

# Package ‘cRegulome’

January 3, 2019

**Type** Package

**Version** 0.3.0

**Title** Obtain and Visualize Regulome-Gene Expression Correlations in Cancer

**Description** Builds a 'SQLite' database file of pre-calculated transcription factor/microRNA-gene correlations (co-expression) in cancer from the Cistrome Cancer Liu et al. (2011) <doi:10.1186/gb-2011-12-8-r83> and 'miRCancerdb' databases (in press). Provides custom classes and functions to query, tidy and plot the correlation data.

**License** GPL-3

**URL** <https://github.com/ropensci/cRegulome>

**BugReports** <https://github.com/ropensci/cRegulome/issues>

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 6.1.1

**Depends** R(>= 2.10.0)

**Imports** DBI, graphics, httr, RSQLite, UpSetR, VennDiagram, ggplot2, ggridges, grid, R.utils, igraph

**VignetteBuilder** knitr

**Suggests** knitr, rmarkdown, testthat, covr, readxl, AnnotationDbi, org.Hs.eg.db, clusterProfiler

**NeedsCompilation** no

**Author** Mahmoud Ahmed [aut, cre] (<<https://orcid.org/0000-0002-4377-6541>>)

**Maintainer** Mahmoud Ahmed <mahmoud.s.fahmy@students.kasralainy.edu.eg>

**Repository** CRAN

**Date/Publication** 2019-01-03 08:20:03 UTC

## R topics documented:

cmicroRNA . . . . .	2
cor_hist . . . . .	3
cor_igraph . . . . .	4
cor_joy . . . . .	5
cor_plot . . . . .	6
cor_prep . . . . .	7
cor_tidy . . . . .	7
cor_upset . . . . .	8
cor_venn_diagram . . . . .	9
cRegulome . . . . .	10
cTF . . . . .	11
get_db . . . . .	11
get_mir . . . . .	12
get_tf . . . . .	14
stat_collect . . . . .	15
stat_collect_targets . . . . .	16
stat_make . . . . .	16
stat_make_targets . . . . .	17
<b>Index</b>	<b>18</b>

---

cmicroRNA

*Construct cmicroRNA object*

---

### Description

Constructs an S3 object called cmicroRNA contains data returned by calling [get\\_mir](#). Used to define methods for printing and visualizing microRNA-gene expression correlations.

### Usage

```
cmicroRNA(dat_mir)
```

### Arguments

dat\_mir            A data.frame such as this returned by calling [get\\_mir](#).

### Value

An S3 object of class cmicroRNA

**Examples**

```
# locate the testset file and connect
fl <- system.file('extdata', 'cRegulome.db', package = 'cRegulome')
conn <- RSQLite::dbConnect(RSQLite::SQLite(), fl)

# enter a custom query with different arguments
dat <- get_mir(conn,
               mir = 'hsa-let-7g',
               study = 'STES',
               min_abs_cor = .3,
               max_num = 5)

# make a cmicroRNA object
cmir <- cmicroRNA(dat)
```

cor\_hist

*A histogram of the correlations of microRNA or tf sets***Description**

Plot a [hist](#) of sets of microRNAs or transcription factors-gene correlations in a TCGA study.

**Usage**

```
cor_hist(ob, study, ...)
```

**Arguments**

ob	A <a href="#">cmicroRNA</a> or <a href="#">cTF</a> object such as this returned by calling <a href="#">cmicroRNA</a> or <a href="#">cTF</a> .
study	A character vector of The Cancer Genome Atlas (TCGA) study identifiers. To view the available studies in TCGA project, <a href="https://tcga-data.nci.nih.gov/docs/publications/tcga">https://tcga-data.nci.nih.gov/docs/publications/tcga</a> . When left to default NULL all available studies will be included.
...	Other options

**Value**

An [hist](#) plot of the correlations values between genes a microRNA or a transcription factor in a TCGA study

**Examples**

```
# locate the testset file and connect
fl <- system.file('extdata', 'cRegulome.db', package = 'cRegulome')
conn <- RSQLite::dbConnect(RSQLite::SQLite(), fl)

# enter a custom query with different arguments
```

```
dat <- get_mir(conn,
               mir = c('hsa-let-7g', 'hsa-let-7i'),
               study = 'STES')

# make a cmicroRNA object
cmir <- cmicroRNA(dat)

# print object
cor_hist(cmir)
```

---

cor\_igraph

*Make an igraph object*

---

## Description

An igraph object of from [cmicroRNA](#) or [cTF](#) objects.

## Usage

```
cor_igraph(ob, directed = FALSE)
```

## Arguments

**ob** A [cmicroRNA](#) or [cTF](#) object such as this returned by calling [cmicroRNA](#) or [cTF](#).  
**directed** A logical when FALSE the graph is undirected

## Value

An igraph object

## Examples

```
# load required libraries
library(RSQLite)
library(cRegulome)

# locate the testset file and connect
fl <- system.file('extdata', 'cRegulome.db', package = 'cRegulome')
conn <- dbConnect(SQLite(), fl)

# enter a custom query with different arguments
dat <- get_mir(conn,
               mir = c('hsa-let-7g', 'hsa-let-7i'),
               study = 'STES')

# make a cmicroRNA object
cmir <- cmicroRNA(dat)
```

```
# print object
cor_igraph(cmir)
```

---

cor_joy	<i>A joy plot of correlation of microRNA or tf sets</i>
---------	---

---

## Description

A `ggridges` joy plot of sets of microRNAs or transcription factors-gene correlations in a TCGA study.

## Usage

```
cor_joy(ob, study, ...)
```

## Arguments

ob	A <code>cmicroRNA</code> or <code>cTF</code> object such as this returned by calling <code>cmicroRNA</code> or <code>cTF</code> .
study	A character vector of The Cancer Genome Atlas (TCGA) study identifiers. To view the available studies in TCGA project, <a href="https://tcga-data.nci.nih.gov/docs/publications/tcga">https://tcga-data.nci.nih.gov/docs/publications/tcga</a> . When left to default NULL all available studies will be included.
...	Other options

## Value

An `ggridges` plot object

## Examples

```
# locate the testset file and connect
fl <- system.file('extdata', 'cRegulome.db', package = 'cRegulome')
conn <- RSQLite::dbConnect(RSQLite::SQLite(), fl)

# enter a custom query with different arguments
dat <- get_mir(conn,
               mir = c('hsa-let-7g', 'hsa-let-7i'),
               study = 'STES')

# make a cmicroRNA object
cmir <- cmicroRNA(dat)

# print object
cor_joy(cmir)
```

---

`cor_plot`*Plot method for `cmicroRNA` and `cTF` objects*

---

### Description

A dot plot of microRNA/TF correlation in a single study of TCGA. When the object `cmicroRNA/cTF` contains more than one TCGA studies, the argument `study` is a requirement.

### Usage

```
cor_plot(ob, study, ...)
```

### Arguments

<code>ob</code>	A <code>cmicroRNA</code> or <code>cTF</code> object such as this returned by calling <code>cmicroRNA</code> or <code>cTF</code> .
<code>study</code>	A character vector of The Cancer Genome Atlas (TCGA) study identifiers. To view the available studies in TCGA project, <a href="https://tcga-data.nci.nih.gov/docs/publications/tcga">https://tcga-data.nci.nih.gov/docs/publications/tcga</a> . When left to default NULL all available studies will be included.
<code>...</code>	Other options

### Value

A ggplot object of a dot plot of the correlation values between genes and microRNAs or transcription factors in a TCGA study.

### Examples

```
# locate the testset file and connect
fl <- system.file('extdata', 'cRegulome.db', package = 'cRegulome')
conn <- RSQLite::dbConnect(RSQLite::SQLite(), fl)

# enter a custom query with different arguments
dat <- get_mir(conn,
               mir = 'hsa-let-7g',
               study = 'STES',
               min_abs_cor = .3,
               max_num = 5)

# make a cmicroRNA object
cmir <- cmicroRNA(dat)

# print object
cor_plot(cmir)
```

---

cor_prep	<i>Prepare correlation data for plotting</i>
----------	--

---

**Description**

Not meant to be called directly by the user.

**Usage**

```
cor_prep(ob, study, add_dir = TRUE, add_corr = TRUE)
```

**Arguments**

ob	A <a href="#">c<b>microRNA</b></a> or <a href="#">c<b>TF</b></a> object such as this returned by calling <a href="#">c<b>microRNA</b></a> or <a href="#">c<b>TF</b></a> .
study	A character vector of The Cancer Genome Atlas (TCGA) study identifiers. To view the available studies in TCGA project, <a href="https://tcga-data.nci.nih.gov/docs/publications/tcga">https://tcga-data.nci.nih.gov/docs/publications/tcga</a> . When left to default NULL all available studies will be included.
add_dir	A logical default TRUE for whether to add a column called Direction that has the direction of the correlation; positive or negative.
add_corr	A logical default TRUE for whether to add a column called Correlation that has the absolute value of the correlation

**Value**

A data.frame

---

cor_tidy	<i>Tidy <a href="#">c<b>microRNA</b></a> and <a href="#">c<b>TF</b></a> objects</i>
----------	---

---

**Description**

Tidy [c\*\*microRNA\*\*](#) and [c\*\*TF\*\*](#) objects

**Usage**

```
cor_tidy(ob)
```

**Arguments**

ob	A <a href="#">c<b>microRNA</b></a> or <a href="#">c<b>TF</b></a> object such as this returned by calling <a href="#">c<b>microRNA</b></a> or <a href="#">c<b>TF</b></a> .
----	---

**Value**

A tidy data.frame of four columns. mirna\_base or tf is the microRNA miRBase IDs, feature is the features/genes, cor is the corresponding expression correlations and study is TCGA study ID.

## Examples

```
# locate the testset file and connect
fl <- system.file('extdata', 'cRegulome.db', package = 'cRegulome')
conn <- RSQLite::dbConnect(RSQLite::SQLite(), fl)

# enter a custom query with different arguments
dat <- get_mir(conn,
  mir = 'hsa-let-7g',
  study = 'STES',
  min_abs_cor = .3,
  max_num = 5)

# make a cmicroRNA object
cmir <- cmicroRNA(dat)

# convert cmicroRNA object to a tidy data.frame
tidy_cmir <- cor_tidy(cmir)
```

---

cor\_upset

*upset plot of microRNA or tf sets*

---

## Description

`upset` of sets of microRNAs or transcription factors and their correlated features in a TCGA study.

## Usage

```
cor_upset(ob, study, ...)
```

## Arguments

ob	A <code>cmicroRNA</code> or <code>cTF</code> object such as this returned by calling <code>cmicroRNA</code> or <code>cTF</code> .
study	A character vector of The Cancer Genome Atlas (TCGA) study identifiers. To view the available studies in TCGA project, <a href="https://tcga-data.nci.nih.gov/docs/publications/tcga">https://tcga-data.nci.nih.gov/docs/publications/tcga</a> . When left to default NULL all available studies will be included.
...	Other options

## Value

An `upset` plot

**Examples**

```
# locate the testset file and connect
fl <- system.file('extdata', 'cRegulome.db', package = 'cRegulome')
conn <- RSQLite::dbConnect(RSQLite::SQLite(), fl)

# enter a custom query with different arguments
dat <- get_mir(conn,
               mir = c('hsa-let-7g', 'hsa-let-7i'),
               study = 'STES')

# make a cmicroRNA object
cmir <- cmicroRNA(dat)

# print object
cor_upset(cmir)
```

---

cor\_venn\_diagram      *Venn Diagram of microRNA or transcription factor correlated features*

---

**Description**

Count and plot the numbers of microRNA correlated features in cmicroRNA object.

**Usage**

```
cor_venn_diagram(ob, study, ...)
```

**Arguments**

ob	A <a href="#">cmicroRNA</a> or <a href="#">cTF</a> object such as this returned by calling <a href="#">cmicroRNA</a> or <a href="#">cTF</a> .
study	A character vector of The Cancer Genome Atlas (TCGA) study identifiers. To view the available studies in TCGA project, <a href="https://tcga-data.nci.nih.gov/docs/publications/tcga">https://tcga-data.nci.nih.gov/docs/publications/tcga</a> . When left to default NULL all available studies will be included.
...	Other options

**Value**

A venn diagram with a circle or an ellipses for each microRNA and the number of correlated features.

## Examples

```
# locate the testset file and connect
fl <- system.file('extdata', 'cRegulome.db', package = 'cRegulome')
conn <- RSQLite::dbConnect(RSQLite::SQLite(), fl)

# enter a custom query with different arguments
dat <- get_mir(conn,
               mir = c('hsa-let-7g', 'hsa-let-7i'),
               study = 'STES')

# make a cmicroRNA object
cmir <- cmicroRNA(dat)

# make graph
cor_venn_diagram(cmir)
```

---

cRegulome

cRegulome *package*

---

## Description

Download, access and visualize Regulome (microRNA and transcription factors) data from miR-Cancer and Cistrome cancer

### cRegulome **functions to download and query the database file**

[get\\_db](#) [get\\_tf](#) [get\\_mir](#)

### cRegulome **functions to create S3 objects**

[cTF](#) [cmicroRNA](#)

### cRegulome **functions to reshape S3 objects**

[cor\\_tidy](#) [cor\\_igraph](#)

### cRegulome **functions to visualize data in S3 objects**

[cor\\_hist](#) [cor\\_joy](#) [cor\\_plot](#) [cor\\_upset](#) [cor\\_venn\\_diagram](#)

---

cTF	<i>Construct cTF object</i>
-----	-----------------------------

---

**Description**

Constructs an S3 object called cTF contains data returned by calling [get\\_tf](#). Used to define methods for printing and visualizing transcription factors-gene expression correlations.

**Usage**

```
cTF(dat_tf)
```

**Arguments**

dat\_tf            A data.frame such as this returned by calling [get\\_tf](#).

**Value**

An S3 object of class cTF

**Examples**

```
# locate the testset file and connect
fl <- system.file('extdata', 'cRegulome.db', package = 'cRegulome')
conn <- RSQLite::dbConnect(RSQLite::SQLite(), fl)

# enter a custom query with different arguments
dat <- get_tf(conn,
              tf = 'LEF1',
              study = 'STES',
              min_abs_cor = .3,
              max_num = 5)

# make a cTF object
ctf <- cTF(dat)
```

---

get_db	<i>Get cRegulome.db file</i>
--------	------------------------------

---

**Description**

This function calls [download.file](#) to download the pre-build database file of cRegulome. Additionally, the function checks the validity of the pre-defined URL and whether the database file exists in the current working directory to avoid re-downloading it. Typically, users would run this function once at the first time they use the package or to update the database to the latest version.

**Usage**

```
get_db(test = FALSE, destfile, ...)
```

**Arguments**

test	A logical, default FALSE. When TRUE downloads a database file with the same structure with a subset of the data for speed.
destfile	A character vector for the desired path for the database file. By default, when not specified, is constructed by using <code>tempdir</code> as a directory and the string <code>cRegulome.db.gz</code>
...	Optional arguments passed to <code>download.file</code>

**Value**

Downloads a compressed `sqlite` file to the current working directory. The file is named `cRegulome.db.gz` by default and it's not advised to change the name to avoid breaking the other functions that calls the database.

**Examples**

```
## Not run:
# download a test set of the database
get_db(test = TRUE)

# download the full database file
get_db(test = FALSE)

## End(Not run)

# load the test db file from shipped with the package
db_file <- system.file("extdata", "cRegulome.db", package = "cRegulome")
file.info(db_file)
```

---

```
get_mir
```

```
Get microRNA correlations from cRegulome.db
```

---

**Description**

This function access the `sqlite` database file which is obtained by running `get_db`. Basically, the function provides ways to query the database to the correlation data of the microRNAs of interest. The function returns an error if the database file `cRegulome.db` is not in the working directory.

**Usage**

```
get_mir(conn, mir, study, min_abs_cor, max_num, targets_only = FALSE,
        targets)
```

**Arguments**

conn	A connection such as this returned by <a href="#">dbConnect</a>
mir	A required character vector of the microRNAs of interest. These are the miR-Base ID which are the official identifiers of the widely used miRBase database, <a href="http://www.mirbase.org/">http://www.mirbase.org/</a> .
study	A character vector of The Cancer Genome Atlas (TCGA) study identifiers. To view the available studies in TCGA project, <a href="https://tcga-data.nci.nih.gov/docs/publications/tcga">https://tcga-data.nci.nih.gov/docs/publications/tcga</a> . When left to default NULL all available studies will be included.
min_abs_cor	A numeric, an absolute correlation minimum between 0 and 1 for each mir.
max_num	An integer, maximum number of features to show for each mir in each study.
targets_only	A logical whether restrict the output to the recognized target features.
targets	A character vector of gene symbol names.

**Value**

A tidy data.frame of four columns. mirna\_base is the microRNA miRBase IDs, feature is the features/genes, cor is the corresponding expression correlations and study is TCGA study ID.

**Examples**

```
# locate the testset file and connect
fl <- system.file('extdata', 'cRegulome.db', package = 'cRegulome')
conn <- RSQLite::dbConnect(RSQLite::SQLite(), fl)

# get microRNA correlations in all studies
get_mir(conn,
        mir = 'hsa-let-7g')

# get correlations in a particular study
get_mir(conn,
        mir = 'hsa-let-7g',
        study = 'STES')

# enter a custom query with different arguments
get_mir(conn,
        mir = 'hsa-let-7g',
        study = 'STES',
        min_abs_cor = .3,
        max_num = 5)
```

---

get\_tf

*Get transcription factor correlations from cRegulome.db*


---

### Description

This function access the sqlite database file which is obtained by running [get\\_db](#). Basically, the function provides ways to query the database to the correlation data of the transcription factors of interest. The function returns an error if the database file `cRegulome.db` is not in the working directory.

### Usage

```
get_tf(conn, tf, study, min_abs_cor, max_num, targets_only = FALSE,
       targets)
```

### Arguments

<code>conn</code>	A connection such as this returned by <a href="#">dbConnect</a>
<code>tf</code>	A required character vector of the transcription factor of interest. These are the HUGO official gene symbols of the genes contains the transcription factor.
<code>study</code>	A character vector of The Cancer Genome Atlas (TCGA) study identifiers. To view the available studies in TCGA project, <a href="https://tcga-data.nci.nih.gov/docs/publications/tcga">https://tcga-data.nci.nih.gov/docs/publications/tcga</a> . When left to default NULL all available studies will be included.
<code>min_abs_cor</code>	A numeric, an absolute correlation minimum between 0 and 1 for each mir.
<code>max_num</code>	An integer, maximum number of features to show for each mir in each study.
<code>targets_only</code>	A logical whether restrict the output to the recognized target features.
<code>targets</code>	A character vector of gene symbol names.

### Value

A tidy data.frame of four columns. `tf` is the official gene symbols of the genes contains the transcription factor, `feature` is the features/genes, `cor` is the corresponding expression correlations and `study` is TCGA study ID.

### Examples

```
# locate the testset file and connect
fl <- system.file('extdata', 'cRegulome.db', package = 'cRegulome')
conn <- RSQLite::dbConnect(RSQLite::SQLite(), fl)

## Not run:
# get transcription factors correlations in all studies
get_tf(conn,
      tf = 'LEF1')
```

```
## End(Not run)

# get correlations in a particular study
get_tf(conn,
        tf = 'LEF1',
        study = 'STES')

# enter a custom query with different arguments
get_tf(conn,
        tf = 'LEF1',
        study = 'STES',
        min_abs_cor = .3,
        max_num = 5)
```

---

stat\_collect

*Collect data from SQLite database*

---

## Description

Not meant to be called directly by the user.

## Usage

```
stat_collect(conn, study, stat, type = "mir")
```

## Arguments

conn	A connection such as this returned by <a href="#">dbConnect</a>
study	A character vector of The Cancer Genome Atlas (TCGA) study identifiers. To view the available studies in TCGA project, <a href="https://tcga-data.nci.nih.gov/docs/publications/tcga">https://tcga-data.nci.nih.gov/docs/publications/tcga</a> . When left to default NULL all available studies will be included.
stat	A string such as this returned by <a href="#">stat_make</a>
type	A character string. Either 'mir' or 'tf'. Used to define columns and tables names.

## Value

A data.frame

---

stat\_collect\_targets    *Collect target features from SQLite database*

---

### Description

Not meant to be called directly by the user.

### Usage

```
stat_collect_targets(conn, stat)
```

### Arguments

conn	A connection such as this returned by <a href="#">dbConnect</a>
stat	A string such as this returned by <a href="#">stat_make</a>

### Value

A character vector

---

stat\_make                    *Make A SQL statement*

---

### Description

Not meant to be called directly by the user.

### Usage

```
stat_make(reg, study, min_abs_cor, max_num, targets, type = "mir")
```

### Arguments

reg	A character vector of one or more regulator ID.
study	A character vector of The Cancer Genome Atlas (TCGA) study identifiers. To view the available studies in TCGA project, <a href="https://tcga-data.nci.nih.gov/docs/publications/tcga">https://tcga-data.nci.nih.gov/docs/publications/tcga</a> . When left to default NULL all available studies will be included.
min_abs_cor	A numeric, an absolute correlation minimum between 0 and 1 for each mir.
max_num	An integer, maximum number of features to show for each mir in each study.
targets	A character vector of gene symbol names.
type	A character string. Either 'mir' or 'tf'. Used to define columns and tables names.

**Value**

A character string

**Examples**

```
stat_make(reg = 'hsa-let-7g',  
          study = 'STES')
```

```
stat_make(reg = 'hsa-let-7g',  
          study = 'STES',  
          min_abs_cor = .3)
```

```
stat_make(reg = 'hsa-let-7g',  
          study = 'STES',  
          min_abs_cor = .3,  
          max_num = 5)
```

---

stat_make_targets	<i>Make A SQL statement to extract target features</i>
-------------------	--

---

**Description**

Not meant to be called directly by the user.

**Usage**

```
stat_make_targets(reg, study, type = "mir")
```

**Arguments**

reg	A character vector of one or more regulator ID.
study	A character vector of The Cancer Genome Atlas (TCGA) study identifiers. To view the available studies in TCGA project, <a href="https://tcga-data.nci.nih.gov/docs/publications/tcga">https://tcga-data.nci.nih.gov/docs/publications/tcga</a> . When left to default NULL all available studies will be included.
type	A character string. Either 'mir' or 'tf'. Used to define columns and tables names.

**Value**

A character string

# Index

`cmicroRNA`, 2, 3–10  
`cor_hist`, 3, 10  
`cor_igraph`, 4, 10  
`cor_joy`, 5, 10  
`cor_plot`, 6, 10  
`cor_prep`, 7  
`cor_tidy`, 7, 10  
`cor_upset`, 8, 10  
`cor_venn_diagram`, 9, 10  
`cRegulome`, 10  
`cRegulome-package (cRegulome)`, 10  
`cTF`, 3–10, 11

`dbConnect`, 13–16  
`download.file`, 11, 12

`get_db`, 10, 11, 12, 14  
`get_mir`, 2, 10, 12  
`get_tf`, 10, 11, 14  
`ggridges`, 5

`hist`, 3

`stat_collect`, 15  
`stat_collect_targets`, 16  
`stat_make`, 15, 16, 16  
`stat_make_targets`, 17

`tempdir`, 12

`upset`, 8