**Package ‘pathifier’**

January 30, 2017

**Type** Package  

**Title** Quantify deregulation of pathways in cancer  

**Version** 1.12.0  

**Date** 2013-06-27  

**Author** Yotam Drier  

**Maintainer** Assif Yitzhaky <assif.yitzhaky@weizmann.ac.il>  

**Description** Pathifier is an algorithm that infers pathway deregulation scores for each tumor sample on the basis of expression data. This score is determined, in a context-specific manner, for every particular dataset and type of cancer that is being investigated. The algorithm transforms gene-level information into pathway-level information, generating a compact and biologically relevant representation of each sample.  

**License** Artistic-1.0  

**Imports** R.oo, princurve  

**biocViews** Network  

**NeedsCompilation** no

### R topics documented:

<table>
<thead>
<tr>
<th>pathifier-package</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>KEGG</td>
<td>2</td>
</tr>
<tr>
<td>quantify_pathways_deregulation</td>
<td>3</td>
</tr>
<tr>
<td>Sheffer</td>
<td>4</td>
</tr>
</tbody>
</table>

**Index**  

| pathifier-package | Quantify deregulation of pathways in cancer |

**Description**  

Pathifier is an algorithm that infers pathway deregulation scores for each tumor sample on the basis of expression data. This score is determined, in a context-specific manner, for every particular dataset and type of cancer that is being investigated. The algorithm transforms gene-level information into pathway-level information, generating a compact and biologically relevant representation of each sample.
Details

Package: pathifier
Type: Package
Version: 1.0
Date: 2013-03-15
License: Artistic-1.0

Author(s)
Yotam Drier <drier.yotam@mgh.harvard.edu> Maintainer: Assif Yitzhaky <assif.yitzhaky@weizmann.ac.il>

References
See more information on: http://www.weizmann.ac.il/pathifier/

Examples
data(KEGG) # Two pathways of the KEGG database
data(Sheffer) # The colorectal data of Sheffer et al.
PDS<-quantify_pathways_deregulation(sheffer$data, sheffer$allgenes,
  kegg$gs, kegg$pathwaynames, shefferNormals, attempts = 100,
  logfile="sheffer.kegg.log", min_exp=sheffer$minexp, min_std=sheffer$minstd)

Description
Two pathways (MISMATCH REPAIR and REGULATION OF AUTOPHAGY) of the KEGG database

Usage
data(KEGG)

Format
pathwaynames  The names of the pathways
gs  The list of genes (by official gene symbol) in each pathway

Source

Examples
data(KEGG)
Description
Pathifier is an algorithm that infers pathway deregulation scores for each tumor sample on the basis of expression data. This score is determined, in a context-specific manner, for every particular dataset and type of cancer that is being investigated. The algorithm transforms gene-level information into pathway-level information, generating a compact and biologically relevant representation of each sample.

Usage
`quantify_pathways_deregulation(data, allgenes, syms, pathwaynames, normals = NULL, ranks = NULL, attempts = 100, maximize_stability = TRUE, logfile = "", samplings = NULL, min_exp = 4, min_std = 0.4)`

Arguments
- **data**: The n x m mRNA expression matrix, where n is the number of genes and m the number of samples.
- **allgenes**: A list of n identifiers of genes.
- **syms**: A list of p pathways, each pathway is a list of the genes it contains (as appear in "allgenes").
- **pathwaynames**: The names of the p pathways.
- **normals**: A list of m logicals, true if a normal sample, false if tumor.
- **ranks**: External knowledge on the ranking of the m samples, if exists (to use initial guess).
- **attempts**: Number of runs to determine stability.
- **maximize_stability**: If true, throw away components leading to low stability of sampling noise.
- **logfile**: Name of the file the log should be written to (use stdout if empty).
- **samplings**: A matrix specifying the samples that should be chosen in each sampling attempt, chooses a random matrix if samplings is NULL.
- **min_exp**: The minimal expression considered as a real signal. Any values below are thresholded to be min_exp.
- **min_std**: The minimal allowed standard deviation of each gene. Genes with lower standard deviation are divided by min_std instead of their actual standard deviation. (Recommended: set min_std to be the technical noise).

Value
- **scores**: The deregulation scores, the main output of pathifier.
- **genesinpathway**: The genes of each pathway used to devise its deregulation score.
- **newmeanstd**: Average standard deviation after omitting noisy components.
- **origmeanstd**: Original average standard deviation, before omitting noisy components.
pathwaysize  The number of components used to devise the pathway score
curves      The principal curve learned for every pathway
curves_order The order of the points of the principal curve learned for every pathway
z           Z-scores of the expression matrix used to learn principal curve
comin       The components not omitted due to noise
xm          The average expression over all normal samples
xs          The standard deviation of expression over all normal samples
center      The centering used by the PCA
rot         The matrix of variable loadings of the PCA
ptaken      The number of principal components used
samplings   A matrix specifying the samples that should be chosen in each sampling attempt
success     Pathways for which a deregulation score was successfully computed
logfile     Name of the file the log was written to

Author(s)
Yotam Drier <drier.yotam@mgh.harvard.edu> Maintainer: Assif Yitzhaky <assif.yitzhaky@weizmann.ac.il>

References

See more information on : http://www.weizmann.ac.il/pathifier/

Examples
data(KEGG)  # Two pathways of the KEGG database
data(Sheffer) # The colorectal data of Sheffer et al.
PDS<-quantify_pathways_deregulation(sheffer$data, sheffer$allgenes, kegg$gs, kegg$pathwaynames, sheffer$normals, attempts = 100, logfile="sheffer.kegg.log", min_exp=sheffer$minexp, min_std=sheffer$minstd)

Ssheffer

Sheffer et al. colorectal dataset

Description
Partial data from Sheffer et al. paper

Usage
data(Sheffer)
**Format**

- `data` the expression data
- `samples` sample names
- `normals` which of the samples is a normal sample
- `minstd` minimal standard deviation allowed
- `minexp` minimal value of expression allowed
- `allgenes` the list of genes (by official gene symbol)

**Source**


**Examples**

- `data(Sheffer)`
Index

∗Topic datasets
   KEGG, 2
   Sheffer, 4
∗Topic package
   pathifier-package, 1

KEGG, 2

pathifier (pathifier-package), 1
pathifier-package, 1

quantify_pathways_deregulation, 3
Sheffer, 4