

# Package ‘MetaVolcanoR’

April 10, 2023

**Type** Package

**Title** Gene Expression Meta-analysis Visualization Tool

**Version** 1.12.0

**Description** MetaVolcanoR combines differential gene expression results.

It implements three strategies to summarize differential gene expression from different studies. i) Random Effects Model (REM) approach, ii) a p-value combining-approach, and iii) a vote-counting approach. In all cases, MetaVolcano exploits the Volcano plot reasoning to visualize the gene expression meta-analysis results.

**Depends** R (>= 4.1.1)

**Imports** methods, data.table, dplyr, tidyr, plotly, ggplot2, cowplot, parallel, metafor, metap, rlang, topconfects, grDevices, graphics, stats, htmlwidgets

**Suggests** knitr, markdown, rmarkdown, testthat

**License** GPL-3

**Encoding** UTF-8

**biocViews** GeneExpression, DifferentialExpression, Transcriptomics, mRNAArray, RNASeq

**LazyData** true

**roxygennote** 6.1.1.9000

**VignetteBuilder** knitr

**RoxygenNote** 6.1.1

**git\_url** <https://git.bioconductor.org/packages/MetaVolcanoR>

**git\_branch** RELEASE\_3\_16

**git\_last\_commit** f1a0ded

**git\_last\_commit\_date** 2022-11-01

**Date/Publication** 2023-04-10

**Author** Cesar Prada [aut, cre],  
Diogenes Lima [aut],  
Helder Nakaya [aut, ths]

**Maintainer** Cesar Prada <cesar.prada.medina@gmail.com>

## R topics documented:

calc_vi . . . . .	2
collapse_deg . . . . .	3
combining_mv . . . . .	3
cum_freq_data . . . . .	4
deg_def . . . . .	5
diffexplist . . . . .	6
draw_cum_freq . . . . .	6
draw_degbar . . . . .	7
draw_forest . . . . .	8
MetaVolcano-class . . . . .	9
plot_mv . . . . .	9
plot_rem . . . . .	10
remodel . . . . .	10
rem_mv . . . . .	11
rename_col . . . . .	12
set_degbar_data . . . . .	13
votecount_mv . . . . .	13
<b>Index</b>	<b>15</b>

---

calc_vi	<i>A function to calculate variance from confidence interval limits</i>
---------	---

---

### Description

This function takes the limits of a confidence interval (95 a calculate a variance

### Usage

```
calc_vi(diffexp, llcol, r1col)
```

### Arguments

diffexp	data.frame/data.table containing differential expression results
llcol	column name of the fold change coinidence interval left limit name <string>
r1col	column name of the fold change coinidence interval left limit name <string>

### Value

data.table/data.frame with a new vi variable

### Examples

```
data(diffexplist)
diffexp <- calc_vi(diffexplist[[1]], "CI.L", "CI.R")
head(diffexp, 3)
```

---

collapse_deg	<i>A function to filter out geneIDs standing for the same gene name</i>
--------------	---

---

**Description**

This function to remove redundant geneIDs standing for the same gene name

**Usage**

```
collapse_deg(diffexp, genenamecol, pcriteria)
```

**Arguments**

diffexp	data.frame/data.table output of the deg.def() function
genenamecol	the column name of the gene name variable <string>
pcriteria	the column name of the pvalue criteria to consider <string>

**Value**

data.table differential expression results with unique gene names

**Examples**

```
data(diffexplist)
diffexp <- collapse_deg(diffexplist[[1]], "Symbol", "pvalue")
head(diffexp, 3)
```

---

combining_mv	<i>A function to draw the 'Combining meta-analysis' MetaVolcano</i>
--------------	---

---

**Description**

This function draws the 'Combining meta-analysis' MetaVolcano

**Usage**

```
combining_mv(diffexp = list(), pcriteria = "pvalue",
  foldchangeacol = "Log2FC", genenamecol = "Symbol", geneidcol = NULL,
  metafc = "Mean", metathr = 0.01, collaps = "FALSE",
  jobname = "MetaVolcano", outputfolder = ".", draw = "HTML")
```

**Arguments**

diffexp	list of data.frame/data.table (s) with DE results where lines are genes
pcriteria	the column name of the Pval criteria to consider c("adj.P.Val", "P.Value") <string>
foldchangecol	the column name of the foldchange variable <string>
genenamecol	the column name of the gene name variable <string>
geneidcol	the column name of the gene ID/probe/oligo/transcript variable <string>
metafc	method for summarizing gene fold-changes across studies c("Mean", "Median") <string>
metathr	top percentage of perturbed genes to be highlighted <double>
collaps	if probes should be collapsed based on the DE direction <logical>
jobname	name of the running job <string>
outputfolder	/path where to write the results/
draw	wheather or not to draw the .pdf or .html visualization <c(NULL, "PDF", "HTML")>

**Value**

MetaVolcano object

**Examples**

```
data(diffexplist)
mv <- combining_mv(diffexplist)
str(mv)
```

---

cum_freq_data	<i>A data formating function for inverse-cummulative DEG distribution</i>
---------------	---

---

**Description**

This function counts how many genes consistly appears as DE along the input studies

**Usage**

```
cum_freq_data(meta_diffexp, nstud)
```

**Arguments**

meta_diffexp	data.frame/data.table containing all the input studies
nstud	the number of inputed GEO2R outputs <integer>

**Value**

data.frame inverse cummulative distribution

**Examples**

```

library(dplyr)
data(diffexplist)
diffexp <- lapply(diffexplist, function(...) deg_def(..., "pvalue",
  "Log2FC", 0.05, 0))
diffexp <- rename_col(diffexp, "Symbol")
meta_diffexp <- Reduce(function(...) merge(..., by = "Symbol", all = TRUE),
  diffexp)
meta_diffexp %>%
dplyr::select(dplyr::matches("deg_")) %>%
  data.matrix -> n_deg
meta_diffexp[['ndeg']] <- rowSums(n_deg^2, na.rm = TRUE)
cfd <- cum_freq_data(meta_diffexp, length(diffexplist))
head(cfd, 3)

```

deg\_def

*A DEG definition function***Description**

This function creates a new variable indicating DEGs as -1, 0, 1 based on the user-defined fold-change and p-value criteria

**Usage**

```
deg_def(diffexp, pcriteria, foldchange_col, pv, fc)
```

**Arguments**

diffexp	data.frame/data.table with differential expression results
pcriteria	column name of the pvalue variable <strings>
foldchange_col	column name of the foldchange variable <string>
pv	pvalue threshold <double>
fc	foldchange threshold <double>

**Value**

data.table/data.frame with a new deg variable

**Examples**

```

data(diffexplist)
diffexp <- deg_def(diffexplist[[1]], "pvalue", "Log2FC", 0.05, 0)
table(diffexp[['deg']])

```

---

`diffexplist`*Differential expression results from five studies*

---

**Description**

A named list with five differential expression results.

**Usage**

```
diffexplist
```

**Format**

A named list with 5 data frames with ~20k genes and 5 variables:

**GSE12050** differential expression result, disease vs healthy

**GSE24883** differential expression result, disease vs healthy ...

**Source**

<https://www.ncbi.nlm.nih.gov/geo/>

---

`draw_cum_freq`*A function to visualize the inverse-cummulative DEG distribution*

---

**Description**

This function create a ggplot object with the inverse-cummulative DEG distribution

**Usage**

```
draw_cum_freq(meta_diffexp, nstud)
```

**Arguments**

`meta_diffexp` data.frame/data.table containing all the input studies

`nstud` the number of inputted GEO2R outputs <integer>

**Value**

ggplot2 object

## Examples

```
library(dplyr)
data(diffexplist)
diffexp <- lapply(diffexplist, function(...) deg_def(..., "pvalue",
  "Log2FC", 0.05, 0))
diffexp <- rename_col(diffexp, "Symbol")
meta_diffexp <- Reduce(function(...) merge(..., by = "Symbol", all = TRUE),
  diffexp)
meta_diffexp %>%
dplyr::select(dplyr::matches("deg_")) %>%
  data.matrix -> n_deg
meta_diffexp[['ndeg']] <- rowSums(n_deg^2, na.rm = TRUE)
gg <- draw_cum_freq(meta_diffexp, length(diffexplist))
plot(gg)
```

---

draw\_degbar

*A function for DEG barplot visualization*

---

## Description

This function visualize as barplots the number of DEGs across the input studies

## Usage

```
draw_degbar(degbar_data)
```

## Arguments

degbar\_data      output of the set\_degbar\_data() function <data.frame/data.table>

## Value

ggplot2 object

## Examples

```
data(diffexplist)
diffexp <- lapply(diffexplist, function(...) deg_def(..., "pvalue",
  "Log2FC", 0.05, 0))
bardat <- set_degbar_data(diffexp)
gg <- draw_degbar(bardat)
plot(gg)
```

---

`draw_forest`*A function to draw a forest plot from the REM MetaVolcano result*

---

## Description

This function draws a forest plot for a given gene based on the REM MetaVolcano result

## Usage

```
draw_forest(remres, gene = "MMP9", genecol = "Symbol",
            foldchangepcol = "Log2FC", llcol = "CI.L", rlcol = "CI.R",
            jobname = "MetaVolcano", outputfolder = ".", draw = "PDF")
```

## Arguments

<code>remres</code>	MetaVolcano object. Output of the <code>rem_mv()</code> function <MetaVolcano>
<code>gene</code>	query gene to plot
<code>genecol</code>	name of the variable with genes <string>
<code>foldchangepcol</code>	the column name of the foldchange variable <string>
<code>llcol</code>	left limit of the fold change confidence interval variable name <string>
<code>rlcol</code>	right limit of the fold change confidence interval variable name <string>
<code>jobname</code>	name of the running job <string>
<code>outputfolder</code>	/path where to write the results/ <string>
<code>draw</code>	either 'PDF' or 'HTML' to save metaolcano as .pdf or .html respectively <string>

## Value

ggplot2 object

## Examples

```
data(diffexplist)
diffexplist <- lapply(diffexplist, function(del) {
  dplyr::filter(del, grepl("MP", Symbol))
})
mv <- rem_mv(diffexplist, metathr = 0.1)
gg <- draw_forest(mv, gene="MMP9")
plot(gg)
```



---

MetaVolcano-class	<i>An S4 class to represent MetaVolcanoR results</i>
-------------------	--

---

**Description**

An S4 class to represent MetaVolcanoR results

**Slots**

input merged differential expression inputs data.frame  
inputnames names of the differential expression inputs character  
metaresult meta-analysis results data.frame  
MetaVolcano plot with meta-analysis results  
degfreq supplementary figure of the vote-counting MetaVolcano

---

plot_mv	<i>A MetaVolcano plotting function</i>
---------	--

---

**Description**

This function plots either the combining- or the vote-counting- MetaVolcanos

**Usage**

```
plot_mv(meta_diffexp, nstud, genecol, comb, metafc)
```

**Arguments**

meta_diffexp	data.frame/data.table containing the differential expression inputs
nstud	the number of differential expression inputs <integer>
genecol	column name of the variable to label genes in the .html file <string>
comb	whether or not the drawing is for the combining-metavolcano <logical>
metafc	method for summarizing gene fold-changes across studies c("Mean", "Median") <string>

**Value**

ggplot2 object

**Examples**

```
data(diffexplist)
mv <- votecount_mv(diffexplist)
gg <- plot_mv(mv@metaresult, length(diffexplist), "Symbol", FALSE, "Mean")
plot(gg)
```

---

plot_rem	<i>A function to plot the Random Effect Model (REM) MetaVolcano</i>
----------	---

---

**Description**

This function plots the REM MetaVolcano using ggplot2

**Usage**

```
plot_rem(meta_diffexp, jobname, outputfolder, genecol, metathr)
```

**Arguments**

meta_diffexp	data.frame/data.table containing the REM results from rem_mv() <data.table/data.frame>
jobname	name of the running job <string>
outputfolder	/path where to write the results/ <string>
genecol	column name of the variable to label genes in the .html file <string>
metathr	top percentage of perturbed genes to be highlighted <double>

**Value**

ggplot2 object

**Examples**

```
data(diffexplist)
diffexplist <- lapply(diffexplist, function(del) {
  dplyr::filter(del, grepl("MP", Symbol))
})
mv <- rem_mv(diffexplist, metathr = 0.1)
gg <- plot_rem(mv@metaresult, "MV", ".", "Symbol", 0.01)
plot(gg)
```

---

remodel	<i>A function to model foldchange variance along several studies This function calculate the REM-summary fold-change</i>
---------	--

---

**Description**

A function to model foldchange variance along several studies This function calculate the REM-summary fold-change

**Usage**

```
remodel(gene, foldchange_col, vcol)
```

**Arguments**

gene                named vector with foldchanges and variances <vector>  
 foldchangeacol    the column name of the foldchange variable <string>  
 vcol                name of the fold change variance variable <string>

**Value**

data.frame with REM results for a gene

**Examples**

```
g <- data.frame('Symbol'="XGENE", 'Log2FC_1'=1.2, 'Log2FC'=0.8,
                'vi_1'=0.01, 'vi_2'=0.1)
remodel(g, 'Log2FC', 'vi')
```

rem\_mv

*A function to perform the Random Effect Model (REM) MetaVolcano***Description**

This function runs the 'Random Effect Model' MetaVolcano section

**Usage**

```
rem_mv(diffexp = list(), pcriteria = "pvalue",
        foldchangeacol = "Log2FC", genenamecol = "Symbol", geneidcol = NULL,
        collaps = FALSE, llcol = "CI.L", rlcol = "CI.R", vcol = NULL,
        cvar = TRUE, metathr = 0.01, jobname = "MetaVolcano",
        outputfolder = ".", draw = "HTML", ncores = 1)
```

**Arguments**

diffexp            list of data.frame/data.table (s) with DE results where lines are genes  
 pcriteria          the column name of the pvalue variable <string>  
 foldchangeacol    the column name of the foldchange variable <string>  
 genenamecol       the column name of the gene name variable <string>  
 geneidcol          the column name of the gene ID/probe/oligo/transcript variable <string>  
 collaps            if probes should be collapsed based on the DE direction <logical>  
 llcol              left limit of the fold change confidence interval variable name <string>  
 rlcol              right limit of the fold change confidence interval variable name <string>  
 vcol                name of the fold change variance variable <string>  
 cvar                weather or not to calculate gene variance from confidence interval limits <logical>  
 metathr            top percentage of perturbed genes to be highlighted <double>

jobname            name of the running job <string>  
 outputfolder      /path where to write the results/  
 draw                wheather or not to draw the .html visualization <logical>  
 ncores             the number of processors the user wants to use <integer>

**Value**

MetaVolcano object

**Examples**

```

data(diffexplist)
diffexplist <- lapply(diffexplist, function(del) {
  dplyr::filter(del, grepl("MP", Symbol))
})
mv <- rem_mv(diffexplist, metathr = 0.1)
str(mv)
  
```

---

rename_col	<i>A column renaming function merged inputs</i>
------------	---

---

**Description**

This function rename the columns of the merged inputs

**Usage**

```
rename_col(diffexp, genecol)
```

**Arguments**

diffexp            list of data.frame/data.table (s) with DE results where lines are genes  
 genecol            the column name of the geneID or gene name variable <string>

**Value**

data.table/data.frame with new colnames

**Examples**

```

data(diffexplist)
lapply(diffexplist, colnames)
diffexp <- rename_col(diffexplist, "Symbol")
lapply(diffexp, colnames)
  
```

---

set_degbar_data	<i>A function setting data format for DEG barplot visualization</i>
-----------------	---

---

**Description**

This function summarize the variable deg from the deg\_def() function to visualize as barplots the number of DEGs per input study

**Usage**

```
set_degbar_data(diffexp)
```

**Arguments**

diffexp            list of data.frame/data.table (s) output of the deg\_def() function <list>

**Value**

data.frame DEG by input

**Examples**

```
data(diffexplist)
diffexp <- lapply(diffexplist, function(...) deg_def(..., "pvalue",
  "Log2FC", 0.05, 0))
bardat <- set_degbar_data(diffexp)
head(bardat, 3)
```

---

votecount_mv	<i>A function to draw the 'Vote-counting meta-analysis' MetaVolcano</i>
--------------	---

---

**Description**

This function draws the vote-counting meta-analysis MetaVolcano

**Usage**

```
votecount_mv(diffexp = list(), pcriteria = "pvalue",
  foldchangepcol = "Log2FC", genenamecol = "Symbol", geneidcol = NULL,
  pvalue = 0.05, foldchange = 0, metathr = 0.01, collaps = FALSE,
  jobname = "MetaVolcano", outputfolder = ".", draw = "HTML")
```

**Arguments**

diffexp	list of data.frame/data.table (s) with DE results where lines are genes
pcriteria	the column name of the Pval criteria to consider <string>
foldchangecol	the column name of the foldchange variable <string>
genenamecol	the column name of the gene name variable <string>
geneidcol	the column name of the gene ID/probe/oligo/transcript variable <string>
pvalue	the Pval to use as threshold c(0:1) <double>
foldchange	the foldchange to use as DE threshold c(-Inf: Inf) <double>
metathr	the proportion of studies a gene has to be DEG to be considered cDEG <double>
collaps	if probes should be collapsed based on the DE direction <logical>
jobname	name of the running job <string>
outputfolder	/path where to write the results/
draw	whether or not to draw a .pdf or .html visualization <c(NULL, 'PDF', 'HTML')>

**Value**

MetaVolcano object

**Examples**

```
data(diffexplist)
mv <- votecount_mv(diffexplist)
str(mv)
```

# Index

- \* **'combining'**
  - combining\_mv, 3
  - rem\_mv, 11
- \* **'vote-counting'**
  - votecount\_mv, 13
- \* **DEG**
  - cum\_freq\_data, 4
  - deg\_def, 5
  - draw\_cum\_freq, 6
  - draw\_degbar, 7
  - set\_degbar\_data, 13
- \* **Probe**
  - collapse\_deg, 3
- \* **REM**
  - plot\_rem, 10
  - remodel, 10
- \* **barplot**
  - draw\_degbar, 7
  - set\_degbar\_data, 13
- \* **collapse**
  - collapse\_deg, 3
- \* **column**
  - rename\_col, 12
- \* **confidence**
  - calc\_vi, 2
- \* **datasets**
  - diffexplist, 6
- \* **data**
  - cum\_freq\_data, 4
  - set\_degbar\_data, 13
- \* **definition**
  - deg\_def, 5
- \* **distribution**
  - cum\_freq\_data, 4
  - draw\_cum\_freq, 6
- \* **draw**
  - draw\_cum\_freq, 6
  - draw\_degbar, 7
  - draw\_forest, 8
- plot\_mv, 9
- \* **forest-plot**
  - draw\_forest, 8
- \* **format**
  - cum\_freq\_data, 4
  - set\_degbar\_data, 13
- \* **from**
  - calc\_vi, 2
- \* **gene**
  - draw\_forest, 8
- \* **interval**
  - calc\_vi, 2
- \* **inverse-cummulative**
  - cum\_freq\_data, 4
  - draw\_cum\_freq, 6
- \* **'meta-analysis'**
  - combining\_mv, 3
  - rem\_mv, 11
  - votecount\_mv, 13
- \* **metavolcano**
  - combining\_mv, 3
  - plot\_mv, 9
  - plot\_rem, 10
  - rem\_mv, 11
  - votecount\_mv, 13
- \* **rename**
  - rename\_col, 12
- \* **set**
  - set\_degbar\_data, 13
- \* **summary**
  - remodel, 10
- \* **variance**
  - calc\_vi, 2
- \* **write**
  - combining\_mv, 3
  - plot\_rem, 10
  - rem\_mv, 11
  - votecount\_mv, 13

calc\_vi, 2

collapse\_deg, 3  
combining\_mv, 3  
cum\_freq\_data, 4  
  
deg\_def, 5  
diffexplist, 6  
draw\_cum\_freq, 6  
draw\_degbar, 7  
draw\_forest, 8  
  
MetaVolcano-class, 9  
  
plot\_mv, 9  
plot\_rem, 10  
  
rem\_mv, 11  
remodel, 10  
rename\_col, 12  
  
set\_degbar\_data, 13  
  
votecount\_mv, 13