Package 'octad.db'

April 8, 2025

Title Open Cancer TherApeutic Discovery (OCTAD) database

Version 1.9.0

Description Open Cancer TherApeutic Discovery (OCTAD) package implies sRGES approach for the drug discovery. The essential idea is to identify drugs that reverse the gene expression signature of a disease by tamping down over-expressed genes and stimulating weakly expressed ones. The following package contains all required precomputed data for whole OCTAD pipeline computation.

License Artistic-2.0

Encoding UTF-8

LazyData false

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.1

Depends R (>= 4.2.0), ExperimentHub

biocViews ExperimentData, CancerData, ExperimentHub, SequencingData, ExpressionData

VignetteBuilder knitr

Suggests knitr, rmarkdown

git_url https://git.bioconductor.org/packages/octad.db

git_branch devel

git_last_commit 8170722

git_last_commit_date 2024-10-29

Repository Bioconductor 3.21

Date/Publication 2025-04-08

Author E. Chekalin [aut, cre],
S. Paithankar [aut],
B. Zeng [aut],
B. Glicksberg [ctb],
P. Newbury [ctb],
J. Xing [ctb],
K. Liu [ctb],

A. Wen [ctb],
D. Joseph [ctb],
B. Chen [aut]

Maintainer E. Chekalin <eygen.chekalin@gmail.com>

Contents

`get_ExperimentHub_data` *Download file from Experimental Hub*

Description

Download file from Experimental Hub.

Usage

```
get_ExperimentHub_data(file=NULL)
```

Arguments

`file`	file id do download it from Experimental Hub archive

Value

Returns a data.frame object depends on the call

See Also

`octad.db`.

Examples

```
phenoDF=get_ExperimentHub_data("EH7274") #load data.frame with samples included in the OCTAD database.
 head(phenoDF)
```

octad.db *Open Cancer TherApeutic Discovery (OCTAD) database package*

Description

This is a support package for the main package octad which can be obtained here

Details

Package includes all required data for drug repurposing OCTAD pipeline. Initialization of the pipeline starts with listing all available samples:

```
phenoDF=.eh[["EH7274"]] #load data.frame with samples included in the OCTAD database. head(phenoDF) #list all data included within the package
```

Besides, the package includes examples of the output from diffExp and runsRGES functions from the octad package along with desctiption of the data: `res=octad.db::res_example` #load example res from octad.db `?res_example` `sRGES=octad.db::sRGES_example` #load example sRGES from octad.db `?sRGES_example`

`CCLE.log2.read.count.matrix` log2-transformed matrix containing expression of 56,318 genes in 1,019 cell lines provided in CCLE database `CCLE.overlaps` log2-transformed matrix containing expression of 56,318 genes in 51 cell lines present in both LINCS and CCLE databases `CCLE.sample.meta` data.frame containing cell lines name and tissue of origin from CCLE database `CTRPv2.sensprof` Cancer Therapeutics Response Portal (CTRP) data.frame containing AUC and IC50 for every drug-cell line pair `EncoderDF` a data.frame (64 features by 19127 samples) computed from autoEncoder for every sample in the OCTAD database `cmpd_sets_ChemCluster` list of compounds and their clusters based on chemical structures `cmpd_sets_mesh` list of compounds and their associations with 662 pharmacological MeSH Terms provided in PubChem `fda_drugs` data.frame of FDA approved drugs containing name, target and clinical phase `lincs_sig_info` data.frame for LINCS signatures including experiment id, cell line used in the experiment, purturbagen name, purturbagen type, dose and time `lincs_signatures` Differential expression of 978 genes in 416,560 experiments `merged_gene_info` data.frame with the annotation of 75,078 genes which could be used for identifier mapping `octad.LINCS.counts` expression matrix of 965 genes and 19127 samples in OCTAD database (965 genes were profiled in LINCS) `phenoDF` data.frame with samples from the OCTAD database including tissue of origin, sample type (tumor-derived or healthy-derived), cancer, orginal source (e.g. TCGA), mutation and stage data if available `random_gsea_score` pre-computed permutated gsea scores for chembl, mesh and ChemCluster databases that are used for drug enrichment analysis `tsne` 2 dimension tSNE for every sample in the OCTAD database along with sample.id, cancer type and source of the initial data `octad.counts.and.tpm.h5` expression (TPM and raw counts) matrix of 60,498 genes and 19,127 samples

The code can be viewed at the GitHub repository, which also lists the contributor code of conduct:

https://bioconductor.org/packages/octad or https://github.com/Bin-Chen-Lab/octad for the pipeline package

References

Zeng, B., Glicksberg, B.S., Newbury, P., Chekalin, E., Xing, J., Liu, K., Wen, A., Chow, C. and Chen, B., 2021. OCTAD: an open workspace for virtually screening therapeutics targeting precise

cancer patient groups using gene expression features. Nature protocols, 16(2), pp.728-753. https://www.nature.com/articles/s41596-020-00430-z

Index