Package 'permPATH'

February 10, 2021

Type Package

Title Permutation Based Gene Expression Pathway Analysis

Version 1.2

Description Can be used to carry out permutation based gene expression pathway analysis. This work was supported by a National Institute of Allergy and Infectious Disease/National Institutes of Health contract (No. HHSN272200900059C).

Depends R (>= 3.4.0), R2HTML (>= 2.3.2), xtable (>= 1.8-2)

VignetteBuilder knitr

Suggests knitr

License GPL-3

LazyLoad yes

NeedsCompilation yes

Author Ivo D. Shterev [aut, cre],
Kouros Owzar [aut],
Gregory D. Sempowski [aut],
Kenneth Wilder [ctb, cph] (wrote original version of ranker.h)

Maintainer Ivo D. Shterev <i.shterev@gmail.com>

Repository CRAN

Date/Publication 2021-02-10 10:50:02 UTC

R topics documented:

permPath-package *Permutation Based Gene Expression Pathway Analysis.*

Description

Can be used to carry out permutation based gene expression pathway analysis. This work was supported by a National Institute of Allergy and Infectious Disease/National Institutes of Health contract (No. HHSN272200900059C).

Details

Package:	permPath
Type:	Package
Version:	0.6
Date:	2016-05-13
License:	GPL-3

Author(s)

I. D. Shterev, K. Owzar and G. D. Sempowski

Maintainer: I. D. Shterev <i.shterev@duke.edu>

References

B. Efron, R. Tibshirani (2007) On Testing the Significance of Sets of Genes. *The Annals of Applied Statistics*. **Vol. 1**, No 1, 107–129.

A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander and J. P. Mesirov (2005), Gene Set Enrichment Analysis: A knowledge-based Approach for Interpreting Genome-Wide Expression Profiles. *Proc. Natl. Acad. Sci. USA*. **Vol. 102**, No 43, 15545–15550.

perm.path *Perform Permutation Based Pathway Analysis*

Description

This is the package main function.

Usage

```
perm.path(expr, y, local.test, global.test="wilcoxon", B, gset, min.num=2, max.num,
imputeval=NULL, transfun=function(x){x}, sort="pval", anno=NULL)
```

Arguments

expr	An $K \times n$ matrix of gene expression data, where K is the number of genes and n is the number of samples.
y	An outcome vector of length n.
local.test	Local test statistic of each gene. Current possible choices are $t-test$, $Wilcoxon$ test, $Pearson$, $Spearman$ and JT test.
global.test	Global test statictic, used to compute the score. Current possible choices are $mean$, $meanabs$ (mean of absolute values) and $maxmean$.
B	specifies the number of random permutations to be performed.
gset	A list of pathways. Each element is a vector of gene names. The list element names are the pathway names.
min.num	Specifies the minimum number of genes that a pathway should have. Pathways with smaller number of genes will be excluded.
max.num	Specifies the maximum number of genes that a pathway should have. Pathways with larger number of genes will be excluded.
imputeval	The gene expression value to be imputed in case of missing values. The default choice is $NULL$ in which case no imputation is done.
transfun	Specifies transformation of the gene expression data. The default option is untransformed gene expression data.
sort	Specifies sorting of the results. If $sort = "pval"$ sorting is done in order of increasing $p - values$. If $sort = "score"$ sorting is done in order of decreasing $scores$.
anno	If $TRUE$ the output contains annotation of each pathway.

Value

This function returns a list consisting of the following elements:

res	Data frame consisting of the pathway names (Pathway), the genes involved in each pathway (Genes), the number of genes in each pathway (Size), the score for each pathway (Score), the permutation raw p-value (pval), the FWER-adjusted permutation p-value (pfwer), the FDR-adjusted permutation p-value, the Bonferroni-adjusted permutation p-value (bonferroni)
stats	The individual test statistic for each gene
scores	A matrix of scores. The matrix is of dimension $(B + 1) \times K$, where K is the number of pathways. The first column contains the unpermuted scores, the remaining B columns contain the scores computed after each permutation.

References

B. Efron, R. Tibshirani (2007) On Testing the Significance of Sets of Genes. *The Annals of Applied Statistics*. **Vol. 1**, No 1, 107–129.

A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander and J. P. Mesirov (2005), Gene Set Enrichment Analysis: A knowledge-based Approach for Interpreting Genome-Wide Expression Profiles. *Proc. Natl. Acad. Sci. USA*. **Vol. 102**, No 43, 15545–15550.

Examples

```
set.seed(1234)

## Generate toy phenotype and gene expression data sets
## This example consists of 40 genes grouped into 5 pathways and 100 patients
## grp is a binary trait (e.g., case vs control)
## bp is a continuous trait (e.g., blood pressure)
## g is a group indicator

n = 100
K = 40
grp = rep(1:0,each=n/2)
bp = rnorm(n)
g = rep(1:(n/20), rep(20,n/20))

pdat = data.frame(grp, bp, g)
rm(grp, bp)
expdat = matrix(rnorm(K*n),K,n)

## Assign marker names g1,...,gK to the expression data set and
## patient ids id1,...,idn to the expression and phenotype data
gnames = paste("g",1:K,sep="")
rownames(expdat) = gnames
patid = paste("id",1:n,sep="")
rownames(pdat) = patid
colnames(expdat) = patid

#Group the K genes into M pathways of sizes n1,...,nM
M = 5
p = runif(M)
p = p/sum(p)
nM = rmultinom(1, size=K, prob=p)
gset = lapply(nM, function(x){gnames[sample(x)]})
names(gset) = paste("pathway",1:M,sep="")

## Carry out permutation analysis with grp as the outcome
## using the two-sample Wilcoxon with B=100 random permutations
perm.path(expdat, y=pdat[["grp"]], local.test="wilcoxon", global.test="maxmean", B=100,
gset=gset, min.num=2, max.num=50, sort="score")

## Carry out permutation analysis with g as the outcome
## using the JT test with B=100 random permutations
perm.path(expdat, y=pdat[["g"]], local.test="jt", global.test="maxmean", B=100,
gset=gset, min.num=2, max.num=50, sort="score")
```

---

permPATH2HTML *This is a function for creating an HTML file*

Description

The function creates an HTML file.

Usage

```
permPATH2HTML(dat, dir, fname, title=NULL, bgcolor="#BBBBEE")
```

Arguments

dat	A data frame.
dir	Directory in which to store the file.
fname	File name.
title	The title of the html file.
bgcolor	Color for the html background.

Examples

```
## Generate toy phenotype and gene expression data sets
## This example consists of 40 genes grouped into 5 pathways and 100 patients
## grp is a binary trait (e.g., case vs control)
## bp is a continuous trait (e.g., blood pressure)
set.seed(1234)
n = 100
K = 40
grp = rep(1:0,each=n/2)
bp = rnorm(n)

pdat = data.frame(grp, bp)
rm(grp, bp)
expdat = matrix(rnorm(K*n),K,n)

## Assign marker names g1,...,gK to the expression data set and
## patient ids id1,...,idn to the expression and phenotype data
gnames = paste("g",1:K,sep="")
rownames(expdat) = gnames
patid = paste("id",1:n,sep="")
rownames(pdat) = patid
colnames(expdat) = patid

#Group the K genes into M pathways of sizes n1,...,nM
M = 5
p = runif(M)
p = p/sum(p)
nM = rmultinom(1, size=K, prob=p)
gset = lapply(nM, function(x){gnames[sample(x)]})
names(gset) = paste("pathway",1:M,sep="")

## Carry out permutation analysis with grp as the outcome
## using the two-sample Wilcoxon with B=100 random permutations
res = perm.path(expdat, y=pdat[["grp"]], local.test="wilcoxon", global.test="maxmean",
B=100, gset=gset, min.num=2, max.num=50, sort="score")

# create an html file
#epermPATH2HTML(rstab, dir="/dir/", fname="tophits")
```

Index