Package ‘BPM’

January 20, 2025

Type Package

Title Bayesian Purity Model to Estimate Tumor Purity

Version 1.0.0

Author Jianzhao Gao, Linghao Shen, Xiaodan Fan

Maintainer Jianzhao Gao and Xiaodan Fan <`gaojz@nankai.edu.cn`>

Description Bayesian purity model to estimate tumor purity using methylation
array data (DNA methylation Infinium 450K array data) without reference samples.

Depends R (>= 2.10)

Imports stats, limma

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 6.1.0

NeedsCompilation no

Repository CRAN

Date/Publication 2018-11-03 16:10:03 UTC

Contents

`annotGeneNames` *gene names of probes in 450K array dat*

Description

gene names of probes in 450K array dat

Format

A vector with length 480457

`ApiGetDMCs` *Get TOPK=500 DMCs and non-DMCs using moderated-t test*

Description

Get TOPK=500 DMCs and non-DMCs using moderated-t test

Usage

```
ApiGetDMCs(betaValue, TOPK = 500, tumorNum = NULL,
  filterProbes = FALSE, userProbes = NULL)
```

Arguments

`betaValue`	A matrix from TCGA array data
`TOPK`	An integer number, default 500. Number of DMCs/non-DMCs.
`tumorNum`	A postive number, First tumorNum columns in betaValue are tumor samples. If tumorNum is NULL, first half of columns are considered as tumor samples,
`filterProbes`	Logistic. defalut is FALSE. The code use all probes in betaValue. If TRUE, you can use default good probes provided in our code. you can also provide your good probes in userProbes.
`userProbes`	A number list. The row numbers in betaValue. These rows are considered as good probes. return DMCs (TOPK DMCs and TOPK non-DMCs row index in betaValue)

Note

User can provide the good probes indexes (row number) to filter the probes. A global variable goodProbes are used in this function. goodProbes: probes with SNPs at the CpG or single base extension sites, and corss-reative probes are removed. More details see the reference paper.

BayPM — *Bayesian Purity Model (BPM) Main functions.*

Description

Bayesian Purity Model (BPM) Main functions.

Usage

```
BayPM(betaValue, TOPK = 500, tumorNum = NULL, filterProbes = FALSE,
  userProbes = NULL)
```

Arguments

`betaValue`	A matrix,TCGA methlation array data. Each row: loci, Tumor1,Tumor2,...,Normal1,Nomral2,...
`TOPK`	A number. Number of DMCs/nonDMCs selected
`tumorNum`	The number of tumor samples. if NULL, the default number is half of column number of dataset.
`filterProbes`	Logistic. defalut is FALSE. The code use all probes in betaValue. If TRUE, you can use default good probes provided in our code. you can also provide your good probes in userProbes.
`userProbes`	A number list. The row numbers in betaValue. These rows are considered as good probes.

Value

tumor purity estimation of tumor samples

Examples

```
### need to install package "limma"
### source("https://bioconductor.org/biocLite.R");biocLite("limma");
BayPM(simUCEC,20,2);
```

BPM — *BPM software package*

Description

Bayesian model for purity estimation using DNA methylation data

Details

The main function is `BayPM`

Author(s)

Jianzhao Gao(gaojz@nankai.edu.cn), Linghao Shen Xiaodan Fan (xfan@cuhk.edu.hk)

References

Jianzhao Gao, Linghao Shen, and Xiaodan Fan, Bayesian model for purity estimation using DNA methylation data.(submitted)

Examples

```
### need to install package "limma"
### source("https://bioconductor.org/biocLite.R");biocLite("limma");
library(BPM);
BayPM(simUCEC,20,2);
```

`estimateNu` *Estimate noise intensity (nv) for non-DMCs, using maximum likelihood estmiation.*

Description

Estimate noise intensity (nv) for non-DMCs, using maximum likelihood estmiation.

Usage

```
estimateNu(z, phi, maxit = 50, beginP = 20)
```

Arguments

`z`	A matrix. Observated mixed turmor samples.
`phi`	mode of beta-values of each row in pure nomral samples y.
`maxit`	A postive integer. The iteration number used in maximum likelihood.
`beginP`	A number, where the method start to search from for root. return estimated nv (noise intensity)

`fullSampler` *Sampling xi and alpha (tumor purity)*

Description

Sampling xi and alpha (tumor purity)

Usage

```
fullSampler(y, z, mstates, xprior = NULL, maxit = 1000,
  burnin = maxit, xpar = FALSE, n_ab0 = NULL, alp0 = NULL,
  xbar0 = NULL, trace = FALSE, verbose = FALSE)
```

Arguments

`y`	A matrix, observed pure normal samples
`z`	A matrix, observed mixed tumor samples
`mstates`	A matrix, hyper/hypo of dataset
`xprior`	A matrix, prior knowledge about purity
`maxit`	A number, maximum iteraction
`burnin`	A number, "burn-in" sample
`xpar`	Logistic, default is FALSE
`n_ab0`	initial value of n_ab
`alp0`	initial value of alpha
`xbar0`	initial value of xbar
`trace`	Logisitc, check the values in code, default is FALSE
`verbose`	Logistic, output the message,default is FALSE

Value

x_bar x_mode, x_last x2 x_sample x_sample xpar xprior2, nab n_ab2, alp alp2

`goodProbes` *good probes in packages*

Description

good probes removed Y chrome.

Format

A vector with length 425698

`simUCEC` *Simulated data to illustrate datasets in packages*

Description

A dataset containing 100 gene and 4 smaples, first two columns are tumor1 tumor2 last two columns are normal1 normal2

- x. the genes
- y. two tumor samples; two normal samples;

Format

A matrix with 100 rows and 4 columns

Index