Package 'BloodGen3Module'

October 14, 2021

Type Package

Title This R package for performing module repertoire analyses and
generating fingerprint representations

Version 1.0.0

Date 2020-11-05

Description The BloodGen3Module package provides functions for R user performing module reper-
toire analyses
and generating fingerprint representations. Functions can perform group comparison or
individual sample analysis and visualization by fingerprint grid plot or fingerprint heatmap.
Module repertoire analyses typically involve determining the percentage of the constitutive genes
for each module that are significantly increased or decreased.
As we describe in de-
tails;https://www.biorxiv.org/content/10.1101/525709v2 and https://pubmed.ncbi.nlm.nih.gov/33624743/,
the results of module repertoire analyses can be represented in a fingerprint format,
where red and blue spots indicate increases or decreases in module activity.
These spots are subsequently represented either on a grid,
with each position being assigned to a given module,
or in a heatmap where the samples are arranged in columns and the modules in rows.

Depends R (>= 4.1)

License GPL-2

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.1

Imports SummarizedExperiment, ExperimentHub, methods, grid, graphics,
stats, grDevices, circlize, testthat, ComplexHeatmap(>=
1.99.8), ggplot2, matrixStats, gtools, reshape2,
preprocessCore, randomcoloR, V8, limma

biocViews Software, Visualization, GeneExpression

Suggests RUnit, devtools, BiocGenerics, knitr, rmarkdown

VignetteBuilder knitr

git_url https://git.bioconductor.org/packages/BloodGen3Module

git_branch RELEASE_3_13

git_last_commit 2f62a72

git_last_commit_date 2021-05-19

Date/Publication 2021-10-14

Author Darawan Rinchai [aut, cre] (<<https://orcid.org/0000-0001-8851-7730>>)

Maintainer Darawan Rinchai <drinchai@gmail.com>

R topics documented:

color	*Color for fingerprint visulization*

Description

Character vector of color for fingerprint grid plot

Usage

```
color
```

Format

A vector of 1134 character

fingerprintplot	*Individual fingerprint visualization The fingerprintplot function will generate fingerprint heatmap plots as a pdf file. The file will be saved in the working directory specified for the analysis. The default cut off for visualization is set at 15%, it can be changed to any value between 0-100%.*

Description

Individual fingerprint visualization The fingerprintplot function will generate fingerprint heatmap plots as a pdf file. The file will be saved in the working directory specified for the analysis. The default cut off for visualization is set at 15%, it can be changed to any value between 0-100%.

Usage

```
fingerprintplot(
  Individual_df,
  sample_info = NULL,
  cutoff = NULL,
  rowSplit = TRUE,
  Ref_group = NULL,
  show_ref_group = FALSE,
  Group_column = NULL,
  Aggregate = NULL,
  filename = NULL,
  height = NULL,
  width = NULL
)
```

Arguments

Individual_df	Dataframe with output generated after running the 'Individualcomparison' function
sample_info	A dataframe with sample annotation.
cutoff	Numeric value specifying the percentage cut off used for fingerprint visualization (0 to 100).
rowSplit	Logical operator (TRUE/FALSE) to indicate if rows of the heatmaps should be split by each aggregate
Ref_group	Characters name of reference group or samples that considered as control
show_ref_group	Character vector specifying value within the group column that will be used as Reference group
Group_column	Name of the columns for the groups used for the analysis
Aggregate	Character vector specifying name of specific module aggregates for heatmap fingerprint plot
filename	Character vector with a name for saving file

height   Sets the height of the graphics region in inches. The default values are 28

width   Sets the width of the graphics region in inches. The default values are 17

Value

A heatmap of % of module response in each single sample

Author(s)

Darawan Rinchai drinchai@gmail.com

Examples

```
## data could be downloaded from ExperimentHub("GSE13015")
library(ExperimentHub)
library(SummarizedExperiment)
dat = ExperimentHub()
res = query(dat , "GSE13015")
GSE13015 = res[["EH5429"]]
Individual_df = Individualcomparison(GSE13015, sample_info = NULL,
                                     FC = 1.5, DIFF = 10, Group_column = "Group_test",
                                     Ref_group = "Control")
fingerprintplot(Individual_df, sample_info = NULL,
                cutoff = 15, rowSplit = TRUE, Ref_group = "Control",
                show_ref_group = FALSE, Group_column = "Group_test",
                Aggregate = c("A28"), filename = tempfile(), height = 5,
                width = 10)
```

---

fold_change   *calculation of Fold-Change*

Description

A function to calculate fold-change between group comparison; "Test_group" vs "Ref_group"

Usage

```
fold_change(
  df_raw = df_raw,
  sample_info = sample_info,
  Group_column = Group_column,
  Test_group = Test_group,
  Ref_group = Ref_group
)
```

Arguments

df_raw	Matrix of normalized expression data (not Log2 transformed). Genes should be in rows and Sample ID in columns. Row names are required to be valid Gene Symbols
sample_info	A dataframe with sample annotation. Sample_info dataframe requires two columns: 1) a column specifying Sample ID (exactly matching Sample ID of data.matrix) and 2) a column specifying group names
Group_column	Character vector identical to the column name from sample_info dataframe that specifies group annotation used for the analysis
Test_group	Character vector specifying values within the group column (Group_column) that will be used as Test group (samples considered as cases or "intervention" group).
Ref_group	Character vector specifying value within the group column (Group_column) that will be used as Reference group

Value

A matrix of the fold change comparison between "Test_group" vs ""Ref_group"

Author(s)

Darawan Rinchai drinchai@gmail.com

Examples

```
## data could be downloaded from ExperimentHub("GSE13015")
library(ExperimentHub)
library(SummarizedExperiment)
dat = ExperimentHub()
res = query(dat , "GSE13015")
GSE13015 = res[["EH5429"]]
data_matrix = assay(GSE13015)
sample_ann = data.frame(colData(GSE13015))

FCgroup = fold_change(df_raw = data_matrix[c(1:5),],
                      sample_info = sample_ann,
                      Group_column = "Group_test",
                      Test_group="Sepsis",
                      Ref_group="Control")
```

Gen3_ann *Functional annotation for visulization*

Description

A dataframe contain module identifer, functional annotation and color in each specific module

Usage

```
Gen3_ann
```

Format

A data.frame with 5 variables: `Modules`: Module identifier, `Function`: Functional annotation, `Position`: position on fingerprint grid plot, `Module_color`: specific color of each module for visulization, `Cluster`: Module cluster membership

---

gridplot   *Fingerprint grid visualization*

Description

The gridplot function will generate a grid plot as a pdf file. Specific working directory for the analysis need to be specified for saving the file. The result of the plot should be return in the same working directory. The default cut off for visualization is set at 15%, it can be changed to any value between 0-100%.

Usage

```
gridplot(Group_df, cutoff = NULL, Ref_group = NULL, filename = NULL)
```

Arguments

`Group_df`	Dataframe with output generated after running the'Groupcomparison' function
`cutoff`	Numeric value specifying the percentage cut off used for fingerprint visualization (from 0 to 100)
`Ref_group`	Character vector specifying value within the group column that will be used as Reference group
`filename`	Character vector with a name for saving file

Value

```
A pdf file of grid plot
```

Author(s)

Darawan Rinchai drinchai@gmail.com

Examples

```
## data could be downloaded from ExperimentHub("GSE13015")
library(ExperimentHub)
library(SummarizedExperiment)
dat = ExperimentHub()
res = query(dat , "GSE13015")
GSE13015 = res[["EH5429"]]
Group_df = Groupcomparison(GSE13015, sample_info = NULL,
                           FC = 0, pval = 0.1, FDR = TRUE, Test_group = "Sepsis",
                           Group_column = "Group_test", Ref_group = "Control")
gridplot(Group_df, cutoff = 15,
         Ref_group = "Control",
         filename= tempfile())
```

Groupcomparison   *Group comparison analysis*

Description

The Groupcomparison function will perform group comparison analyses and the results are ex-pressed "at the module level" as percent of genes increased or decreased.

- Expression matrix and sample annotation files are required to perform this analysis.

- The sample annotation file must be loaded using a specific name = "sample_info".

- The names of the columns for the conditions used in the analysis must be specified.

Usage

```
Groupcomparison(
  data.matrix,
  sample_info = NULL,
  FC = NULL,
  pval = NULL,
  FDR = TRUE,
  Group_column = NULL,
  Test_group = "Test_group",
  Ref_group = "Control",
  SummarizedExperiment = TRUE
)
```

Arguments

data.matrix   Matrix of normalized expression data (not Log2 transformed).Row names are re-quired to be valid Gene Symbols. Columns names are sample IDs or data.matrix can also be given a summarizedexperiment object and assigned data.matrix and sample_info accordingly from the object.

sample_info	A dataframe with sample annotation. Sample_info dataframe requires two columns: 1) a column specifying Sample ID (exactly matching the Sample ID of data.matrix) and 2) a column specifying group names
FC	Numeric value specifying the foldchange cut off that will be applied to define increase or decrease of a given transcript compared to the reference group
pval	Numeric value specifying p-value cut off or False discovery rate when FDR = TRUE
FDR	Logical operator to specify whether False discovery rate cut off (using BH-method) should be used
Group_column	Character vector identical to the column name from sample_info dataframe that specifies group annotation used for the analysis
Test_group	Character vector specifying values within the group column (Group_column) that will be used as Test group (samples considered as cases or "intervention" group).
Ref_group	Character vector specifying value within the group column (Group_column) that will be used as Reference group
SummarizedExperiment	
	Output data as the SummarizedExperiment class when SummarizedExperiment = TRUE

Value

A matrix of the percentahe of module response in each group comparison

Author(s)

Darawan Rinchai drinchai@gmail.com

Examples

```
## data could be downloaded from ExperimentHub("GSE13015")
library(ExperimentHub)
library(SummarizedExperiment)
dat = ExperimentHub()
res = query(dat , "GSE13015")
GSE13015 = res[["EH5429"]]
Group_df = Groupcomparison(GSE13015, sample_info = NULL,
                        FC = 0, pval = 0.1, FDR = TRUE, Test_group = "Sepsis",
                        Group_column = "Group_test", Ref_group = "Control")
```

Groupcomparisonlimma   *Group comparison analysis using limma*

Description

The Groupcomparisonlimma function will perform group comparison analyses using "limma" function from "limma R package" and the results are expressed "at the module level" as percent of genes increased or decreased.

- Expression matrix and sample annotation files are required to perform this analysis.
- The sample annotation file must be loaded using a specific name = "sample_info".
- The names of the columns for the conditions used in the analysis must be specified.

Usage

```
Groupcomparisonlimma(
  data.matrix,
  sample_info = NULL,
  FC = NULL,
  pval = NULL,
  FDR = TRUE,
  Group_column = NULL,
  Test_group = "Test_group",
  Ref_group = "Control",
  SummarizedExperiment = TRUE
)
```

Arguments

data.matrix	Matrix of normalized expression data (not Log2 transformed).Row names are required to be valid Gene Symbols. Columns names are sample IDs or data.matrix can also be given a summarizedexperiment object and assigned data.matrix and sample_info accordingly from the object.
sample_info	A dataframe with sample annotation.
FC	Numeric value specifying the foldchange cut off that will be applied to define increase or decrease of a given transcript compared to the reference group
pval	Numeric value specifying p-value cut off or False discovery rate when FDR = TRUE
FDR	Logical operator to specify whether False discovery rate cut off (using BH-method) should be used
Group_column	Character vector identical to the column name from sample_info dataframe that specifies group annotation used for the analysis
Test_group	Character vector specifying value within the group column that will be used as Test group

Ref_group   Character vector specifying value within the group column that will be used as
                   Reference group

SummarizedExperiment

                   Output data as the SummarizedExperiment class when SummarizedExperiment
                   = TRUE

Value

              A matrix of the percentahe of module response in each group comparison

Author(s)

Darawan Rinchai drinchai@gmail.com

Examples

```
## data could be downloaded from ExperimentHub("GSE13015")
library(ExperimentHub)
library(SummarizedExperiment)
dat = ExperimentHub()
res = query(dat , "GSE13015")
GSE13015 = res[["EH5429"]]
Group_limma <- Groupcomparisonlimma(GSE13015, sample_info = NULL,
FC = 1.5, pval = 0.1, FDR = TRUE, Group_column = "Group_test",
Test_group = "Sepsis", Ref_group = "Control")
```

Individualcomparison   *Individual single sample analysis*

Description

The Individualcomparison function will perform individual sample comparison analysis in refer-
ence to a control sample or group of samples, with the results are expressed "at the module level"
as percent of genes increased or decreased.

Usage

```
Individualcomparison(
  data.matrix,
  sample_info = NULL,
  FC = NULL,
  DIFF = NULL,
  Group_column = NULL,
  Ref_group = NULL,
  SummarizedExperiment = TRUE
)
```

Arguments

`data.matrix`	Matrix of normalized expression data (not Log2 transformed).Row names are required to be valid Gene Symbols. Columns names are sample IDs or data.matrix can also be given a summarizedexperiment object and assigned data.matrix and sample_info accordingly from the object.
`sample_info`	A dataframe with sample annotation.
`FC`	Numeric value specifying the foldchange cut off that will be applied to define increase or decrease of a given transcript compared to the reference group
`DIFF`	Numeric value specifying the difference cut off that will be applied to define increase or decrease of a given transcript compared to the reference group
`Group_column`	Character vector identical to the column name from sample_info dataframe that specifies group annotation used for the analysis
`Ref_group`	Character vector specifying value within the group column that will be used as Reference group
`SummarizedExperiment`	
	Output data as the SummarizedExperiment class when SummarizedExperiment = TRUE

Details

- Expression matrix and sample annotation file are required in order to perform this analysis.

- The sample annotation file must be loaded using a specific name = "sample_info".

- The names of the columns for the conditions used in the analysis must be specified

- The default cutoff is set at FC =1.5 and DIFF =10

Value

A matrix of the percentahe of module response at individual level and SummarizedExperiment object

Author(s)

Darawan Rinchai drinchai@gmail.com

Examples

```
## data could be downloaded from ExperimentHub("GSE13015")
library(ExperimentHub)
library(SummarizedExperiment)
dat = ExperimentHub()
res = query(dat , "GSE13015")
GSE13015 = res[["EH5429"]]
Individual_df = Individualcomparison(GSE13015, sample_info = NULL,
                                    FC = 1.5, DIFF = 10, Group_column = "Group_test",
                                    Ref_group = "Control")
```

Module_listGen3   *Module identifer and list membership in each module*

Description

A dataframe contain gene member of 3rd generantion of blood module repertoire construction

Usage

```
Module_listGen3
```

Format

A data.frame with 14168 rows by 5 variables: `Module:` Module identifier, `Gene symbol:` gene membership, `Module_gene:` gene specific module membership, `Function:` Functional annotation, `position:` position on fingerprint grid plot

Index