Package 'miMediation'

December 20, 2022

Version 0.1

Date 2022-12-13

Title Mediation Test for Microbime Data

Description A phylogeny-based test of mediation effect in microbiome.

Author Qilin Hong

Maintainer Qilin Hong <qhong8@wisc.edu>

Depends R (>= 3.5.0)

Encoding UTF-8

License GPL-3

LazyData true

biocViews

Imports ape,
  data.table,
  data.tree,
  fdrtool,
  ggplot2,
  ggtree,
  harmonicmeanp,
  MASS,
  phyloseq,
  SKAT,
  TreeTools

Suggests knitr,
  prettydoc,
  rmarkdown

RoxygenNote 7.2.2

VignetteBuilder knitr,
  prettydoc

R topics documented:

1

---

data.cecal   *Real data*

Description

This data is derived from a real microbiome dataset (Cho et al., 2012), which includes 48 samples (38 in antibiotics vs. 10 in controls).

Usage

```
data(data.cecal)
```

Format

data.cecal contains the following components:

treatment treatment indicator: Antibiotics group and control group are coded as 1 and 0, respectively.

mediators an abundance matrix with the top 100 most abundant taxa that have at least 20% non-zero observations.

outcome body fat percentage.

tree a phylogenetic tree.

Source

https://doi.org/10.1038/nature11400

References

Cho, I. et al. (2012). Antibiotics in early life alter the murine colonic microbiome and adiposity. *Nature* 488:621-626.

Examples

```
data(data.cecal)
```

---

data.zeeviD   *Real data*

Description

This data is derived from a real microbiome dataset (Zeevi, D. et al., 2015), which includes microbiome samples from 200 healthy subjects.

Usage

```
data(data.zeeviD)
```

Format

data.zeeviD contains the following components:

treatment treatment indicator.

mediators an abundance matrix with the top 100 most abundant taxa.

outcome continuous outcome.

tree a taxonomy table.

Source

https://doi.org/10.1016/j.cell.2015.11.001

References

Zeevi, D. et al. (2015). Personalized nutrition by prediction of glycemic responses. *Cell* 163:1079-1094.

Examples

```
data(data.zeeviD)
```

---

phyloMed   *Phylogeny-based test of mediation effect in microbiome (PhyloMed)*

Description

phyloMed enables us to test the mediation effect in high-dimensional microbial composition. The method leverages the hierarchical phylogeny relationship among different microbial taxa to decompose the complex mediation model on the full microbial composition into multiple simple independent local mediation models on subcompositions. The phyloMed function (a) performs the mediation test for the subcomposition at each internal node of the phylogenetic tree or taxonomic tree and pinpoint the mediating nodes with significant test p-values; and (b) combine all subcomposition p-values to assess the overall mediation effect of the entire microbial community.

Usage

```
phyloMed(
  treatment,
  mediators,
  outcome,
  tree,
  pi.method = "product",
  confounders = NULL,
  interaction = FALSE,
  fdr.alpha = 0.05,
  n.perm = NULL,
  verbose = FALSE,
  graph = FALSE
)
```

Arguments

treatment	A numeric vector of the treatment.
mediators	A named numeric matrix containing microbiome abundance. Each row is a subject and each column is an OTU or a taxon. Row name contains the subject ID and column name contains the taxon name.
outcome	A numeric vector of continuous or binary outcome.
tree	A phylogenetic tree (phylo-class object) or a taxonomy table (matrix-class). The tip labels in the phylogenetic tree or the row names in the taxonomy table should overlap with the column names in the mediators matrix. The column names in the taxonomy table should start from the higher level to lower level, e.g., from kingdom to genus.
pi.method	An optional character string denotes the method to used in estimate proportion of null. Default method is "product", an alternative method is "maxp".
confounders	An optional numeric vector or matrix containing confounders that may affect the treatment, mediators and outcome. Each row is a subject and each column is a specific confounder, e.g., age or sex. Default is NULL.
interaction	An optional logical value. If TRUE, the interaction term between treatment and mediator will be taken into account. Default is FALSE.
fdr.alpha	An optional numeric value for the desired FDR significance level in identifying mediating nodes on the tree. Default is 0.05.
n.perm	An optional numeric value for the maximum number of permutations. Default is NULL. See Details.
verbose	An optional logical value. If TRUE, information of the test on each node will be printed. Default is FALSE.
graph	An optional logical value. If TRUE, generate a graph that contains a phylogenetic tree or taxonomic tree with identified mediating nodes highlighted. Default is FALSE. See Details.

Details

PhyloMed uses the treatment-mediator association test p-value and mediation-outcome association test p-value to construct the subcomposition mediation test statistic at each local model (Hong et al., Manuscript). The two p-values can come from either the asymptotic test or the permutation test. Asymptotic test is faster but less accurate when the study sample size is small. By default (n.perm=NULL), only asymptotic test will be performed. Otherwise, if n.perm is set to a positive number, results from two versions of PhyloMed will be output, one based on the asymptotic p-value and the other based on the permutation p-value. Graph only highlights the mediating nodes identified from permutation version when both versions are performed.

Value

A phyloseq-class object named clean.data and a list named rslt.

clean.data contains the following components:

sample_data	Input treatment, outcome and confounders.
otu_table	The abundance data for the taxa that are present on the tips of the phy_tree.
tax_table	The taxonomy table with rows exactly match the taxa in the otu_table.
phy_tree	The binary and rooted phylogenetic tree with tips exactly match the taxa in the otu_table.

If `n.perm` is not NULL, the function will return two lists in `rslt` named `PhyloMed.A` and `PhyloMed.P`, respectively. Otherwise, only one list named `PhyloMed.A` will be returned.

Each list contains the following components:

`node.pval`	A numeric vector of subcomposition mediation p-values for all internal nodes.
`sig.clade`	A list of significant nodes with their descendants.
`null.prop`	A vector of the estimated proportion of different types of null hypotheses across all local mediation tests.
`global.pval`	A global test p-value using harmonic mean.

If graph is TRUE, the phylogenetic or taxonomic tree will be plot. The size of the circle at each internal node is proportional to $-\log_{10}$(subcomposition p-value), the larger circle indicates a smaller p-value. The significant nodes are highlighted by blue rectangle.

Author(s)

Qilin Hong <qhong8@wisc.edu>

References

Hong, Q., Chen G., and Tang Z-Z.. PhyloMed: a phylogeny-based test of mediation effect in microbiome. Manuscript.

See Also

prepareTree

Examples

```
# Load real data
data(data.cecal)
# Run test
Trt = data.cecal$treatment
M = data.cecal$mediators
Y = data.cecal$outcome
tree = data.cecal$tree
rslt.phylomed = phyloMed(Trt, M, Y, tree, graph = TRUE)
# Run test
Trt = data.zeeviD$treatment
M = data.zeeviD$mediators
Y = data.zeeviD$outcome
tree = data.zeeviD$tree
rslt.phylomed = phyloMed(Trt, M, Y, tree, graph = TRUE)
```

prepareTree   *Prepare phylogenetic tree to be used in* phyloMed *function*

Description

The phyloMed function requires a rooted and binary tree as input. The prepareTree is an utility function that allows users to preprocess their tree if it is unrooted and/or non-binary. The tree output from prepareTree function can be directly fed into phyloMed function.

Usage

```
prepareTree(tree, verbose = FALSE)
```

Arguments

tree	A phylo-class object.
verbose	An optional logical value. If TRUE, tree manipulation process will be printed. Default is FALSE.

Value

A binary and rooted tree with nodes and tips conform with following phylo standards:

- The internal nodes are numbered with value larger than the number of tips.
- The internal nodes are numbered sequentially, with values increasing away from the root.

Author(s)

Qilin Hong <qhong8@wisc.edu>

Examples

```
# Load real data
data(data.cecal)
tree = prepareTree(data.cecal$tree)
```

Index