

# Package ‘ADVOCATE’

June 25, 2018

**Type** Package

**Title** The ADVOCATE algorithm uses a Machine Learning based algorithm to combine weak evidence derived from the expression of individual genes into a model that provides an optimal estimate of the compartment-specific composition of a heterogeneous bulk tissue.

**Version** 0.1.0

**Author** Jing He, Pasquale Laise and Andrea Califano.

**Maintainer** Pasquale Laise <pl2659@cumc.columbia.edu>

**Description**

**License** file LICENSE

**LazyData** TRUE

**RoxygenNote** 6.0.1

**NeedsCompilation** no

## R topics documented:

calCellTypeExpression . . . . .	1
calCellTypeExpression_3comp . . . . .	2
calGeneWiseProbability . . . . .	3
hedgeBeta_deg . . . . .	3
predict_bulk . . . . .	4
predict_bulk_3comp . . . . .	5
<b>Index</b>	<b>6</b>

---

calCellTypeExpression	<i>Inferring Epithelium, Stromal virtual Expression of give bulk PDA samples Expression</i>
-----------------------	---

---

## Description

calCellTypeExpression uses LCM compartment specific expression from experiments, output of predict\_bulk to infer virtual epithelium/stromal compartment specific gene expression.

## Usage

```
calCellTypeExpression(lcmexp, deg, fc, pval, sampleInfo, inputexpmat, fracAll,  
  method = "lcm")
```

**Arguments**

lcmexp	LCM sample expression
deg,	result from hedgeBeta_deg
fc,	output from hedgeBeta_deg
pval,	output from hedgeBeta_deg
sampleInfo,	output from hedgeBeta_deg
inputexpmat	input bulk PDA sample normalized expression (VST and Z transformed)
fracAll	result from predict_bulk
method	method for inference, currently support 'lcm'

**Value**

a matrix of n\_gene \* (nsample \* 2)

**Author(s)**

J.HE

---

calCellTypeExpression\_3comp

*Inferring Epithelium, Stromal virtual Expression of give bulk PDA samples Expression*

---

**Description**

calCellTypeExpression uses LCM compartment specific expression from experiments, output of predict\_bulk\_3comp to infer virtual epithelium/stromal compartment specific gene expression.

**Usage**

```
calCellTypeExpression_3comp(lcmexp, deg, fc, pval, sampleInfo, inputexpmat,
  fracAll, method = "lcm")
```

**Arguments**

lcmexp	LCM sample expression
deg,	result from hedgeBeta_deg
fc,	output from hedgeBeta_deg
pval,	output from hedgeBeta_deg
sampleInfo,	output from hedgeBeta_deg
inputexpmat	input bulk PDA sample normalized expression (VST and Z transformed)
fracAll	result from predict_bulk_3comp
method	method for inference, currently support 'lcm'

**Value**

a matrix of n\_gene \* (nsample \* 2)

**Author(s)**

J.HE

---

`calGeneWiseProbability`*Calculate each genes probability of coming from Epithelium and Stromal compartment*

---

**Description**

`calGeneWiseProbability` uses LCM gene expression from epithelium and stromal to calculate the bulk sample genes probability of epithelium/stromal compartment

**Usage**

```
calGeneWiseProbability(trainfile, bulkexp, n_comp = 2, epsilon = 0.01,
  betaFunc = "pval")
```

**Arguments**

<code>trainfile</code>	train file generated using <code>hedgeBeta_deg</code> function.
<code>bulkexp</code>	input bulk PDA sample normalized expression (VST and Z transformed)
<code>n_comp</code>	number of compartments, default 2
<code>epsilon</code>	accuracy of estimation, default 0.01
<code>betaFunc</code>	weighting metric for learning rate, default p-value

**Value**

a matrix of `number_gene * (number_sample*2)`

**Author(s)**

J.HE

---

`hedgeBeta_deg`*Calculate DEG and its weighting metric for ADVOCATE training*

---

**Description**

`hedgeBeta_deg` uses raw LCM samples from epi and stroma to calculate DEG and the weight metric for ADVOCATE model training.

**Usage**

```
hedgeBeta_deg(rawcounts, idxE, idxS, outfile = "input_advocate_train.rda")
```

**Arguments**

rawcounts	raw counts matrix from LCM RNAseq
idxE	group1 (Epithelium) sample name
idxS	group2 (stromal) sample name
outfile	rda file name to store all metrics needed for training

**Author(s)**

J.HE

---

predict_bulk	<i>Predict Epithelium and Stromal fraction of give bulk PDA samples expression</i>
--------------	--

---

**Description**

predict\_bulk uses probabilistic model to calculate the fraction of epithelium/stromal compartment using gene expression.

**Usage**

```
predict_bulk(trainfile, bulkexp, epsilon = 0.01, betaFunc = "pval")
```

**Arguments**

trainfile	train file generated using hedgeBeta_deg function.
bulkexp	input bulk PDA sample normalized expression (VST and Z transformed)
epsilon	accuracy of estimation, default 0.01
betaFunc	weighting metric for learning rate, default p-value

**Value**

a matrix of number\_inputbulkexp \* 7, first 2 columns are the epithelium and stromal fractions

**Author(s)**

J.HE

---

predict_bulk_3comp	<i>Predict Epithelium, Stromal and Other compartment fractions of give bulk PDA samples expression</i>
--------------------	--

---

**Description**

predict\_bulk\_3comp uses probabilistic model to calculate the fraction of epithelium/stromal/other compartment using gene expression.

**Usage**

```
predict_bulk_3comp(trainfile, bulkexp, epsilon = 0.01, betaFunc = "pval",  
  type = "p3l2fast")
```

**Arguments**

trainfile	train file generated using hedgeBeta_deg function.
bulkexp	input bulk PDA sample normalized expression (VST and Z transformed)
epsilon	accurray of estimation, default 0.01
betaFunc	weighting metric for learning rate, defaul p-value

**Value**

a matrix of number\_inputbulkexp \* 7, first 3 columns are the epithelium, stromal and other fractions

**Author(s)**

J.HE

# Index

calCellTypeExpression, [1](#)  
calCellTypeExpression\_3comp, [2](#)  
calGeneWiseProbability, [3](#)  
  
hedgeBeta\_deg, [3](#)  
  
predict\_bulk, [4](#)  
predict\_bulk\_3comp, [5](#)