

Package ‘InterSIM’

July 16, 2018

Type Package

Title Simulation of Inter-Related Genomic Datasets

Version 2.2.0

Date 2018-07-13

Author Prabhakar Chalise, Rama Raghavan, Brooke Fridley

Maintainer Prabhakar Chalise <pchalise@kumc.edu>

Description Generates three inter-related genomic datasets : methylation, gene expression and protein expression.

License GPL

Depends R (>= 3.5.0), MASS, NMF, tools

NeedsCompilation no

Repository CRAN

RoxygenNote 6.0.1

Date/Publication 2018-07-16 17:40:16 UTC

R topics documented:

InterSIM-package	1
InterSIM	2

Index	4
--------------	----------

InterSIM-package	<i>Simulation of inter-related genomic datasets</i>
------------------	---

Description

InterSIM is an R package that generates three inter-related data set with realistic inter- and intra-relationships based on the DNA methylation, mRNA expression and protein expression from the TCGA ovarian cancer study.

Details

Package: InterSIM
 Type: Package
 Version: 2.2.0
 Date: 2018-07-13
 License: GPL>=2

Author(s)

Prabhakar Chalise, Rama Raghavan, Brooke Fridley; Maintainer: Prabhakar Chalise

InterSIM

InterSIM

Description

This function simulates three inter-related genomic datasets : DNA methylation, gene expression and protein expression.

Usage

```
InterSIM(n.sample=500, cluster.sample.prop=c(0.30,0.30,0.40), delta.methyl=2.0,
delta.expr=2.0, delta.protein=2.0, p.DMP=0.2,
p.DEG=NULL, p.DEP=NULL, sigma.methyl=NULL, sigma.expr=NULL, sigma.protein=NULL,
cor.methyl.expr=NULL, cor.expr.protein=NULL, do.plot=FALSE, sample.cluster=TRUE,
feature.cluster=TRUE)
```

Arguments

n.sample	Number of subjects to simulate
cluster.sample.prop	Proportion of samples in the clusters. The number of proportions entered is used to determine the number of clusters in the simulated data. e.g. if (0.3,0.4,0.3) is entered then the number of clusters will be 3.
delta.methyl	Cluster mean shift for methylation data
delta.expr	Cluster mean shift for expression data
delta.protein	Cluster mean shift for protein data
p.DMP	proportion of DE CpGs (DE = Differentially Expressed)
p.DEG	proportion of DE mRNA, if NULL (default) mRNAs mapped by DE CpGs will be selected
p.DEP	proportion of DE protein, if NULL (default) proteins mapped by DE mRNAs will be selected

<code>sigma.methyl</code>	Covariance structure methylation data, if NULL (default) precomputed values will be used. "indep" gives covariance structure with diagonal elements only (Independent features)
<code>sigma.expr</code>	Covariance structure mRNA data, if NULL (default) precomputed values will be used. "indep" gives covariance structure with diagonal elements only (Independent features)
<code>sigma.protein</code>	Covariance structure Protein data, if NULL (default) precomputed values will be used. "indep" gives covariance structure with diagonal elements only (Independent features)
<code>do.plot</code>	TRUE to generate heatmap, default is FALSE
<code>sample.cluster</code>	TRUE (default), if clustering should be done on samples for heatmap. This option will be applicable only if <code>do.plot=TRUE</code> .
<code>feature.cluster</code>	TRUE (default), if clustering should be done on genomic features for heatmap. This option will be applicable only if <code>do.plot=TRUE</code> .
<code>cor.methyl.expr</code>	Correlation between methylation and mRNA, if NULL (default) precomputed values will be used
<code>cor.expr.protein</code>	Correlation between mRNA and protein, if NULL (default) precomputed values will be used

Value

This function returns three datasets as matrices - DNA methylation, gene expression and protein expression. It also returns a vector that has true cluster assignment for each subject in the generated data.

Author(s)

Prabhakar Chalise <pchalise@kumc.edu>, Rama Raghavan <rraghavan@kumc.edu>, Brooke Fridley <bfridley@kumc.edu>

Examples

```
#
prop <- c(0.20,0.30,0.27,0.23)
effect <- 5
sim.data <- InterSIM(n.sample=500, cluster.sample.prop = prop,
delta.methyl=effect, delta.expr=effect, delta.protein=effect,
p.DMP=0.2, p.DEG=NULL, p.DEP=NULL,
sigma.methyl=NULL, sigma.expr=NULL, sigma.protein=NULL,
cor.methyl.expr=NULL, cor.expr.protein=NULL,
do.plot=FALSE, sample.cluster=TRUE, feature.cluster=TRUE)
sim.methyl <- sim.data$dat.methyl
sim.expr <- sim.data$dat.expr
sim.protein <- sim.data$dat.protein
```

Index

`cov.expr` (InterSIM), [2](#)
`cov.M` (InterSIM), [2](#)
`cov.protein` (InterSIM), [2](#)
`CpG.gene.map.for.DEG` (InterSIM), [2](#)

InterSIM, [2](#)
InterSIM-package, [1](#)

`logit` (InterSIM), [2](#)

`mean.expr` (InterSIM), [2](#)
`mean.M` (InterSIM), [2](#)
`mean.protein` (InterSIM), [2](#)
`methyl.gene.level.mean` (InterSIM), [2](#)

`protein.gene.map.for.DEP` (InterSIM), [2](#)

`rev.logit` (InterSIM), [2](#)
`rho.expr.protein` (InterSIM), [2](#)
`rho.methyl.expr` (InterSIM), [2](#)