

Package ‘metapone’

October 18, 2022

Type Package

Title Conducts pathway test of metabolomics data using a weighted permutation test

Version 1.2.0

Date 2022-2-24

Description The package conducts pathway testing from untargetted metabolomics data. It requires the user to supply feature-level test results, from case-control testing, regression, or other suitable feature-level tests for the study design. Weights are given to metabolic features based on how many metabolites they could potentially match to. The package can combine positive and negative mode results in pathway tests.

Depends R (>= 4.1.0), BiocParallel, fields, markdown, fdrtool, fgsea, ggplot2, ggrepel

Imports methods

biocViews Technology, MassSpectrometry, Metabolomics, Pathways

License Artistic-2.0

LazyLoad yes

NeedsCompilation no

Suggests rmarkdown, knitr

VignetteBuilder knitr

git_url <https://git.bioconductor.org/packages/metapone>

git_branch RELEASE_3_15

git_last_commit 6164c92

git_last_commit_date 2022-04-26

Date/Publication 2022-10-18

Author Leqi Tian [aut],
Tianwei Yu [aut],
Tianwei Yu [cre]

Maintainer Tianwei Yu <yutianwei@cuhk.edu.cn>

R topics documented:

metapone-package	2
bbplot1d	3
bbplot2d	4
fable	5
hmdbCompMZ	6
hmdbCompMZ.metapone	6
metapone	7
metaponeResult-class	10
neg	11
pa	11
pos	12
ptable	13
Index	14

metapone-package	<i>Conducts pathway test of metabolomics data using a weighted permutation test</i>
------------------	---

Description

The package conducts pathway testing from untargetted metabolomics data. It requires the user to supply feature-level test results, from case-control testing, regression, or other suitable feature-level tests for the study design. Weights are given to metabolic features based on how many metabolites they could potentially match to. The package can combine positive and negative mode results in pathway tests. The package contains two types of statistical testing that considers matching uncertainty - (1) a permutation test that is based on the hypergeometric test and (2) a GSEA type test with weighted features/metabolites.

Details

The package conducts (1) a weighted hypergeometric test using permutations on metabolomics data. The weights are assigned based on how many metabolites each data feature can match to, (2) a GSEA type test based on an estimation of importance of metabolites/features. The importance is evaluated by the size of matching for each metabolite/feature and the p-value of features.

The user can tune a parameter to change the penalty for multiple-matched features and choose the type of pathway testing.

Author(s)

Tianwei Yu (<yutianwei@cuhk.edu.cn>)

bbplot1d *Plot of metapone result.*

Description

The function `bbplot1d()` select important pathways with their P-value less than a threshold and returns ranked bubble plot showing important pathways names and their corresponding $-\log_{10}(\text{Pvalue})$.

Usage

```
bbplot1d(res, p_thres = 0.05, sig_metab_thres = 1)
```

Arguments

<code>res</code>	The result matrix obtained from metapone with columns: "p_value", "n_significant metabolites", "n_mapped_metabolites", "n_metabolites", "significant metabolites", "mapped_metabolites", "fdr".
<code>p_thres</code>	The threshold of P-value for pathways to be shown in the bubble plot. The default threshold is 0.05.
<code>sig_metab_thres</code>	The threshold of fractional matched significant metabolite count for pathways to be shown in the bubble plot. The default is 1.

Author(s)

Leqi Tian (<leqitian@link.cuhk.edu.cn>)

See Also

[metapone](#)

Examples

```
data(hmdbCompMZ.metapone)
data(pa)
data(pos)
data(neg)
r<-metapone(pos, neg, pa, hmdbCompMZ=hmdbCompMZ.metapone, p.threshold=0.05,
  n.permu=100, fractional.count.power=0.5, max.match.count=10)
bbplot1d(ptable(r)) # p_thres = 0.05
```

`bbplot2d`*Plot of metapone result.*

Description

The function `bbplot2d()` select important pathways with their P-value less than a threshold and returns a 2-D bubble plot with $-\log_{10}(\text{Pvalue})$ and the number of significant metabolites as coordinate axes.

Usage

```
bbplot2d(res, p_thres = 0.05, sig_metab_thres = 1)
```

Arguments

<code>res</code>	The result matrix obtained from metapone with columns: "p_value", "n_significant metabolites", "n_mapped_metabolites", "n_metabolites", "significant metabolites", "mapped_metabolites", "fdr".
<code>p_thres</code>	The threshold of P-value for pathways to be shown in the bubble plot. The default threshold is 0.05.
<code>sig_metab_thres</code>	The threshold of fractional matched significant metabolite count for pathways to be shown in the bubble plot. The default is 1.

Author(s)

Leqi Tian (<leqitian@link.cuhk.edu.cn>)

See Also

[metapone](#)

Examples

```
data(hmdbCompMZ.metapone)
data(pa)
data(pos)
data(neg)
r<-metapone(pos, neg, pa, hmdbCompMZ=hmdbCompMZ.metapone, p.threshold=0.05,
  n.permu=100, fractional.count.power=0.5, max.match.count=10)
bbplot2d(htable(r)) # p_thres = 0.05
```

ftable	<i>Accessor functions for the feature mapping table in a metaponeResult object.</i>
--------	---

Description

Returns a list containing the mapped features in each pathway.

Usage

```
## S4 method for signature 'metaponeResult'  
ftable(object)
```

Arguments

object A metaponeResult object.

Details

Each pathway is represented by a `data.frame` as an item in the list object. The dataframe include information of `m.z`, `retention.time`, `p.value`, `statistic`, `HMDB_ID`, `theoretical m.z`, `ion.type`, `fractional counts`.

Value

The method returns a list. Each item is for a pathway. Matched significant metabolites are included.

Author(s)

Tianwei Yu <yutianwei@cuhk.edu.cn>

See Also

`ptable`

Examples

```
data(hmdbCompMZ.metapone)  
data(pa)  
data(pos)  
data(neg)  
r<-metapone(pos, neg, pa, hmdbCompMZ=hmdbCompMZ.metapone, p.threshold=0.05,  
  n.permu=100, fractional.count.power=0.5, max.match.count=10)  
ftable(r)[1:6]
```

hmdbCompMZ

the m/z values of common adduct ions of HMDB metabolites

Description

Monoisotopic mass of common adduct ions.

Usage

```
data("hmdbCompMZ")
```

Format

A data frame with 5704350 observations on the following 3 variables.

HMDB_ID HMDB ID.

ion.type Adduct ion type.

m.z the m/z of the adduct ion.

Source

<https://hmdb.ca/>

References

<https://hmdb.ca/>

Examples

```
data(hmdbCompMZ)
```

hmdbCompMZ.metapone

the m/z values of common adduct ions of metapone metabolites

Description

Monoisotopic mass of common adduct ions, limited to those included in the pathways in metapone.

Usage

```
data("hmdbCompMZ.metapone")
```

Format

A data frame with 79350 observations on the following 3 variables.

HMDB_ID HMDB ID.

ion.type Adduct ion type.

m.z the m/z of the adduct ion.

Details

The main difference of using this dataset vs using hmdbCompMZ, is the metabolite universe in testing is limited to those metabolites matched to metapone pathways, not all HMDB metabolites.

Source

[The Human Metabolome Database](#)

References

[The Human Metabolome Database](#)

Examples

```
data(hmdbCompMZ)
```

metapone	<i>METAbolic pathway testing using both POSitive and NEgative mode data</i>
----------	---

Description

Metapone conducts pathway tests for untargeted metabolomics data. It has three main characteristics: (1) expanded database combining SMPDB and Mummichog databases, with manual cleaning to remove redundancies; (2) A new weighted testing scheme to address the issue of metabolite-feature matching uncertainties; (3) Can consider positive mode and negative mode data in a single analysis.

Usage

```
metapone(pos=NULL, neg=NULL, pa, hmdbCompMZ, pos.adductlist = c("M+H",
  "M+NH4", "M+Na", "M+ACN+H", "M+ACN+Na", "M+2ACN+H", "2M+H", "2M+Na",
  "2M+ACN+H"), neg.adductlist = c("M-H", "M-2H", "M-2H+Na", "M-2H+K",
  "M-2H+NH4", "M-H2O-H", "M-H+Cl", "M+Cl", "M+2Cl"),
  use.fractional.count=TRUE, match.tol.ppm=5, p.threshold=0.05,
  n.permu=200, fractional.count.power=0.5, max.match.count=10,
  use.fgsea = FALSE, use.meta = FALSE)
```

Arguments

<code>pos</code>	The positive ion mode test results. A matrix with four columns: m/z, retention time, p-value, test statistic. The package doesn't require both <code>pos</code> and <code>neg</code> to be present. One ion mode result is sufficient.
<code>neg</code>	The negative ion mode test results. A matrix with four columns: m/z, retention time, p-value, test statistic. The package doesn't require both <code>pos</code> and <code>neg</code> to be present. One ion mode result is sufficient.
<code>pa</code>	Pathway information. A data frame with five columns: database pathway ID, pathway name, HMDB ID, KEGG ID, category of pathway.
<code>hmdbCompMZ</code>	the m/z values of common adduct ions of HMDB metabolites. See the help file of <code>hmdbCompMZ</code> for details.
<code>pos.adductlist</code>	The vector of positive adduct ions to be considered.
<code>neg.adductlist</code>	The vector of negative adduct ions to be considered.
<code>use.fractional.count</code>	A lot of features match to multiple metabolites by m/z. Whether to discount such matches by using fractional counts.
<code>match.tol.ppm</code>	The ppm level when conducting m/z match.
<code>p.threshold</code>	The threshold of p-values of metabolic features to be considered significant.
<code>n.permu</code>	The number of permutations in permutation test.
<code>fractional.count.power</code>	The fractional counts are taken to this power to transform the weights.
<code>max.match.count</code>	When calculating fractional counts, some features might be matched to too many. In that case the number of matches is capped by the value of <code>max.match.count</code> .
<code>use.fgsea</code>	Whether to use a GSEA type test when performing pathway testing. When it is <code>FALSE</code> , a permutation-based weighted hypergeometric test is performed.
<code>use.meta</code>	Whether to perform a GSEA type test with weighted metabolites. When it is <code>FALSE</code> , a GSEA type test is performed on weighted features.

Value

The method returns a generic S4 object of class "metapone.result":

`@test.results` A matrix with 8 columns: "p_value", "n_significant metabolites", "n_mapped_metabolites", "n_metabolites", "significant metabolites", "mapped_metabolites", "lfdr", "adjusted.p". Each row is for a pathway. When using GSEA test, "ES", "NES", "nMoreExtreme" are returned additionally.

`@mapped.features`

A list. Each item is for a pathway. The item lists matched significant metabolites.

The columns in `test.result` are the following:

`p_value` The p-value for each enrichment.

n_significant_metabolites	The number of weighted significant metabolites associated with the enrichment.
n_mapped_metabolites	The number of weighted metabolites associated with the enrichment.
n_metabolites	The number of metabolites associated with the enrichment.
significant_metabolites	A string with the names of significant metabolites that drive the enrichment.
mapped_metabolites	A string with the names of metabolites that drive the enrichment.
lfdr	The local fdr value for each enrichment.
adjust.p	The enrichment BH-adjusted p-value for each enrichment.
ES	The enrichment score (Avaliable in GSEA test).
NES	The enrichment score normalized to mean enrichment of random samples of the same size (Avaliable in GSEA test).
nMoreExtreme	The number of times a random metabolite set had a more extreme enrichment score value (Avaliable in GSEA test).

Author(s)

Tianwei Yu (<yutianwei@cuhk.edu.cn>) Leqi Tian (<leqitian@link.cuhk.edu.cn>)

References

[Small Molecule Pathway Database](#)

[Mummichog](#)

See Also

[pa](#), [hmdbCompMZ](#)

Examples

```
data(hmdbCompMZ.metapone)
data(pa)
data(pos)
data(neg)
# Permutation-based weighted hypergeometric test
r<-metapone(pos, neg, pa, hmdbCompMZ=hmdbCompMZ.metapone, p.threshold=0.05,
  n.permu=100, fractional.count.power=0.5, max.match.count=10)
hist(ptable(r)[,1])

# Metabolites based GSEA test
r<-metapone(pos, neg, pa, hmdbCompMZ=hmdbCompMZ.metapone, p.threshold=0.05,
  n.permu=100, fractional.count.power=0.5, max.match.count=10, use.fgsea = TRUE, use.meta = TRUE)
hist(ptable(r)[,1])

# Features based GSEA test
r<-metapone(pos, neg, pa, hmdbCompMZ=hmdbCompMZ.metapone, p.threshold=0.05,
```

```
n.permu=100, fractional.count.power=0.5, max.match.count=10, use.fgsea = FALSE, use.meta = FALSE)  
hist(ptable(r)[,1])
```

metaponeResult-class *Class "metaponeResult"*

Description

This class represents the results of pathway testing. The testing result contain two major components: the significant level of each pathway, and the features matched to each pathway.

Objects from the Class

Objects can be created by calls of the form `new("metaponeResult", ...)`.

Slots

test.result: a dataframe containing `p_value`, `n_significant` metabolites, `n_mapped_metabolites`, `n_metabolites`, `significant` metabolites, `mapped_metabolite` IDs, `lfr` and pathway name.

mapped.features: A list containing `n` entries, where `n` is the number of pathways. Each entry is a data frame, containing the features mapped to this pathway. The information include `m.z`, `retention.time`, `p.value`, `statistic`, `HMDB_ID`, `theoretical m.z`, `ion.type`, `fractional` counts.

Methods

ptable signature(object = "metaponeResult"): return the data.frame of test statistics for each pathway, including `p_value`, `n_significant` metabolites, `n_mapped_metabolites`, `n_metabolites`, `significant` metabolites, `mapped_metabolite` IDs `lfr` and and pathway name.

ftable signature(object = "metaponeResult"): Returns a list containing the mapped features in each pathway. Each pathway is represented by a data.frame as an item in the list object. The dataframe include information of `m.z`, `retention.time`, `p.value`, `statistic`, `HMDB_ID`, `theoretical m.z`, `ion.type`, `fractional` counts.

Author(s)

Tianwei Yu

neg	<i>Example negative mode data from the Metabolome Atlas of the Aging Mouse Brain</i>
-----	--

Description

The data is generated from the hippocampus data of the Metabolome Atlas of the Aging Mouse Brain (ST001888) dataset. The p-values and test statistics were obtained by contrasting mouse hippocampus metabolome between prime-age mice and aging mice.

Usage

```
data("neg")
```

Format

A data frame with 6947 observations on the following 4 variables.

`m.z` a numeric vector. The mass-to-charge ratio of the features.

`retention.time` a numeric vector. The retention time of the features.

`p.value` a numeric vector. The p-values of the features.

`statistic` a numeric vector. The test statistics of the features.

References

<https://www.metabolomicsworkbench.org/data/DRCCMetadata.php?Mode=Study&DataMode=FactorsData&StudyID=ST001888>

Examples

```
data(neg)
```

pa	<i>Pathway-metabolite match file.</i>
----	---------------------------------------

Description

mapps pathways with metabolites.

Usage

```
data("pa")
```

Format

A data frame with 5395 observations on the following 5 variables.

database a character vector
pathway.name a character vector
HMDB.ID a character vector
KEGG.ID a character vector
category a character vector

Source

Small Molecule Pathway Database
Mummichog

Examples

```
data(pa)
```

pos	<i>Example positive mode data from the Metabolome Atlas of the Aging Mouse Brain</i>
-----	--

Description

The data is generated from the hippocampus data of the Metabolome Atlas of the Aging Mouse Brain (ST001888) dataset. The p-values and test statistics were obtained by contrasting mouse hippocampus metabolome between prime-age mice and aging mice.

Usage

```
data("pos")
```

Format

A data frame with 10085 observations on the following 4 variables.

m.z a numeric vector. The mass-to-charge ratio of the features.
retention.time a numeric vector. The retention time of the features.
p.value a numeric vector. The p-values of the features.
statistic a numeric vector. The test statistics of the features.

References

<https://www.metabolomicsworkbench.org/data/DRCCMetadata.php?Mode=Study&DataMode=FactorsData&StudyID=ST001888>

Examples

```
data(pos)
```

ptable	<i>Accessor functions for the test result table in a metaponeResult object.</i>
--------	---

Description

return the data.frame of test statistics for each pathway.

Usage

```
## S4 method for signature 'metaponeResult'  
ptable(object)
```

Arguments

object A metaponeResult object.

Details

Includes p_value, n_significant metabolites, n_mapped_metabolites, n_metabolites, significant metabolites, mapped_metabolite IDs and pathway name.

Value

The method returns a data frame with 6 columns: "p_value", "n_significant metabolites", "n_mapped_metabolites", "n_metabolites", "significant metabolites", "mapped_metabolites".

Author(s)

Tianwei Yu <yutianwei@cuhk.edu.cn>

See Also

fable

Examples

```
data(hmdbCompMZ.metapone)  
data(pa)  
data(pos)  
data(neg)  
r<-metapone(pos, neg, pa, hmdbCompMZ=hmdbCompMZ.metapone,  
  p.threshold=0.05,n.permu=100,fractional.count.power=0.5, max.match.count=10)  
head(ptable(r))
```

Index

- * **classes**
 - metaponeResult-class, [10](#)
- * **datasets**
 - hmdbCompMZ, [6](#)
 - hmdbCompMZ.metapone, [6](#)
 - neg, [11](#)
 - pa, [11](#)
 - pos, [12](#)
- * **package**
 - metapone-package, [2](#)

bbplot1d, [3](#)
bbplot2d, [4](#)

fable, [5](#)
fable, metaponeResult-method
(metaponeResult-class), [10](#)
fable,metaponeResult-method (fable), [5](#)

hmdbCompMZ, [6](#), [9](#)
hmdbCompMZ.metapone, [6](#)

metapone, [3](#), [4](#), [7](#)
metapone-package, [2](#)
metaponeResult (metaponeResult-class),
[10](#)
metaponeResult-class, [10](#)
metaponeResult-method
(metaponeResult-class), [10](#)

neg, [11](#)

pa, [9](#), [11](#)
pos, [12](#)
ptable, [13](#)
ptable, metaponeResult-method
(metaponeResult-class), [10](#)
ptable,metaponeResult-method (ptable),
[13](#)