

# Package ‘pathifier’

November 17, 2025

**Type** Package

**Title** Quantify deregulation of pathways in cancer

**Version** 1.49.0

**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 (>= 2.0.4)

**biocViews** Network

**git\_url** <https://git.bioconductor.org/packages/pathifier>

**git\_branch** devel

**git\_last\_commit** f57c93e

**git\_last\_commit\_date** 2025-10-29

**Repository** Bioconductor 3.23

**Date/Publication** 2025-11-16

## Contents

|                                          |   |
|------------------------------------------|---|
| 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**

|                    |                                                                                                                                                                                                                   |
|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 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                |
| compin         | 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                                     |
| pctaken        | The number of principal components used                                        |
| samplings      | A matrix specifying the samples that should be chosen in each sampling attempt |
| sucess         | 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**

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

- \* **datasets**

- KEGG, [3](#)

- Sheffer, [5](#)

- \* **package**

- pathifier-package, [2](#)

KEGG, [3](#)

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

pathifier-package, [2](#)

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

Sheffer, [5](#)