

# Package ‘iPath’

March 24, 2023

**Type** Package

**Title** iPath pipeline for detecting perturbed pathways at individual level

**Version** 1.5.0

**Description** iPath is the Bioconductor package used for calculating personalized pathway score and test the association with survival outcomes. Abundant single-gene biomarkers have been identified and used in the clinics. However, hundreds of oncogenes or tumor-suppressor genes are involved during the process of tumorigenesis. We believe individual-level expression patterns of pre-defined pathways or gene sets are better biomarkers than single genes. In this study, we devised a computational method named iPath to identify prognostic biomarker pathways, one sample at a time. To test its utility, we conducted a pan-cancer analysis across 14 cancer types from The Cancer Genome Atlas and demonstrated that iPath is capable of identifying highly predictive biomarkers for clinical outcomes, including overall survival, tumor subtypes, and tumor stage classifications. We found that pathway-based biomarkers are more robust and effective than single genes.

**License** GPL-2

**Encoding** UTF-8

**Suggests** rmarkdown, BiocStyle, knitr

**VignetteBuilder** knitr

**Imports** Rcpp (>= 1.0.5), matrixStats, ggpubr, ggplot2, survminer, stats

**biocViews** Pathways, Software, GeneExpression, Survival

**NeedsCompilation** yes

**SystemRequirements** C++11

**LinkingTo** Rcpp, RcppArmadillo

**Depends** R (>= 4.1), mclust, BiocParallel, survival

**RoxygenNote** 7.1.1

**BugReports** <https://github.com/suke18/iPath/issues>

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

**git\_branch** devel

**git\_last\_commit** 7f5a990

**git\_last\_commit\_date** 2022-11-01

**Date/Publication** 2023-03-24

**Author** Kenong Su [aut, cre],  
Zhaohui Qin [aut]

**Maintainer** Kenong Su <kenong.su@emory.edu>

## R topics documented:

density_fall . . . . .	2
GSDB_example . . . . .	3
GSEA . . . . .	4
iES_cal2 . . . . .	4
iES_surv . . . . .	5
iES_survPlot . . . . .	5
prad_cli . . . . .	6
prad_exprs . . . . .	7
prad_inds . . . . .	7
rem_data . . . . .	8
setUp_BPPARAM . . . . .	8
water_fall . . . . .	9

<b>Index</b>	<b>10</b>
--------------	-----------

---

density_fall	<i>density fall plot</i>
--------------	--------------------------

---

### Description

This function allows you to express your love of cats.

### Usage

```
density_fall(iES_mat, gs_str, indVec, title = TRUE)
```

### Arguments

iES_mat, gs_str	is the iES_mat with tumor and normal and gs name.
indVec	the binary indicator for normal(0) and tumor (1) patients.
title	boolean true or false for including the title in the ggplot.

**Value**

ggplot object containing the KM plot.

**Examples**

```
data(PRAD_data)
data(GSDB_example)
iES_mat = iES_cal2(prad_exprs, GSDB = GSDB_example)
density_fall(iES_mat, gs_str = "SimPathway1", indVec = prad_inds)
```

---

GSDB_example	<i>example gene set database (GSDB)</i>
--------------	-----------------------------------------

---

**Description**

includes geneset.names, genesets.

**Usage**

```
data("GSDB_example")
```

**Format**

A list of gene set database

**Source**

<https://www.gsea-msigdb.org/gsea/msigdb/>

**References**

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3106198/>

**Examples**

```
data("GSDB_example")
GSDB_example$geneset.names
```

---

GSEA *GSEA calculation*

---

**Description**

This function calculates the GSEA enrichment score.

**Usage**

```
GSEA(gene_list, gene_set, stats_vector)
```

**Arguments**

gene\_list is a list of genes.  
 gene\_set is a set of genes.  
 stats\_vector a vector quantify the level of genes in the gene list.

**Value**

the original GSEA score.

---

iES\_cal2 *iES calculation Function*

---

**Description**

This function calculates the iES matrix which is the core of iPath.

**Usage**

```
iES_cal2(Y, GSDB, BPPARAM = NULL, nPro = 0)
```

**Arguments**

Y is the expression matrix.  
 GSDB is the gene set database.  
 BPPARAM parameters from the BiocParallel.  
 nPro number of processors (default = 0).

**Value**

a matrix with rows corresponding to the pathways and columns corresponding to the patients.

**Examples**

```
data(PRAD_data)
data(GSDB_example)
iES_mat = iES_cal2(prad_exprs, GSDB = GSDB_example)
```

---

iES_surv	<i>iES calculation Function</i>
----------	---------------------------------

---

**Description**

This function allows to investigate on one specific pathway.

**Usage**

```
iES_surv(iES_mat, cli, indVec = NULL, npatsThre = 5)
```

**Arguments**

iES_mat	is iES matrix with rows corresponding to the pathway and columns corresponding to the patients.
cli	clinical data associated to the gene expression data.
indVec	binary vector indicating normal (0) and tumor (1).
npatsThre	the threshold of number of patients for survival analysis.

**Value**

a matrix of survival analysis from coxph.

**Examples**

```
data(PRAD_data)
data(GSDB_example)
iES_mat = iES_cal2(prad_exprs, GSDB = GSDB_example)
iES_surv(iES_mat, cli = prad_cli, indVec = prad_inds)
```

---

iES_survPlot	<i>iES survival for a certain pathway</i>
--------------	-------------------------------------------

---

**Description**

This function allows you to express your love of cats.

**Usage**

```
iES_survPlot(iES_mat, cli, gs_str, indVec = NULL, npatsThre = 5, title = TRUE)
```

**Arguments**

`iES_mat`, `gs_str` is the GSDB `iES_mat` with tumor and normal and `gs` name.

`cli` clinical data corresponding to the expression data.

`indVec` the binary indicator for normal(0) and tumor (1) patients.

`npatsThre` the threshold of number of patients for survival analysis.

`title` boolean true or false for including the title (`gs_str`) in the `ggplot`.

**Value**

`ggplot` object containing the KM plot.

**Examples**

```
data(PRAD_data)
data(GSDB_example)
iES_mat = iES_cal2(prad_exprs, GSDB = GSDB_example)
iES_survPlot(iES_mat, cli = prad_cli, gs_str = "SimPathway1", indVec = prad_inds)
```

---

<code>prad_cli</code>	<i>simulated clinical data for PRAD cancer patients</i>
-----------------------	---------------------------------------------------------

---

**Description**

`prad_cli` is the clinical data containing three variables `times`, `bcr_patient_barcode`, and `patient.vital_status`.

**Usage**

```
data("PRAD_data")
```

**Format**

An object of "matrix" class contains the clinical outcomes

**Source**

<https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>

**References**

Kosinski M, Biecek P (2021). RTCGA: The Cancer Genome Atlas Data Integration. R package version 1.22.0, <https://rtcga.github.io/RTCGA>.

**Examples**

```
data("PRAD_data")
prad_cli[1:10,]
```

---

prad_exprs	<i>expression matrix for PRAD cancer patients in TCGA</i>
------------	-----------------------------------------------------------

---

**Description**

prad\_exprs is the RPKM expression matrix which belongs to "matrix" class. The data includes 102 samples about human preimplantation embryos and embryonic stem cells. It contains 19304 genes after removing genes with extreme high dropout rate.

**Usage**

```
data("PRAD_data")
```

**Format**

An object of "matrix" class contains the mRNA expressions

**Source**

<https://www.bioconductor.org/packages/release/bioc/html/RTCGA.html>

**References**

Kosinski M, Biecek P (2021). RTCGA: The Cancer Genome Atlas Data Integration. R package version 1.22.0, <https://rtcga.github.io/RTCGA>.

**Examples**

```
data("PRAD_data")
prad_exprs[1:10, 1:4]
```

---

prad_inds	<i>normal (0) and tumor (1) classes associated with PRAD expression data</i>
-----------	------------------------------------------------------------------------------

---

**Description**

normal (0) and tumor (1) classes associated with PRAD expression data.

**Usage**

```
data("PRAD_data")
```

**Format**

A character vector contains the class label

**Source**

<https://www.bioconductor.org/packages/release/bioc/html/RTCGA.html>

**References**

Kosinski M, Biecek P (2021). RTCGA: The Cancer Genome Atlas Data Integration. R package version 1.22.0, <https://rtega.github.io/RTCGA>.

**Examples**

```
data("PRAD_data")
table(prad_inds)
```

---

rem_data	<i>remove genes with 0 sd</i>
----------	-------------------------------

---

**Description**

This function helps remove non-informative genes.

**Usage**

```
rem_data(Y)
```

**Arguments**

Y is the expression matrix.

**Value**

a processed matrix

---

setUp_BPPARAM	<i>set up for the parallel computing for biocParallel.</i>
---------------	------------------------------------------------------------

---

**Description**

This function sets up the environment for parallel computing.

**Usage**

```
setUp_BPPARAM(nproc = 0, BPPARAM = NULL)
```

**Arguments**

nproc	number of processors
BPPARAM	bpparameter from bpparam



**Value**

BAPPARAM settings

---

water_fall	<i>water fall plot</i>
------------	------------------------

---

**Description**

This function allows you to express your love of cats.

**Usage**

```
water_fall(iES_mat, gs_str, indVec, title = TRUE)
```

**Arguments**

iES_mat, gs_str	is the iES_mat with tumor and normal and gs name.
indVec	the binary indicator for normal(0) and tumor (1) patients.
title	boolean true or false for including the title (gs_str) in the ggplot.

**Value**

ggplot object containing the KM plot.

**Examples**

```
data(PRAD_data)
data(GSDB_example)
iES_mat = iES_cal2(prad_exprs, GSDB = GSDB_example)
water_fall(iES_mat, gs_str = "SimPathway1", indVec = prad_inds)
```

# Index

- \* **analysis**
  - iES\_surv, 5
- \* **and**
  - density\_fall, 2
  - iES\_surv, 5
  - iES\_survPlot, 5
  - water\_fall, 9
- \* **calclatioin.**
  - iES\_cal2, 4
- \* **datasets**
  - GSDB\_example, 3
  - prad\_cli, 6
  - prad\_exprs, 7
  - prad\_inds, 7
- \* **densityfall**
  - density\_fall, 2
  - iES\_survPlot, 5
- \* **for**
  - density\_fall, 2
  - iES\_surv, 5
  - iES\_survPlot, 5
  - water\_fall, 9
- \* **groups**
  - iES\_surv, 5
- \* **iES**
  - iES\_cal2, 4
- \* **iPath**
  - iES\_surv, 5
- \* **normal-like.**
  - iES\_surv, 5
- \* **normal**
  - density\_fall, 2
  - iES\_survPlot, 5
  - water\_fall, 9
- \* **of**
  - iES\_surv, 5
- \* **patients:**
  - iES\_surv, 5
- \* **perturbed**
  - iES\_surv, 5
- \* **plot**
  - density\_fall, 2
  - iES\_survPlot, 5
  - water\_fall, 9
- \* **sample.**
  - density\_fall, 2
- \* **sample**
  - iES\_survPlot, 5
  - water\_fall, 9
- \* **statistics**
  - iES\_cal2, 4
- \* **survival**
  - iES\_surv, 5
- \* **tumor**
  - density\_fall, 2
  - iES\_survPlot, 5
  - water\_fall, 9
- \* **two**
  - iES\_surv, 5
- \* **waterfall**
  - water\_fall, 9
- density\_fall, 2
- GSDB\_example, 3
- GSEA, 4
- iES\_cal2, 4
- iES\_surv, 5
- iES\_survPlot, 5
- prad\_cli, 6
- prad\_exprs, 7
- prad\_inds, 7
- rem\_data, 8
- setUp\_BPPARAM, 8
- water\_fall, 9