

Package ‘HTGM’

May 14, 2025

Version 1.2

Date 2025-05-14

Title High Throughput 'GoMiner'

Maintainer Barry Zeeberg <barryz2013@gmail.com>

Author Barry Zeeberg [aut, cre]

Depends R (>= 4.2.0)

Imports minimalistGODB, GoMiner, grDevices, stats, gplots, vprint

LazyData true

LazyDataCompression xz

Description Two papers published in the early 2000's (Zeeberg, B.R., Feng, W., Wang, G. et al. (2003) <[doi:10.1186/gb-2003-4-4-r28](https://doi.org/10.1186/gb-2003-4-4-r28)>) and (Zeeberg, B.R., Qin, H., Narashimhan, S., et al. (2005) <[doi:10.1186/1471-2105-6-168](https://doi.org/10.1186/1471-2105-6-168)>) implement 'GoMiner' and 'High Throughput GoMiner' ('HTGM') to map lists of genes to the Gene Ontology (GO) <<https://geneontology.org>>. Until recently, these were hosted on a server at The National Cancer Institute (NCI). In order to continue providing these services to the bio-medical community, I have developed stand-alone versions. The current package 'HTGM' builds upon my recent package 'GoMiner'. The output of 'GoMiner' is a heatmap showing the relationship of a single list of genes and the significant categories into which they map. 'High Throughput GoMiner' ('HTGM') integrates the results of the individual 'GoMiner' analyses. The output of 'HTGM' is a heatmap showing the relationship of the significant categories derived from each gene list. The heatmap has only 2 axes, so the identity of the genes are unfortunately ``integrated out of the equation." Because the graphic for the heatmap is implemented in Scalable Vector Graphics (SVG) technology, it is relatively easy to hyperlink each picture element to the relevant list of genes. By clicking on the desired picture element, the user can recover the ``lost" genes.

License GPL (>= 2)

Encoding UTF-8

VignetteBuilder knitr

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

RoxygenNote 7.3.2

Config/testthat/edition 3
NeedsCompilation no
Repository CRAN
Date/Publication 2025-05-14 17:00:02 UTC

Contents

Housekeeping_Genes	2
HTGM	3
htgmM	4
hyperGenes	5
hyperlinks	5
pasteHyperlinks	6
x_cn	7
x_cnc	7
x_l	7
x_m	8
x_rn	8
x_rnr	8
x_svg	8
x_svgr	9
Index	10

Housekeeping_Genes	<i>HTGM data set</i>
--------------------	----------------------

Description

HTGM data set

Usage

data(Housekeeping_Genes)

HTGM

*HTGM***Description**

driver to invoke GoMiner for multiple studies, and integrate the results in a categories versus study hyperlinked heatmap

Usage

```
HTGM(  
  title = NULL,  
  dir = tempdir(),  
  sampleLists,  
  GOGOA3,  
  ONT,  
  enrichThresh = 2,  
  countThresh = 5,  
  fdrThresh = 0.1,  
  nrand = 100,  
  mn = 2,  
  mx = 200,  
  opt = 0,  
  verbose = 1  
)
```

Arguments

title	character string descriptive title
dir	character string full pathname to the directory acting as result repository
sampleLists	list of character vector of user-supplied genes of interest
GOGOA3	return value of subsetGOGOA()
ONT	c("molecular_function","cellular_component","biological_process")
enrichThresh	numerical acceptance threshold for enrichment passed to GoMiner
countThresh	numerical acceptance threshold for gene count passed to GoMiner
fdrThresh	numerical acceptance threshold for fdr passed to GoMiner
nrand	integer number of randomizations passed to GoMiner
mn	integer param passed to trimGOGOA3, min size threshold for a category
mx	integer param passed to trimGOGOA3, max size threshold for a category
opt	integer 0:1 parameter used to select randomization method
verbose	integer parameter passed to vprint()

Value

returns the matrix of significant categories versus study

Examples

```
## Not run:
# GOGOA3.RData is too large to include in the R package
# so I need to load it from a file that is not in the package.
# Since this is in a file in my own file system, I could not
# include this as a regular example in the package.
# you can generate it using the package 'minimalistGODB'
# or you can retrieve it from https://github.com/barryzee/GO/tree/main/databases
load("/Users/barryzeeberg/personal/GODB_RDATA/goa_human/GOGOA3_goa_human.RData")

# load("data/Housekeeping_Genes.RData")
sampleList<-unique(as.matrix(Housekeeping_Genes[, "Gene.name"]))
n<-nrow(sampleList)
sampleLists<-list()
# test the effect of random sampling of the entire gene set
# this can give an idea of the quality of the GoMiner results
# when the complete gene set is yet to be determined
sampleLists[["1"]]<-sampleList[sample(n,n/2)]
sampleLists[["2"]]<-sampleList[sample(n,n/2)]
sampleLists[["3"]]<-sampleList[sample(n,n/2)]
sampleLists[["4"]]<-sampleList[sample(n,n/2)]
sampleLists[["5"]]<-sampleList[sample(n,n/2)]
sampleLists[["ALL"]]<-sampleList
m<-HTGM(title=NULL,dir=tempdir(),sampleLists,GOGOA3,ONT="biological_process",
  enrichThresh=2,countThresh=5,fdrThresh=0.10,nrand=100)

## End(Not run)
```

htgmM

htgmM

Description

generate FDR matrix of id versus cat

Usage

```
htgmM(l, fdrThresh)
```

Arguments

<code>l</code>	list of return values of GoMiner()
<code>fdrThresh</code>	numerical acceptance threshold for fdr

Value

returns numeric matrix `m` containing FDR values

Examples

```
# load("data/x_1.RData")
m<-htgmM(x_1,.1)
```

hyperGenes	<i>hyperGenes</i>
------------	-------------------

Description

populate subdirectory of hyperlinked gene lists

Usage

```
hyperGenes(l, dir)
```

Arguments

- l return value of GoMiner()
- dir character string containing path name of results directory

Value

returns no value but has side effect of populating subdirectory of hyperlinked gene lists

Examples

```
# x_1<-load("data/x_1.RData")
dir<-tempdir()
print(dir)
hyperGenes(x_1,dir)
```

hyperlinks	<i>hyperlinks</i>
------------	-------------------

Description

driver to add gene list hyperlinks to the HTGM heatmap

Usage

```
hyperlinks(s, rownames, colnames)
```

Arguments

s	character path name of the file containing the HTGM svg
rownames	character vector of row names
colnames	character vector of column names

Value

returns the path name of the file containing the hyperlinked HTGM svg

Examples

```
#load("data/x_rn.RData")
#load("data/x_cn.RData")
#load("data/x_svg.RData")
s<-system.file("extdata","x_htgm.svg",package="HTGM")
# need to avoid writing to "extdata"
dir<-tempdir()
file.copy(from=s, to=dir)
hyperlinkedFileName<-hyperlinks(sprintf("%s/%s",dir,"x_htgm.svg"),x_rn,x_cn)
print("hyperlinkedFileName")
print(hyperlinkedFileName)
```

pasteHyperlinks	<i>pasteHyperlinks</i>
-----------------	------------------------

Description

add gene list hyperlinks to the HTGM heatmap

Usage

```
pasteHyperlinks(str, c1, c2)
```

Arguments

str	character a line from the svg that is to have a hyperlink inserted
c1	character list of row names
c2	character list of column names

Value

returns a line of code to insert into svg

Examples

```
#load("data/x_svgr.RData")
#load("data/x_rnr.RData")
#load("data/x_cnc.RData")
hl<-pasteHyperlinks(x_svgr,x_rnr,x_cnc)
```

x_cn	HTGM data set
------	---------------

Description

HTGM data set

Usage

```
data(x_cn)
```

x_cnc	HTGM data set
-------	---------------

Description

HTGM data set

Usage

```
data(x_cnc)
```

x_l	HTGM data set
-----	---------------

Description

HTGM data set

Usage

```
data(x_l)
```

x_m	<i>HTGM data set</i>
-----	----------------------

Description

HTGM data set

Usage

data(x_m)

x_rn	<i>HTGM data set</i>
------	----------------------

Description

HTGM data set

Usage

data(x_rn)

x_rnr	<i>HTGM data set</i>
-------	----------------------

Description

HTGM data set

Usage

data(x_rnr)

x_svg	<i>HTGM data set</i>
-------	----------------------

Description

HTGM data set

Usage

data(x_svg)

`x_svgr`*HTGM data set*

Description

HTGM data set

Usage`data(x_svgr)`

Index

Housekeeping_Genes, [2](#)

HTGM, [3](#)

htgmM, [4](#)

hyperGenes, [5](#)

hyperlinks, [5](#)

pasteHyperlinks, [6](#)

x_cn, [7](#)

x_cnc, [7](#)

x_l, [7](#)

x_m, [8](#)

x_rn, [8](#)

x_rnr, [8](#)

x_svg, [8](#)

x_svgr, [9](#)