Package 'rcellminerData'

December 3, 2024

Title rcellminerData: Molecular Profiles and Drug Response for the NCI-60 Cell Lines

Version 2.28.0

Date 2024-10-15

Author Augustin Luna, Vinodh Rajapakse, Fabricio Sousa

Maintainer Augustin Luna <lunaa@cbio.mskcc.org>, Vinodh Rajapakse <vinodh.rajapakse@nih.gov>, Fathi Elloumi <fathi.elloumi@nih.gov>

Description The NCI-60 cancer cell line panel has been used over the course of several decades as an anti-cancer drug screen. This panel was developed as part of the Developmental Therapeutics Program (DTP, http://dtp.nci.nih.gov/) of the U.S. National Cancer Institute (NCI). Thousands of compounds have been tested on the NCI-60, which have been extensively characterized by many platforms for gene and protein expression, copy number, mutation, and others (Reinhold, et al., 2012). The purpose of the CellMiner project (http://discover.nci.nih.gov/cellminer) has been to integrate data from multiple platforms used to analyze the NCI-60 and to provide a powerful suite of tools for exploration of NCI-60 data.

License LGPL-3 + file LICENSE

Depends R (>= 3.5.0), Biobase

Suggests knitr, testthat, BiocStyle, rcellminer, rmarkdown

LazyData true

VignetteBuilder knitr

biocViews CancerData, CopyNumberVariationData, ExpressionData, SNPData, NCI, MicroarrayData, miRNAData

Encoding UTF-8

RoxygenNote 7.1.2

git_url https://git.bioconductor.org/packages/rcellminerData

git_branch RELEASE_3_20

git_last_commit f54392c

git_last_commit_date 2024-10-29

Repository Bioconductor 3.20

Date/Publication 2024-12-03

Contents

`drugData` *CellMiner compound data: NCI-60 activity and metadata*

Description

The drugData object, of rcellminer package class DrugData, organizes data for 20861 compounds screened over the NCI-60 cancer cell lines by the NCI Developmental Therapeutics Program. These include activity profiles, as well as other information such as chemical structure, and where appropriate, name, mechanism of action, clinical testing status, etc.

Usage

```
data(drugData)
```

Details

The activity data has two forms. A standardized 'z-score' data set provides a single NCI-60 profile for each compound, possibly averaging data derived from multiple experiments (subject to quality standards described in Reinhold et al.). A repeat activity data set provides data at the level of individual NCI-60 profiling experiments.

Value

the drugData object that contains Cellminer drug information as a DrugData object

Source

CellMiner Website: <URL: http://discover.nci.nih.gov/cellminer/>

References

Reinhold, W.C., et al. (2012) CellMiner: a web-based suite of genomic and pharmacologic tools to explore transcript and drug patterns in the NCI-60 cell line set, Cancer Research, 72, 3499-3511

`molData` *CellMiner molecular profiling data for the NCI-60 cell lines.*

Description

The molData object, of rcellminer package class MolData, organizes CellMiner's molecular profiling data for the NCI-60 cancer cell lines. Also provided are cell line annotations, including tissue of origin, detailed histology, and source information.

Usage

```
data(molData)
```

Details

Data matrices for the following types of data are provided. NOTE: Missing values are represented as NA. Pubmed IDs (PMID) are provided.

- cop: Copy Number (Combined_aCGH): The gene copy data set values are derived from log 2 probe intensities averaged over 4 platforms: Agilent Human Genome CGH Microarray 44A, Nimblegen HG19 CGH 385K WG Tiling v2.0, Affymetrix GeneChip Human Mapping 500k Array Set, and Illumina Human1Mv1_C Beadchip. For a given gene and cell line, the indicated value is the deviation from the median average intensity value (with the median taken over all genes in that cell line). Values close to zero thus correspond to a copy number of 2, while positive and negative deviations from zero indicate copy gains and losses, respectively. Citation: PMID: 24670534.
- mut: Binary Gene Level Summarized Mutation: The binary gene mutation data are derived from the variant level data in exo using the rcellminer function getBinaryMutationData (see data generation script make_rcellminerdata.R for details). The binary gene mutation profiles indicate, for each gene and NCI-60 cell line, the presence or absence of a deleterious (and presumably somatic) mutation. Deleterious mutations are frameshift, nonsense, splice site, or missense mutations predicted to alter function by the SIFT and POLYPHEN2 algorithms. Potential germline mutations are filtered by excluding variants with an estimated frequency greater than 0.005 in the normal population (based the 1000 Genomes and ESP5400 databases).
- exo: Exome: Exome regions captured using Agilent SureSelect All Exon v1.0 Kit and sequenced using Illumina Genome Analyzer IIx. Values are percent conversion of genetic variant. This data captures mutations in the NCI-60. The file "exome_gene_probeIds_mapping.txt" is provided, since the CellMiner data dump does not provide a mapping between genes and mutations. Citation: PMID: 23856246. NOTE: 0 indicates no variant allele
- exp: Expression: The mRNA transcription level values are average intensity values combined from 4 platforms: Affy HG-U133(A-B), Affy HuEx 1.0, Affy HG-U133 Plus 2.0, Affy HG-U95(A-E). Average intensity values for given gene are presented if all probe expression patterns are well correlated. Citation: PMID: 22802077.
- xai: Expression (Avg. log2 Intensity).
- pro: Protein (RPPA; Protein Lysate): Selected values are derived from 176 antibodies associated with 94 proteins. For some proteins, expression data is provided from multiple antibodies. Averaging the latter is not appropriate, but proteins with a single antibody can be selected by excluding probe names with an underscore (yielding a 74 protein data set). Values are provided as dose interpolation values using a method described in Nishizuka et al., 2003. PMID: 14623978

- mir: MicroRNA: 15,000 probes for 723 human and 76 human viral miRNA's. Each slide contains 8 arrays. PMID:20442302.
- mda: Metadata (Phenotype Data). Further information: http://discover.nci.nih.gov/cellminer/celllineMetadata.do
  - age: Patient Age (Years)
  - is_epithelial: Is cell derived from epithelial source?
  - is_p53_mut: Is TP53 mutated?
  - mdr: Multi-Drug Resistance Assay. PMID: 7969041
  - doublingtime: Cell doubling time in hours

Value

the molData object that contains Cellminer molecular profiling information as a MolData object

Source

CellMiner Website: URL: http://discover.nci.nih.gov/cellminer Data Set Information: URL: http://discover.nci.nih.gov/ce

References

Reinhold, W.C., et al. (2012) CellMiner: a web-based suite of genomic and pharmacologic tools to explore transcript and drug patterns in the NCI-60 cell line set, Cancer research, 72, 3499-3511

Index