

GSVAdata

October 25, 2011

GSVAdata-package *Data employed in the vignette of the GSVA package.*

Description

This package contains data employed in the vignette of the GSVA package.

Data sets

- [leukemia](#) Leukemia data by Armstrong et al. (2002) from the Broad Institute.
- [c2BroadSets](#) C2 canonical pathways from the MSigDB 3.0 database of gene sets at the Broad Institute.

Author(s)

S. Haenzelmann, J. Guinney and R. Castelo

References

Scott A. Armstrong, Jane E. Staunton, Lewis B. Silverman, Rob Pieters, Monique L. den Boer, Mark D. Minden, Stephen E. Sallan, Eric S. Lander, Todd R. Golub and Stanley J. Korsmeyer. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet* 30:41-47, 2002.

brainTxDbSets *Gene sets signatures of brain cell types*

Description

Gene sets signatures specific to four different brain cell types (astrocytes, oligodendrocytes, neurons and cultured astroglial cells) derived from murine models (Cahoy et al. 2008).

Usage

```
data(brainTxDbSets)
```

Details

The data is contained in an `list` object called `brainTxDbSets` obtained from the Brain Transcriptome Database (Cahoy et al., 2008).

Source

Cahoy, J.D., Emery, B., Kaushal, A., Foo, L.C., Zamanian, J.L. et al. A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J Neurosci*, 28:264-278, 2008.

Examples

```
data(brainTxDbSets)
head(lapply(brainTxDbSets, head))
```

c2BroadSets	<i>C2 collection of canonical pathways from MSigDB 3.0</i>
-------------	--

Description

C2 Broad Sets.

Usage

```
data(c2BroadSets)
```

Details

The data is contained in an `GeneSetCollection` object called `c2BroadSets` obtained by parsing the file `c2.all.v3.0.entrez.gmt`, downloaded from <http://www.broadinstitute.org/gsea>, using the `getGmt()` function from the `GSEABase` package.

Source

Subramanian, Tamayo, et al. *PNAS*, 102:15545-15550, 2005.

Mootha, Lindgren, et al. *Nat Genet*, 34:267-273, 2003.

Examples

```
data(c2BroadSets)
c2BroadSets
```

`gbm_eset`*Glioblastoma Multiforme (GBM) Data by Verhaak et al. (2010)*

Description

Microarray data from Glioblastoma multiforme (GBM) downloaded from the TCGA website (<http://cancergenome.nih.gov>). The data is provided as an `ExpressionSet` object containing RMA-processed expression values.

Usage

```
data(gbm_VerhaakEtAl)
```

Details

The data is contained in an `ExpressionSet` object called `gbm_eset` and was obtained using RMA (Irizarry et al. 2003).

Source

Scott A. Armstrong, Jane E. Staunton, Lewis B. Silverman, Rob Pieters, Monique L. den Boer, Mark D. Minden, Stephen E. Sallan, Eric S. Lander, Todd R. Golub and Stanley J. Korsmeyer. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet* 30:41-47, 2002.

References

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, 4(2):249–64, 2003.

Verhaak, R.G.W., Hoadley, K.A., Purdom, E., Wang, V., Qi, Y., et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*, 17:98–110.

Examples

```
data(gbm_VerhaakEtAl)
gbm_eset
head(pData(gbm_eset))
table(gbm_eset$subtype)
```

`leukemia_eset`*Leukemia Data by Armstrong et al. (2002) from the Broad Institute*

Description

Microarray data hybridized on the Affymetrix Human Genome U95 Set chip, for 37 different individuals with human acute leukemias, where 20 of them had conventional childhood acute lymphoblastic leukemia (ALL) and the other 17 were affected with the MLL (mixed-lineage leukemia gene) translocation. The data is provided as an `ExpressionSet` object containing RMA-processed expression values.

Usage

```
data(leukemia)
```

Details

The data is contained in an `ExpressionSet` object called `leukemia_eset` obtained as follows:

- Raw CEL files corresponding to the data of the entire study (72 individuals) were downloaded from http://www.broadinstitute.org/cgi-bin/cancer/publications/pub_paper.cgi?mode=view&paper_id=63
- 41 ALL and MLL samples with the same scanning date were kept and the rest were discarded.
- Based on quality assessments by NUSE and RLE diagnostics (Bolstad et al., 2005), 4 additional samples were discarded such that 20 ALL and 17 MLL samples were finally kept.
- Probe-level data from these 37 samples were background corrected, normalized and summarized using RMA (Irizarry et al., 2003) providing this final `ExpressionSet` object.

Source

Scott A. Armstrong, Jane E. Staunton, Lewis B. Silverman, Rob Pieters, Monique L. den Boer, Mark D. Minden, Stephen E. Sallan, Eric S. Lander, Todd R. Golub and Stanley J. Korsmeyer. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet* 30:41-47, 2002.

References

Bolstad, B.M., Collin, F., Brettschneider, K., Simpson, L., Irizarry, R., and Speed, T.P. Quality assessment of Affymetrix GeneChip data. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, pg. 33–48, Springer, 2005.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, 4(2):249–64, 2003.

Examples

```
data(leukemia)
leukemia_eset
head(pData(leukemia_eset))
```

Index

*Topic **datasets**

brainTxDbSets, 1

c2BroadSets, 2

gbm_eset, 3

leukemia_eset, 4

*Topic **dataset**

GSVAdata-package, 1

brainTxDbSets, 1

c2BroadSets, 1, 2

gbm_eset, 3

gbm_VerhaakEtAl (*gbm_eset*), 3

GSVAdata (*GSVAdata-package*), 1

GSVAdata-package, 1

leukemia, 1

leukemia (*leukemia_eset*), 4

leukemia_eset, 4