

Community structures

@rvosa

2024-12-04

Bray-Curtis dissimilarity

Here we define a function to read an ASV table and compute Bray-Curtis dissimilarity:

```
# Read the TSV files
df_el <- read_tsv('../data/elements_ESV_table.tsv/elements_ESV_table.tsv')

## Rows: 15354 Columns: 486
## -- Column specification -----
## Delimiter: "\t"
## chr   (3): temporary_ID, unique_ID, Seq
## dbl (483): S1084, S1302, S605, S1280, S179, S1110, S1113, S222, S159, S1258,...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
df_il <- read_tsv('../data/NextSeq_ESV_table.tsv/NextSeq_ESV_table.tsv')

## Rows: 34535 Columns: 486
## -- Column specification -----
## Delimiter: "\t"
## chr   (3): temporary_ID, unique_ID, Seq
## dbl (483): S661, S682, S767, S1118, S681, S1133, S113, S1132, S277, S1125, S...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.

# Filter columns that sum to 0
df_el <- df_el[, sapply(df_el, function(x) is.numeric(x) && sum(x) != 0)]
df_il <- df_il[, sapply(df_il, function(x) is.numeric(x) && sum(x) != 0)]

# Get common column names
common_samples <- intersect(colnames(df_el), colnames(df_il))

# Subset both DFs
df_el <- df_el[, common_samples]
df_il <- df_il[, common_samples]

# Extract just the sample columns (those starting with "ITS")
sample_cols <- grep("^S[0-9]+", colnames(df_el), value = TRUE)

# Subset to just the samples
el_mat <- as.matrix(df_el[, sample_cols])
il_mat <- as.matrix(df_il[, sample_cols])
```

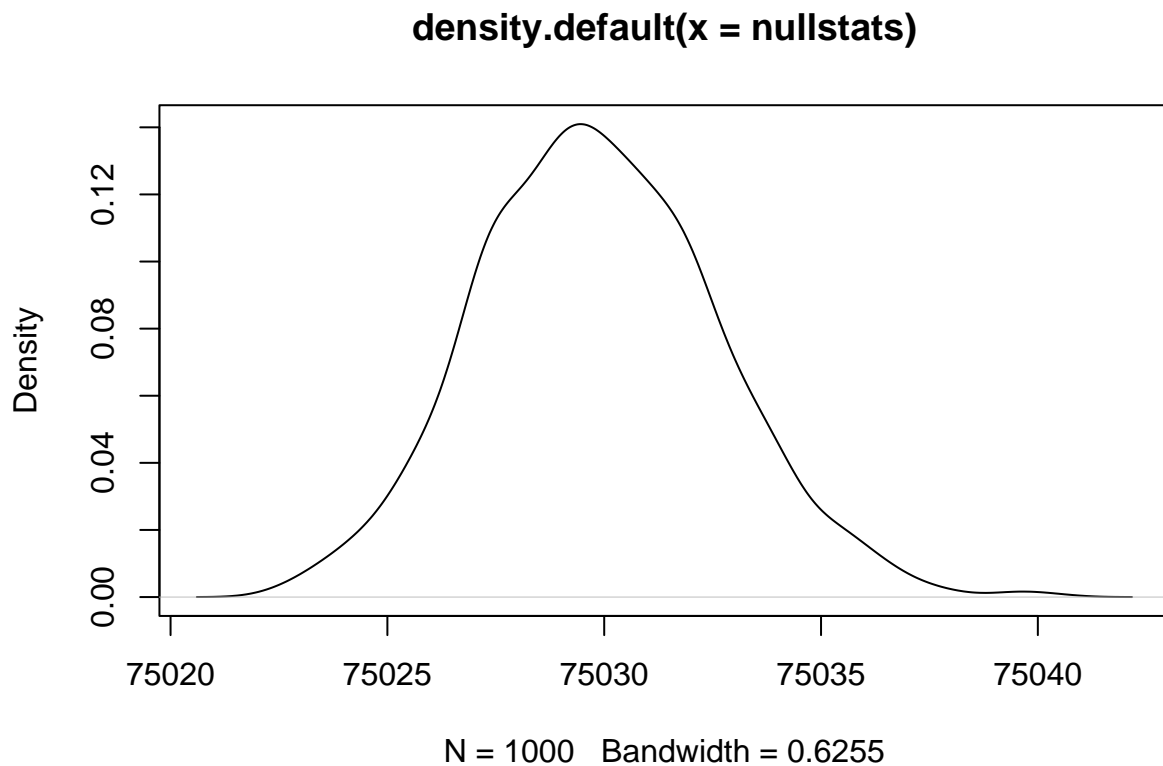
```

# Convert to proper format for vegan
# Rows should be samples, columns should be ASVs
el_mat <- t(el_mat)
il_mat <- t(il_mat)

# Calculate Bray-Curtis dissimilarity
el_dist <- as.matrix(vegdist(el_mat, method = "bray"))
il_dist <- as.matrix(vegdist(il_mat, method = "bray"))

# DO the test
mantel.test(el_dist,il_dist, nperm = 1000,graph = T)

```



```

## $z.stat
## [1] 75716.38
##
## $p
## [1] 0.000999001
##
## $alternative
## [1] "two.sided"

```