Package ‘msig’

June 29, 2021

Title  An R Package for Exploring Molecular Signatures Database
Version  1.0
Description  The Molecular Signatures Database ('MSigDB') is one of the most widely used and comprehensive databases of gene sets for performing gene set enrichment analysis <doi:10.1016/j.cels.2015.12.004>. The 'msig' package provides you with powerful, easy-to-use and flexible query functions for the 'MsigDB' database.

There are 2 query modes in the 'msig' package: online query and local query. Both queries contain 2 steps: gene set name and gene.

The online search is divided into 2 modes: registered search and non-registered browse. For registered search, email that you registered should be provided.

Local queries can be made from local database, which can be updated by msig_update() function.

License  GPL-2
Encoding  UTF-8
RoxygenNote  7.1.1
Imports  do, rvest, xml2, set, plyr, httr, jsonlite, utils, kableExtra,
tmcn, crayon, dplyr, stringr, sqldf
Depends  R (>= 4.1.0)
NeedsCompilation  no
Author  Jing Zhang [aut, cre],
        Zhi Jin [aut]
Maintainer  Jing Zhang <zj391120@163.com>
Repository  CRAN
Date/Publication  2021-06-29 07:50:02 UTC

R topics documented:

  browse_msig .................................................. 2
  browse_show_collection ..................................... 3
  local_msig .................................................... 4
  local_version ................................................ 5
browse_msig

Retrieve Gene set Names from MSigDB database Retrieve gene set names from MSigDB database by the gene set name and collection. the search filed is gene name.

Description

Retrieve Gene set Names from MSigDB database Retrieve gene set names from MSigDB database by the gene set name and collection. the search filed is gene name.

Usage

browse_msig(geneSetName = "", collection = "")

Arguments

  geneSetName       one keyword for gene set name, default is empty
  collection       one collection, default is empty

Value

  gene set names
browse_show_collection

Show collection of MSigDB database

Description

Show collection of MSigDB database

Usage

browse_show_collection()

Value

show all collection in MSigDB in web page http://www.gsea-msigdb.org/gsea/msigdb/genesets.jsp. For chromosome, we should treat as collection together.

Examples

browse_show_collection()
local_msig  

Query gene set names from local data

Description

Query gene set names from local data

Usage

local_msig(
  geneset = NULL,
  description = NULL,
  collection = NULL,
  sub_collection = NULL,
  organism = NULL,
  contributor = NULL,
  contributor_org = NULL,
  author = NULL,
  chip = NULL,
  gene = NULL,
  geneset_fuzzy = NULL,
  collection_fuzzy = NULL,
  sub_collection_fuzzy = NULL,
  organism_fuzzy = NULL,
  contributor_fuzzy = NULL,
  contributor_org_fuzzy = NULL,
  author_fuzzy = NULL,
  chip_fuzzy = NULL,
  gene_fuzzy = NULL
)

Arguments

geneset  one sql format character for exact match
description  one sql format character for exact match
collection  one sql format character for exact match
sub_collection  one sql format character for exact match
organism  one sql format character for exact match
contributor  one sql format character for exact match
contributor_org  one sql format character for exact match
author  one sql format character for exact match
chip  one sql format character for exact match
gene  one sql format character for exact match
### local_version

<table>
<thead>
<tr>
<th>geneset_fuzzy</th>
<th>fuzzy match</th>
</tr>
</thead>
<tbody>
<tr>
<td>collection_fuzzy</td>
<td>fuzzy match</td>
</tr>
<tr>
<td>sub_collection_fuzzy</td>
<td>fuzzy match</td>
</tr>
<tr>
<td>organism_fuzzy</td>
<td>fuzzy match</td>
</tr>
<tr>
<td>contributor_fuzzy</td>
<td>fuzzy match</td>
</tr>
<tr>
<td>contributor_org_fuzzy</td>
<td>fuzzy match</td>
</tr>
<tr>
<td>author_fuzzy</td>
<td>fuzzy match</td>
</tr>
<tr>
<td>chip_fuzzy</td>
<td>fuzzy match</td>
</tr>
<tr>
<td>gene_fuzzy</td>
<td>fuzzy match</td>
</tr>
</tbody>
</table>

**Value**

one dataframe with attribute of msig_local.

**Examples**

```r
x <- local_msig('IMMUNE_RESPONSE')
x <- local_msig('IMMUNE_RESPONSE|IMMUNE_SYSTEM_PROCESS')
```

---

### local_version

<table>
<thead>
<tr>
<th>local_version</th>
<th>Local database version</th>
</tr>
</thead>
</table>

**Description**

Local database version

**Usage**

local_version()

**Value**

version of local database

**Examples**

local_version()
msig_download

msig_detail Retrieve detail information of gene set

Description
Retrieve detail information of gene set

Usage
msig_detail(...)

Arguments
... one or more gene set names, which can be little or capital.

Value
Print detail information about the geneset, number of genes and return all gene names.

Examples

d <- msig_detail('izadpanah_stem_cell_adipose_vs_bone_dn',
                 'AAACCAC_MIR140')

msig_download Download MsigDB database

Description
Download MsigDB database

Usage
msig_download(version)

Arguments
version version

Value
download the data to local PC
**msig_filt**

**Filt data by key words Case insensitive**

**Description**

Filt data by key words Case insensitive

**Usage**

```r
msig_filt(x, ...)
```

**Arguments**

- `x`: data from msig package
- `...`: one or more key words

**Value**

Filted results with high light information.

**Examples**

```r
browse_msig('immune') |>  
  msig_filt('response')
```

---

**msig_gene**

**Retrieve gene by Gene set Name**

**Description**

Retrieve gene by Gene set Name

**Usage**

```r
msig_gene(..., list = TRUE, info = TRUE)
```

# S3 method for class 'character'
```r
msig_gene(..., list = TRUE, info = TRUE)
```

# S3 method for class 'data.frame'
```r
msig_gene(..., list = TRUE, info = TRUE)
```
Arguments

... one or more geneset names, which can be little or capital.
list logical, default is FALSE, whether to show result by list.
info logical, whether to show information about gene set.

Value

Print detail information about the geneset, number of genes and return all gene names.

Examples

```r
genes <- msig_gene('izadorpanah_stem_cell_adipose_vs_bone_dn',
                     'REACTOME_DEGRADATION_OF_AXIN')
genes |> msig_view()
```

---

### msig_geneSymbol

**Retrieve gene by Gene Symbol from MsigDB**

**Description**

Retrieve gene by Gene Symbol from MsigDB

**Usage**

```r
msig_geneSymbol(..., local = FALSE)
```

```r
## S3 method for class 'list'
msig_geneSymbol(..., local = FALSE)
```

```r
## S3 method for class 'data.frame'
msig_geneSymbol(..., local = FALSE)
```

```r
## S3 method for class 'character'
msig_geneSymbol(..., local = FALSE)
```

**Arguments**

... one or more geneset names, which can be little or capital.
local logical, default is FALSE, whether to extract gene symbol from local database

**Value**

gene symbol
msig_update

Examples

```r
genes <- msig_geneSymbol('izadpanah_stem_cell_adipose_vs_bone_dn')
genes <- msig_geneSymbol('izadpanah_stem_cell_adipose_vs_bone_dn',
                          'REACTOME_DEGRADATION_OF_AXIN')
```

---

**msig_update**

*Update local MsigDB database*

---

**Description**

Update local MsigDB database

**Usage**

```r
msig_update(xml = NULL, version = NULL)
```

**Arguments**

- `xml` msigdb xml file
- `version` version, if missing, the latest version will be used

**Value**

update local MsigDB database

---

**msig_version**

*version information of MSigDB database*

---

**Description**

version information of MSigDB database

**Usage**

```r
msig_version()
```

**Value**

version dataframe
msig_view  View data in viewer panel

Description
View data in viewer panel

Usage
msig_view(x, ...)

Arguments
x   dataframe
...   one or more highlighted words

Value
open data in view panel in rstudio

Examples
#' browse_msig('immune') |>  
msig_view('response')

NewMsigDB  Create NewMsigDB object for new versions of MsigDB database

Description
Create NewMsigDB object for new versions of MsigDB database

Usage
NewMsigDB(xml)

Arguments
xml   path of xml msigdb file path

Value
dataframe which can be used inner package
### read_msigdb_xml

**Description**
read MSigDB xml data

**Usage**

```
read_msigdb_xml(xml)
```

**Arguments**

- `xml` xml data path

**Value**

one dataframe contains gene information

### related_geneset

**Description**
Query related gene sets

**Usage**

```
related_geneset(geneSetName)
```

**Arguments**

- `geneSetName` one gene set name

**Value**

related gene sets from gene set detailed information table

**Examples**

```
x <- related_geneset('AAANWWTGC_UNKNOWN')
x |>  
    msig_filt('unknown') |>  
    msig_view('ttt')
```
search_msig

*Query MSigDB database by cookie*

**Description**

Query MSigDB database by cookie

**Usage**

```r
search_msig(keywords, collection = "", organism = "", contributor = ", email)
```

**Arguments**

- `keywords`: one keywords see Detail field
- `collection`: one or more collections
- `organism`: one or more organisms
- `contributor`: one or more contributors
- `email`: email that registered for MSigDB database.

**Value**

dataframe contains name, description and so on.

**Examples**

```r
email <- 'your email'
x <- search_msig('immune & response')
x |>
    msig_filt('system') |> 
    msig_view('C2')
```

search_show_collection

*Show collections for msigdb_search()*

**Description**

Show collections for msigdb_search()

**Usage**

```r
search_show_collection(email)
```
**Arguments**

email email that registered for MSigDB database.

**Value**

collections from MsigDB website.

**Examples**

```
search_show_collection("your email")
# or
email <- 'your email'
search_show_collection()
```

---

**search_show_contributor**

*Show contributor for msigdb_search()*

---

**Description**

Show contributor for msigdb_search()

**Usage**

```
search_show_contributor(email)
```

**Arguments**

email email that registered for MSigDB database.

**Value**

contributors from MsigDB website.

**Examples**

```
search_show_contributor("your email")
# or
email <- 'your email'
search_show_contributor()
```
**search_show_organism**  
*Show organism for msigdb_search()*

**Description**
Show organism for msigdb_search()

**Usage**

```r
search_show_organism(email)
```

**Arguments**

- `email`  
  email that registered for MSigDB database.

**Value**
organisms from MsigDB website.

**Examples**

```r
search_show_organism("your email")  
# or
email <- 'your email'
search_show_organism()
```

---

**show_local_collection**  
*Show collections of local MsigDB database*

**Description**
Show collections of local MsigDB database

**Usage**

```r
show_local_collection()
```

**Value**
A dataframe contains 2 columns. The first column is the name of the collection. The second column is the number of frequencies it has.

**Examples**

```r
show_local_collection()
```
show_local_contributor

*Show contributors of local MsigDB database*

**Description**

Show contributors of local MsigDB database

**Usage**

`show_local_contributor()`

**Value**

A dataframe contains 2 columns. The first column is the name of the contributor. The second column is the number of frequencies it has.

**Examples**

`show_local_contributor()`

---

show_local_contributor_org

*Show contributor_orgs of local MsigDB database*

**Description**

Show contributor_orgs of local MsigDB database

**Usage**

`show_local_contributor_org()`

**Value**

A dataframe contains 2 columns. The first column is the name of the contributor_org. The second column is the number of frequencies it has.

**Examples**

`show_local_contributor_org()`
show_local_data  

*Show local data used in this package*

**Description**
Show local data used in this package

**Usage**
```r
show_local_data()
```

**Value**
data used inner this package

**Examples**
```r
show_local_data()
```

---

show_local_sub_collection  

*Show sub_collections of local MsigDB database*

**Description**
Show sub_collections of local MsigDB database

**Usage**
```r
df <- show_local_sub_collection()
df
```

**Value**
A dataframe contains 2 columns. The first column is the name of the sub_collection. The second column is the number of frequencies it has.

**Examples**
```r
df <- show_local_sub_collection()
df
```
**similarity_geneset**

Query similarity gene sets

**Description**
Query similarity gene sets

**Usage**
similarity_geneset(geneSetName)

**Arguments**
geneSetName one gene set name

**Value**
similarity gene sets

**Examples**

```r
x <- similarity_geneset('REACTOME_DEGRADATION_OF_AXIN')
x |> msign_view()
```
Index

browse_msig, 2
browse_show_collection, 3

local_msig, 4
local_version, 5

msig_detail, 6
msig_download, 6
msig_filt, 7
msig_gene, 7
msig_geneSymbol, 8
msig_update, 9
msig_version, 9
msig_view, 10

NewMsigDB, 10

read_msigdb_xml, 11
related_geneset, 11

search_msig, 12
search_show_collection, 12
search_show_contributor, 13
search_show_organism, 14
show_local_collection, 14
show_local_contributor, 15
show_local_contributor_org, 15
show_local_data, 16
show_local_sub_collection, 16
similarity_geneset, 17