

# Package ‘pathifier’

October 12, 2016

**Type** Package

**Title** Quantify deregulation of pathways in cancer

**Version** 1.10.0

**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:

pathifier-package . . . . .	2
KEGG . . . . .	3
quantify_pathways_deregulation . . . . .	3
Sheffer . . . . .	5

<b>Index</b>	<b>6</b>
--------------	----------

---

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

Drier Y, Sheffer M, Domany E. Pathway-based personalized analysis of cancer. *Proceedings of the National Academy of Sciences*, 2013, vol. 110(16) pp:6388-6393. ([www.pnas.org/cgi/doi/10.1073/pnas.1219651110](http://www.pnas.org/cgi/doi/10.1073/pnas.1219651110))

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)
```

---

KEGG

*Two pathways of the KEGG database*

---

**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**

Kanehisa M, Goto S, Sato Y, Furumichi M and Tanabe M. KEGG for integration and interpretation of large-scale molecular datasets. *Nucleic Acids Res*, 2012, Vol 40(Database issue):D109-D114.

**Examples**

```
data(KEGG)
```

---

```
quantify_pathways_deregulation
```

*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.

**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**

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

**Value**

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

Drier Y, Sheffer M, Domany E. Pathway-based personalized analysis of cancer. *Proceedings of the National Academy of Sciences*, 2013, vol. 110(16) pp:6388-6393. ([www.pnas.org/cgi/doi/10.1073/pnas.1219651110](http://www.pnas.org/cgi/doi/10.1073/pnas.1219651110))

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)
```

---

Sheffer

*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 standart deviation allowed
minexp minimal value of experssion allowed
allgenes the list of genes (by official gene symbol)
```

**Source**

Sheffer et.\ al. Association of survival and disease progression with chromosomal instability: A genomic exploration of colorectal cancer. *PNAS*, 2009, Vol 106(17) pp: 7131-7136.

**Examples**

```
data(Sheffer)
```

# Index

\*Topic **datasets**

KEGG, [3](#)

Sheffer, [5](#)

\*Topic **package**

pathifier-package, [2](#)

KEGG, [3](#)

pathifier (pathifier-package), [2](#)

pathifier-package, [2](#)

quantify\_pathways\_deregulation, [3](#)

Sheffer, [5](#)