Package ‘phyloseqGraphTest’
February 7, 2020

Title Graph-Based Permutation Tests for Microbiome Data
Version 0.1.0
Author Julia Fukuyama [aut, cre]
Maintainer Julia Fukuyama <julia.fukuyama@gmail.com>
Description Provides functions for graph-based multiple-sample testing and visualization of microbiome data, in particular data stored in ‘phyloseq’ objects. The tests are based on those described in Friedman and Rafsky (1979) <http://www.jstor.org/stable/2958919>, and the tests are described in more detail in Callahan et al. (2016) <doi:10.12688/f1000research.8986.1>.
Imports ggnetwork (>= 0.5.1), igraph (>= 1.1.2), intergraph (>= 2.0.2)
Depends R (>= 3.5.0), ggplot2 (>= 2.2.1), phyloseq (>= 1.24.0)
License CC0
LazyData true
Suggests knitr, rmarkdown
VignetteBuilder knitr
URL https://github.com/jfukuyama/phyloseqGraphTest
biocViews
RoxygenNote 6.0.1
NeedsCompilation no
Repository CRAN
Date/Publication 2020-02-07 16:30:02 UTC

R topics documented:

phyloseqGraphTest-package .................................................. 2
graph_perm_test ......................................................... 2
plot_permutations .................................................... 3
plot_test_network ..................................................... 4
print.psgraphtest ...................................................... 4
**graph_perm_test**

**Description**

This package lets you test for differences between groups of samples with a graph-based permutation test.

**Details**

The main function in the package is `graph_perm_test`, which takes a phyloseq object. The graph used in the test can be visualized using `plot_test_network`. The permutation distribution and the test statistic can be visualized with `plot_permutations`.

**Usage**

```r
graph_perm_test(physeq, sampletype, grouping = 1:nsamples(physeq),
                distance = "jaccard", type = c("mst", "knn", "threshold.value",
                                            "threshold.nedges"), max.dist = 0.4, knn = 1, nedges = nsamples(physeq),
                keep.isolates = TRUE, nperm = 499)
```

**Arguments**

- **physeq**
  - A phyloseq object.

- **sampletype**
  - A string giving the column name of the sample to be tested. This should be a factor with two or more levels.

- **grouping**
  - Either a string with the name of a sample data column or a factor of length equal to the number of samples in physeq. These are the groups of samples whose labels should be permuted and are used for repeated measures designs. Default is no grouping (each group is of size 1).

- **distance**
  - A distance, see `distance` for a list of the possible methods.

- **type**
  - One of "mst", "knn", "threshold". If "mst", forms the minimum spanning tree of the sample points. If "knn", forms a directed graph with links from each node to its k nearest neighbors. If "threshold", forms a graph with edges between every pair of samples within a certain distance.
max.dist For type "threshold", the maximum distance between two samples such that we put an edge between them.
knn For type "knn", the number of nearest neighbors.
nedges If using "threshold.nedges", the number of edges to use.
keep.isolates In the returned network, keep the unconnected points?
nperm The number of permutations to perform.

Value
A list with the observed number of pure edges, the vector containing the number of pure edges in each permutation, the permutation p-value, the graph used for testing, and a vector with the sample types used for the test.

Examples
library(phyloseq)
data(enterotype)
gt = graph_perm_test(enterotype, sampletype = "SeqTech", type = "mst")

plot_permutations

Description
Plots a histogram of the permutation distribution of the number of pure edges and a mark showing the observed number of pure edges.

Usage
plot_permutations(graphtest, bins = 30)

Arguments
graphtest The output from graph_perm_test.
bins The number of bins to use for the histogram.

Value
A ggplot object.

Examples
library(phyloseq)
data(enterotype)

plot_permutations(graph_perm_test(enterotype, sampletype = "SeqTech"))
plot_test_network

Plots the graph used for testing

Description
When using the graph_perm_test function, a graph is created. This function will plot the graph used for testing with nodes colored by sample type and edges marked as pure or mixed.

Usage
plot_test_network(graphtest)

Arguments
graphtest The output from graph_perm_test.

Value
A ggplot object created by ggnetwork.

Examples
library(phyloseq)
data(enterotype)
gt = graph_perm_test(enterotype, sampletype = "SeqTech")
plot_test_network(gt)

print.psgraphtest

Print psgraphtest objects

Description
Print psgraphtest objects

Usage
## S3 method for class 'psgraphtest'
print(x, ...)

Arguments
x psgraphtest object.
... Not used
Index

distance, 2

graph_perm_test, 2, 2

phyloseq, 2
phyloseqGraphTest-package, 2
plot_permutations, 2, 3
plot_test_network, 2, 4
print.psgraphtest, 4