Package 'PathNetData'

April 25, 2024

Title Experimental data for the PathNet package

Date 24 Sep 2012

Version 1.39.0

Author Bhaskar Dutta <bhaskar.dutta@gmail.com>, Anders Wallqvist <awallqvist@bhsai.org>, and Jaques Reifman <jreifman@bhsai.org>

Maintainer Ludwig Geistlinger <ludwig.geistlinger@gmail.com>

Depends R (>= 1.14.0)

Description This package contains the data employed in the vignette of the PathNet package. These data belong to the following publication: PathNet: A tool for pathway analysis using topological information. Dutta B, Wallqvist A, and Reifman J., Source Code for Biology and Medicine 2012 Sep 24;7(1):10.

License GPL-3

biocViews ExperimentData, PathwayInteractionDatabase

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

git_branch devel

git_last_commit 05e9d15

git_last_commit_date 2023-10-24

Repository Bioconductor 3.19

Date/Publication 2024-04-25

Contents

PathNetData-package *Experimental data for the PathNet package.*

Description

This package contains the experimental data used by the **PathNet** package.

List of datasets:

- `brain_regions`: Dataset containing the effects of Alzheimer's disease in six different brain regions.
- `disease_progression`: Expression profiling of brain hippocampi from 22 postmortem subjects with Alzheimer's disease at various stages of severity.
- `pathway`: Pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database
- `A`: Connectivity information among genes in the pooled pathway.

List of files in "extdata":

- '`brain_regions_data.txt`': text file for the brain_regions dataset
- '`disease_progression_data.txt`': : text file for the disease_progression dataset
- '`pathway_data.txt`': text file for the pathway dataset
- '`adjacency_data.txt`': text file for the A dataset

References

Dutta B, Wallqvist A, and Reifman J. PathNet: A tool for pathway analysis using topological information. Source Code for Biology and Medicine 2012 Sep 24;7(1):10.

A *Adjacency matrix*

Description

The connectivity information among genes in the pooled pathway is represented by the adjacency matrix A

Usage

```
data(A)
```

Details

Pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database available in November 2010 were downloaded as KEGG Markup Language files. Each of the 130 non-metabolic pathways present in the KEGG database were represented as directed graphs, where the nodes and edges of a graph were characterized by unique gene IDs and interactions in the pathway, respectively. KEGG interactions representing processes, such as phosphorylation, dephosphorylation, activation, inhibition, and repression, were accounted for by directed edges, whereas bidirectional edges were used to represent binding/association events. The complete mapping between edge directionality and KEGG protein interaction attributes is provided in the Supplementary File 1 of the Dutta et al. manuscript. All 130 pathways were combined to create a pooled pathway, and the R interface to boost graph library package from Bioconductor (http://www.bioconductor.org) was used to convert this information into the adjacency matrix A. The adjacency matrix is a non-symmetric square matrix, where the number of rows (and columns) represents the number of genes present in the pooled pathway. The diagonal elements of matrix A were set to zero to exclude self interactions.

Source

The data was downloaded from KEGG database.

References

- Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T. et al. (2008) KEGG for linking genomes to life and the environment. Nucleic Acids Res, 36, D480-484.
- Dutta B, Wallqvist A, and Reifman J. PathNet: A tool for pathway analysis using topological information. Source Code for Biology and Medicine 2012 Sep 24;7(1):10.

Examples

```
data(A)
A[100:110,100:110]
```

---

brain_regions *Direct evidence from brain regions dataset*

Description

This dataset examined the effect of Alzheimer's disease (AD) in six different brain regions: the entorhinal cortex (EC), hippocampus (HIP), middle temporal gyrus (MTG), posterior cingulate cortex (PC), superior frontal gyrus (SFG), and primary visual cortex (VCX). Microarray platform used was Affymetrix Human Genome U133 Plus 2.0 Array

Usage

```
data(brain_regions)
```

Details

The direct evidence, i.e., association of each gene with the disease, was calculated by comparing gene expression data in control patients with disease patients. Here, we used t-test to identify the significance of association (p-value) of each gene with the disease separately for each brain region. If multiple probes are present corresponding to a gene, the probe with minimum p-value was selected. The negative log10 transformed p-value of the significance of association was used as direct evidence.

Source

NCBI GEO database (GEO ID: GSE5281).

References

Liang, W.S., Dunckley, T., Beach, T.G., Grover, A., Mastroeni, D., Walker, D.G., Caselli, R.J., Kukull, W.A., McKeel, D., Morris, J.C. et al. (2007) Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. Physiol Genomics, 28, 311-322.

Examples

```
data(brain_regions)
head(brain_regions)
```

---

disease_progression *Direct evidence from disease progression dataset*

Description

Expression profiling of brain hippocampi from 22 postmortem subjects with Alzheimer's disease (AD) at various stages of severity. Seven, 8, and 7 subjects diagnosed with incipient, moderate, and severe AD respectively. Results provide insight into mechanisms underlying the early pathogenesis of AD.

Usage

```
data(disease_progression)
```

Details

The direct evidence, i.e., association of each gene with the disease, was calculated by comparing gene expression data in control patients with incipient, moderate, and severe AD, respectively. Here, we used t-test to identify the significance of association (p-value) of each gene with the disease. If multiple probes are present corresponding to a gene, the probe with minimum p-value was selected. The negative log10 transformed p-value of the significance of association was used as direct evidence.

Source

NCBI GEO database (GEO ID: GDS810).

References

Blalock, E.M., Geddes, J.W., Chen, K.C., Porter, N.M., Markesbery, W.R. and Landfield, P.W. (2004) Incipient Alzheimer's disease: microarray correlation analyses reveal major transcriptional and tumor suppressor responses. Proc Natl Acad Sci U S A, 101, 2173-2178.

Examples

```
data(disease_progression)
head(disease_progression)
```

---

pathway *Pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database*

Description

Pathways from the KEGG database

Usage

```
data(pathway)
```

Details

We used regulatory pathways from the KEGG database. KEGG Markup Language files containing pathway information were downloaded from KEGG server and they were converted and combined to a text file. Each row in the pathway data represents an edge in the pooled pathway. The first column is the row index. Second and third columns denote the gene IDs connected by an edge in the pathway. The fourth column contains the name of the pathway where the edge is present.

Source

www.kegg.com

References

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T. et al. (2008) KEGG for linking genomes to life and the environment. Nucleic Acids Res, 36, D480-484.

Examples

```
data(pathway)
head(pathway)
```

Index