**Package ‘InterSIM’**

July 30, 2016

**Type** Package

**Title** Simulation of Inter-Related Genomic Datasets

**Version** 2.1

**Date** 2015-12-11

**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 (>= 2.14), MASS, NMF, tools

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2016-07-30 00:09:16

**R topics documented:**

<table>
<thead>
<tr>
<th>InterSIM-package</th>
<th>InterSIM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

**Index**

<table>
<thead>
<tr>
<th>InterSIM-package</th>
<th>Simulation of inter-related genomic datasets</th>
</tr>
</thead>
</table>

**Description**

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

**Details**
Author(s)
Prabhakar Chalise, Rama Raghavan, Brooke Fridley Maintainer: Rama Raghavan

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

Usage
`intersim(nSample=500, clusterSampleProp=c(0.30, 0.30, 0.40), deltaMethyl=5.0, deltaExpr=5.0, deltaProtein=5.0, pDMP=0.2, pDEG=NULL, pDEP=NULL, sigmaMethyl=NULL, sigmaExpr=NULL, sigmaProtein=NULL, corMethylExpr=NULL, corExprProtein=NULL, doPlot=FALSE, sampleCluster=TRUE, featureCluster=TRUE)`

Arguments
- `nSample`: Number of subjects to simulate
- `clusterSampleProp`: 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.
- `deltaMethyl`: Cluster mean shift for methylation data
- `deltaExpr`: Cluster mean shift for expression data
- `deltaProtein`: Cluster mean shift for protein data
- `pDMP`: proportion of DE Cpgs (DE = Differentially Expressed)
- `pDEG`: proportion of DE mRNA, if NULL (default) mRNAs mapped by DE Cpgs will be selected
- `pDEP`: proportion of DE protein, if NULL (default) proteins mapped by DE mRNAs will be selected
InterSIM

sigma.methyl: Covariance structure methylation data, if NULL precomputed values will be used. "indep" gives covariance structure with diagonal elements only (Independent features)

sigma.expr: Covariance structure mRNA data, if NULL precomputed values will be used. "indep" gives covariance structure with diagonal elements only (Independent features)

sigma.protein: Covariance structure Protein data, if NULL precomputed values will be used. "indep" gives covariance structure with diagonal elements only (Independent features)

do.plot: TRUE to generate heatmap, default is FALSE

sample.cluster: TRUE, if clustering should be done on samples

feature.cluster: TRUE, if clustering should be done on genomic features

cor.methyl.expr: Correlation between methylation and mRNA, if NULL precomputed values will be used

cor.expr.protein: Correlation between mRNA and protein, if NULL precomputed values will be used

Value

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 <pclalise@kumc.edu>, Rama Raghavan <rraghavan@kumc.edu>, Brooke Fridley <bfridley@kumc.edu>

Examples

```r
# Example:

prop <- c(0.20, 0.30, 0.27, 0.23)
effect <- 2.5
sim.data <- InterSIM(n.sample=500, cluster.sample.prop = c(0.30, 0.30, 0.40),
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