# Package 'funOmics'

November 12, 2026

**Title** Aggregating Omics Data into Higher-Level Functional Representations

**Version** 1.6.0

**Description** The 'funOmics' package ggregates or summarizes omics data into higher level functional representations such as GO terms gene sets or KEGG metabolic pathways. The aggregated data matrix represents functional activity scores that facilitate the analysis of functional molecular sets while allowing to reduce dimensionality and provide easier and faster biological interpretations. Coordinated functional activity scores can be as informative as single molecules!

**License** MIT + file LICENSE

**Encoding** UTF-8

**RoxygenNote** 7.3.1

**URL** https://github.com/elisagdelope/funomics

**BugReports** https://github.com/elisagdelope/funomics/issues

**Depends** R (>= 4.4.0), NMF

**Imports** NMF, pathifier, stats, KEGGREST, AnnotationDbi, org.Hs.eg.db, dplyr, stringr

**Suggests** knitr, rmarkdown, testthat (>= 3.0.0), MultiAssayExperiment, SummarizedExperiment, airway

**VignetteBuilder** knitr

**VignetteIndexEntry** funomics_vignette Vignette

**biocViews** Software, Transcriptomics, Metabolomics, Proteomics, Pathways, GO, KEGG

**Config/testthat/edition** 3

**git_url** https://git.bioconductor.org/packages/funOmics

**git_branch** RELEASE_3_23

**git_last_commit** 0852938

**git_last_commit_date** 2026-04-28

**Repository** Bioconductor 3.23

**Date/Publication** 2026-11-11

**Author** Elisa Gomez de Lope [aut, cre] (ORCID:
<https://orcid.org/0000-0002-7115-7393>),
Enrico Glaab [ctb] (ORCID: <https://orcid.org/0000-0003-3977-7469>)

**Maintainer** Elisa Gomez de Lope <elisa.gomezdelope@uni.lu>

# Contents

funOmics-package . . . . . 2
get_kegg_sets . . . . . 3
short_sets_detail . . . . . 4
summarize_pathway_level . . . . . 5

**Index** **7**

---

`funOmics-package` *Aggregating Omics Data into Higher-Level Functional Representations*

---

## Description

The 'funOmics' package aggregates or summarizes omics data into higher-level functional representations such as GO terms, gene sets, or KEGG metabolic pathways. The aggregated data matrix represents functional activity scores that facilitate the analysis of functional molecular sets while allowing dimensionality reduction and providing easier and faster biological interpretations. Coordinated functional activity scores can be as informative as single molecules!

## Details

The package offers functionalities for:

- Data aggregation into functional representations
- Dimensionality reduction of omics data sets
- Analysis of coordinated functional activity scores

Use function 'get_kegg_sets' to get KEGG pathway sets for a given organism and geneid type:

get_kegg_sets(organism = "hsa", geneid_type = "entrez")

Use function 'summarize_pathway_level' to summarize omics data into higher-level functional representations that can be interpreted as functional activity scores or measures:

summarize_pathway_level(omicsmat, sets = NULL, type = "mean", minsize = 10)

## Author(s)

Elisa Gomez de Lope, Enrico Glaab

Maintainer: Elisa Gomez de Lope (elisa.gomezdelope@uni.lu)

## See Also

summarize_pathway_level get_kegg_sets

---

get_kegg_sets *Retrieves KEGG pathway gene sets for a specified organism and gene ID type.*

---

## Description

This function retrieves KEGG pathway gene sets for a specified organism. It fetches all pathways available for the specified organism from the KEGG database and maps the genes involved in each pathway. Currently, the function only supports choice of gene identifiers (entrez IDs, gene symbols or Ensembl IDs) for Homo sapiens (organism = "hsa") using the org.Hs.eg.db package.

## Usage

```
get_kegg_sets(organism = "hsa", geneid_type = "entrez")
```

## Arguments

| | |
|---|---|
| organism | The organism abbreviation for which KEGG pathway gene sets are to be retrieved (e.g., "ecj" for E. coli). Default is "hsa" (Homo sapiens). |
| geneid_type | The type of gene IDs to provide. Default is "entrez"; options are "entrez", "symbol", or "ensembl". This parameter is only used when the organism is "hsa" (Homo sapiens). |

## Value

A list where each element represents a KEGG pathway gene set. The names of the list correspond to the pathway names.

## See Also

summarize_pathway_level

keggLink, keggList

mapIds

## Examples

```
# Retrieve KEGG pathway gene sets for Homo sapiens with entrez IDs (default)
hsa_kegg_sets_entrez <- get_kegg_sets()

# Retrieve KEGG molecular sets using gene symbols
hsa_kegg_sets_symbol <- get_kegg_sets(geneid_type = "symbol")

# Retrieve KEGG molecular sets using Ensembl IDs
hsa_kegg_sets_ensembl <- get_kegg_sets(geneid_type = "ensembl")
```

```
# Retrieve KEGG pathway gene sets for another organism (e.g., Escherichia coli)
ecoli_kegg_sets <- get_kegg_sets(organism = "ecj")
```

---

`short_sets_detail` *Retrieve details and information about short molecular sets*

---

## Description

This function identifies molecular sets with sizes less than a specified threshold and returns information about these sets.

## Usage

```
short_sets_detail(sets, minsize)
```

## Arguments

| | |
|---|---|
| `sets` | A list of molecular sets. |
| `minsize` | The minimum size threshold for sets. |

## Details

This function identifies molecular sets in the input list that have sizes less than the specified minimum size (`minsize`). It returns a list containing the names, lengths, and molecules of these short molecular sets.

## Value

A list containing information about short molecular sets:

| | |
|---|---|
| `short_sets` | Names of the short molecular sets. |
| `lengths` | Lengths of the short molecular sets. |
| `genes` | Short molecular sets themselves. |

## Examples

```
ex_sets <- list(set1 = c("mol1", "mol2"), set2 = c("mol3", "mol4", "mol5"))
short_sets_info <- short_sets_detail(ex_sets, minsize = 3)
```

summarize_pathway_level *Aggregates or summarizes omics data into higher-level functional representations that can be interpreted as functional activity scores or measures.*

## Description

Given an omics matrix and a list of functional molecular sets, this function aggregates or summarizes the omics data into higher-level functional representations such as GO terms gene sets or KEGG metabolic pathways, facilitating the analysis of functional molecular sets that allow reducing dimensionality and providing easier and faster biological interpretations. Coordinated functional activity scores can be as informative as single molecules.

## Usage

```
summarize_pathway_level(omicsmat, sets = NULL, type = "mean", minsize = 10)
```

## Arguments

| | |
|---|---|
| omicsmat | A matrix or data frame representing omics data. Rows correspond to molecular identifiers, and columns correspond to samples. |
| sets | A list of functional sets. Each element in the list should represent a molecular set, and the elements of the set should match the row names of the omics matrix. |
| type | The type of pooling operator to be applied for each set. Possible values include "mean" (default), "median", "sd", "min", "max", "pca", "mds", "pathifier", "nmf", "ttest", "wilcox", "kolmogorov". |
| minsize | The minimum size per molecular set (default is 10). |

## Details

Notes:

- Different aggregation operators can be used, including summary statistics such as median (default), mean, sd, min, max, dimensionality reduction scores such as pca, mds, pathifier, or nmf, and statistical tests such as ttest, wilcoxon test, kolmogorov test.

- The minimum size per molecular set is by default 10 molecules (e.g. genes or metabolites) and can be changed with the parameter minsize.

- If "pathifier" is chosen as pooling type, the ‘aggby_pathifier‘ function internally generates a log file named ’pathifierlog.txt’ during its execution.

This log file may contain additional information that could be useful for troubleshooting or advanced analysis.

Users typically do not need to interact with this file directly, but it is mentioned here for informational purposes. For more details, this function utilizes the Pathifier package.

## Value

A matrix-like table with the activity measures for each group or set of molecules, i.e., sxn matrix, for s molecular sets and n samples.

## Author(s)

Elisa Gomez de Lope

## Examples

```
# Example usage:
g <- 10000
s <- 20
X <- matrix(abs(rnorm(g * s)), nrow = g, dimnames = list(paste0("g", 1:g), paste0("s", 1:s)))
pathways <- as.list(sample(10:100, size = 100, replace = TRUE))
pathways <- lapply(pathways, function(s, g) paste0("g", sample(1:g, size = s, replace = FALSE)), g)
names(pathways) <- paste0("pathway", seq_along(pathways))
pathway_activity <- summarize_pathway_level(X, pathways, type = "mean", minsize = 12)
```

# Index


∗ **activity-scores**
    funOmics-package, 2
    summarize_pathway_level, 5
∗ **aggregation**
    funOmics-package, 2
    summarize_pathway_level, 5
∗ **functional**
    funOmics-package, 2
    summarize_pathway_level, 5
∗ **omics**
    funOmics-package, 2
    summarize_pathway_level, 5
∗ **package**
    funOmics-package, 2
∗ **pathway**
    funOmics-package, 2
    summarize_pathway_level, 5
∗ **representation**
    funOmics-package, 2
    summarize_pathway_level, 5
∗ **summary**
    funOmics-package, 2
    summarize_pathway_level, 5

funOmics *(*funOmics-package*)*, 2
funOmics-package, 2

get_kegg_sets, *3*, 3

keggLink, *3*
keggList, *3*

mapIds, *3*

short_sets_detail, 4
summarize_pathway_level, *3*, 5