Package 'MPAC'

December 2, 2025

Title Multi-omic Pathway Analysis of Cells

Version 1.4.0

Description Multi-omic Pathway Analysis of Cells (MPAC), integrates multi-omic data for understanding cellular mechanisms. It predicts novel patient groups with distinct pathway profiles as well as identifying key pathway proteins with potential clinical associations. From CNA and RNA-seq data, it determines genes' DNA and RNA states (i.e., repressed, normal, or activated), which serve as the input for PARADIGM to calculate Inferred Pathway Levels (IPLs). It also permutes DNA and RNA states to create a background distribution to filter IPLs as a way to remove events observed by chance. It provides multiple methods for downstream analysis and visualization.

```
License GPL-3
Encoding UTF-8
Roxygen list(markdown = TRUE)
RoxygenNote 7.3.2
Depends R (>= 4.4.0)
URL https://github.com/pliu55/MPAC
```

BugReports https://github.com/pliu55/MPAC/issues

```
Imports data.table (>= 1.14.2), SummarizedExperiment (>= 1.30.2), BiocParallel (>= 1.28.3), fitdistrplus (>= 1.1), igraph (>= 1.4.3), BiocSingular (>= 1.10.0), S4Vectors (>= 0.32.3), SingleCellExperiment (>= 1.16.0), bluster (>= 1.4.0), fgsea (>= 1.20.0), scran (>= 1.22.1), ComplexHeatmap (>= 2.16.0), circlize (>= 0.4.16), scales (>= 1.3.0), stringr (>= 1.5.1), viridis (>= 0.6.5), ggplot2 (>= 3.5.1), ggraph (>= 2.2.1), survival (>= 3.7), survminer (>= 0.4.9), grid, stats
```

Suggests rmarkdown, knitr, svglite, bookdown(>= 0.34), testthat (>= 3.0.0)

Config/testthat/edition 3

VignetteBuilder knitr

biocViews Software, Technology, Sequencing, RNASeq, Survival, Clustering, ImmunoOncology

2 Contents

SystemsRequirements The `runPrd()` function requires an external software named PARADIGM. For details, please see the 'Required external software' section in vignette's 'Run PARADIGM: runPrd()'.

git_url https://git.bioconductor.org/packages/MPAC
git_branch RELEASE_3_22
git_last_commit_778e22d
git_last_commit_date 2025-10-29

Repository Bioconductor 3.22

Date/Publication 2025-12-01

Author Peng Liu [aut, cre] (ORCID: https://orcid.org/0000-0001-5655-2259), Paul Ahlquist [aut], Irene Ong [aut], Anthony Gitter [aut]

Contents

Index

Samp	3
olPermIPL	3
olRealIPL	4
onMtf	5
tByPerm	5
etSignifOvrOnCl	6
vrGMT	7
ltConMtf	8
ltMtfPrtIPL	8
ltNeiStt	9
ltOvrHm	10
ltSttKM	11
pCnInp	12
pPermInp	12
pRealInp	13
pRnaInp	14
pRunPrd	14
nnPermPrd	15
ınPrd	16
ıbNtw	17

18

clSamp 3

clSamp	Cluster samples by pathway over-representation
	7 · · · · · · · · · · · · · · · · · · ·

Description

Cluster samples by pathway over-representation

Usage

```
clSamp(ovrmat, n_neighbors = 10, n_random_runs = 200, threads = 1)
```

Arguments

ovrmat A matrix of gene set over-representation adjusted p-values with rows as gene

sets and columns as samples. It is the output from ovrGMT().

size of samples is large. Default: 10.

n_random_runs Number of random runs. Due to randomness introduced to the Louvain algo-

rithm in R igraph 1.3.0 (https://github.com/igraph/rigraph/issues/539), a large number of runs are recommended to evaluate randomness in the clustering results. Default: 200, which shall be safe for sample size < 50. Please increase it

accordingly for a larger sample size.

threads Number of threads to run in parallel. Default: 1

Value

A data table with each row representing one clustering result, and the first column denotes the number of occurrences of a clustering result and the rest of columns indicating each sample's cluster index. Rows are ordered by the number of occurrences from high to low.

Examples

```
fovr = system.file('extdata/clSamp/ovrmat.rds', package='MPAC')
ovrmat = readRDS(fovr)
clSamp(ovrmat)
```

colPermIPL Collect Inferred Pathway Levels (IPLs) from PARADIGM runs on permuted data

Description

Collect Inferred Pathway Levels (IPLs) from PARADIGM runs on permuted data

```
colPermIPL(indir, n_perms, sampleids = NULL, threads = 1)
```

4 colRealIPL

Arguments

indir Input folder that saves PARADIGM results. It should be set as the same as

outdir as in runPrd().

n_perms Number of permutations to collect.

sample ids Sample IDs for which IPLs to be collected. If not provided, all files with suffix

'_ipl.txt' in indir will be collected. Default: NULL.

threads Number of threads to run in parallel. Default: 1

Value

A data.table object with columns of permutation index, pathway entities and their IPLs.

Examples

```
indir = system.file('/extdata/runPrd/', package='MPAC')
n_perms = 3
colPermIPL(indir, n_perms)
```

colRealIPL Collect Inferred Pathway Levels (IPLs) from PARADIGM runs on real

data

Description

Collect Inferred Pathway Levels (IPLs) from PARADIGM runs on real data

Usage

```
colRealIPL(indir, sampleids = NULL, file_tag = NULL)
```

Arguments

indir Input folder that saves PARADIGM results. It should be set as the same as

outdir as in runPrd().

sampleids Sample IDs for which IPLs to be collected. If not provided, all files with suffix

'_ipl.txt' in indir will be collected. Default: NULL.

file_tag A string of output file name tag. Default: NULL

Value

A data.table object with columns of pathway entities and their IPLs.

```
indir = system.file('/extdata/runPrd/', package='MPAC')
colRealIPL(indir)
```

conMtf 5

			_
CO	 N 4	_	£

Find consensus pathway motifs from a list of pathways

Description

Find consensus pathway motifs from a list of pathways

Usage

```
conMtf(subntwl, omic_genes = NULL, min_mtf_n_nodes = 5)
```

Arguments

subntwl A list of igraph objects representing input pathways from different samples. It

is the output from subNtw()

omic_genes A vector of gene symbols to narrow down over-representation calculation to

only those with input genomic data. If not provided, all genes in the GMT file

will be considered. Default: NULL.

min_mtf_n_nodes

Number of minimum nodes in a motif. Default: 5

Value

A list of igraph objects representing consensus pathway motifs

Examples

```
fsubntwl = system.file('extdata/conMtf/subntwl.rds', package='MPAC')
subntwl = readRDS(fsubntwl)

fomic_gns = system.file('extdata/TcgaInp/inp_focal.rds', package='MPAC')
omic_gns = rownames(readRDS(fomic_gns))

conMtf(subntwl, omic_gns, min_mtf_n_nodes=50)
```

fltByPerm

Filter IPLs from real data by distribution from permuted data

Description

Filter IPLs from real data by distribution from permuted data

```
fltByPerm(realdt, permdt, threads = 1)
```

6 getSignifOvrOnCl

Arguments

realdt A data table object containing entities and their IPLs from real data. It is the

output from colRealIPL().

permdt A data.table object containing permutation index, entities and their IPLs from

permuted data. It is the output from colPermIPL().

threads Number of threads to run in parallel. Default: 1

Value

A matrix of filtered IPLs with rows as entities and columns as samples. Entities with IPLs observed by chance are set to NA.

Examples

```
freal = system.file('extdata/fltByPerm/real.rds', package='MPAC')
fperm = system.file('extdata/fltByPerm/perm.rds', package='MPAC')
realdt = readRDS(freal)
permdt = readRDS(fperm)

fltByPerm(realdt, permdt)
```

getSignifOvrOnCl

Get significantly over-represented gene sets for clustered samples

Description

Get significantly over-represented gene sets for clustered samples

Usage

```
getSignifOvrOnCl(ovrmat, cldt, min_frc = 0.8)
```

Arguments

ovrmat A matrix containing over-representation adjusted P with rows as gene set names

and columns as sample IDs. It is the output of the ovrGMT() function.

cldt A data table with each row representing one clustering result, and the first col-

umn denotes the number of occurrences of a clustering result and the rest of columns indicating each sample's cluster index. It is the output of the clSamp()

function. Only the most frequent clustering result will be used to plot.

min_frc A minimum fraction of samples in a cluster that have a gene set significantly

over-represented (adjusted P < 0.05). This is used to select gene sets to plot.

Default: 0.8

Value

A list of a matrix and a data.table object. The matrix has rows as over-represented gene sets, columns as samples, and each cell stores an adjusted P for over-representation. The data.table has the clustering informations with samples in the same order as the matrix's column.

ovrGMT 7

Examples

```
ovrmat <- system.file('extdata/pltOvrHm/ovr.rds',package='MPAC') |> readRDS()
cldt <- system.file('extdata/pltOvrHm/cl.rds', package='MPAC') |> readRDS()
getSignifOvrOnCl(ovrmat, cldt)
```

ovrGMT

Calculate over-representation of gene sets in each sample by genes from sample's largest sub-pathway

Description

Calculate over-representation of gene sets in each sample by genes from sample's largest sub-pathway

Usage

```
ovrGMT(subntwlist, fgmt, omic_genes = NULL, threads = 1)
```

Arguments

subntwlist A list of igraph objects represented the largest sub-pathway for each sample. It

is the output of subNtw().

fgmt A gene set GMT file. This will be the same file used for the gene set over-

representation calculation in the next step. It is used here to ensure output sub-

pathway contains a minimum number of genes from to-be-used gene sets.

omic_genes A vector of gene symbols to narrow down over-representation calculation to

only those with input genomic data. If not provided, all genes in the GMT file

will be considered. Default: NULL.

threads Number of threads to run in parallel. Default: 1

Value

A matrix containing over-representation adjusted P with rows as gene set names and columns as sample IDs.

```
fsubntwl = system.file('extdata/subNtw/subntwl.rds', package='MPAC')
fgmt = system.file('extdata/ovrGMT/fake.gmt', package='MPAC')
fomic_gns = system.file('extdata/TcgaInp/inp_focal.rds', package='MPAC')
subntwl = readRDS(fsubntwl)
omic_gns = rownames(readRDS(fomic_gns))
ovrGMT(subntwl, fgmt, omic_gns)
```

8 pltMtfPrtIPL

_				_
- ⊸ 7	+ (`~~	· M +	·
() I	1.0	.()[۱Mt	. Т

Plot consensus pathway submodules

Description

Plot consensus pathway submodules

Usage

```
pltConMtf(grphl, proteins = NULL)
```

Arguments

grphl A list of igraph objects representing consensus pathway submodules. It is the

output from conMtf().

proteins A vector of protein symbols to highlight in the plot. Default: no protein will be

highlighted.

Value

a plot of consensus pathway submodules

Examples

pltMtfPrtIPL

Plot a heatmap of IPLs on proteins from consensus pathway submodules

Description

Plot a heatmap of IPLs on proteins from consensus pathway submodules

```
pltMtfPrtIPL(fltmat, cldt, grphl, proteins = NULL)
```

pltNeiStt 9

Arguments

fltmat	A matrix contains filterd IPL with rows as entity and column as samples. This is the output from fltByPerm(). Entity with NA value will be set to 0 and plotted as in 'normal' state.
cldt	A data table with each row representing one clustering result, and the first column denotes the number of occurrences of a clustering result and the rest of columns indicating each sample's cluster index. It is the output of the clSamp() function. Only the most frequent clustering result will be used to plot.
grphl	A list of igraph objects representing consensus pathway submodules. It is the output from conMtf().
proteins	A vector of proteins, of which IPLs to plot. Default: all proteins that in both grphl and fltmat.

Value

A heatmap of IPLs o proteins from consensus pathway submodules

Examples

pltNeiStt	Plot a heatmap of pathway and omic states of a protein and its path-
	way neighbors

Description

Plot a heatmap of pathway and omic states of a protein and its pathway neighbors

Usage

```
pltNeiStt(real_se, fltmat, fpth, protein = "")
```

Arguments

O	
real_se	A SummarizedExperiment object of PARADIGM CNA and RNA states. It is the output fromm ppRealInp() and must contain the omic states for the one defined in the protein argument.
fltmat	A matrix contains filterd IPL with rows as entity and column as samples. This is the output from fltByPerm(). Entity with NA value will be set to 0 and plotted as in 'normal' state.
fpth	Name of a pathway file for PARADIGM.
protein	Name of the protein to plot. It requires to have CN and RNA state data, as well as pathway data from the input. Default: "

10 pltOvrHm

Value

A heatmap of pathway and omic states of a protein and its pathway neighbors

Examples

```
fpth = system.file('extdata/Pth/tiny_pth.txt', package='MPAC')
freal = system.file('extdata/pltNeiStt/inp_real.rds', package='MPAC')
fflt = system.file('extdata/pltNeiStt/fltmat.rds', package='MPAC')
real_se = readRDS(freal)
fltmat = readRDS(fflt)
protein = 'CD86'
pltNeiStt(real_se, fltmat, fpth, protein)
```

plt0vrHm

Plot a heatmap of over-represented gene sets for clustered samples

Description

Plot a heatmap of over-represented gene sets for clustered samples

Usage

```
pltOvrHm(ovrmat, cldt, min_frc = 0.8)
```

Arguments

ovrmat	A matrix containing over-representation adjusted P with rows as gene set names and columns as sample IDs. It is the output of the ovrGMT() function.
cldt	A data table with each row representing one clustering result, and the first column denotes the number of occurrences of a clustering result and the rest of columns indicating each sample's cluster index. It is the output of the clSamp() function. Only the most frequent clustering result will be used to plot.
min_frc	A minimum fraction of samples in a cluster that have a gene set significantly over-represented (adjusted $P < 0.05$). This is used to select gene sets to plot. Default: 0.8

Value

A heatmap with rows as over-represented gene sets and columns as samples splited by clusters.

```
ovrmat <- system.file('extdata/pltOvrHm/ovr.rds',package='MPAC') |> readRDS()
cldt <- system.file('extdata/pltOvrHm/cl.rds', package='MPAC') |> readRDS()
pltOvrHm(ovrmat, cldt)
```

pltSttKM 11

pltSttKM	Plot a Kaplan-Meier curve for samples stratified by given protein(s)' pathway states

Description

Plot a Kaplan-Meier curve for samples stratified by given protein(s)' pathway states

Usage

```
pltSttKM(
   cdrmat,
   fltmat,
   event = "OS",
   time = "OS_days",
   proteins = NULL,
   strat_func = ">0"
)
```

Arguments

cdrmat	A matrix containing survival data with rows as patient samples and columns as survival event and time.
fltmat	A matrix contains filterd IPL with rows as entity and column as samples. This is the output from fltByPerm(). Entity with NA value will be set to 0 and plotted as in 'normal' state.
event	The column name in cdrmat to indicate survival event. Default: 'OS'.
time	The column name in cdrmat to indicate survival time. Default: 'OS_days'.
proteins	Rowname(s) in fltmat. Its/their pathway states will be used to stratify patient samples. Default: all proteins in fltmat will be used.
strat_func	A function applied on protein(s) pathway states to stratify patient samples. Available options: '>0', '<0', Default: '>0', i.e., IPL >0 vs. the rest.

Value

A Kaplan-Meier plot

ppPermInp

ppCnInp

Prepare input copy-number (CN) alteration data to run PARADIGM

Description

Prepare input copy-number (CN) alteration data to run PARADIGM

Usage

```
ppCnInp(cn_tumor_mat)
```

Arguments

cn_tumor_mat

A matrix of tumor CN focal data with rows as genes and columns as samples. A value of 0 means normal CN, > 0 means amplification, and < 0 means deletion.

Value

A SummarizedExperiment object of CN state for PARADIGM

Examples

```
fcn = system.file('extdata/TcgaInp/focal_tumor.rds', package='MPAC')
cn_tumor_mat = readRDS(fcn)
ppCnInp(cn_tumor_mat)
```

ppPermInp

Permute input genomic state data between genes in the same sample

Description

Permute input genomic state data between genes in the same sample

Usage

```
ppPermInp(real_se, n_perms=100, threads=1)
```

Arguments

real_se A SummarizedExperiment object of CN and RNA states from real samples with

rows as genes and columns as samples. It is the output from ppRealInp().

n_perms Number of permutations. Default: 100

threads Number of threads to run in parallel. Default: 1

Value

A list of SummarizedExperiment objects of permuted CN and RNA states. The metadata i in each object denotes its permutation index.

ppRealInp 13

Examples

```
freal = system.file('extdata/TcgaInp/inp_real.rds', package='MPAC')
real_se = readRDS(freal)

ppPermInp(real_se, n_perms=3)
```

ppRealInp

Prepare input copy-number (CN) alteration and RNA data to run PARADIGM

Description

Prepare input copy-number (CN) alteration and RNA data to run PARADIGM

Usage

```
ppRealInp(
  cn_tumor_mat,
  rna_tumor_mat,
  rna_normal_mat,
  rna_n_sd = 2,
  threads = 1
)
```

Arguments

cn_tumor_mat A matrix of tumor CN focal data with rows as genes and columns as samples. A value of 0 means normal CN, > 0 means amplification, and < 0 means deletion.

rna_tumor_mat A matrix of RNA data from tumor samples with rows as genes and columns as samples

rna_normal_mat A matrix of RNA data from normal samples with rows as genes and columns as samples

rna_n_sd Standard deviation range from fitted normal samples to define RNA state. Default: 2, i.e. 2*sd

threads Number of threads to run in parallel. Default: 1

Value

A SummarizedExperiment object of CN and RNA state for PARADIGM

```
fcn = system.file('extdata/TcgaInp/focal_tumor.rds', package='MPAC')
ftumor = system.file('extdata/TcgaInp/log10fpkmP1_tumor.rds', package='MPAC')
fnorm = system.file('extdata/TcgaInp/log10fpkmP1_normal.rds', package='MPAC')
cn_tumor_mat = readRDS(fcn)
rna_tumor_mat = readRDS(ftumor)
rna_norm_mat = readRDS(fnorm)
ppRealInp(cn_tumor_mat, rna_tumor_mat, rna_norm_mat)
```

14 ppRunPrd

	_	_		
gg	Dп	ъT	n	n
DD	ıvı	a_{\perp}		IJ

Prepare input RNA data to run PARADIGM

Description

Prepare input RNA data to run PARADIGM

Usage

```
ppRnaInp(rna_tumor_mat, rna_normal_mat, rna_n_sd = 2, threads = 1)
```

Arguments

rna_tumor_mat A matrix of RNA data from tumor samples with rows as genes and columns as samples

rna_normal_mat A matrix of RNA data from normal samples with rows as genes and columns as samples

rna_n_sd Standard deviation range from fitted normal samples to define RNA state. Default: 2, i.e. 2*sd

Value

threads

A SummarizedExperiment of RNA state for PARADIGM

Examples

```
ftumor = system.file('extdata/TcgaInp/log10fpkmP1_tumor.rds', package='MPAC')
fnorm = system.file('extdata/TcgaInp/log10fpkmP1_normal.rds', package='MPAC')
rna_tumor_mat = readRDS(ftumor)
rna_norm_mat = readRDS(fnorm)

ppRnaInp(rna_tumor_mat, rna_norm_mat, threads=2)
```

Number of threads to run in parallel. Default: 1

ppRunPrd

Prepare required files to run PARADIGM

Description

Prepare required files to run PARADIGM

```
ppRunPrd(pat, cnmat, rnamat, outdir, file_tag=NULL)
```

runPermPrd 15

Arguments

pat Sample ID
cnmat CN matrix
rnamat RNA matrix

outdir Output folder to save all results.

file_tag A string of output file name tag. Default: NULL

Value

None

runPermPrd Run PARADIGM on permuted data

Description

Run PARADIGM on permuted data

Usage

```
runPermPrd(perml, fpth, outdir,
    PARADIGM_bin=NULL, nohup_bin=NULL, sampleids=NULL, threads=1)
```

Arguments

perml A list of SummarizedExperiment objects of permuted CNA and RNA states gen-

erated by ppPermInp().

fpth Name of a pathway file for PARADIGM.

outdir Output folder to save all results.

PARADIGM_bin PARADIGM binary, which can be downloaded from https://github.com/sng87/paradigm-

scripts/tree/master/public/exe. Note that the binary is only available for Linux

or MacOS. Default: NULL

nohup_bin nohup binary, which is used for long running PARADIGM jobs. Default: NULL

sampleids A vector of sample IDs to run PARADIGM on. If not provided, all the samples

that exist in both copy-number alteration and RNA files will be ran. Default:

NULL

threads Number of threads to run in parallel. Default: 1

Value

None

16 runPrd

Examples

```
fperm = system.file('extdata/TcgaInp/inp_perm.rds', package='MPAC')
perml = readRDS(fperm)
fpth = system.file('extdata/Pth/tiny_pth.txt', package='MPAC')
outdir = tempdir()
paradigm_bin = '/path/to/PARADIGM' ## change to binary location
pat = 'TCGA-CV-7100'

# depends on external PARADIGM binary, do not run
runPermPrd(perml, fpth, outdir, paradigm_bin, sampleids=c(pat))
```

runPrd

Run PARADIGM on multi-omic data

Description

Run PARADIGM on multi-omic data

Usage

```
runPrd(real_se, fpth, outdir, PARADIGM_bin=NULL, nohup_bin=NULL,
    sampleids=NULL, file_tag=NULL, threads=1)
```

Arguments

real_se	A SummarizedExperiment object of PARADIGM CNA and RNA states. It is the same matrix as the output from ppRealInp().
fpth	Name of a pathway file for PARADIGM.
outdir	Output folder to save all results.
PARADIGM_bin	PARADIGM binary, which can be downloaded from https://github.com/sng87/paradigm-scripts/tree/master/public/exe. Note that the binary is only available for Linux or MacOS. Default: NULL
nohup_bin	nohup binary, which is used for long running PARADIGM jobs. Default: NULL
sampleids	A vector of sample IDs to run PARADIGM on. If not provided, all the samples that exist in both copy-number alteration and RNA files will be ran. Default: NULL
file_tag	A string of output file name tag. Default: NULL
threads	Number of threads to run in parallel. Default: 1

Value

None

subNtw 17

Examples

```
freal = system.file('extdata/TcgaInp/inp_real.rds', package='MPAC')
real_se = readRDS(freal)

fpth = system.file('extdata/Pth/tiny_pth.txt', package='MPAC')
outdir = tempdir()
paradigm_bin = '/path/to/PARADIGM' ## change to binary location

# depends on external PARADIGM binary
runPrd(real_se, fpth, outdir, paradigm_bin, sampleids=c('TCGA-CV-7100'))
```

subNtw

Subset pathways by IPL results

Description

Subset pathways by IPL results

Usage

```
subNtw(fltmat, fpth, fgmt, min_n_gmt_gns = 2, threads = 1)
```

Arguments

fltmat	A matrix contains filterd IPL with rows as 'entity' and column as samples. This is the output from fltByPerm(). Entity with NA in all columns will be ignored.
fpth	Name of a pathway file for PARADIGM.
fgmt	A gene set GMT file. This will be the same file used for the gene set over- representation calculation in the next step. It is used here to ensure output sub- pathway contains a minimum number of genes from to-be-used gene sets.
min_n_gmt_gns	Minimum number of genes from the GMT file in the output sub-pathway. Default: 2.
threads	Number of threads to run in parallel. Default: 1

Value

A list of igraph objects representing the largest sub-pathway for each sample.

```
fflt = system.file('extdata/fltByPerm/flt_real.rds', package='MPAC')
fltmat = readRDS(fflt)
fpth = system.file('extdata/Pth/tiny_pth.txt', package='MPAC')
fgmt = system.file('extdata/ovrGMT/fake.gmt', package='MPAC')
subNtw(fltmat, fpth, fgmt, min_n_gmt_gns=1)
```

Index

```
clSamp, 3
colPermIPL, 3
colRealIPL, 4
conMtf, 5
fltByPerm, 5
getSignifOvrOnCl, 6
ovrGMT, 7
pltConMtf, 8
{\tt pltMtfPrtIPL}, {\color{red} 8}
pltNeiStt, 9
pltOvrHm, 10
pltSttKM, 11
ppCnInp, 12
ppPermInp, 12
ppRealInp, 13
ppRnaInp, 14
ppRunPrd, 14
runPermPrd, 15
runPrd, 16
subNtw, 17
```