Package: moal (via r-universe)

November 4, 2026

Title Multi Omic Analysis at Lab

Version 1.2.3

Description Multi Omic Analysis at Lab.

License GPL-3

Depends foreach, ggplot2, igraph, magrittr, dplyr, R (>= 4.3)

Imports broom, colourvalues, dendextend, doParallel, edgeR, fgsea, forcats, ggforce, ggpubr, ggrepel, gplots, graphics, grDevices, gridExtra, limma, moalannotensg, moalannotensp, moalannotenst, moalannotgene, moalstringdbdr, moalstringdbhs, moalstringdbmm, moalstringdboa, moalstringdbrn, moalstringdbss, parallel, plyr, Rgraphviz, rlang, scales, stats, stringr, tidyselect, utils

Suggests knitr, rmarkdown

VignetteBuilder knitr

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 8.0.0

Config/pak/sysreqs
cmake libfontconfig1-dev libfreetype6-dev libglpk-dev make libicu-dev libxml2-dev zlib1g-dev

Repository https://fdumbioinfo.r-universe.dev

Date/Publication 2026-11-04 01:51:12 UTC

RemoteUrl https://github.com/fdumbioinfo/moal

RemoteRef HEAD

RemoteSha 3103b455464adcd69bd9722f87a615d8ff2c101b

Contents

annotdb 2
ena 3
heatmap 5
input 7

norm 7
omic 8
output 13
qc 14
replacena 15
venn 16
volcanoplot 17

Index **19**

---

annotdb *Annotation function for Symbol, NCBI or Ensembl IDs*

Description

Annotation function for Symbol, NCBI or Ensembl IDs

Usage

```
annotdb(
  symbollist = NULL,
  species = NULL,
  ortholog = F,
  dboutput = "ncbi",
  idtype = NULL
)
```

Arguments

symbollist	character list of IDs or Symbols
species	character available species: hs mm rn ss pt bt oa dr gg xt dm ce
ortholog	logical if TRUE return homo sapiens ortholog for choosen species
dboutput	character database used for output 'ncbi'(default) or 'ebi'
idtype	character database ID accepted: 'SYMBOL'(default), 'GENE', 'ENST', 'ENSG', 'ENSP'

Details

Use moal:::orthoinfo to see complete species list

Value

data.frame

Author(s)

Florent Dumont florent.dumont@universite-paris-saclay.fr

Examples

```
# not run
# annot(Symbol)
```

---

ena *Gene set enrichment analysis and interaction network*

Description

Gene set enrichment analysis and interaction network

Usage

```
ena(
  omicdata = NULL,
  gmtfiles = NULL,
  species = "hs",
  dat = NULL,
  factor = NULL,
  filtergeneset = NULL,
  threshold = 1,
  topdeg = 100,
  rangedeg = NULL,
  topena = 50,
  topgeneset = 50,
  intmaxdh = 5000,
  nodesize = 0.6,
  bg = 25000,
  doena = TRUE,
  gsearank = "logfc",
  gseatail = "twotail",
  layout = 1,
  mings = 5,
  maxgs = 700,
  overlapmin = 2,
  addratioena = TRUE,
  addenarankbarplot = TRUE,
  dotopnetwork = TRUE,
  dotopgenesetnetwork = FALSE,
  dogmtgenesetnetwork = FALSE,
  dotopheatmap = TRUE,
  dotopgenesetheatmap = TRUE,
  dogmtgenesetheatmap = TRUE,
  path = NULL,
  dirname = NULL,
  dopar = TRUE
)
```

Arguments

omicdata	character data.frame see details
gmtfiles	character gmt files list path
species	character available species: hs mm rn ss pt bt oa dr gg xt dm ce
dat	data.frame file paths
factor	factor factor for heatmap color
filtergeneset	character list to filter MSigDB geneset collection
threshold	numeric pval 0.05 fc 1.5 by default see details
topdeg	numeric top feature to plot on network
rangedeg	numeric top DEGs from 1 to topdeg by rangedeg to plot on network
topena	numeric top geneset for ena plot
topgeneset	numeric top geneset number to plot on network
intmaxdh	numeric maximum number of interaction to use for Davidson and Harel algorithm layout
nodesize	numeric change Symbol size
bg	numeric background used for functional analysis over-representation test
doena	logical do MSigDB enrichment analysis
gsearank	character to choose gsea rank type among fc (by default) logration logfc sqrt
gseatail	character to choose gsea twotail (by default) or onetail
layout	numeric for layout neetwork 1 fr by default 2 dh 3 tree 4 circle 5 grid 6 sphere
mings	numeric minimal size of a gene set
maxgs	numeric maximal size of a gene set
overlapmin	numeric minimal overlap to keep for gene set analysis
addratioena	logical if TRUE add overlap and geneset size on enrichment barplot
addenarankbarplot	logical if TRUE add ena barplot ranked by NES score
dotopnetwork	logical do top networks
dotopgenesetnetwork	logical do geneset networks
dogmtgenesetnetwork	logical do keyword networks
dotopheatmap	logical do top heatmap
dotopgenesetheatmap	logical do geneset heatmap
dogmtgenesetheatmap	logical do keyword heatmap
path	character for relative path of output directory
dirname	character name for output
dopar	logical TRUE for parallelization

Details

omicdata needs a data.frame with at list 4 column: rowID, (p-values,fold-change) x N and Symbol annotation.

Symbol list are accepted to make ORA enrichment analysis.

To generate heatmap dat and factor parameter are needed. dat accepted complete matrix with rowID for first column.

dat row IDs must match with omicdata row IDs.

Make MSigDB enrichment analysis using GSEA method for non filtering list as input (> 2000)

Make MSigDB Over-Representation enrichment analysis (ORA) using Fisher exact test for list < 2000

Generate STRINGDB interaction network and heatmap for top geneset according to topena par (80 by default)

Only features with p-values < 0.05 et fold-change > 1.1 are displayed on geneset heatmaps (threshold = 1 by default).

See omic function details to display all threshold

Value

file with enrichment analysis results

Author(s)

Florent Dumont florent.dumont@universite-paris-saclay.fr

Examples

```
# not run
# ena( omicdata , species = "mm")
```

`heatmap` *Heatmap*

Description

To make a heatmap

Usage

```
heatmap(
  dat = NULL,
  factor = NULL,
  method = "complete",
  dendrogram = "both",
  k = NULL,
  labCol = "",
```

```
  cexCol = 0.5,
  labRow = "",
  cexRow = NULL,
  cexlegend = 0.65,
  keysize = 0.9,
  keycolor = c("darkgreen", "orange", "darkred"),
  parmar = c(5, 4, 5, 6),
  scale = "row"
)
```

Arguments

dat	matrix numeric
factor	factor
method	character
dendrogram	character to display 'none', 'row', 'column' or 'both' (by default) dendrograms
k	numeric number of clusters to colorize for rows
labCol	character
cexCol	numeric
labRow	Character
cexRow	numeric
cexlegend	numeric
keysize	numeric
keycolor	character of 3 for low mid high value of the key
parmar	numeric 4 values for margin sizes
scale	numeric standardize row by defaut and column or none accepted

Details

To make a heatmap from a matrix or a data.frame

Value

no returned value

Author(s)

Florent Dumont florent.dumont@universite-paris-saclay.fr

Examples

```
# not run
# library(magrittr)
# data(sif1)
# data(mat1)
# mat1 %>% heatmap(sif1$F3)
```

input *import tab file in data.frame*

Description

import tab file in data.frame

Usage

```
input(filename, sep = "\t", quote = "")
```

Arguments

filename	character path to the file to read
sep	character for field separator
quote	character for field quote

Details

wrapper of read.table function for tabular separated files

Value

data.frame

Author(s)

Florent Dumont florent.dumont@universite-paris-saclay.fr

Examples

```
# not run
# input( "filename" ) -> dt
```

norm *Normalization function*

Description

Apply normalization method and log2 on numerical table

Usage

```
norm(dat = NULL, method = "quantile", log = TRUE)
```

Arguments

dat	data.frame data.frame without row ID column
method	character apply quantile normalization by default see details for all methods
log	logical apply log base 2

Details

Normalization method available: "TMM","TMMwsp","RLE","upperquartile" from EdgeR and "quantile","scale","cyclicloess" from limma "none" for no normalization

Value

data.frame

Author(s)

Florent Dumont florent.dumont@universite-paris-saclay.fr

Examples

```
# not run
# norm(dt)
```

---

omic *Omic bioanalysis workflow*

Description

Omic function workflow description:

- Quality controls and unsupervised analysis: histogram, box plot, PCA and sample clustering.
- Supervised analysis: analysis of variance (ANOVA) and filter application.
- Unsupervised analysis for selected features: row clustering, PCA and pattern search across factor levels.
- Graph generation for selected feature: volcanoplots, heatmaps, lineplots, boxplots, PCA
- Functional analysis: MSigDB enrichment analysis and STRINGDB interaction network

See help("omic") section to test workflow with internal GEO data set GSE65055 and reproduce enrichment results for chromosome cytogenetic bands (doi: 10.1111/cge.12731)

Usage

```
omic(
  dat = NULL,
  sif = NULL,
  annot = NULL,
  species = "hs",
  model = NULL,
  paired = NULL,
  nested = NULL,
  batch = NULL,
  addfactor = NULL,
  doqc = TRUE,
  threshold = c(1, 2, 3, 4, 9, 10, 11, 12),
  padj = "none",
  anovafc = "anovafc",
  logratio = FALSE,
  dopattern = TRUE,
  dovenn = TRUE,
  docluster = TRUE,
  nc = c(2, 3, 6, 12),
  maxclusterheatmap = 5000,
  doheatmap = TRUE,
  heatmapcluster = "row",
  maxheatmap = 2000,
  minheatmap = 3,
  dovolcanoplot = TRUE,
  nbgenevolc = 5,
  dolineplot = TRUE,
  doboxplotrow = TRUE,
  doena = TRUE,
  gsearank = "logfc",
  gseatail = "twotail",
  topdeg = 100,
  topena = 50,
  doenaora = FALSE,
  maxenaoralist = 2000,
  minenaoralist = 5,
  gmtfiles = NULL,
  filtergeneset = NULL,
  bg = 25000,
  dotopnetwork = TRUE,
  dotopheatmap = TRUE,
  layout = 2,
  mings = 5,
  maxgs = 700,
  overlapmin = 2,
  addenarankbarplot = TRUE,
  dotopgenesetnetwork = FALSE,
```

```
  dotopgenesetheatmap = TRUE,
  dogmtgenesetnetwork = FALSE,
  dogmtgenesetheatmap = TRUE,
  crosscompint = FALSE,
  sample = NULL,
  seed = 123679,
  dopar = TRUE,
  nbcore = NULL,
  nrowsim = 2000,
  libloc = NULL,
  savehistory = TRUE,
  path = ".",
  dirname = NULL,
  zip = FALSE,
  remove = FALSE
)
```

Arguments

dat	data.frame normalize data table with rowID for first column
sif	data.frame sample information file including model factors
annot	data.frame annotation with Symbol column for functional analysis
species	character available species: hs mm rn ss pt bt oa dr gg xt dm ce
model	character anova model factors (see details)
paired	character factor for paired design
nested	character factor for nested design
batch	character factor for batch effect design
addfactor	character additionnal factors
doqc	logical quality controls
threshold	numeric vector from 1 to 24 (see details)
padj	character fdr by defaut for Benjamini-Hochberg false discovery correction
anovafc	character fold-change calculate with anova model (by default) use meanfc or meanfcdelog
logratio	logical change fc (by default) in log2ratio
dopattern	logical search relevant pattern across levels factor
dovenn	logical venn diagram
docluster	logical row hierarchical clustering using pearson correlation
nc	numeric number of clusters to cut in dendrogramm
maxclusterheatmap	numeric max row for cluster analysis
doheatmap	logical do heatmaps for all lists
heatmapcluster	character row clustering only by default both accepted

maxheatmap	numeric max rows for heatmap
minheatmap	numeric min rows for heatmap
dovolcanoplot	logical make volcanoplot for each threshold
nbgenevolc	numeric number of Symbol to display in volcanoplot
dolineplot	logical do lineplot for significant features
doboxplotrow	logical do boxplot for significant features with Kruskal
doena	logical msigdb enrichement analysis using gsea method without filtering
gsearank	character gsea rank type among logfc (by default) logration fc sqrt
gseatail	character gsea twotail test (by default) or onetail
topdeg	numeric top DEGs number to plot on network
topena	numeric top geneset for ena plot
doenaora	logical msigdb enrichement analysis using ora method for diff list
maxenaoralist	numeric max diff list size to do enrichment analysis with ora fir diff list
minenaoralist	numeric max diff list size to do enrichment analysis with ora fir diff list
gmtfiles	character gmt files list path
filtergeneset	character regular expression to filter collection geneset (e.g. "reactomeltft")
bg	numeric background used for functional analysis over-representation test
dotopnetwork	logical do top networks
dotopheatmap	logical do top heatmap
layout	numeric for layout neetwork 1 fr by default 2 dh 3 tree 4 circle 5 grid 6 sphere
mings	numeric minimal size of a gene set
maxgs	numeric maximal size of a gene set
overlapmin	numeric minimal overlap to keep for gene set analysis
addenarankbarplot	logical if TRUE add ena barplot ranked by NES score
dotopgenesetnetwork	logical do geneset networks
dotopgenesetheatmap	logical do geneset heatmap
dogmtgenesetnetwork	logical do keyword networks
dogmtgenesetheatmap	logical do keyword heatmap
crosscompint	logical add cross comparison to results for interaction model
sample	numeric analysis using random subset
seed	numeric seed for random function
dopar	logical if TRUE do parallel computation
nbcore	numerical core number all available core if NULL
nrowsim	numeric row number for normal distribution simulate matrix analysis

`libloc`	character path to R local library if need
`savehistory`	logical save r session history
`path`	character results directory path
`dirname`	character results directory name
`zip`	logical compress results directory if TRUE
`remove`	logical remove uncompress results directory if TRUE

Details

Use moal::env() to load required libraries before moal::omic() (see example)

Use input() function to import and analyse your own data starting from tsv file (or csv with sep = ",")

dat must have one IDs columns in the same order than annotations.

Use annot() function for annotation with Symbol, NCBI, Ensembl IDs.

sif must contains column with description sample corresponding to anova factor analysis.

sif rows must have the same number of samples in the same order that in the dat table.

Experimental design examples for model parameters:

- 1-way anova: model = "TREATMENT"
- 2-ways anova: model = "PHENOTYPE+TREATMENT"
- 2-ways anova with interaction: model = "TREATMENT+TIME+TREATMENT*TIME"
- 2-ways anova with paired factor: model = "TREATMENT", paired = "CASE"
- 2-ways anova with batch factor: model = "TREATMENT", batch = "BATCH"
- 2-ways anova with nested factor: model = "TREATMENT", nested = "CASEinTREATMENT"
- 3-ways or 4-ways anova (without interaction): model = "PHENOTYPE+TREATMENT+AGE"

For paired, batch and nested design, remove batch effect from limma package are used to calculate fold-change

Use dopar = 2 to decrease computing resources.

Use sample for random subset analysis.

To see complete threshold list: moal:::thresholdlist %>% lapply("[",c(1,2)) %>% unlist %>% matrix(ncol=2,byrow = T) %>% data.frame %>% setNames(c("pval","fc"))

Annotation updates: 22-04-2025 for gene and ensembl, MSigDB 2024.1.Hs, StringDB 12.0

Value

omic results directory

Author(s)

Florent Dumont florent.dumont@universite-paris-saclay.fr

Examples

```
# # Test workflow with internal GEO data set GSE65055
# # and reproduce enrichment results for chromosome cytogenetic bands (doi: 10.1111/cge.12731)
# # loading libraries:
# library(moal);moal::env()
# # loading data:
# moal:::GSE65055normdata -> dat
# moal:::GSE65055sampledata -> sif
# # Ordering factors for pairwise comparisons which compute contrast p-values and fold-changes.
# sif$ANEUPLOIDY %>% ordered(c("Control","T13","T18","T21")) -> sif$ANEUPLOIDY
# sif$TISSUE %>% as.factor -> sif$TISSUE
# # annotation
# dat$rowID %>% moal::annot(species= "hs",idtype="GENE",dboutput="ncbi") -> annot
# # omic analysis
# moal::omic(dat,sif,annot,species="hs",model="ANEUPLOIDY",batch="TISSUE",dirname="GSE65055")
```

`output` *export data.frame in tab file*

Description

export data.frame in tab file

Usage

```
output(dt, filename)
```

Arguments

`dt`	data.frame
`filename`	character

Author(s)

Florent Dumont florent.dumont@universite-paris-saclay.fr

Examples

```
# not run
# output( dt )
```

qc *Quality Controls*

Description

Descriptive analysis applied on column:

- histogram, boxplot, hierarchical clustering and PCA for column

Usage

```
qc(
  dat,
  sif = NULL,
  dooutputinput = FALSE,
  dohisto = TRUE,
  doboxplot = TRUE,
  dohc = TRUE,
  dopca = TRUE,
  breaks = 70,
  dirname = NULL,
  path = "."
)
```

Arguments

dat	data.frame first column for rowID column
sif	data.frame sample information file
dooutputinput	logical if TRUE (by default) export input data
dohisto	logical if TRUE (by default) do histogram
doboxplot	logical if TRUE (by default) do boxplot
dohc	logical if TRUE (by default)do hierarchical clustering
dopca	logical if TRUE (by default) do PCA
breaks	numeric break number for histogramm function
dirname	character
path	character

Details

dat row must be equal to sif row

Value

directory including analysis pdf plots

Author(s)

Florent Dumont florent.dumont@universite-paris-saclay.fr

Examples

```
# not run
# qc(dat,sif)
```

`replacena` *replace 0 value by background noise estimate*

Description

replace 0 value by background noise estimate

Usage

```
replacena(
  dat = NULL,
  value = NULL,
  factor = NULL,
  ns = NULL,
  pval = 0.05,
  quantv = 0.01,
  logbgvalue = FALSE
)
```

Arguments

`dat`	character vector of mzxml file
`value`	numeric value to substitute
`factor`	character factor for remove non random 0 value distribution across factor levels
`ns`	number 0 number to remove row
`pval`	number fisher test p-value threshold for 0 distribution accross
`quantv`	number quantile probability for background noise value (1st percentile by default)
`logbgvalue`	logical background noise value logged if TRUE

Details

replace value by column median if group size is one and replace by row group median if group size > 1

Value

data.frame

Author(s)

Florent Dumont florent.dumont@universite-paris-saclay.fr

Examples

```
# not run
# replacegroupmed(dat)
```

---

venn *Venn diagramm*

Description

To make a Venn diagramm of 2, 3 or 4 lists

Usage

```
venn(
  list = NULL,
  listnames = NULL,
  returnlist = F,
  title = "Venn Diagram",
  plot = T,
  export = F,
  path = ".",
  dirname = "venn"
)
```

Arguments

list	list of 2 , 3 or 4 character vector
listnames	character list names to display on graph
returnlist	logical
title	character title to display on graph
plot	logical to display the plot or not
export	logical export lists in a directory
path	character
dirname	character name of the directory created when export = T

Value

venn plot and new lists generated by venn.

Author(s)

Florent Dumont florent.dumont@universite-paris-saclay.fr

Examples

```
# library(magrittr)
# list(
#   c(letters[6:20] , letters[25] ) ,
#   letters[1:15] ,
#   c( letters[2:5] , letters[8:23] ) ) %>% moal::venn(.)
```

`volcanoplot` *Volcanoplot*

Description

Do volcanoplot

Usage

```
volcanoplot(
  dat = NULL,
  pval = 0.05,
  fc = 1.5,
  topgenename = TRUE,
  topgenenamen = 5,
  genenamelist = NULL,
  genenamesize = 2,
  title = "Volcanoplot"
)
```

Arguments

`dat`	data.frame table with 4 columns (see details)
`pval`	numeric p-value threshold
`fc`	numeric fold-change threshold
`topgenename`	logical display gene label TRUE by default
`topgenenamen`	numeric increase number of gene label
`genenamelist`	character vector of gene list to label
`genenamesize`	numeric label size for gene name
`title`	character

Details

dat parameter must have 4 columns: rowID , p_AvsB , fc_AvsF and Symbol.

Value

no returned value

Author(s)

Florent Dumont florent.dumont@universite-paris-saclay.fr

Examples

```
# not run
# data.frame(rowID,p_AvsB,fc_AvsB,Symbol) -> dat
# volcanoplot(dat)
```

Index

annotdb, 2

ena, 3

heatmap, 5

input, 7

norm, 7

omic, 8
output, 13

qc, 14

replacena, 15

venn, 16
volcanoplot, 17