Package ‘SkeletalVis’

May 6, 2025

Type Package

Title Exploration and Visualisation of Skeletal Transcriptomics Data

Version 0.1.1

Description Allows search and visualisation of a collection of uniformly processed skeletal transcriptomic datasets. Includes methods to identify datasets where genes of interest are differentially expressed and find datasets with a similar gene expression pattern to a query dataset Soul J, Hardingham TE, Boot-Handford RP, Schwartz JM (2019) <doi:10.1093/bioinformatics/bty947>.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.3.2

Suggests mockery, rmarkdown, knitr, testthat (>= 3.0.0), withr

Config/testthat/edition 3

Imports arrow, cowplot, dplyr, DT, ggplot2, ggrepel, httr, igraph, miniUI, pbapply, plotly, rlang, shiny, stats, tidyr, utils, visNetwork

VignetteBuilder knitr

URL `https://github.com/soulj/SkeletalVis`

BugReports `https://github.com/soulj/SkeletalVis/issues`

Depends R (>= 4.0)

NeedsCompilation no

Author Jamie Soul [aut, cre, cph]

Maintainer Jamie Soul <`jamie.soul@liverpool.ac.uk`>

Repository CRAN

Date/Publication 2025-05-06 08:30:02 UTC

Contents

`browse_skeletalvis` *Browse the skeletalvis database metadata*

Description

Opens an interactive table showing the experiments available for analysis. Row can be selected to return the datasetID for the comparison of interest within that experiment.

Usage

```
browse_skeletalvis(skeletalvis)
```

Arguments

`skeletalvis`	The path to the SkeletalVis data folder.

Value

The datasetID of the selected gene expression profile from the interactive table.

Examples

```
if(interactive()){
skeletalvis <- load_skeletalvis(demo=TRUE)
selected_id <- browse_skeletalvis(skeletalvis)
}
```

`experiment_similarity` *Get cosine similarity for a query dataset against the skeletalvis database*

Description

Computes the cosine similarity of the log2 fold changes of a given query dataset against the SkeletalVis fold changes, returning a data frame of experiment similarities.

Usage

```
experiment_similarity(skeletalvis, dataset, add_meta_data = TRUE)
```

Arguments

`skeletalvis`	The path to the SkeletalVis data folder.
`dataset`	A dataframe with human gene symbols and log2 fold changes.
`add_meta_data`	Add metadata such as species, tissue, description of overall experiment and specific comparison

Value

A data frame containing cosine similarity values, IDs, and z-scores.

Examples

```
skeletalvis <- load_skeletalvis(demo=TRUE)

# Create a query dataset (this should be a data frame with the first column as gene IDs)
query_dataset <- data.frame(ID = c("SOX9", "ACAN"), fold_change = c(2.5, -1.8))

# Get cosine similarities
similarity_results <- experiment_similarity(
  skeletalvis = skeletalvis,
  dataset = query_dataset,
)

# View results
head(similarity_results)
```

`get_comparisons` *Retrieve metadata for the SkeletalVis database*

Description

‘get_comparisons‘Retieves the metadata file containing accession and comparison information for the skeletalvis database.

Usage

```
get_comparisons(skeletalvis, accession = NULL)
```

Arguments

`skeletalvis`	The path to the SkeletalVis data folder.
`accession`	Optionally an experiment accession number for the experiment of interest to filter the metadata by.

Value

A dataframe containing metadata.

Examples

```
# Get the path to the directory with the accessions.txt file
skeletalvis_path <- load_skeletalvis(demo=TRUE)

# Retrieve the metadata
metadata <- get_comparisons(skeletalvis_path)

# View the metadata
head(metadata)

metadata <- get_comparisons(skeletalvis_path, "GSE85761")
```

`get_experiment` *Get the gene differential expression results for an experiment of interest*

Description

This function loads the differential expression data (log2 foldchanges and FDR values) for a given dataset id

Usage

```
get_experiment(skeletalvis, dataset_id)
```

Arguments

`skeletalvis`	The path to the SkeletalVis folder.
`dataset_id`	The dataset ID to extract results for.

Value

A data frame containing differential expression results for the specified dataset ID.

Examples

```
skeletalvis <- load_skeletalvis(demo=TRUE)
experiment_results <- get_experiment(skeletalvis, "GSE12860_6")
```

`get_exp_table` *Retrieve the experiment table for the SkeletalVis database*

Description

Retrieves the experiment table file containing experiment accessions and descriptions from the skeletalvis database.

Usage

```
get_exp_table(skeletalvis)
```

Arguments

`skeletalvis`	The path to the skeletalvis data folder.

Value

A dataframe containing the experiment information

Examples

```
# Get the path to the directory with the accessions.txt file
skeletalvis_path <- load_skeletalvis(demo=TRUE)

# Retrieve the metadata
exptable <- get_exp_table(skeletalvis_path)

# View the metadata
head(exptable)
```

`get_gene_fold_changes` *Get gene differential expression results for genes of interest*

Description

Extracts the fold change and p-values across the SkeletalVis database for the specified genes.

Usage

```
get_gene_fold_changes(
  skeletalvis,
  gene_symbols,
  return_fdr = TRUE,
  add_meta_data = TRUE
)
```

Arguments

`skeletalvis`	The path to the SkeletalVis data folder.
`gene_symbols`	The human gene symbols to extract fold change and FDR values for.
`return_fdr`	Return FDR values (FALSE by default).
`add_meta_data`	Add metadata such as species, tissue, description of overall experiment and specific comparison

Value

A tibble containing gene expression results for the specified genes.

Examples

```
skeletalvis <- load_skeletalvis(demo=TRUE)

gene_results <- get_gene_fold_changes(skeletalvis, c("SOX9","ACAN"))

head(gene_results)
```

`load_skeletalvis` *Load SkeletalVis data*

Description

Checks the default SkeletalVis data location for the differential expression data and metadata files. If any files are missing, it lists them and ask permission to download the missing files.

Usage

```
load_skeletalvis(verbose = TRUE, ask = TRUE, demo = FALSE)
```

Arguments

`verbose`	Logical. If TRUE, prints messages about file existence and download status. Default is TRUE
`ask`	Logical. If TRUE, prompts the user for confirmation before downloading missing files. Default is TRUE.
`demo`	Logical. If TRUE, uses built in demo data suitable for testing the package functions. Default is FALSE

Value

The path to the SkeletalVis data folder.

Examples

```
# Load the demo SkeletalVis data.
skeletalvis <- load_skeletalvis(demo = TRUE)

# Download full dataset (only if running interactively)
if (interactive()) {
  skeletalvis <- load_skeletalvis()
}
```

`plot_similarity` *Plot ranked cosine similarity scores with top n labels*

Description

Takes a similarity table with columns for ID, cosine, and zscore values, and generates a plot showing the rank of the z-scores on a log scale. Labels for the top n IDs with the highest z-scores are displayed on the plot.

Usage

```
plot_similarity(similarity_table, top_n = 10)
```

Arguments

`similarity_table`	A data frame with columns: ‘ID‘ (identifier for each item), ‘cosine‘ (cosine similarity score), and ‘zscore‘ (z-score).
`top_n`	An integer indicating the number of top IDs by zscore to label in the plot. Default is 10.

Value

A ggplot object displaying the ranked z-scores on a log scale, with labels for the top n items by zscore.

Examples

```
data(query)
skeletalvis <- load_skeletalvis(demo=TRUE)
similarity_results <- experiment_similarity(skeletalvis, query)

# Plot similarity table with labels for top 5 items
plot_similarity(similarity_results, top_n = 5)
```

`query` *Example gene expression dataset*

Description

An example gene expression dataset comparing damaged versus intact osteoarthritic cartilage.

Usage

```
query
```

Format

‘query‘ A data frame with 42,459 rows and 2 columns:

ID Human Gene Symbol

‘E-MTAB-4304_1‘ log2 fold change

Source

<https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-4304>

Examples

```
head(query)
```

`search_skeletalvis` *Search for experiments in SkeletalVis*

Description

Search for experiments matching a search term in the SkeletalVis database. The search can be done across the entire database or specific metadata columns.

Usage

```
search_skeletalvis(skeletalvis, search_term, columns = NULL)
```

Arguments

`skeletalvis`	Folder with the SkeletalVis data
`search_term`	A string representing the term to search for.
`columns`	A character vector of column names to limit the search. If NULL (default), the entire table is searched.

Value

A data frame containing experiments that match the search term in the specified columns or the whole table.

Examples

```
skeletalvis <- load_skeletalvis(demo=TRUE)
# Search across all columns for rows containing "SOX9"
result <- search_skeletalvis(skeletalvis, "SOX9")

# Search only in the 'Perturbation' and 'Description' columns
result <- search_skeletalvis(skeletalvis, "SOX9", columns = c("Perturbation", "Description"))
```

`view_curated_oagenes` *View curated osteoarthritis associated genes*

Description

Retrieves the OATargets table of literature curated genes associated with osteoarthritis joint damage in animals.

Usage

```
view_curated_oagenes(skeletalvis)
```

Arguments

`skeletalvis`	The path to the SkeletalVis folder.

Value

A data frame with the following columns:

PMID The PubMed ID of the paper describing the observation.

Gene The human gene symbol of the gene modulated in the study.

Effect on gene product The effect of the gene modulation on the gene product.

Model The type of OA model where the gene has been modulated.

Susceptibility observed Increase (detrimental) or decrease (protective) in OA severity upon gene modulation.

Inferred gene effect Based on the modulation effect, whether the gene is protective or detrimental in OA.

Delivery The tissue specificity of the gene modulation.

Species The species under study.

pub_date The date of publication (epub).

LastAuthor The last author of the publication.

Type Genetic or exogenous modulation.

Intervention For exogenous modulations, the method of gene modulation.

simpleModel The broad type of OA model where the gene has been modulated.

effectConsensus Consensus inferred gene effect across observations; ’ambiguous’ if discrepancies exist.

NumStudies The number of publications studying each gene.

Examples

```
skeletalvis <- load_skeletalvis(demo=TRUE)
oatargets <- view_curated_oagenes(skeletalvis)
```

`view_network` *View the OATarget gene network*

Description

This function visualises the protein-protein interaction network surrounding a specified gene, highlighting those genes that have been experimentally pertubated and the resulting effect of OA.

Usage

```
view_network(skeletalvis, query, hide_unannotated = TRUE, physics = TRUE)
```

Arguments

`skeletalvis`	A string specifying the path to the directory containing the "network.RDS" file.
`query`	The name of the gene to visualise within the network.
`hide_unannotated`	A logical value indicating whether to hide nodes with no measured effect (default: TRUE).
`physics`	A logical value indicating whether to enable physics for the network (default: TRUE).

Details

The function extracts the subnetwork surrounding the specified gene. Nodes are coloured based on their effect ("Protective", "Detrimental", "Ambiguous", "No effect", or "Not measured"). If ‘hide_unannotated = TRUE‘, nodes not yet studied are removed unless they are the target gene.

Value

A visNetwork interactive network’

Examples

```
# Specify the path to the skeletalvis directory
skeletalvis <- load_skeletalvis(demo=TRUE)

# Visualize the network for a specific gene
vis <- view_network(skeletalvis, "COL2A1")

# Show unannotated nodes
vis_filtered <- view_network(skeletalvis, "COL2A1", hide_unannotated = FALSE)
```

`view_prioritised_oagenes`

View prioritised osteoarthritis associated genes

Description

Retrieves the machine learning model prioritised table of osteoarthritis joint damage in animals. Genes are ranked by their probability to cause an OA damage phenotype when experimentally perturbed. The predicted direction of effect and druggability of that gene is also provided.

Usage

```
view_prioritised_oagenes(skeletalvis)
```

Arguments

`skeletalvis`	The path to the SkeletalVis folder.

Value

A data frame with the following columns:

Gene The human gene symbol.

Rank The rank of the priortisation scores from the machine learning model

PredictedEffect The predicted effect this gene

Category_sm The small molecular tractibility of this gene

Category_ab The antibody tractibility of this gene

Examples

```
skeletalvis <- load_skeletalvis(demo=TRUE)
oatargets <- view_prioritised_oagenes(skeletalvis)
```

`volcano_plot` *Volcano Plot of Gene Expression Data*

Description

Creates a volcano plot showing the log2 fold change and FDR values with optional labelling of points.

Usage

```
volcano_plot(
  data,
  number_points = 5,
  selected_points = NULL,
  interactive = FALSE,
  logFC_threshold = log2(1.5),
  FDR_threshold = 0.05,
  point_size = 2,
  lab_size = 4
)
```

Arguments

`data`	A data frame containing gene expression data, containing an ID, log2 fold-change and FDR columns.
`number_points`	Number of top up and down regulated points to label
`selected_points`	Character vector of the IDs to label
`interactive`	Should an interactive plotly graph be made?

`logFC_threshold`	The foldchange threshold to define up and down regulated points
`FDR_threshold`	The pval threshold to define up and down regulated points
`point_size`	Size for the points
`lab_size`	Size for the labels

Value

A ggplot object representing the volcano plot.

Examples

```
skeletalvis <- load_skeletalvis(demo=TRUE)

experiment_results <- get_experiment(skeletalvis, "GSE155118_1")

volcano_plot(experiment_results)
```

Index