

Package ‘TFactSR’

August 31, 2023

Type Package

Title Enrichment Approach to Predict Which Transcription Factors are Regulated

Version 0.99.0

Description R implementation of 'TFactS' to predict which are the transcription factors (TFs), regulated in a biological condition based on lists of differentially expressed genes (DEGs) obtained from transcriptome experiments. This package is based on the 'TFactS' concept by Essaghir et al. (2010) <[doi:10.1093/nar/gkq149](https://doi.org/10.1093/nar/gkq149)> and expands it. It allows users to perform 'TFactS'-like enrichment approach. The package can import and use the original catalogue file from the 'TFactS' as well as users' defined catalogues of interest that are not supported by 'TFactS' (e.g., Arabidopsis).

License GPL-3

URL <https://afukushima.github.io/TFactSR/>

Depends R (>= 4.3.0)

Imports qvalue, stats

Suggests BiocStyle, knitr, rmarkdown, spelling, testthat

VignetteBuilder knitr

biocViews Network, Software, DifferentialExpression, GeneTarget, GeneExpression, Microarray, RNASeq, Transcription, NetworkEnrichment

Encoding UTF-8

RoxygenNote 7.2.3

Language en-US

NeedsCompilation no

Author Atsushi Fukushima [aut, cre] (<<https://orcid.org/0000-0001-9015-1694>>)

Maintainer Atsushi Fukushima <afukushima@gmail.com>

Repository CRAN

Date/Publication 2023-08-31 14:00:08 UTC

R topics documented:

AtCatalog	2
calculateEnrichmentTest	3
calculateEvalue	4
calculateFDRBH	5
calculateQvalue	5
calculateRC	6
calculateTFactS	7
catalog	9
DEGs	9
DEGs39	10
example.df	10
example.list	11
extractTFTG	11
FASTcalculateRC	12
formatET	13
formatRC	14
GenesUp_SH1H	15
TFactSR	15
Index	16

AtCatalog	<i>A list of TF-target genes (TG) in Arabidopsis thaliana.</i>
-----------	--

Description

A catalogue of annotated target genes in Arabidopsis thaliana

Details

The original regulatory network datasets were downloaded from AtRegNet - AGRIS Arabidopsis (<http://arabidopsis.med.ohio-state.edu/moreNetwork.html>) and ATRM (<http://atrm.cbi.pku.edu.cn/>), respectively. Then we integrated them based on AGI code.

The catalogues version is 0.0.1.

Value

None

Author(s)

Atsushi Fukushima

References

Yilmaz A et al. NAR 39:D1118-1122 (2011)

```
calculateEnrichmentTest
  performs enrichment analysis
```

Description

This function performs enrichment test (ET) based on Fisher's exact test

Usage

```
calculateEnrichmentTest(
  DEGs,
  catalog,
  TFs,
  TF.col = "TF..OFFICIAL_TF_CODING_GENE_NAME."
)
```

Arguments

DEGs	a character vector of DEGs (differentially expressed genes)
catalog	a data frame of TFactS catalog (ver. 2)
TFs	a character vector of transcription factor
TF.col	the name of the column that contains the TF names

Value

data.frame

Author(s)

Atsushi Fukushima Definition: m is the number of target genes annotated for the TF under consideration n is the number of query genes N is the number of regulations in the catalog k is the number of query genes that are annotated as regulated by TF (i.e., the intersection between the query and the TF signature)

References

Essaghir A et al. Nucleic Acids Res. 2010 Jun;38(11):e120.

Examples

```
data(DEGs)
data(catalog)

tftg <- extractTFTG(DEGs, catalog)
TFs <- tftg$TFs
```

```
res <- calculateEnrichmentTest(DEGs, catalog, TFs)
```

calculateEvalue	<i>calculates E-value</i>
-----------------	---------------------------

Description

This function calculates E-value based on .

Usage

```
calculateEvalue(df, TFs)
```

Arguments

df	a data frame containng p-values
TFs	a character vector of transcription factor

Value

data.frame

Author(s)

Atsushi Fukushima

References

Essaghir A et al. Nucleic Acids Res. 2010 Jun;38(11):e120.

Examples

```
data(DEGs)
data(catalog)

tftg <- extractTFTG(DEGs, catalog)
TFs <- tftg$TFs

p.value <- runif(10)/(1:10)
df <- data.frame(p.value = p.value)
res <- calculateEvalue(df, TFs)
```

calculateFDRBH	<i>calculates FDR by Benjamini and Hochberg method</i>
----------------	--

Description

This function calculates FDR based on BH.

Usage

```
calculateFDRBH(df)
```

Arguments

df a data frame containng p-values

Value

data.frame

Author(s)

Atsushi Fukushima

References

Benjamini Y and Hochberg Y, J Roy Stat Soc B 57: 289?300 (1995)

Examples

```
p.value <- runif(10)/(1:10)
df <- data.frame(p.value = p.value)
res <- calculateFDRBH(df)
```

calculateQvalue	<i>calculates Q-value</i>
-----------------	---------------------------

Description

This function calculates Q-value based on Storey.

Usage

```
calculateQvalue(df, lambda = seq(0.05, 0.5, 0.01))
```

Arguments

df a data frame containng p-values
lambda a vector of the lambda values utilized to obtain pi0.lambda

Value

data.frame

Author(s)

Atsushi Fukushima

References

Storey JD, The Annals of Statistics 31:2013-2035 (2003)

Examples

```
data(example.df)
p.value <- example.df$p.value
df <- data.frame(p.value = p.value)
res <- calculateQvalue(df)
```

calculateRC *calculates Random Control (RC)*

Description

This function calculates Random Control (RC)

Usage

```
calculateRC(
  df,
  DEGs,
  catalog,
  TFs,
  all.targets,
  TF.col = "TF..OFFICIAL_TF_CODING_GENE_NAME.",
  lambda = 0.05,
  nRep = 100
)
```

Arguments

df	a data frame containng p-values
DEGs	a character vector of DEGs (differentially expressed genes)
catalog	a data frame of TFactS catalog (ver. 2)
TFs	a character vector of transcription factor
all.targets	a character vector of all target genes
TF.col	the name of the column that contains the TF names
lambda	a user-specified threshold of E-value (default: 0.05)
nRep	number of random selections (default: 100)

Value

data.frame

Author(s)

Atsushi Fukushima Definition: m is the number of target genes annotated for the TF under consideration n is the number of query genes N is the number of regulations in the catalog k is the number of query genes that are annotated as regulated by TF (i.e., the intersection between the query and the TF signature)

References

Essaghir A et al. Nucleic Acids Res. 2010 Jun;38(11):e120.

Examples

```
data(example.df)
data(catalog)
data(DEGs)

tftg <- extractTFTG(DEGs, catalog)
TFs <- tftg$TFs
all.targets <- tftg$all.targets

res <- calculateRC(example.df, DEGs, catalog, TFs, all.targets)
```

calculateTFactS

calculates TFactS

Description

This function calculates TFactS

Usage

```
calculateTFactS(  
  DEGs,  
  catalog,  
  TFs,  
  all.targets,  
  Q.value = FALSE,  
  lambda1 = seq(0.05, 0.5, 0.01),  
  lambda2 = 0.05,  
  nRep = 100,  
  TF.col = "TF..OFFICIAL_TF_CODING_GENE_NAME."  
)
```

Arguments

DEGs	a character vector of DEGs (differentially expressed genes)
catalog	a data frame of TFactS catalog (ver. 2)
TFs	a character vector of transcription factor
all.targets	a character vector of all target genes
Q.value	logical. If it is TRUE, Q.value by Storey method.
lambda1	a vector of the lambda values utilized to obtain pi0.lambda
lambda2	a user-specified threshold of E-value (default: 0.05)
nRep	number of random selections (default: 100)
TF.col	the name of the column that contains the TF names

Value

data.frame

Author(s)

Atsushi Fukushima

References

Essaghir A et al. Nucleic Acids Res. 2010 Jun;38(11):e120.

Examples

```
data(DEGs)  
data(catalog)  
  
tftg <- extractTFTG(DEGs, catalog)  
TFs <- tftg$TFs  
all.targets <- tftg$all.targets  
  
res <- calculateTFactS(DEGs, catalog, TFs, all.targets)
```

catalog	<i>A list of TF-target genes (TG) in TFactS.</i>
---------	--

Description

A catalogue of annotated target genes based on TFactS.

Details

This was downloaded from TFactS (<http://www.tfacts.org/>).

The catalogues version is 2 (<http://www.tfacts.org/TFactS-new/TFactS-v2/tfacts/data/Catalogues.xls>).

Value

None

Author(s)

Atsushi Fukushima

References

Essaghir A et al. Nucleic Acids Res. 2010 Jun;38(11):e120.

DEGs	<i>An example list of DEGs from DEMO data in TFactS.</i>
------	--

Description

An example list of differentially expressed genes (DEGs).

Details

This was from DEMO data in TFactS (<http://www.tfacts.org/>). The list corresponds to 18 up-regulated genes.

Value

None

Author(s)

Atsushi Fukushima

References

Essaghir A et al. Nucleic Acids Res. 2010 Jun;38(11):e120.

DEGs39

An example list of 39 DEGs from DEMO data in TFactS.

Description

A list of 39 differentially expressed genes (DEGs).

Details

This was from DEMO data in TFactS (<http://www.tfacts.org/>). The list corresponds to 18 up-regulated genes and 21 down-regulated genes.

Value

None

Author(s)

Atsushi Fukushima

References

Essaghir A et al. Nucleic Acids Res. 2010 Jun;38(11):e120.

example.df

An example result of enrichment analysis

Description

An example result of enrichment analysis

Details

The size of data frame is 345 TFs and 6 columns.

Value

None

Author(s)

Atsushi Fukushima

References

Essaghir A et al. Nucleic Acids Res. 2010 Jun;38(11):e120.

example.list	<i>An example list of result of enrichment analysis</i>
--------------	---

Description

An example list of result of enrichment analysis

Details

The length of the list is 345 (TFs).

Value

None

Author(s)

Atsushi Fukushima

References

Essaghir A et al. Nucleic Acids Res. 2010 Jun;38(11):e120.

extractTFTG	<i>extracts transcription factor (TF) and target gene (TG) information</i>
-------------	--

Description

This function extracts information about transcription factor (TF) and target gene (TG) with TFactS Catalogue (v2).

Usage

```
extractTFTG(
  DEGs,
  catalog,
  TF.col = "TF..OFFICIAL_TF_CODING_GENE_NAME.",
  TG.col = "Target.gene..OFFICIAL_GENE_NAME."
)
```

Arguments

DEGs	a character vector of DEGs (differentially expressed genes)
catalog	a data frame of TFactS catalog (ver. 2)
TF.col	the name of the column that contains the TF names
TG.col	the name of the column that contains the TG names

Value

list

Author(s)

Atsushi Fukushima

Examples

```
data(DEGs)
data(catalog)

res <- extractTFTG(DEGs, catalog)
head(res$TFs)
```

FASTcalculateRC *calculates Random Control (RC) fastly?*

Description

This function calculates Random Control (RC)

Usage

```
FASTcalculateRC(
  df,
  DEGs,
  catalog,
  TFs,
  all.targets,
  TF.col = "TF.OFFICIAL_TF_CODING_GENE_NAME.",
  lambda = 0.05,
  nRep = 100
)
```

Arguments

df	a data frame containng p-values
DEGs	a character vector of DEGs (differentially expressed genes)
catalog	a data frame of TFactS catalog (ver. 2)
TFs	a character vector of transcription factor
all.targets	a character vector of all target genes
TF.col	the name of the column that contains the TF names
lambda	a user-specified threshold of E-value (default: 0.05)
nRep	number of random selections (default: 100)

Value

data.frame

Author(s)

Atsushi Fukushima

References

Essaghir A et al. Nucleic Acids Res. 2010 Jun;38(11):e120.

Examples

```
data(example.df)
data(catalog)
data(DEGs)

tftg <- extractTFTG(DEGs, catalog)
TFs <- tftg$TFs
all.targets <- tftg$all.targets

res <- FASTcalculateRC(example.df, DEGs, catalog, TFs, all.targets)
```

formatET

formats the result of enrichment test (ET)

Description

This function formats the result of enrichment test (ET) based on Fisher's exact test

Usage

```
formatET(list)
```

Arguments

`list` a list of the result of enrichment test (ncol = 6)

Value

data.frame

Author(s)

Atsushi Fukushima

Examples

```
data(example.list)
res <- formatET(example.list)
```

formatRC*formats the result of Random Control (RC)*

Description

This function formats the result of Random Control (RC) with random simulation based on Fisher's exact test

Usage

```
formatRC(df, list, nRep)
```

Arguments

df	a data frame of ET including E-values, FDR-BH, and Q-values
list	a list of the result of RC (ncol = 2)
nRep	the number of random selections (negative control)

Value

data.frame

Author(s)

Atsushi Fukushima

Examples

```
data(example.df)
data(example.list)
nRep <- 100
res <- formatRC(example.df, example.list, nRep)
```

GenesUp_SH1H

An example list of 74 DEGs from AtGenExpress stress dataset.

Description

A list of 74 differentially expressed genes (DEGs) from AtGenExpress

Details

This was from DEMO data in the Supplemental data (Table S4. Genes regulated by UV-B light stress in roots and shoots; filename: TPJ_3052_sm_TabS4.xls) of the paper by Killian et al. TPJ (2008). The list corresponds to 74 genes up-regulated in roots and shoots 1 hour after onset of UV-B stress treatment.

Value

None

Author(s)

Atsushi Fukushima

References

Killian J et al. Plant J. 2007 Apr;50(2):347-363.

TFactSR

TFactSR: A package for identifying critical transcription factors.

Description

TFactS is to predict which are the transcription factors (TFs), regulated in a biological condition based on lists of differentially expressed genes (DEGs) obtained from transcriptome experiments. This package is based on the TFactS concept and expands it. It allows users to performe TFactS-like enrichment approach. The package can import and use the original catalogue file from the TFactS website (<http://www.tfacts.org/>) as well as users' defined catalogues of interest that are not supported by TFactS (e.g., Arabidopsis).

Index

* datasets

- AtCatalog, [2](#)
- catalog, [9](#)
- DEGs, [9](#)
- DEGs39, [10](#)
- example.df, [10](#)
- example.list, [11](#)
- GenesUp_SH1H, [15](#)

AtCatalog, [2](#)

- calculateEnrichmentTest, [3](#)
- calculateEvalue, [4](#)
- calculateFDRBH, [5](#)
- calculateQvalue, [5](#)
- calculateRC, [6](#)
- calculateTFactS, [7](#)
- catalog, [9](#)

- DEGs, [9](#)
- DEGs39, [10](#)

- example.df, [10](#)
- example.list, [11](#)
- extractTFTG, [11](#)

- FASTcalculateRC, [12](#)
- formatET, [13](#)
- formatRC, [14](#)

- GenesUp_SH1H, [15](#)

- TFactSR, [15](#)