Package ‘famat’

May 29, 2024

Title Functional analysis of metabolic and transcriptomic data
Version 1.14.0

Description Famat is made to collect data about lists of genes and metabolites provided by user, and to visualize it through a Shiny app.
Information collected is:
- Pathways containing some of the user's genes and metabolites (obtained using a pathway enrichment analysis).
- Direct interactions between user’s elements inside pathways.
- Information about elements (their identifiers and descriptions).
- Go terms enrichment analysis performed on user's genes.
The Shiny app is composed of:
- information about genes, metabolites, and direct interactions between them inside pathways.
- an heatmap showing which elements from the list are in pathways (pathways are structured in hierarchies).
- hierarchies of enriched go terms using Molecular Function and Biological Process.

License GPL-3
LazyData false
Depends R (>= 4.0)
Imports KEGGREST, mgcv, stats, BiasedUrn, dplyr, gprofiler2, rWikiPathways, reactome.db, stringr, GO.db, ontologyIndex, tidyr, shiny, shinydashboard, shinyBS, plotly, magrittr, DT, clusterProfiler, org.Hs.eg.db

VignetteBuilder knitr
Suggests BiocStyle, knitr, rmarkdown, testthat, BiocManager
biocViews FunctionalPrediction, GeneSetEnrichment, Pathways, GO, Reactome, KEGG

Encoding UTF-8
Roxygen list(markdown = TRUE)
RoxygenNote 7.1.1

BugReports https://github.com/emiliesecherre/famat/issues
compl_data

Description

Complete and prepare data obtained with interactions function, to use it in the Shiny interface. GO terms enrichment analysis is performed using clusterProfiler.

Usage

compl_data(listparam)

Arguments

listparam Output from interactions function
**Value**

A list containing:

- **heatmap**: Dataframe heatmap-like, with in abscissa elements of pathways ("X" is written if an element is present in a pathway), and with in ordinate hierarchies of pathways.
- **meta_list**: User’s metabolites given in path_enrich function.
- **allResBP**: Results of Go BP terms enrichment analysis performed by clusterProfileR (20 best).
- **go_genelist**: Dataframe containing enriched GO terms per genes of user’s list.
- **allResMF**: Results of Go MF terms enrichment analysis performed by clusterProfileR (20 best).
- **types**: Dataframe containing categories of pathways per pathways.
- **genetype**: List of genes categories, containing genes of user’s list.
- **metatab**: Dataframe of metabolites and informations about them, as names and chebi ids.
- **genetab**: Dataframe of genes and informations about them, as gene symbols and names.
- **intetab**: Dataframe of direct interactions between elements of user’s list, and informations about them as elements in the interaction, how they interact, and which pathways are concerned.
- **gompf_tab**: Dataframe of GO MF terms hierarchies containing our enriched GO terms, plus description of these GO terms and genes of user’s list concerned by enriched GO terms.
- **gobp_tab**: Dataframe of GO BP terms hierarchies containing our enriched GO terms, plus description of these GO terms and genes of user’s list concerned by enriched GO terms.
- **gene_list**: User’s genes given in path_enrich function.
- **gomf_list**: List containing GO MF terms hierarchies, with indices in the joliMF dataframe and genes concerned by the hierarchy.
- **gobpplist**: List containing GO BP terms hierarchies, with indices in the joliMF dataframe and genes concerned by the hierarchy.
- **hierabrite**: List of pathways categories, containing pathways concerned by a category and their indices in "trait" dataframe.
- **hierapath**: List of hierarchies of pathways, containing indices of pathways in "trait" dataframe and elements contained in the hierarchy.
- **save_cluster_elem**: Vector of clustered elements.
- **centrality**: Matrix trait-like with values of centralities (number of direct interactions between an element of user’s list and other elements of the pathway) instead of “X”.
- **inter_values**: Matrix trait-like with values representing direct interactions (3/2/1, respectively for genes/metabolites implicated in a direct interaction, and for elements not implicated in a direct interaction) instead of "X".
- **gene_notin**: Dataframe of genes which aren’t in pathways and informations about them, as gene symbols and names.
Matrix trait-like with pathway and element informations instead of "X"

This list is used by rshiny function.

Author(s)
Emilie Secherre <emisecherre@gmail.com>

References
Guangchuang Yu, Li-Gen Wang, Yanyan Han and Qing-Yu He. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS: A Journal of Integrative Biology 2012, 16(5):284-287

See Also
interactions rshiny

Examples
## load example data
data(interactions_result)
compl_data_result=compl_data(interactions_result)

compl_data_result  Output of compl_data function

Description
Enriched and structured informations about genes, metabolites, their interactions, pathways and enriched GO terms.

Usage
data("compl_data_result")

Format
List of 22 items.

heatmap Dataframe heatmap-like, with abscissa elements of pathways ("X" is written if an element is present in a pathway), and with in ordinate hierarchies of pathways. A data frame with 16 rows and 9 variables.

path_name Hierarchies of pathways obtained by pathways enrichment analysis. Pathways are given by their name.

path_id Identifiers of pathways in hierarchies.

meta_ratio Metabolites ratio, so the part of user’s metabolites in the total number of metabolites in the pathway.
**gene_ratio**  Genes ratio, so the part of user’s genes in the total number of genes in the pathway.
**blank** Just an empty column to separate heatmap data from pathways data (NA).
**SLC6A12** The value "1" shows that the element SLC6A12 is in the pathway on the row, "0" shows it is not in this pathway (0–1).
**Betaine** The value "1" shows that the element Betaine is in the pathway on the row, "0" shows it is not in this pathway (0–1).
**ATP** The value "1" shows that the element ATP is in the pathway on the row, "0" shows it is not in this pathway (0–1).
**Betaine / SLC6A12** The value "1" shows that the interaction Betaine / SLC6A12 is in the pathway on the row, "0" shows it is not in this pathway (0–1).

**meta_list** Vector containing user’s metabolites (C00002, C00719)

**allResBP** Results of Go BP terms enrichment analysis performed by clusterProfileR (20 best). A data frame with 20 rows and 9 variables.
- **ID** Identifiers of enriched GO terms.
- **Description** Names of enriched GO terms.
- **GeneRatio** Number of user’s genes concerned by the enriched GO term, by total number of user’s genes.
- **BgRatio** Number of genes concerned by the enriched GO term, by the total number of annotated genes in the database.
- **pvalue** Pvalue of the go term enrichment analysis (0.001392161–0.011536208).
- **p.adjust** Adjusted pvalue of the go term enrichment analysis (0.01713291–0.02670275).
- **qvalue** Qvalue of the go term enrichment analysis (0.001803464–0.002810816).
- **geneID** Entrez Gene identifiers of all genes concerned by the enriched GO term (30, 6539).
- **Count** Number of genes concerned by the enriched GO term described by an Entrez Gene identifier(1).

**go_genelist** Dataframe containing enriched GO terms per genes of user’s list. A data frame with 13 rows and 2 variables.
- **hgnc_symbol** Gene symbol of the gene concerned by an enriched GO term (SLC6A12, ACAA1).
- **go_id** Identifier of the GO term concerning the gene (GO:0005328, GO:0015293, GO:0003333, GO:0015812, GO:0015171, GO:0042165, GO:0008028, GO:0006635, GO:0006625, GO:0033540, GO:0036109, GO:0008206, GO:0000038)

**allResMF** Results of Go MF terms enrichment analysis performed by clusterProfileR (20 best) A data frame with 20 rows and 9 variables.
- **ID** Identifiers of enriched GO terms.
- **Description** Names of enriched GO terms.
- **GeneRatio** Number of user’s genes concerned by the enriched GO term, by total number of user’s genes.
- **BgRatio** Number of genes concerned by the enriched GO term, by the total number of annotated genes in the database.
- **pvalue** Pvalue of the go term enrichment analysis (0.001392161–0.011536208).
- **p.adjust** Adjusted pvalue of the go term enrichment analysis (0.01713291–0.02670275).
- **qvalue** Qvalue of the go term enrichment analysis (0.001803464–0.002810816).
**geneID**  Entrez Gene identifiers of all genes concerned by the enriched GO term (30, 6539).
**Count**  Number of genes concerned by the enriched GO term described by an Entrez Gene identifier (1).

**types**  Dataframe containing categories of pathways per pathways. A data frame with 12 rows and 2 variables.
- **id**  Identifier of pathways from pathways enrichment analysis.
- **root**  Name of the pathway category concerning a pathway.

**genetype**  List of genes categories, containing genes of user’s list. A list of 3 items.

**metatab**  Dataframe of metabolites and informations about them, as names and chebi ids. A data frame with 2 rows and 2 variables.
- **id**  Name of user’s metabolites (ATP, Betaine).
- **name**  Chebi identifier of user’s metabolites (CHEBI:15422, CHEBI:17750).

**genetab**  Dataframe of genes and informations about them, as gene symbols and names. A data frame with 1 rows and 2 variables.
- **id**  Gene symbols of user’s genes contained in pathways (SLC6A12).
- **name**  Name of user’s genes contained in pathways (solute carrier family 6 member 12).

**intetab**  Dataframe of direct interactions between elements of user’s list, and informations about them as elements in the interaction, how they interact, and which pathways are concerned A data frame with 1 row and 8 variables.
- **tag**  Summary of elements concerned by the interaction (Betaine / SLC6A12)
- **first_item**  First element of the direct interaction (Betaine)
- **link**  Description of how the two elements interact (Control(In: ACTIVATION of BiochemicalReaction), controls-transport-of-chemical).
- **sec_item**  Second element of the direct interaction (SLC6A12)
- **go**  Value "1" means that a gene of the interaction is concerned by an enriched GO term, "0" means no element is concerned by an enriched GO term (1).
- **path**  Pathways containing the direct interaction ("R-HSA-112310, R-HSA-112315, R-HSA-112316, R-HSA-382551, R-HSA-425366, R-HSA-425393, R-HSA-425407, R-HSA-888590, R-HSA-352230, R-HSA-442660, R-HSA-888593")
- **type**  Interaction type, can be gene/gene, metabolite/metabolite, or gene/metabolite (g/m)
- **cat**  Categories of pathways containing the direct interaction (Neuronal System, Transport of small molecules)

**gomf_tab**  Dataframe of Go MF terms hierarchies containing our enriched GO terms, plus description of these GO terms and genes of user’s list concerned by enriched GO terms. A data frame with 93 rows and 3 variables.
- **goterm**  Hierarchies of enriched GO terms.
- **go_name**  Names of GO terms.
- **genes**  Genes concerned by GO terms.

**gobp_tab**  Dataframe of Go BP terms hierarchies containing our enriched GO terms, plus description of these GO terms and genes of user’s list concerned by enriched GO terms. A data frame with 107 rows and 3 variables.
- **goterm**  Hierarchies of enriched GO terms.
go_name  Names of GO terms.
genes  Genes concerned by GO terms.
gene_list  Vector containing user’s genes (ACAA1, SLC6A12)
gomflist  List containing GO MF terms hierarchies, with indices in the joliMF dataframe and genes concerned by the hierarchy. A list of 3 items.
gobplist  List containing GO BP terms hierarchies, with indices in the joliMF dataframe and genes concerned by the hierarchy. A list of 5 items.
hierabrite  List of pathways categories, containing pathways concerned by a category and their indices in "trait" dataframe. A list of 3 items.
hierapath  List of hierarchies of pathways, containing indices of pathways in "trait" dataframe and elements contained in the hierarchy. A list of 3 items.
save_cluster_elem  Vector of clustered elements
centrality  Matrix heatmap-like with values of centralities (number of direct interactions between an element of user’s list and other elements of the pathway) instead of "X". Other cells contain the value "0" (0–65). An integer matrix with 16 rows and 9 columns.
inter_values  Matrix heatmap-like with values representing direct interactions (3/2/1, respectively for genes/metabolites implicated in a direct interaction, and for elements not implicated in a direct interaction) instead of "X". Other cells contain the value "0" (0–3). An integer matrix with 16 rows and 9 columns.
gene_notin  Dataframe of genes which aren’t in pathways and informations about them, as gene symbols and names. A data frame with 1 row and 2 variables.
id  Gene symbols of genes (ACAA1).
name  Names of genes (acyl-CoA acyltransferase 1).
sub  Matrix heatmap-like with pathway and element informations instead of "X". Cells with no informations contain only "". A character matrix with 16 rows and 9 columns.

Source
compl_data function

<table>
<thead>
<tr>
<th>genes</th>
<th>List of genes.</th>
</tr>
</thead>
</table>

Description
Example of a list of genes that can be provided by an user.

Usage
data("genes")

Format
A vector with 2 observations (ACAA1, SLC6A12).
Source
Sample of data from a study on chickens, under heat-stress condition.

interactions  
Interactions between genes and metabolites

Description
Gather informations about direct interactions between genes and metabolites inside a pathway, and about pathways themselves. These informations are direct interactions between these two elements and number of relations between an element from the list provided by the user and other elements of the pathway (centrality). Direct interactions extraction was performed using BioPax, KGML and GPML files parsed with PaxtoolsR, graphite and author’s parsers.

Usage
interactions(listk, listr, listw)

Arguments
listk  Output from path_enrich function, with "KEGG" argument.
listr  Output from path_enrich function, with "REAC" argument.
listw  Output from path_enrich function, with "WP" argument.

Value
A list containing:

size  Dataframe containing pathways, genes and metabolites in pathways (from the list or not), and number of elements in pathways
pathtot  Dataframe containing pathways names and ids from pathway enrichment analysis on Reactome, Kegg and Wikipathways pathways
tagged  Dataframe containing direct interactions between elements from the user’s list per pathways
keggchebiname  Dataframe containing all human metabolites ids (kegg and chebi) and names
central  List of pathways, each pathway containing the number of direct interactions between an element of user’s list and other elements in the pathway
no_path  Dataframe containing direct interactions between elements from the user’s list, but not per pathways
genes  User’s genes given in path_enrich function
meta  User’s metabolites given in path_enrich function

This list is used by compl_data function.
Author(s)
Emilie Secherre <emisecherre@gmail.com>

References

See Also
path_enrich compl_data

Examples

```r
## load example data
data(listk)
data(listr)
data(listw)

interactions_result=interactions(listk,listr,listw)
```

interactions_result  
Output of interactions function

Description
List containing informations about interactions between genes and metabolites, centrality and pathways. Direct interactions extraction was performed using BioPax, KGML and GPML files parsed with PaxtoolsR, graphite and author’s parsers.

Usage
data("interactions_result")

Format
List of 8 items.

- **size**  Description on which elements (from user’s list or not) are contained in pathways from pathway enrichment analysis. A data frame with 286 rows and 9 variables.
- **path**  Pathways obtained through pathways enrichment analysis on KEGG, Reactome and Wikipathways pathways.
- **nb_gene_query**  Number of user’s genes contained in the pathway (0–2).
- **gene_que**  User’s genes contained in the pathway (ACAA1, SLC6A12, ACAA1 # SLC6A12).
**nb_gene_tot** Total number of genes contained in the pathway (0–2075).

**genes** All the genes contained in the pathway.

**nb_meta_query** Number of user’s metabolites contained in the pathway (0–2).

**meta_que** User’s metabolites contained in the pathway (Betaine, ATP, Betaine # ATP, ATP # Betaine).

**nb_meta_tot** Total number of metabolites contained in the pathway (0–915).

**meta** All the metabolites contained in the pathway.

**pathtot** All results of pathways enrichment analysis performed on Kegg, Reactome and Wikipathways pathways. A data frame with 286 rows and 2 variables.

**name** Name of pathways resulting in genes pathway enrichment analysis performed on Kegg, Reactome and Wikipathways.

**id** Identifiers of pathways resulting in genes pathway enrichment analysis performed on Kegg, Reactome and Wikipathways.

**tagged** Description of all direct interactions between user’s elements in pathways. A data frame with 11 rows and 6 variables.

**from** First element of the direct interaction (Betaine, SLC6A12)

**link** Description of how the two elements interact (Control(In: ACTIVATION of BiochemicalReaction), controls-transport-of-chemical).

**to** Second element of the direct interaction (Betaine, SLC6A12)

**path** Pathway containing the direct interaction (R-HSA-112310, R-HSA-112315, R-HSA-112316, R-HSA-382551, R-HSA-425366, R-HSA-425393, R-HSA-425407, R-HSA-888590, R-HSA-352230, R-HSA-442660, R-HSA-888593)

**tag** Summary of elements concerned by the interaction (Betaine / SLC6A12, SLC6A12 / Betaine)

**type** Interaction type, can be gene/gene, metabolite/metabolite, or gene/metabolite (g/m)

**keggchebiname** Dataframe containing all human metabolites ids (kegg and chebi) and names. A data frame with 16075 rows and 3 variables.

**kegg** Kegg_compound identifiers of all human metabolites.

**chebi** Chebi identifiers of all human metabolites.

**name** Names of all human metabolites.

**central** List of pathways, each pathway containing the number of direct interactions between an element of user’s list and other elements in the pathway. A list of 138 items.

**no_path** Dataframe containing direct interactions between elements from the user’s list, but not per pathways. A data frame with 1 rows and 6 variables.

**from** First element of the direct interaction (Betaine, SLC6A12)

**link** Description of how the two elements interact (Control(In: ACTIVATION of BiochemicalReaction), controls-transport-of-chemical).

**to** Second element of the direct interaction (Betaine, SLC6A12)

**path** Pathways containing the direct interaction ("R-HSA-112310, R-HSA-112315, R-HSA-112316, R-HSA-382551, R-HSA-425366, R-HSA-425393, R-HSA-425407, R-HSA-888590, R-HSA-352230, R-HSA-442660, R-HSA-888593")

**tag** Summary of elements concerned by the interaction (Betaine / SLC6A12)

**type** Interaction type, can be gene/gene, metabolite/metabolite, or gene/metabolite (g/m)

**genes** Vector containing user’s genes (ACAA1, SLC6A12)

**meta** Vector containing user’s metabolites (C00002, C00719)
Pathway enrichment analysis results for KEGG pathways.

Description
Results of pathways enrichment analysis on the list of genes and metabolites, using KEGG pathways. Pathways enrichment analysis is performed using MPINet for metabolites and gprofiler2 for genes.

Usage
data("listk")

Format
A list of 4 items.

resmeta Pathway enrichment analysis results for metabolites. A data frame with 7 rows and 2 variables.
  name Name of pathways resulting in metabolites pathway enrichment analysis.
  id Identifiers of pathways resulting in metabolites pathway enrichment analysis.

resgene Pathway enrichment analysis results for genes. A data frame with 11 rows and 2 variables.
  name Name of pathways resulting in genes pathway enrichment analysis.
  id Identifiers of pathways resulting in genes pathway enrichment analysis.

gened Vector containing user's genes (ACAA1, SLC6A12)
metad Vector containing user's metabolites (C00002, C00719)

Source
path_enrich function.
References


listr  
*Pathway enrichment analysis results for Reactome pathways.*

Description
Results of pathways enrichment analysis on the list of genes and metabolites, using Reactome pathways. Pathways enrichment analysis is performed using MPINet for metabolites and gprofiler2 for genes.

Usage
data("listr")

Format
A list of 4 items.

resmeta  Pathway enrichment analysis results for metabolites. A data frame with 278 rows and 2 variables.

  name  Name of pathways resulting in metabolites pathway enrichment analysis.
  id  Identifiers of pathways resulting in metabolites pathway enrichment analysis.

resgene  Pathway enrichment analysis results for genes. A data frame with 27 rows and 2 variables.

  name  Name of pathways resulting in genes pathway enrichment analysis.
  id  Identifiers of pathways resulting in genes pathway enrichment analysis.

gened  Vector containing user’s genes (ACAA1, SLC6A12)

metad  Vector containing user’s metabolites (C00002, C00719)

Source
path_enrich function.

References

**Description**

Results of pathways enrichment analysis on the list of genes and metabolites, using Wikipathways pathways. Pathways enrichment analysis is performed using MPINet for metabolites and gprofiler2 for genes.

**Usage**

```r
data("listw")
```

**Format**

A list of 4 items.

- **resmeta** Pathway enrichment analysis results for metabolites. A data frame with 48 rows and 2 variables.
  - `name` Name of pathways resulting in metabolites pathway enrichment analysis.
  - `id` Identifiers of pathways resulting in metabolites pathway enrichment analysis.

- **resgene** Pathway enrichment analysis results for genes. A data frame with 8 rows and 2 variables.
  - `name` Name of pathways resulting in genes pathway enrichment analysis.
  - `id` Identifiers of pathways resulting in genes pathway enrichment analysis.

- **gened** Vector containing user’s genes (ACAA1, SLC6A12)

- **metad** Vector containing user’s metabolites (C00002, C00719)

**Source**

`path_enrich` function.

**References**

Yanjun Xu, Chunquan Li and Xia Li (2013). MPINet: The package can implement the network-based metabolite pathway identification of pathways.. R package version 1.0. https://CRAN.R-project.org/package=MPINet

**meta**

*List of metabolites.*

---

**Description**

Example of a list of metabolites that can be provided by an user.

**Usage**

```r
data("meta")
```

**Format**

A vector with 2 observations (C00002, C00719).

**Source**

Sample of data from a study on chickens, under heat-stress condition.

---

**MPINetData**

*The variables in the environment variable MPINetData of the system*

---

**Description**

The variables in the environment variable MPINetData of the system.

**Format**

An environment variable

**Author(s)**

Yanjun Xu <tonghua605@163.com>, Chunquan Li <lcqbio@aliyun.com.cn> and Xia Li <lixia@hrbmu.edu.cn>
Description

Perform a pathway enrichment analysis using a list of genes and a list of metabolites. Pathways enrichment analysis is performed using MPINet for metabolites and gprofiler2 for genes.

Usage

`path_enrich(source, metabo, genes)`

Arguments

- `source`: Pathways database used, either Kegg ("KEGG"), Reactome ("REAC") or Wikipathays ("WP")
- `metabo`: Dataframe with three columns: the first column contain the list of metabolites, the second some quantitative data about the metabolites, the last one words "DOWN" or "UP" depending on the metabolites concentration behavior in a certain condition. Last two columns can contain only/some NAs. All metabolites ids are KEGG Compound ids.
- `genes`: Dataframe with three columns: the first column contain the list of genes, the second some quantitative data about the genes, the last one words "DOWN" or "UP" depending on the genes expression behavior in a certain condition. Last two columns can contain only/some NAs. All genes ids are gene symbol.

Value

A list containing:

- `resmeta`: Results of metabolites pathway enrichment analysis
- `resgene`: Results of genes pathway enrichment analysis
- `genes`: Vector containing genes
- `metabo`: Vector containing metabolites

This list is used by `interactions` function.

Author(s)

Emilie Secherre <emisecherre@gmail.com>

References

Yanjun Xu, Chunquan Li and Xia Li (2013). MPINet: The package can implement the network-based metabolite pathway identification of pathways.. R package version 1.0. https://CRAN.R-project.org/package=MPINet

See Also

interactions

Examples

```r
## load example data
data(genes)
data(meta)

## perform pathway enrichment analysis on Reactome pathways
lstr=path_enrich("REAC", meta, genes)
```

---

**rshiny**

*Shiny interface*

Description

Visualize and filter all functional informations gathered by famat using a Shiny interface.

Usage

```r
rshiny(listdata)
```

Arguments

| listdata       | Output from compl_data function |

Value

Shiny interface

Author(s)

Emilie Secherre <emisecherre@gmail.com>

References


See Also

compl_data

Examples

```r
## Not run: rshiny(compl_data_result)
```

```r
## load example data
data(compl_data_result)
```
Index

* datasets
  genes, 7
  interactions_result, 9
  listk, 11
  listr, 12
  listw, 13
  meta, 14

* file
  MPINetData, 14

compl_data, 2, 9, 16
compl_data_result, 4
consensusPath (MPINetData), 14

genes, 7
getBackground (MPINetData), 14
getnodeseq (MPINetData), 14
getStr (MPINetData), 14

interactions, 4, 8, 16
interactions_result, 9

listk, 11
listr, 12
listw, 13

meta, 14
MPINetData, 14

path_enrich, 9, 15

rshiny, 4, 16