Package 'GOGANPA'

February 19, 2015

Type Package

Title GO-Functional-Network-based Gene-Set-Analysis

Version 1.0

Date 2011-12-30

Author Billy Chang

Maintainer Billy Chang <billy.chang@utoronto.ca>

Description
Accounting for genes' functional-non-equivalence within pathways in classical Gene-set-analysis.

Depends R (>= 2.10), GANPA, WGCNA

Suggests GANPA, WGCNA, GANPAdata

License GPL-2

Repository CRAN

Date/Publication 2012-01-03 07:44:00

NeedsCompilation no

R topics documented:

GOGANPA-package *GO-Functional-Network-based Gene-Set-Analysis*

Description

Accounting for genes' functional-non-equivalence within pathways in classical Gene-set-analysis.

Details

Package:	GOGANPA
Type:	Package
Version:	1.0
Depends:	GANPA, WGCNA
Suggests:	GANPA, WGCNA, GANPAdata
Date:	2011-12-30
License:	GPL-2

Author(s)

Billy Chang

Maintainer: Billy Chang <billy.chang@utoronto.ca>

References

Chang, B, Kustra, R, and Tian, WD, (2012) Functional-Network-Based Gene-Set-Analysis using Gene Ontology. Submitted.

Fang, ZY, Tian, WD, and Ji, HB. (September 6, 2011) A network-based gene-weighting approach for pathway analysis. Cell Research. Advanced Publication.

---

getGNET *Gene Network Construction by Similarity-Thresholding*

Description

Construct a gene network by linking gene-pairs with GO similarity above a chosen threshold.

Usage

```
getGNET(simMat, rho)
```

Arguments

`simMat`	The GO-similairty matrix. Missing and negative entries are not allowed. The gene names should be assigned to the row and column names.
`rho`	The threshold, chosen e.g. by `selectRho`. Gene-pairs with similarity above the threshold will be linked.

Value

A list, where each element contains the names of the genes connected to the corresponding gene indicated by the element-header.

Note

Note that certain GO-similarity measures are unbounded (e.g. the Resnik similarity). This code will not normalize the similarity matrix, and rho should therefore be chosen according to the range of the GO-similarity values inside simMat.

Author(s)

Billy Chang

References

Chang, B., Kustra, R. and Tian, WD (2012) Functional-Network-based Gene Set Analysis using Gene Ontology. Submitted.

Zhang, B. and Horvath, S. (2005) A General Framework for Weighted Gene Co-Expression Network Analysis. Statistical Applications in Genetics and Molecular Biology. 4:1:A17.

See Also

selectRho

Examples

```
#Not to Run
data("simMatSmall",package="GOGANPA")
gNET <- getGNET(simMatSmall,rho=0.7)
hist(sapply(gNET,length)) # network connectivities (excluding unconnected genes)
```

GOGANPA *GO-based Gene-Set-Analysis*

Description

A Gene-Set-Analysis method that accounts of functional-linkages among Genes, as indicated by a GO-functional network constructed from a GO-similarity matrix.

Usage

```
GOGANPA(gExprs.obj, gsets, gNET = NULL, simMat = NULL, rho = NULL, msp.groups,
check.exprs = TRUE, msp.correction = TRUE, size.min = 15, size.max = 500,
permN = 2000, randN = 30, permFDR.cutoff = 0.15, output.label = "GOGANPAResult")
```

Arguments

gExprs.obj	Gene expression experiment data object
gsets	A list of gene sets.
gNET	A gene association network stored in a list.
simMat	The GO-similairty matrix. Missing and negative entries are not allowed. The gene names should be assigned to the row and column names.
rho	The threshold, chosen e.g. by selectRho. If NULL, then chosen automatically by selectRho.
msp.groups	A list of multi-subunit-proteins-coding genes.
check.exprs	Logical (TRUE by default). Check and correct the missing values and scaling in the gExprs.obj. If the scale is natural, it will be converted to log2.
msp.correction	Logical (TRUE). Whether to do a correction for multi-subunit proteins in gene weighting.
size.min	Minimum size of gene sets used for analysis. By default 15 genes.
size.max	Maximum size of gene sets used for analysis. By default 500 genes.
permN	Sample permutation times. By default 2000 times.
randN	Gene randomization times. Can be set smaller (say, 30) if you do not care randomization-based significance so as to be faster.
permFDR.cutoff	Sample permutation FDR cutoff. A number between 0 and 1. Set it larger if wish to see the significance of more gene sets.
output.label	A label to name output files.

Details

Exactly one of gNET and simMat must be NULL. If simMat and rho are provided, getGNET will be called to obtain the gene network. If simMat is provided but rho is missing, then selectRho will also be called to provide an automatic choice of rho. This code is based on GANPA (Fang et. al. 2011), the gene network, gNET, whether supplied or derived from simMat, will be fed into GSE.Test.Main in the package GANPA for weighted Gene-Set-Analysis.

Value

A .csv file containing various statistics.

Author(s)

Billy Chang

References

Chang, B., Kustra, R. and Tian, WD (2012) Functional-Network-based Gene Set Analysis using Gene Ontology. Submitted.

Fang, ZY, Tian, WD, and Ji, HB. (September 6, 2011) A network-based gene-weighting approach for pathway analysis. Cell Research. Advanced Publication.

See Also

getGNET, selectRho

Examples

```
#Not to Run
require(GANPA)
data("simMatSmall",package="GOGANPA")
data("gExprs.p53", "gsets.msigdb.pnas", "msp.groups", package="GANPAdata")
set.seed(1000)
GOGANPA(gExprs.obj=gExprs.p53, gsets=gsets.msigdb.pnas, gNET=NULL, simMat=simMatSmall, rho=NULL,
                    msp.groups=msp.groups, check.exprs=TRUE, msp.correction=TRUE,
                    size.min=15, size.max=500, permN=2000, randN=30,
                    permFDR.cutoff=0.15,output.label="GOGANPAResult")
```

---

selectRho *Choosing a threshold based on the Scale-Free-Topology-Criterion*

Description

Determine the threshold parameter which will result in a network with optimal scale-free fitness.

Usage

```
selectRho(simMat, rhovec = NULL)
```

Arguments

simMat	The GO-similairty matrix. Missing and negative entries are not allowed. The gene names should be assigned to the row and column names.
rhovec	a vector of candidate thresholds, or if NULL, a set of thresholds chosen according to the range of the similarity matrix.

Details

The scale-free fitness measure is based on linear-regression-based R-squared goodness-of-fit measure.

Value

A list, with elements:

criterion	a summary table of the candidate thresholds' resulting fits.
bestrho	The candidate threshold with the highest R-squared.

Note

Note that certain GO-similarity measures are unbounded (e.g. the Resnik similarity). This code will not normalize the similarity matrix, and rhovec, if supplied, should be chosen according to the range of the GO-similarity values inside simMat.

Author(s)

Billy Chang

References

Chang, B., Kustra, R. and Tian, WD (2012) Functional-Network-based Gene Set Analysis using Gene Ontology. Submitted.

Zhang, B. and Horvath, S. (2005) A General Framework for Weighted Gene Co-Expression Network Analysis. Statistical Applications in Genetics and Molecular Biology. 4:1:A17.

See Also

getGNET

Examples

```
#Not to Run
data("simMatSmall",package="GOGANPA")
fit <- selectRho(simMatSmall)
plot(fit$criterion[,1],fit$criterion[,2])
abline(v=fit$bestrho,col=2)
```

---

simMatSmall *A Resnik Similarity Matrix.*

Description

A Resnik Similarity Matrix (normalized) for 2000 human genes sampled from 14173 annotated human genes.

Details

The similarity matrix was computed using the R package csbl.go (http://csbi.ltdk.helsinki.fi/csbl.go/), using a GO term specificity table computed using GO BP annotations for all human Entrez Genes available in the Bioconductor package org.Hs.eg.db, version 2.6.4 (not the default table provided csbl.go).

Note

This matrix is provided for test-running GOGANPA only. Although it is sampled from the similarity matrix used in Chang et. al. (2012), it cannot be used to reproduce the results presented in Chang et. al. (2012).

Author(s)

Billy Chang

References

Ovaska, K, Laakso M, and Hautaniemi, S. (2008) Fast Gene Ontology based clustering for microarray experiments. BioData Mining. 1:11.

Chang, B., Kustra, R. and Tian, WD (2012) Functional-Network-based Gene Set Analysis using Gene Ontology. Submitted.

Examples

```
#Not to Run
data("simMatSmall",package='GOGANPA')
```

Index