

Package ‘DYNATE’

October 14, 2022

Title Dynamic Aggregation Testing

Version 0.1

Description A multiple testing procedure aims to find the rare-variant association regions. When variants are rare, the single variant association test approach suffers from low power. To improve testing power, the procedure dynamically and hierarchically aggregates smaller genome regions to larger ones and performs multiple testing for disease associations with a controlled node-level false discovery rate. This method are members of the family of ancillary information assisted recursive testing introduced in Pura, Li, Chan and Xie (2021) <[arXiv:1906.07757v2](#)> and Li, Sung and Xie (2021) <[arXiv:2103.11085v2](#)>.

License GPL-3

Encoding UTF-8

RoxygenNote 7.1.2.9000

Imports data.table, tidyverse, Matrix, reshape2, stats, methods,
tibble, dplyr

Suggests knitr, rmarkdown

VignetteBuilder knitr

Depends R (>= 3.5.0)

LazyData true

NeedsCompilation no

Author Xuechan Li [aut, cre],
Jichun Xie [ctb]

Maintainer Xuechan Li <xuechanli0530@gmail.com>

Repository CRAN

Date/Publication 2022-10-14 10:05:21 UTC

R topics documented:

DYNATE	2
p_leaf	3
snp_dat	3
Test_Leaf	3

DYNATE	<i>DYNATE Function to conduct hierarchical multiple testing based on the leaf p-values</i>
--------	--------------------------------------------------------------------------------------------

Description

DYNATE Function to conduct hierarchical multiple testing based on the leaf p-values

Usage

```
DYNATE(struct_map, L = 5, alpha = 0.05)
```

Arguments

struct_map	a data frame with both leaf information and P-value information.
L	maximum number of layers
alpha	desired FDR

Value

a data frame with testing results.

References

Li, Xuechan, Anthony Sung, and Jichun Xie. "Distance Assisted Recursive Testing." arXiv preprint arXiv:2103.11085 (2021). Pura, John, et al. "TEAM: A Multiple Testing Algorithm on the Aggregation Tree for Flow Cytometry Analysis." arXiv preprint arXiv:1906.07757 (2019).

Examples

```
data("p_leaf")

# Set tuning parameters
L <- 3 # layer number
alpha <- 0.05 # desired FDR

# conduct dynamic and hierarchical testing based on the leaf level p values.
out <- DYNATE(struct_map=p_leaf,L=L,alpha=alpha)
summary(out)
```

p_leaf	<i>p_leaf</i>
--------	---------------

Description

p_leaf

Usage

p_leaf

Format

A data frame with 16281 rows and 5 variables. Each row links to a SNP that belongs to a leaf with testing p-value<1.

snp_dat	<i>snp_dat</i>
---------	----------------

Description

snp_dat

Usage

snp_dat

Format

A data frame with 210454 rows and 6 variables.

Test_Leaf	<i>Test_Leaf</i> The function is used to generate Leaf P-values for case-control study. Users can input the leaf information through argument <i>struct_map</i> . If there is not leaf information e.g. <i>struct_map=NULL</i> , <i>Test_Leaf</i> will automatically construct leaf. Argument <i>thresh_val</i> specifies the leaf size constructed from the function. When the argument <i>Gmat_case</i> and <i>Gmat_ctrl</i> is null, <i>Test_Leaf</i> will automatically generate those matrices.
-----------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Description

Test_Leaf The function is used to generate Leaf P-values for case-control study. Users can input the leaf information through argument struct_map. If there is not leaf information e.g. struct_map=NULL, Test_Leaf will automatically construct leaf. Argument thresh_val specifies the leaf size constructed from the function. When the argument Gmat_case and Gmat_ctrl is null, Test_Leaf will automatically generate those matrices.

Usage

```
Test_Leaf(snp_dat = NULL, thresh_val = 10, covars = NULL, teststat = "FET")
```

Arguments

snp_dat	an optional data frame containing patients mutation information. If snp_dat=NULL, the mutation information should be taken from Gmat_case, Gmat_ctrl and glm_input. See vignettes for detail.
thresh_val	a positive integer for leaf size.
covars	an optional vector about the name of covariates to be considered in the fitting process. Should be NULL (default) or a character vector.
teststat	the statistic used to derive p-value. Must be one of "FET" (default) or "score".

Value

a dataframe of rejected leafs with snp information.

Examples

```
data("snp_dat")

# Set leaf size M
M <- 5

#Construct leaves and generate leaf p-value.
p.leaf <- Test_Leaf(snp_dat=snp_dat,thresh_val=M)
summary(p.leaf)
```

Index

* datasets

p_leaf, 3

snp_dat, 3

DYNATE, 2

p_leaf, 3

snp_dat, 3

Test_Leaf, 3