovia

20/11/2020

## NMDS and PERMANOVAs comparison across samples types

```
#### load packages ####
library(QsRutils)
library(EcolUtils)
library(phyloseq)
library(ggplot2)
library(dplyr)
library(phylosmith)

### importing rarefied phyloseq object ###
phylo_merge_rare <- readRDS("data/phylo_merge_rarefied_16S_paper.rds")
phylo_merge_rare

## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 1206 taxa and 149 samples ]
## sample_data() Sample Data: [ 149 samples by 4 sample variables ]
## tax_table() Taxonomy Table: [ 1206 taxa by 7 taxonomic ranks ]
## phy_tree() Phylogenetic Tree: [ 1206 tips and 1204 internal nodes ]

### disabling scientific notation ###
options(scipen = 999)

### set.seed for reproducibility ###
set.seed(3245)

### root tree ###
phylo_merge_rare <- root_phyloseq_tree(phylo_merge_rare)
tree1 = phy_tree(phylo_merge_rare)
```

## whole dataset *excluding artificial*

```
#####
### NMDS plot ###
#####
### Remove artificial seagrass ###
```

```

phylo_NO_artificial_1 <- subset_samples(phylo_merge_rare, sample_growth!="artificial")

### IMPORTANT !!! ###
### Re-order phyloseq object to run PERMANOVA ###
phylo_NO_artificial <- set_sample_order(phylo_NO_artificial_1, c('sample_growth'))
# check if re-ordering worked, set same order for metadata later
SampleID_reordered <- as.data.frame(sample_data(phylo_NO_artificial)$SampleID)
#View(SampleID_reordered) # "seawater", "zostera_new", "zostera_old"

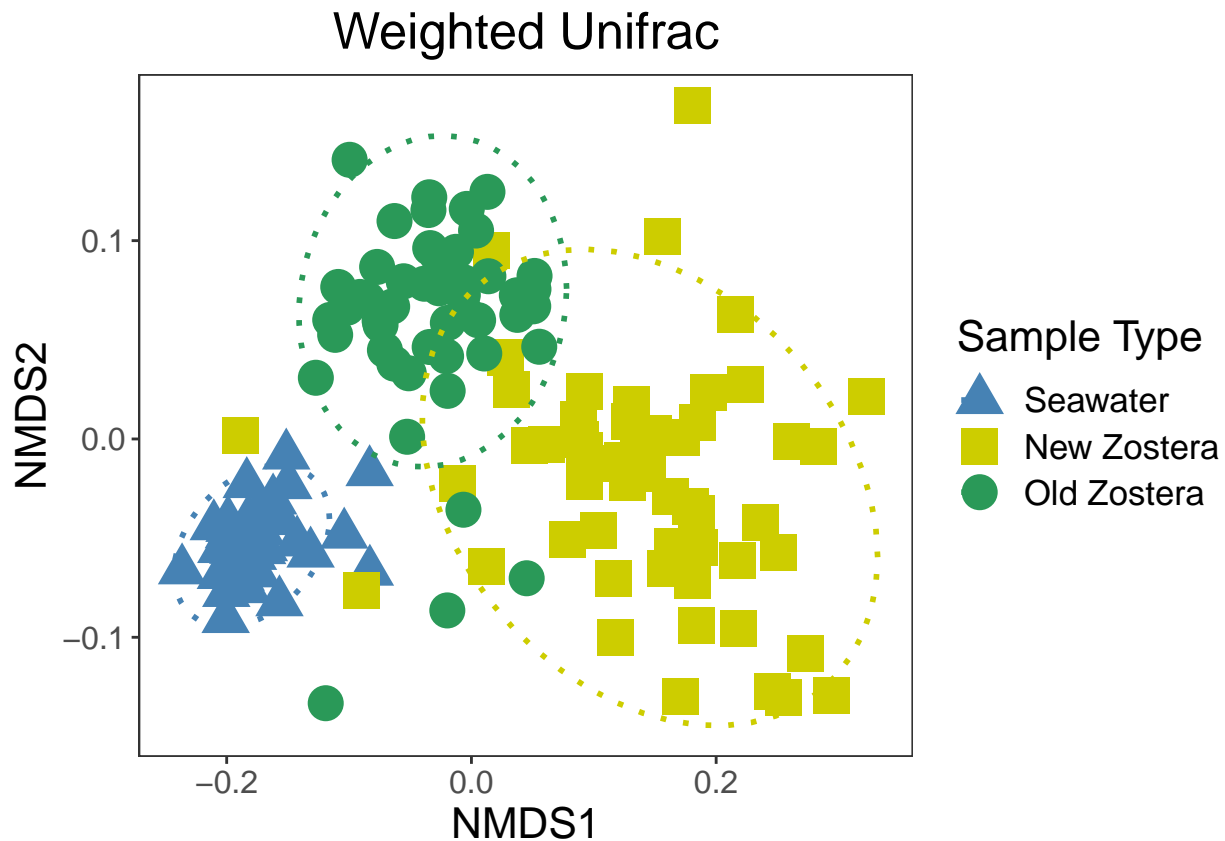
### ordinate using weighted Unifrac, Bray-Curtis and Jaccard dissimilarity
NMDS_bray_NO_artificial <- ordinate(phylo_NO_artificial, "NMDS", "bray")
NMDS_UNI_NO_artificial <- ordinate(phylo_NO_artificial, "NMDS", "wunifrac")
NMDS_jacc_NO_artificial <- ordinate(phylo_NO_artificial, "NMDS", "jaccard")

### NMDS plot graph ###
colours_no_artificial <- c(zostera_new="yellow3",
                           zostera_old="#2a9958",
                           seawater="steelblue")

NMDS_all <- plot_ordination(phylo_NO_artificial, NMDS_UNI_NO_artificial,
                           color = "sample_growth",
                           shape = "sample_growth",
                           title = "Weighted Unifrac") +

  geom_point(size=6) +
  scale_colour_manual(name = "Sample Type", values=colours_no_artificial,
                     labels=c("Seawater","New Zostera","Old Zostera")) +
  scale_shape_manual(name = "Sample Type", values=c(17,15,16),
                     labels=c("Seawater","New Zostera","Old Zostera")) +
  stat_ellipse(type = "t", linetype = 3, size = 1) +
  theme_bw() +
  theme (text = element_text(size=16),
        axis.text.x=element_text(size=12),
        axis.text.y = element_text(size=12),
        plot.title = element_text(hjust = 0.5),
        panel.grid.major = element_blank(), #remove major grid
        panel.grid.minor = element_blank())#remove minor grid
NMDS_all

```



```
#####
### PERMANOVAs ###
#####
#### Importing master table ####
allsamples <- read.csv("data/Bact_3000_metadata_clean.csv", header = T)

### get metadata only ###
metadata <- allsamples[,1:4]

### Remove artificial seagrass ###
remove_artificials <- c("artificial")
noartificials <- allsamples %>%
  dplyr::filter(!sample_growth %in% remove_artificials)
x <- c("seawater", "zostera_new", "zostera_old")
noartificials <- noartificials %>%
  mutate(sample_growth = factor(sample_growth, levels = x)) %>%
  arrange(sample_growth)

### Get metadata without artificial seagrass ###
metadata_no_artificial <- metadata %>%
  dplyr::filter(!sample_growth %in% remove_artificials)
metadata_no_artificial <- metadata_no_artificial %>%
  mutate(sample_growth = factor(sample_growth, levels = x)) %>%
  arrange(sample_growth)

### Calculate weighted Unifrac, Bray-Curtis and Jaccard dissimilarity ###
phylo_NO_artificial.df <- data.frame(sample_data(phylo_NO_artificial))
```

```
sampledist_all_bray <- phyloseq::distance(phylo_NO_artificial, method = "bray")
sampledist_all_UNI <- phyloseq::distance(phylo_NO_artificial, method = "wunifrac")
sampledist_all_jacc <- phyloseq::distance(phylo_NO_artificial, method = "jaccard")
```

```
### Convert distances from phyloseq objects to matrices ###
```

```
bray <- as.matrix(sampledist_all_bray)
UNI <- as.matrix(sampledist_all_UNI)
jacc <- as.matrix(sampledist_all_jacc)
dist_bray_all_mtx <- as.dist(as(bray, "matrix"))
dist_UNI_all_mtx <- as.dist(as(UNI, "matrix"))
dist_jacc_all_mtx <- as.dist(as(jacc, "matrix"))
```

```
### test for heteroscedasticity in the data ###
```

```
homogeneity <- betadisper(dist_bray_all_mtx ,metadata_no_artificial$sample_growth)
permutest(homogeneity, pairwise = TRUE)
```

```
##
## 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      2 0.97793 0.48897 71.88    999 0.001 ***
## Residuals 137 0.93194 0.00680
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Pairwise comparisons:
## (Observed p-value below diagonal, permuted p-value above diagonal)
##
##           seawater          zostera_new zostera_old
## seawater                0.001000000000000000208167      0.001
## zostera_new 0.000000000000000000000029418              0.007
## zostera_old 0.000000000000023608193926018 0.0028683187047296209229441
```

```
#boxplot(homogeneity)
```

```
homogeneity <- betadisper(dist_UNI_all_mtx ,metadata_no_artificial$sample_growth)
permutest(homogeneity, pairwise = TRUE)
```

```
##
## 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      2 0.20542 0.102708 38.308    999 0.001 ***
## Residuals 137 0.36731 0.002681
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Pairwise comparisons:
## (Observed p-value below diagonal, permuted p-value above diagonal)
```

```
##          seawater      zostera_new zostera_old
## seawater              0.00100000000000000    0.001
## zostera_new 0.000000000000038717            0.010
## zostera_old 0.0000000002885092 0.0055906415960946

#boxplot(homogeneity)

homogeneity <-betadisper(dist_jacc_all_mtx ,metadata_no_artificial$sample_growth)
permutest(homogeneity, pairwise = TRUE)

##
## 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     2 0.55741 0.278706 81.308   999 0.001 ***
## Residuals 137 0.46961 0.003428
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Pairwise comparisons:
## (Observed p-value below diagonal, permuted p-value above diagonal)
##               seawater      zostera_new zostera_old
## seawater             0.001000000000000000208167      0.001
## zostera_new 0.0000000000000000000000022538      0.002
## zostera_old 0.00000000000003699179346949 0.0020638498401832337282580

#boxplot(homogeneity)

### Because we have heteroscedasticity, we should have the same n so PERMANOVA is robust ###
# GET LIST OF RANDOMLY SUBSET OF DATA FOR PHYLOSEQ FILTERING / PERMANOVA #
sum(noartificials $sample_growth == "seawater") #35 samples

## [1] 35

sum(noartificials $sample_growth == "zostera_new") #54 samples

## [1] 54

sum(noartificials $sample_growth == "zostera_old") #51 samples

## [1] 51

### Randomly get 35 samples from water and zostera so they have same n: ###
bact_random_noartificials <- noartificials %>%
  group_by(sample_growth) %>%
  sample_n(size = 35)

## See if it worked ###
sum(bact_random_noartificials$sample_growth == "seawater")

## [1] 35

sum(bact_random_noartificials$sample_growth== "zostera_new")

## [1] 35
```

```

sum(bact_random_noartificials$sample_growth == "zostera_old")

## [1] 35

### list of randomly subset samples ###
bact_random_list_noartificials <- bact_random_noartificials$SampleID
length(bact_random_list_noartificials)

## [1] 105

### Now use the list above to subsample data in phyloseq object ###
phylo_random_noartificials <- subset_samples(phylo_NO_artificial,
                                              sample_names(phylo_NO_artificial) %in%
                                              bact_random_list_noartificials)

### Calculate weighted Unifrac, Bray-Curtis and Jaccard dissimilarity again, but for random samples ###
phylo_random_noartificials.df <- data.frame(sample_data(phylo_random_noartificials))
sampledist_random_bray <- phyloseq::distance(phylo_random_noartificials,
                                              method = "bray")
sampledist_random_UNI <- phyloseq::distance(phylo_random_noartificials,
                                              method = "wunifrac")
sampledist_random_jacc <- phyloseq::distance(phylo_random_noartificials,
                                              method = "jaccard")

### Convert distances from phyloseq objects to matrices ###
bray_random <- as.matrix(sampledist_random_bray)
UNI_random <- as.matrix(sampledist_random_UNI)
jacc_random <- as.matrix(sampledist_random_jacc)
dist_bray_random_mtx <- as.dist(as(bray_random , "matrix"))
dist_UNI_random_mtx <- as.dist(as(UNI_random, "matrix"))
dist_jacc_random_mtx <- as.dist(as(jacc_random, "matrix"))

#### get list of subsamples from metadata ###
noartificials_subsample <- metadata_no_artificial %>%
  dplyr::filter(SampleID %in% bact_random_list_noartificials)
noartificials_subsample$sample_growth <- factor(noartificials_subsample$sample_growth)

#####
### RANDOM samples (35 samples for each sample type) ###
#####
### with geographical regions as a random effect ###

#####
#### PERMANOVA Bray Curtis ####
#####
adonis2(dist_bray_random_mtx ~ noartificials_subsample$sample_growth + region,
        data= noartificials_subsample , permutations=10000, by = "margin")

## Permutation test for adonis under reduced model
## Marginal effects of terms
## Permutation: free
## Number of permutations: 10000
##
## adonis2(formula = dist_bray_random_mtx ~ noartificials_subsample$sample_growth + region, data = noar
##
Df SumOfSqs      R2      F      Pr(>F)

```

```
## noartificials_subsample$sample_growth  2   13.665 0.34418 29.398 0.00009999
## region                                3    2.514 0.06331  3.605 0.00009999
## Residual                             99   23.009 0.57953
## Total                               104   39.704 1.00000
##
## noartificials_subsample$sample_growth ***
## region                                ***
## Residual
## Total
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# pairwise comparisons
set.seed(3245)
adonis.pair(dist_brays_random_mtx, noartificials_subsample$sample_growth,
            nper = 10000, corr.method = "BH")
```

```
##               combination SumsOfSqs MeanSqs F.Model      R2      P.value
## 1   seawater <-> zostera_new 8.527401 8.527401 36.59432 0.3498691 0.00009999
## 2   seawater <-> zostera_old 8.962933 8.962933 42.11586 0.3824686 0.00009999
## 3 zostera_new <-> zostera_old 3.780699 3.780699 12.40253 0.1542554 0.00009999
##   P.value.corrected
## 1           0.00009999
## 2           0.00009999
## 3           0.00009999
```

```
#####
#### PERMANOVA Weighted Unifrac ####
#####
adonis2(dist_UNI_random_mtx ~ noartificials_subsample$sample_growth + region,
        data= noartificials_subsample , permutations=10000, by = "margin")
```

```
## Permutation test for adonis under reduced model
## Marginal effects of terms
## Permutation: free
## Number of permutations: 10000
##
```

```
## adonis2(formula = dist_UNI_random_mtx ~ noartificials_subsample$sample_growth + region, data = noart.
##               Df SumOfSqs      R2      F      Pr(>F)
## noartificials_subsample$sample_growth  2   3.0965 0.55081 70.7544 0.00009999
## region                                3   0.2369 0.04214  3.6091   0.0004
## Residual                             99   2.1663 0.38535
## Total                               104   5.6217 1.00000
##
## noartificials_subsample$sample_growth ***
## region                                ***
## Residual
## Total
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# pairwise comparisons
set.seed(3245)
adonis.pair(dist_UNI_random_mtx, noartificials_subsample$sample_growth,
            nper = 10000, corr.method = "BH")
```

```
##              combination SumsOfSqs  MeanSqs  F.Model      R2    P.value
## 1    seawater <-> zostera_new  2.401254  2.401254  104.65861  0.6061592  0.00009999
## 2    seawater <-> zostera_old  1.351031  1.351031   79.83888  0.5400398  0.00009999
## 3 zostera_new <-> zostera_old  1.075378  1.075378   34.89440  0.3391283  0.00009999
##    P.value.corrected
## 1          0.00009999
## 2          0.00009999
## 3          0.00009999

#####
#### PERMANOVA Jaccard ####
#####
adonis2(dist_jacc_random_mtx ~ noartificials_subsample$sample_growth + region,
        data= noartificials_subsample , permutations=10000, by = "margin")

## Permutation test for adonis under reduced model
## Marginal effects of terms
## Permutation: free
## Number of permutations: 10000
##
## adonis2(formula = dist_jacc_random_mtx ~ noartificials_subsample$sample_growth + region, data = noar
##              Df SumOfSqs      R2      F    Pr(>F)
## noartificials_subsample$sample_growth  2   10.223  0.23180  16.2904  0.00009999
## region                                3    2.462  0.05583   2.6155  0.00009999
## Residual                             99   31.064  0.70435
## Total                               104   44.103  1.00000
##
## noartificials_subsample$sample_growth ***
## region                                ***
## Residual
## Total
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# pairwise comparisons
set.seed(3245)
adonis.pair(dist_jacc_random_mtx, noartificials_subsample$sample_growth,
            nper = 10000, corr.method = "BH")

##              combination SumsOfSqs  MeanSqs  F.Model      R2
## 1    seawater <-> zostera_new  6.356585  6.356585  20.389251  0.23067568
## 2    seawater <-> zostera_old  6.739848  6.739848  22.728630  0.25051222
## 3 zostera_new <-> zostera_old  2.769003  2.769003   7.329975  0.09730489
##    P.value P.value.corrected
## 1 0.00009999      0.00009999
## 2 0.00009999      0.00009999
## 3 0.00009999      0.00009999
```

### choked dataset *including artificial*

```
#####
### NMDS plot ###
#####
### Select only Choked region ###
phylo_choked <- subset_samples(phylo_merge_rare,
```



```

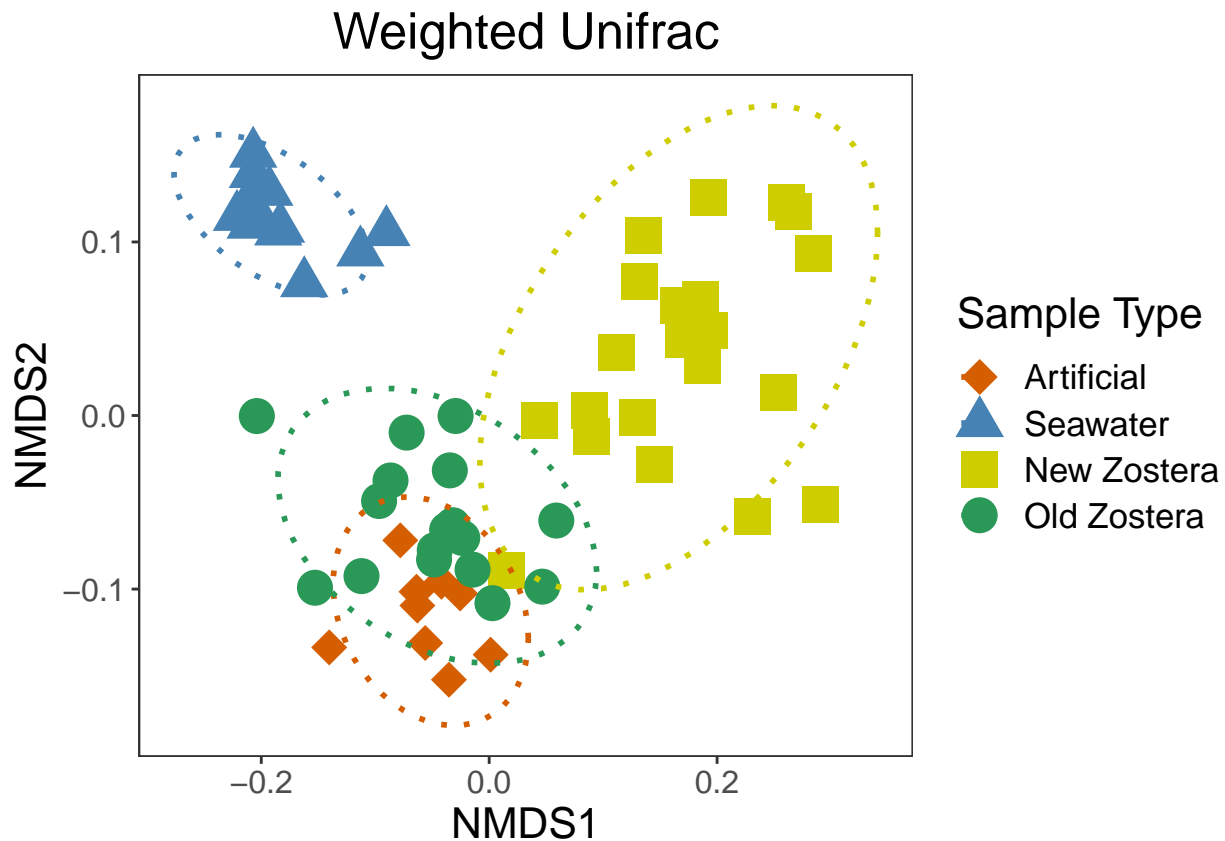
region=="choked")

### IMPORTANT !!! ###
### Re-order phyloseq object to run PERMANOVA ###
phylo_choked <- set_sample_order(phylo_choked, c('sample_growth'))
# check if re-ordering worked, set same order for metadata later
SampleID_reordered_choked <- as.data.frame(sample_data(phylo_choked)$SampleID)
#View(SampleID_reordered_choked) # "artificial". "seawater", "zostera_new", "zostera_old"

### ordinate using weighted Unifrac, Bray-Curtis and Jaccard dissimilarity
NMDS_bray_CHOKED <- ordinate(phylo_choked, "NMDS", "bray")
NMDS_UNI_choked <- ordinate(phylo_choked, "NMDS", "wunifrac")
NMDS_jacc_choked <- ordinate(phylo_choked, "NMDS", "jaccard")

### NMDS plot graph ###
colours_choked <- c(zostera_new="yellow3",
                    zostera_old="#2a9958",
                    artificial="#D55E00",
                    seawater="steelblue")
NMDS_choked <- plot_ordination(phylo_choked, NMDS_UNI_choked,
                              color = "sample_growth",
                              shape = "sample_growth",
                              title = "Weighted Unifrac") +
  geom_point(size = 6) +
  scale_colour_manual(name = "Sample Type", values=colours_choked,
                     labels=c("Artificial","Seawater", "New Zostera","Old Zostera")) +
  scale_shape_manual(name = "Sample Type", values=c(18,17,15,16),
                     labels=c("Artificial","Seawater","New Zostera","Old Zostera")) +
  stat_ellipse(type = "t", linetype = 3, size = 1) +
  theme_bw() +
  theme (text = element_text(size=16),
        axis.text.x=element_text(size=12),
        axis.text.y = element_text(size=12),
        plot.title = element_text(hjust = 0.5),
        panel.grid.major = element_blank(), #remove major grid
        panel.grid.minor = element_blank())
NMDS_choked

```



```
#####
### PERMANOVAs ###
#####

#### Importing master table ####
allsamples <- read.csv("data/Bact_3000_metadata_clean.csv", header = T)

### get metadata only ###
metadata <- allsamples[,1:4]

### Remove other regions ###
remove_other_regions <- c("triquet", "mcmullin", "goose")
metadata_choked <- metadata %>%
  dplyr::filter(!region %in% remove_other_regions)
x <- c("artificial", "seawater", "zostera_new", "zostera_old")
metadata_choked <- metadata_choked %>%
  mutate(sample_growth = factor(sample_growth, levels = x)) %>%
  arrange(sample_growth)

### Calculate weighted Unifrac, Bray-Curtis and Jaccard dissimilarity ###
phylo_choked.df <- data.frame(sample_data(phylo_choked))
sampledist_choked_bray <- phyloseq::distance(phylo_choked, method = "bray")
sampledist_choked_UNI <- phyloseq::distance(phylo_choked, method = "wunifrac")
sampledist_choked_jacc <- phyloseq::distance(phylo_choked, method = "jaccard")

### Convert distances from phyloseq objects to matrices ###
bray_choked <- as.matrix(sampledist_choked_bray)
UNI_choked <- as.matrix(sampledist_choked_UNI)
```

```
jacc_choked <- as.matrix(sampledist_choked_jacc)
dist_bray_choked_mtx <- as.dist(as(bray_choked, "matrix"))
dist_UNI_choked_mtx <- as.dist(as(UNI_choked, "matrix"))
dist_jacc_choked_mtx <- as.dist(as(jacc_choked, "matrix"))

### test for heteroscedasticity in the data ###
homogeneity <- betadisper(dist_bray_choked_mtx, metadata_choked$sample_growth)
permutest(homogeneity, pairwise = TRUE)
```

```
##
## 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   3 0.42253 0.140844 11.243   999 0.001 ***
## Residuals 55 0.68897 0.012527
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Pairwise comparisons:
## (Observed p-value below diagonal, permuted p-value above diagonal)
##      artificial      seawater zostera_new zostera_old
## artificial              0.967000000 0.001000000      0.005
## seawater      0.969629502              0.001000000      0.004
## zostera_new 0.000020651 0.000067519              0.376
## zostera_old 0.002064833 0.003334703 0.375445337
```

```
#boxplot(homogeneity)
```

```
homogeneity <- betadisper(dist_UNI_choked_mtx, metadata_choked$sample_growth)
permutest(homogeneity, pairwise = TRUE)
```

```
##
## 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   3 0.062009 0.0206696 7.6437   999 0.003 **
## Residuals 55 0.148727 0.0027041
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Pairwise comparisons:
## (Observed p-value below diagonal, permuted p-value above diagonal)
##      artificial      seawater zostera_new zostera_old
## artificial              0.259000000 0.006000000      0.016
## seawater      0.27229566              0.001000000      0.001
## zostera_new 0.00812457 0.00031821              0.675
## zostera_old 0.02327334 0.00155221 0.68578051
```

```

#boxplot(homogeneity)

homogeneity <-betadisper(dist_jacc_choked_mtx ,metadata_choked$sample_growth)
permutest(homogeneity, pairwise = TRUE)

##
## 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      3 0.26594 0.088646 12.398   999  0.001 ***
## Residuals  55 0.39326 0.007150
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Pairwise comparisons:
## (Observed p-value below diagonal, permuted p-value above diagonal)
##           artificial      seawater  zostera_new zostera_old
## artificial              0.8330000000 0.0010000000      0.002
## seawater      0.8445627460              0.0010000000      0.005
## zostera_new 0.0000023338 0.0000309397              0.302
## zostera_old 0.0016194344 0.0030342657 0.2985231996

#boxplot(homogeneity)

### Because we have heteroscedasticity, we should have the same n so PERMANOVA is robust ###
# GET LIST OF RANDOMLY SUBSET OF DATA FOR PHYLOSEQ FILTERING / PERMANOVA #
sum(metadata_choked$sample_growth == "seawater") #12 samples

## [1] 12
sum(metadata_choked$sample_growth == "zostera_new") #21 samples

## [1] 21
sum(metadata_choked$sample_growth == "zostera_old") #17 samples

## [1] 17
sum(metadata_choked$sample_growth == "artificial") # 9 samples

## [1] 9
### Randomly get 9 samples from all sample types so they have same n: ###
bact_random_choked <- metadata_choked %>%
  group_by(sample_growth) %>%
  sample_n(size = 9)

## See if it worked ###
sum(bact_random_choked$sample_growth == "seawater")

## [1] 9
sum(bact_random_choked$sample_growth== "zostera_new")

## [1] 9

```

```

sum(bact_random_choked$sample_growth == "zostera_old")

## [1] 9

sum(bact_random_choked$sample_growth == "artificial")

## [1] 9

### list of randomly subset samples ###
bact_random_list_choked <- bact_random_choked$SampleID
length(bact_random_list_choked)

## [1] 36

### Now use the list above to subsample data in phyloseq object ###
phylo_random_choked <- subset_samples(phylo_choked, sample_names(phylo_choked) %in% bact_random_list_choked)

### Calculate weighted Unifrac, Bray-Curtis dissimilarity ###
phylo_random_choked.df <- data.frame(sample_data(phylo_random_choked))
sampledist_random_bray <- phyloseq::distance(phylo_random_choked, method = "bray")
sampledist_random_UNI <- phyloseq::distance(phylo_random_choked, method = "wunifrac")
sampledist_random_jacc <- phyloseq::distance(phylo_random_choked, method = "jaccard")

### Convert distances from phyloseq objects to matrices ###
bray_random <- as.matrix(sampledist_random_bray)
UNI_random <- as.matrix(sampledist_random_UNI)
jacc_random <- as.matrix(sampledist_random_jacc)
dist_bray_random_mtx <- as.dist(as(bray_random, "matrix"))
dist_UNI_random_mtx <- as.dist(as(UNI_random, "matrix"))
dist_jacc_random_mtx <- as.dist(as(jacc_random, "matrix"))

#### get list of subsamples from metadata ###
choked_subsample <- metadata_choked %>%
  filter(SampleID %in% bact_random_list_choked)
choked_subsample$sample_growth <- factor(choked_subsample$sample_growth)

#####
### RANDOM samples (9 samples for each sample type) ###
#####

#####
#### PERMANOVA Bray Curtis ####
#####
adonis2(dist_bray_random_mtx ~ sample_growth,
        data= choked_subsample, permutations=10000, by = "margin")

## Permutation test for adonis under NA model
## Marginal effects of terms
## Permutation: free
## Number of permutations: 10000
##
## adonis2(formula = dist_bray_random_mtx ~ sample_growth, data = choked_subsample, permutations = 10000)
##              Df SumOfSqs      R2      F      Pr(>F)
## sample_growth  3   6.4184 0.49084 10.283 0.00009999 ***
## Residual      32   6.6579 0.50916
## Total        35  13.0763 1.00000

```

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

# pairwise comparisons
set.seed(3245)
adonis.pair(dist_bray_random_mtx, choked_subsample$sample_growth,
            nper = 10000, corr.method = "BH")

##              combination SumsOfSqs  MeanSqs  F.Model      R2
## 1    artificial <-> seawater 3.2314321 3.2314321 25.927859 0.6183922
## 2    artificial <-> zostera_new 2.2237379 2.2237379 10.559932 0.3975888
## 3    artificial <-> zostera_old 0.8470396 0.8470396  3.960743 0.1984266
## 4      seawater <-> zostera_new 2.6631364 2.6631364 13.166946 0.4514338
## 5      seawater <-> zostera_old 2.4766169 2.4766169 12.049588 0.4295816
## 6 zostera_new <-> zostera_old 1.3947721 1.3947721  4.785034 0.2302154
##      P.value P.value.corrected
## 1 0.00009999      0.000149985
## 2 0.00009999      0.000149985
## 3 0.00019998      0.000239976
## 4 0.00009999      0.000149985
## 5 0.00009999      0.000149985
## 6 0.00049995      0.000499950

#####
#### PERMANOVA Weighted Unifrac ####
#####
adonis2(dist_UNI_random_mtx ~ sample_growth,
        data= choked_subsample , permutations=10000, by = "margin")

## Permutation test for adonis under NA model
## Marginal effects of terms
## Permutation: free
## Number of permutations: 10000
##
## adonis2(formula = dist_UNI_random_mtx ~ sample_growth, data = choked_subsample, permutations = 10000
##              Df SumOfSqs      R2      F      Pr(>F)
## sample_growth  3  1.21427 0.64027 18.985 0.00009999 ***
## Residual      32  0.68224 0.35973
## Total         35  1.89651 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# pairwise comparisons
set.seed(3245)
adonis.pair(dist_UNI_random_mtx, choked_subsample$sample_growth,
            nper = 10000, corr.method = "BH")

##              combination SumsOfSqs  MeanSqs  F.Model      R2
## 1    artificial <-> seawater 0.4789498 0.4789498 44.560071 0.7357995
## 2    artificial <-> zostera_new 0.4854919 0.4854919 22.577663 0.5852522
## 3    artificial <-> zostera_old 0.0686746 0.0686746  3.081452 0.1614894
## 4      seawater <-> zostera_new 0.6763421 0.6763421 33.229715 0.6749930
## 5      seawater <-> zostera_old 0.3560962 0.3560962 16.847228 0.5128965
## 6 zostera_new <-> zostera_old 0.3629810 0.3629810 11.381724 0.4156686
##      P.value P.value.corrected
## 1 0.00009999      0.000149985
```

```
## 2 0.00009999      0.000149985
## 3 0.01209879      0.012098790
## 4 0.00009999      0.000149985
## 5 0.00009999      0.000149985
## 6 0.00049995      0.000599940

#####
#### PERMANOVA Jaccard ####
#####
adonis2(dist_jacc_random_mtx ~ sample_growth,
        data= choked_subsample , permutations=10000, by = "margin")

## Permutation test for adonis under NA model
## Marginal effects of terms
## Permutation: free
## Number of permutations: 10000
##
## adonis2(formula = dist_jacc_random_mtx ~ sample_growth, data = choked_subsample, permutations = 10000)
##              Df SumOfSqs      R2      F      Pr(>F)
## sample_growth  3   5.3671 0.36833 6.2197 0.00009999 ***
## Residual      32   9.2044 0.63167
## Total         35  14.5714 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# pairwise comparisons
set.seed(3245)
adonis.pair(dist_jacc_random_mtx, choked_subsample$sample_growth,
            nper = 10000, corr.method = "BH")

##              combination SumsOfSqs  MeanSqs  F.Model      R2
## 1 artificial <-> seawater 2.6575711 2.6575711 12.486156 0.4383237
## 2 artificial <-> zostera_new 1.8203658 1.8203658 6.185789 0.2788176
## 3 artificial <-> zostera_old 0.8805755 0.8805755 2.976182 0.1568378
## 4 seawater <-> zostera_new 2.1522277 2.1522277 7.703042 0.3249812
## 5 seawater <-> zostera_old 2.0539797 2.0539797 7.309743 0.3135918
## 6 zostera_new <-> zostera_old 1.1694048 1.1694048 3.226545 0.1678172
##      P.value P.value.corrected
## 1 0.00009999      0.000149985
## 2 0.00009999      0.000149985
## 3 0.00019998      0.000239976
## 4 0.00009999      0.000149985
## 5 0.00009999      0.000149985
## 6 0.00049995      0.000499950
```

## new and old growth seagrass leaves

to account for sampling from the same plant

```
phylo_NO_artificial
```

```
## phyloseq-class experiment-level object
## otu_table() OTU Table:      [ 1206 taxa and 140 samples ]
## sample_data() Sample Data:  [ 140 samples by 4 sample variables ]
## tax_table() Taxonomy Table: [ 1206 taxa by 7 taxonomic ranks ]
## phy_tree() Phylogenetic Tree: [ 1206 tips and 1205 internal nodes ]
```

```

### Remove seawater samples from phyloseq object ###
phylo_old_new <- subset_samples(phylo_NO_artificial,
                               sample_growth!="seawater")

### IMPORTANT !!! ###
### Re-order phyloseq object to run PERMANOVA ###
phylo_old_new <- set_sample_order(phylo_old_new, c('sample_growth'))
# check if re-ordering worked, set same order for metadata later
SampleID_reordered_old_new <- as.data.frame(sample_data(phylo_old_new)$SampleID)
#View(SampleID_reordered_old_new) #"zostera_new", "zostera_old"

### Remove seawater samples from metadata ###
remove_water <- c("seawater")
metadata_new_old <- metadata_no_artificial %>%
  dplyr::filter(!sample_growth %in% remove_water)
x <- c("zostera_new", "zostera_old")
metadata_new_old <- metadata_new_old %>%
  mutate(sample_growth = factor(sample_growth, levels = x)) %>%
  arrange(sample_growth)

### Calculate weighted Unifrac, Bray-Curtis dissimilarity ###
phylo_new_old.df <- data.frame(sample_data(phylo_old_new ))
sampledist_new_old_bray <- phyloseq::distance(phylo_old_new,
                                              method = "bray")
sampledist_new_old_UNI <- phyloseq::distance(phylo_old_new,
                                             method = "wunifrac")

### Convert distances from phyloseq objects to matrices ###
bray_new_old <- as.matrix(sampledist_new_old_bray)
UNI_new_old <- as.matrix(sampledist_new_old_UNI)
dist_bray_new_old_mtx <- as.dist(as(bray_new_old , "matrix"))
dist_UNI_new_old_mtx <- as.dist(as(UNI_new_old, "matrix"))

### test for heteroscedasticity in the data ###
homogeneity <- betadisper(dist_bray_new_old_mtx ,metadata_new_old$sample_growth)
permutest(homogeneity, pairwise = TRUE)

##
## 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      1 0.06017 0.060165 9.3292   999 0.003 **
## Residuals 103 0.66426 0.006449
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Pairwise comparisons:
## (Observed p-value below diagonal, permuted p-value above diagonal)
##           zostera_new zostera_old
## zostera_new                0.002
## zostera_old      0.002871

```



```

#boxplot(homogeneity)

homogeneity <-betadisper(dist_UNI_new_old_mtx ,metadata_new_old$sample_growth)
permutest(homogeneity, pairwise = TRUE)

##
## 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      1 0.023806 0.0238060 8.0018   999 0.011 *
## Residuals 103 0.306435 0.0029751
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Pairwise comparisons:
## (Observed p-value below diagonal, permuted p-value above diagonal)
##           zostera_new zostera_old
## zostera_new                0.011
## zostera_old    0.0056176

#boxplot(homogeneity)

### Because we have heteroscedasticity, we should have the same n so PERMANOVA is robust ###
# GET LIST OF RANDOMLY SUBSET OF DATA FOR PHYLOSEQ FILTERING / PERMANOVA #
sum(metadata_new_old$sample_growth == "zostera_new") #54 samples

## [1] 54

sum(metadata_new_old$sample_growth == "zostera_old") #51 samples

## [1] 51

###Randomly get 51 samples from new and old zostera so they have same n:
bact_random_new_old <- metadata_new_old %>%
  group_by(sample_growth) %>%
  sample_n(size = 51)

## See if it worked ###
sum(bact_random_new_old$sample_growth== "zostera_new")

## [1] 51

sum(bact_random_new_old$sample_growth == "zostera_old")

## [1] 51

### list of randomly subset samples ###
bact_random_new_old_list <- bact_random_new_old$SampleID
length(bact_random_new_old_list)

## [1] 102

### Now use the list above to subsample data in phyloseq object ###
phylo_random_new_old <- subset_samples(phylo_old_new,
                                       sample_names(phylo_old_new) %in%

```

```

bact_random_new_old_list)

### Calculate weighted Unifrac, Bray-Curtis dissimilarity ###
phylo_random_new_old.df <- data.frame(sample_data(phylo_random_new_old))
sampledist_random_new_old_bray <- phyloseq::distance(phylo_random_new_old,
                                                    method = "bray")
sampledist_random_new_old_UNI <- phyloseq::distance(phylo_random_new_old,
                                                    method = "wunifrac")

### Convert distances from phyloseq objects to matrices ###
bray_random_new_old <- as.matrix(sampledist_random_new_old_bray)
UNI_random_new_old <- as.matrix(sampledist_random_new_old_UNI)
dist_bray_random_new_old_mtx <- as.dist(as(bray_random_new_old , "matrix"))
dist_UNI_random_new_old_mtx <- as.dist(as(UNI_random_new_old , "matrix"))

#### get list of subsamples from metadata ###
new_old_subsample <- metadata_new_old %>%
  filter(SampleID %in% bact_random_new_old_list)

#####
### RANDOM samples (51 samples for each sample type) ###
#####

#####
#### PERMANOVA Bray Curtis ####
#####
adonis2(dist_bray_random_new_old_mtx ~ sample_growth + individuals,
        data= new_old_subsample,
        permutations=10000, by = "margin")

## Permutation test for adonis under reduced model
## Marginal effects of terms
## Permutation: free
## Number of permutations: 10000
##
## adonis2(formula = dist_bray_random_new_old_mtx ~ sample_growth + individuals, data = new_old_subsample)
##           Df SumOfSqs      R2      F      Pr(>F)
## sample_growth  1    4.326 0.11857 15.3717 0.00009999 ***
## individuals   57   19.049 0.52217  1.1876  0.0032 **
## Residual      43   12.100 0.33169
## Total        101   36.481 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#####
#### PERMANOVA Weighted Unifrac ####
#####
adonis2(dist_UNI_random_new_old_mtx ~ sample_growth + individuals,
        data= new_old_subsample,
        permutations=10000, by = "margin")

## Permutation test for adonis under reduced model
## Marginal effects of terms
## Permutation: free

```

```

## Number of permutations: 10000
##
## adonis2(formula = dist_UNI_random_new_old_mtx ~ sample_growth + individuals, data = new_old_subsampl
##           Df SumOfSqs      R2      F      Pr(>F)
## sample_growth  1   1.3999 0.29250 49.8261 0.00009999 ***
## individuals   57   1.9229 0.40179  1.2007   0.09199 .
## Residual      43   1.2081 0.25243
## Total        101   4.7858 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```