Package 'DeMAND'

April 14, 2017

Title DeMAND

Version 1.4.0

Date 2015-07-13

Author Jung Hoon Woo <jw2853@columbia.edu>, Yishai Shimoni <ys2559@columbia.edu>

Maintainer Jung Hoon Woo <jw2853@columbia.edu>, Mariano Alvarez <reef103@gmail.com>

Description DEMAND predicts Drug MoA by interrogating a cell context specific regulatory network with a small number (N >= 6) of compound-induced gene expression signatures, to elucidate specific proteins whose interactions in the network is dysregulated by the compound.

Depends R (>= 2.14.0), KernSmooth, methods

License file LICENSE

biocViews SystemsBiology, NetworkEnrichment, GeneExpression, StatisticalMethod, Network

NeedsCompilation no

R topics documented:

---

bcellAnno   *Annotation for the expression data*

Description

Annotation information for the probes of the gene expression matrix

---

| bcellExp | *B cell expression data* |

Description

subest of a gene expression profiles from DLBCL cells treated by Geldanamycin and by DMSO as control.

| bcellNetwork | *B cell network* |

Description

A subset of a molecular interaction network of Bcell assembled by the ARACNe*Margolin2006* algorithm for protein-DN interactions and Bayesian method*Lefebvre2010* for protein- protein interactions.

| caseIndex | *Case sample index* |

Description

Column indices of the gene expression matrix for the samples treated by Geldanamycin.

| controlIndex | *Control sample index* |

Description

Column indices of the gene expression matrix for the samples treated by DMSO

---

demand-instance   *The DeMAND instance*

Description

This instance stores parameters and results of theDeMAND algorithm

Arguments

exp
: A N by M numeric matrix and the rows are N probes and the columns are M samples.

anno
: A N by 2 character matrix. The rows are probes but the order should be the same with the *demand* matrix. The first column can be anything (usually probe IDs) but the second column should includes Official Gene Symbol information for each probe.

network
: A K by 4 character matrix which contains K interactions. The 1st column and the 2nd column contain pairs of interacted genes. The 3rd and 4th columns indicate whether the interactions are pr otein-protein interaction (ppi) or protein-DNA (pdi) interaction. Column name should be as follows: c("Gene1", "Gene2", "ppi", "pdi")

moa
: A data frame contains DeMAND MoA predictions (e.g. Gene, p-value, adjusted p-values)

KLD
: A matrix containing the KL-divergence of the interactions that were analysed, the KL-divergence that was evaluated, and the p-value associated with the divergenece

---

demandClass   *The demand class constructor*

Description

This function generates demand class instances

Usage

```
demandClass(exp, anno, network, moa=NULL, KLD=NULL)
```

Arguments

exp
: A N-by-M numeric matrix where the rows represent N probes (or genes) and the columns represent M samples.

anno
: A N-by-2 character matrix where the rows represent probes or genes in the same order as the exp matrix. The first column must hold the probe id or gene name as appears in the exp matrix, and the second column should hold their corresponding names (e.g gene symbol) as appears in the network matrix

network   A K-by-L (L>1) character matrix containing K interactions. The 1st column
                 and the 2nd column contain the names of the interacting genes. If the following
                 columns include a column called "ppi" then genes connected by ppi will be used
                 without estimating the residuals in the Brown correction method. The rest of the
                 columns are used only to distinguish duplicates.

moa   Filled my the `runDeMAND` function. A data frame containing DeMAND MoA
                 predictions (i.e. Gene, p-value, adjusted p-values)

KLD   Filled my the `runDeMAND` function. A matrix containing the interactions that
                 were analysed, their KL-divergence, and the p-value associated with the diver-
                 genece

Value

Instance of class demand

Examples

```
## Load toy example
data(inputExample)
dobj <- demandClass(exp=bcellExp, anno=bcellAnno, network=bcellNetwork)
printDeMAND(dobj)
```

---

printDeMAND   *Basic methods for class demand*

Description

This document lists a series of basic methods for the class DeMAND

Usage

```
printDeMAND(x)
```

Arguments

x   An instance of class demand which includes: a gene expression data, annotation
                 information, and a molecular interaction network.

Value

printDeMAND returnssummary information about the diggit object

Examples

```
data(inputExample)
dobj <- demandClass(exp=bcellExp, anno=bcellAnno, network=bcellNetwork)
printDeMAND(dobj)
```

---

runDeMAND *Run DeMAND algorithm*

Description

DeMAND.

This function is based on the realization that drugs affect the protein activity of their targets, but not necessarily their mRNA expression levels. In contrast, the change in protein activity directly affects the mRNA expression levels of downstream genes. Based on this hypothesis, DeMAND identifies drug MoA by comparing gene expression profiles following drug perturbation with control samples, and computing the change in the individual interactions within a pre-determined integrated transcriptional and post-translational regulatory model (interactome).

Usage

```
runDeMAND(x, fgIndex, bgIndex, verbose=TRUE, method="bandwidth",
          keepLeaves=FALSE, alpha=0.05)
```

Arguments

x	An instance of class demand which includes: a gene expression data, annotation information, and a molecular interaction network.
fgIndex	A numeric vector contains indices of columns which represent case samples (e.g. drug treated). The sample size should be greater than 3.
bgIndex	A numeric vector contains indices of columns which represent control samples (e.g. drug treated). The sample size should be greater than 3.
verbose	A boolean value (TRUE by defalut) indicating whether to print progression outputs
method	A string value indicating whether to evaluate the KL-divergence using grid points based on `'bandwidth'` (default) or on the `'integers'` space
keepLeaves	A boolean value indicating whether to return a p-value for genes the have only a single neighbor in the network (default is FALSE, which returns 1 for such genes)
alpha	The cutoff for estimating a p-value using pareto fitting (default=0.05)

Details

For each edge in the interactome we determine the two-dimensional probability distribution of the gene expression levels both in the control state, and following drug treatment. Any changes in the probability distribution are estimated using the Kullback-Leibler (KL) divergence, from which we determine the statistical significance of the dysregulation of each edge. In the second step of DeMAND, we interrogate each gene independently to determine whether its interactions are enriched in dysregulated ones, suggesting that it is a candidate mechanism of action.

Value

Returns a DeMAND class object holding the same exp, anno, and network slots as the input, and where the moa and KLD slots hold the results of DeMAND algorithm. The moa slot is a matrix containing a list of genes, corresponding p-value, and adjusted p-value. The KLD slot is a matrix with the gene names of the edges that were evaluated, their KL-divergence, and the p-value assigned to that divergence.

Examples

```
## Load toy example
data(inputExample)
dobj <- demandClass(exp=bcellExp, anno=bcellAnno, network=bcellNetwork, moa=NULL)
dobj <- runDeMAND(dobj, fgIndex=caseIndex, bgIndex=controlIndex)
## results (head)
printDeMAND(dobj)
## results (all)
print(dobj@moa)
print(dobj@KLD)
```

Index