Package 'bionetdata'

September 7, 2014

Type Package

Title Biological and chemical data networks

Version 1.0.1

Date 2014-09-07

Author Matteo Re, Giorgio Valentini -- Universita' degli Studi di Milano

Maintainer Matteo Re<re@di.unimi.it>

Description Data Package that includes several examples of chemical and biological data networks, i.e. data graph structured.

License GPL (>= 2)

LazyLoad yes

Depends R (>= 2.10)

R topics documented:

---

bionetdata-package   *bionetdata: Biological and chemical data networks*

Description

Data Package that includes several examples of chemical and biological data networks represented through adjacency matrices of a graph

Details

Package:	bionetdata
Type:	Package
Version:	1.0.1
Date:	2014-09-07
License:	GPL (>= 2)
LazyLoad:	yes

Author(s)

Giorgio Valentini and *Matteo Re*

DI, Dipartimento di Scienze dell'Informazione

Universita' degli Studi di Milano

<{valentini,re}@di.unimi.it>

Maintainer: *Matteo Re*

CGM.Cat   *Cancer Gene Modules*

Description

Cancer Gene Modules classes for the genes included in `FIN.data`. Annotations are taken from the GSEA MSigDB (Molecular Signatures Database) public repository:

http://www.broadinstitute.org/gsea/msigdb

The annotations are available in MSigDB C4 (computational gene sets collections) CM (cancer modules), file: c4.cgn.v3.0.symbols.gmt . This collection of gene sets is taken from a work published by Segal and colleagues aimed at the definition of a Cancer Modules Maps. According to the definition of the authors Cancer Gene Modules are groups of genes that act in concert to carry out a biological function/process. Cancer Gene Modules have been used to describe expression profiles in different tumors types in terms of the behavior of modules. Further information about specific

modules can be found at: http://robotics.stanford.edu/~erans/cancer/browse_by_modules.html It is worth noting that the Cancer Gene Modules annotations contained in bionetdata covers only the 2033 genes included in `FIN.data`.

Usage

```
data(CGM.Cat)
```

Format

A 2033 x 10 named matrix where rows refer to human gene symbols, columns to 10 Cancer Modules classes.

Source

Original annotations (c4.cm.v3.0.symbols.gmt) are available at:

http://www.broadinstitute.org/gsea/msigdb

References

Segal E., Friedman N., Koller D. and Regev A., A module map showing conditional activity of expression modules in cancer, Nature Genetics, 36(10), 2004

Examples

```
data(CGM.Cat);
CGM.Cat[1:10,1:10];
```

DD.chem.data	*Drug-drug chemical similarities data*

Description

Chemical structure similarities between 1253 FDA approved drugs obtained from DrugBank 3.0.

Usage

```
data(DD.chem.data)
```

Format

A 1253 x 1253 named matrix where both rows and columns refer to DrugBank drugs. Drugs names are DrugBank 3.0 identifiers.

Details

This matrix contains the Tanimoto chemical structure similarity scores between 1253 DrugBank drugs. Canonical Simplified molecular-input line-entry specifications (SMILES) of the drugs were extracted from the DrugBank Drugcards. The SMILES were then converted into molecular extended fingerprints (1024 bits) using the rcdk package. The set of fingerprints was finally converted into a Tanimoto similarity matrix using the fp.sim.matrix rcdk function. All the real values contained in this matrix represent the chemical structure similarity between each possible pair of drugs and are comprised between 0 (completely different chimical structures) and 1 (identical chemical structures).

Source

The SMILES representations of the drugs were extracted by the Smiles_String_canonical fields of the DrugCards contained in the file drugbank.txt available at:

http://www.drugbank.ca/system/downloads/current/drugbank.txt.zip

References

Wishart, D., Knox, C., Guo, A., Shrivastava, S., Hassanali, M., Stothard, P., Chang, Z., Woolsey, J.: DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res. 34(Jan), D668-D672 (2006)

Guha, R.: Chemical Informatics Functionality in R. Journal of Statistical Software 18(6), (2007)

Examples

```
data(DD.chem.data)
DD.chem.data[1:10,1:10]
```

DrugBank.Cat   *DrugBank categories*

Description

DrugBank 3.0 drugs categories of the 1253 drugs contained in the DD.chem.data matrix.

Usage

```
data(DrugBank.Cat)
```

Format

A 1253 x 45 named binary matrix where rows refer to DrugBank drugs, columns to 45 Drug categories extracted from the DrugBank DrugCards. This matrix contains a 1 if the drug associated to the considered row is annotated in the category represented by the considered column and a 0 otherwise.

Details

The drug categories were extracted by parsing the Drug_Category field of the DrugBank 3.0 Drug-Cards of the 1253 FDA approved drugs contained in the `DD.chem.data` matrix. The DrugBank Drug categories are produced by manual curators. On the contrary of other drugs functional annotation schemes these drugs categories are not restricted to specific deseases. Indeed it is possible to find in this functional annotation scheme categories associated to drugs involved in the treatment of specific deseases (i.e. 'AntiParkinsons_Agents') or to the treatment of symptoms (i.e. 'Anticonvulsants') associated to many deseases. This functional labelling scheme contains also categories based on the Mode of Action (MOA) of the drugs as in the case of the 'Adrenergic_Uptake_Inhibitors' category.

Source

The drug categories were extracted by parsing the Drug_Category field of the DrugCards contained in the file drugbank.txt available at:

http://www.drugbank.ca/system/downloads/current/drugbank.txt.zip

References

Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbou, Y., Eisner, R., Guo, A., Wishart, D.: DrugBank 3.0: a comprehensive resource for "omics" research on drugs. Nucleic Acids Res. 39(Jan), D1035-41 (2011)

Examples

```
data(DrugBank.Cat)
DrugBank.Cat[1:10,1:10];
dimnames(DrugBank.Cat)[[2]];
```

FIN.data	*Functional Interaction Network data for human.*

Description

Functional Interaction data obtained from the supplemental materials of a paper written by Wu and colleagues: 'A human functional protein interaction network and its application to cancer data analysis'. Data are represented through a binary named matrix and represent the presence or absence of gene-gene functional interactions. The original network is composed by more than 9000 genes. The network contained in the bionetdata R package is a reduced version containing 2033 nodes and obtained using the walktrap.community function of the R package igraphi. The walktrap.community function is an implementation of a community detection method developed by Pons and Latapy based on a between vertices similarity measure computed by means of a random walk.

Usage

```
data(FIN.data)
```

Format

Binary named matrix. Entry `FIN.data[i,j] = 1` if there is a functional interaction between gene i and j, otherwise `FIN.data[i,j] = 0`.

Source

Wu and colleagues original data are available as : http://genomebiology.com/content/supplementary/gb-2010-11-5-r53-s3.zip

References

Wu G., Feng X. and Stein L. A human functional protein interaction network and its application to cancer data analysis, Genome Biology 11:R53, 2010.

Pons P. and Latapy M. Computing communities in large networks using random walks, J. of Graph Alg. and App. bf, 10:284-293, 2004.

Examples

```
data(FIN.data);
FIN.data[1:10,1:10];
```

Yeast.Biogrid.data   *Yeast BioGRID data*

Description

Protein-protein interaction (PPI) data (BioGRID) of yeast have been downloaded from the BioGRID database, that collects PPI data from both high-throughput studies and conventional focused studies (Stark et al. 2006). Data are represented through a binary named matrix and represent the presence or absence of protein-protein interactions. Names correspond to systematic names of yeast genes.

Usage

```
data(Yeast.Biogrid.data)
```

Format

Binary named matrix. Entry `Yeast.Biogrid.data[i,j] = 1` if there is an interaction between gene i and j, otherwise `Yeast.Biogrid.data[i,j] = 0`.

Source

BioGRID data base: http://thebiogrid.org

References

Stark, C., Breitkreutz, B., Reguly, T., Boucher, L., Breitkreutz, A., and Tyers, M. (2006). BioGRID: a general repository for interaction datasets. Nucleic Acids Res., 34, D535-D539.

Examples

```
data(Yeast.Biogrid.data);
Yeast.Biogrid.data[1:10,1:5];
```

Yeast.Biogrid.FunCat   *Yeast FunCat classes for* BioGRID *data.*

Description

FunCat classes for the genes included in `Yeast.Biogrid.data`. Annotations refer the funcat-2.1 scheme, and funcat-2.1 data 20070316 data, available from the MIPS web site.

Usage

```
data(Yeast.Biogrid.FunCat)
```

Format

A named matrix where rows refer to yeast genes, columns to FunCat classes. Names of yeast genes are systematic names. Names of columns correspond to FunCat IDs.

Source

http://mips.gsf.de/projects/funcat

References

Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Guldener, U., Mannhaupt, G., Munsterkotter, M., and Mewes, H. (2004). The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Research, 32(18), 5539-5545.

Examples

```
data(Yeast.Biogrid.FunCat)
Yeast.Biogrid.FunCat[1:10,1:6]
```

Yeast.STRING.data   *Yeast interactions from STRING*

Description

Binary protein-protein interactions from the STRING data base (von Mering et al. 2002), representing interaction data from yeast two-hybrid assay, mass-spectrometry of purified complexes, correlated mRNA expression and genetic interactions.

Usage

```
data(Yeast.STRING.data)
```

Format

Binary named matrix. Entry `Yeast.Biogrid.STRING[i,j]` = 1 if there is an interaction between gene i and j, otherwise `Yeast.STRING.data[i,j]` = 0.

Source

<http://string-db.org>

References

von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S., Fields, S., and Bork, P. (2002). Comparative assessment of large-scale data sets of protein-protein interactions. Nature, 417, 399-403.

Examples

```
data(Yeast.STRING.data)
```

Yeast.STRING.FunCat   *Yeast FunCat classes for* STRING *data.*

Description

FunCat classes for the genes included in `Yeast.STRING.data`. Annotations refer the funcat-2.1 scheme, and funcat-2.1 data 20070316 data, available from the MIPS web site.

Usage

```
data(Yeast.STRING.FunCat)
```

Format

A named matrix where rows refer to yeat genes, columns to FunCat classes. Names of yeast genes are systematic names. Names of columns correspond to FunCat IDs.

Source

<http://mips.gsf.de/projects/funcat>

References

Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Guldener, U., Mannhaupt, G., Munsterkotter, M., and Mewes, H. (2004). The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Research, 32(18), 5539-5545.

Examples

```
data(Yeast.STRING.FunCat)
```

Index