Package ‘GeneReg’
February 19, 2015

Type Package
Title Construct time delay gene regulatory network
Version 1.1.2
Date 2010-09-01
Author Tao Huang
Maintainer Tao Huang <huangtao@sibs.ac.cn>
Description GeneReg is an R package for inferring time delay gene regulatory network using time course gene expression profiles. The main idea of time delay linear model is to fit a linear regression model using a set of putative regulators to estimate the transcription pattern of a specific target gene.
Depends splines, igraph
License LGPL
URL http://www.biomedcentral.com/1756-0500/3/142
Citation Tao Huang, Lei Liu, Ziliang Qian, Kang Tu, Yixue Li, Lu Xie: Using GeneReg to construct time delay gene regulatory networks. BMC Res Notes 2010, 3(1):142.
Repository CRAN
Date/Publication 2012-10-29 08:57:07
NeedsCompilation no

R topics documented:

GeneReg-package .................................................. 2
mut.expr.data ...................................................... 3
plot ................................................................. 4
plot.GeneReg ....................................................... 5
tf.list ............................................................. 5
timedelay.lm ..................................................... 6
timedelay.lm.batch ............................................. 7
timedelay.univariate.lm ....................................... 7
ts.bspline ....................................................... 8
wt.expr.data .................................................... 9
Description

**GeneReg** is an R package for inferring time delay gene regulatory network using time course gene expression data.

Details

The goal of time delay linear model is to fit a linear regression model using a set of putative regulators to estimate the transcription pattern of a specific target gene.

\[ E_g = a_1 \ast E_{tf1}(−t_1) + a_2 \ast E_{tf2}(−t_2) + \ldots + a_i \ast E_{tfi}(−t_i) + \ldots + a_n \ast E_{tfn}(−t_n) \]

Where \( E_g \) is the relative expression level of target gene, \( E_{tfi} \) is the relative expression level of \( TFi \), \( t_i \) is time delay of \( TFi \), \( a_i \) is the regression coefficient of \( TFi \).

The regression models were optimized by the AIC forward selection procedure.

In the following example, we will demonstrate how to use this package to analyze a cell cycle datasets which is publicly available at GEO [http://www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo) under accession number GSE8799. The data set included the gene expression profiles of wild-type cells and cyclin-mutant cells with 15 time points during two cell cycles. 1271 periodic genes acquired from wild-type cells were analyzed to investigate how yeast regulators regulate target gene expression. These periodic genes formed the list of target genes. Then, a candidate pool of potential regulators of target genes was constructed by intersecting the periodic genes and the transcription factors in YEASTRACT [http://www.yeastract.com/](http://www.yeastract.com/).

First, B spline interpolation was applied to estimate 100 time points according to the original 15 time points.

```r
> data(wt.expr.data)
> wt.bspline.data<- ts.bspline(wt.expr.data,
+ ts.point=as.numeric(colnames(wt.expr.data)), data.predict=100)
```

Then, the time delay linear model was carried out based on the interpolated expression data. single.adj.r.squared and multiple.adj.r.squared were the cutoffs of adjusted R squared of single regulator regression and multiple regulator regression, respectively. tfs.list were the candidate pool of regulators. min.coef means the regression coefficient should be larger than 1/4 and smaller than 4. maxdelay means the time points delay should not exceed 1/4 of all time points.

```r
> data(tf.list)
```
> dir.create('wt.model')
> setwd('wt.model')

> wt.models<-timedelay.lm.batch(bspline.data=wt.bspline.data,
  + expr.data=wt.expr.data, regulator.list=tf.list,
  + target.list=rownames(wt.bspline.data),
  + single(adj.r.squared=0.8, multiple(adj.r.squared=0.9,
  + maxdelay=ncol(wt.bspline.data)*0.1, min.coef=0.25, max.coef=4,
  + output=T, topdf=T, xlab='Time point (lifeline)',
  + ylab='Relative expression level (in log ratio')
> setwd('..')

Finally, the gene regulatory network can be plotted according to the series of time delay linear models.

> pdf('wt.network.pdf',width=70, height=70)
> plot.GeneReg(wt.models, vertex.size=2, layout=layout.fruchterman.reingold)
> dev.off()

The analysis of cyclin-mutant cells is similar.

Author(s)
Tao Huang <<huangtao@sibs.ac.cn>>

References

mut.expr.data is from a yeast cell cycle datasets which is publicly available at GEO [http://www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo) under accession number GSE8799. The time course gene expression profiles included 15 time points at 16 min resolution in cyclin-mutant Saccharomyces cerevisiae.

### Usage

```r
data(mut.expr.data)
```

### Format

1267 genes expression profiles at 15 time points

### Source


### References


---

**plot**

*Plot gene regulatory network*

### Description

Plot gene regulatory network based on a series of time delay linear models.

### Usage

```r
plot(edge, ...)
```

### Arguments

- `edge`: output of `timedelay.lm.batch`, which has 4 columns. The first column is the regulator, the second column is the target gene, the third column is the regression coefficient, the fourth column is time delay.
- `...`: Additional arguments, passed to `plot.igraph`. 
**plot.GeneReg**  

**Plot gene regulatory network**

**Description**
Plot gene regulatory network based on a series of time delay linear models.

**Usage**

```r
## S3 method for class 'GeneReg'
plot(edge, ...)
```

**Arguments**
- `edge`: output of `timedelay.lm.batch`, which has 4 columns. The first column is the regulator, the second column is the target gene, the third column is the regression coefficient, the fourth column is time delay.
- `...`: Additional arguments, passed to `plot.igraph`.

**References**


---

**tf.list**  

**The candidate pool of potential regulators**

**Description**

tf.list was constructed by intersecting the periodic genes and the transcription factors in YEASTRACT [http://www.yeastract.com/](http://www.yeastract.com/).

**Usage**

```r
data(tf.list)
```

**Format**
A list of potential regulators

**Source**


References


Description
timedelay.lm is for time delay linear model.

Usage
timedelay.lm(bspline.data, expr.data, target, regulator, maxdelay=ncol(bspline.data)*0.25, single.adj.r.squared=0.8, multiple.adj.r.squared=0.9, min.coef=0.25, max.coef=4, output=FALSE, topdf=FALSE, xlab='time point', ylab='log ratio')

Arguments

bspline.data an interpolated expression matrix with genes in the rows, predicted time points in the columns
expr.data an observed expression matrix with genes in the rows, predicted time points in the columns
target target gene
regulator potential regulators
maxdelay max time points delay
single.adj.r.squared the cutoff of adjusted R squared of univariate regression
multiple.adj.r.squared the cutoff of adjusted R squared of multivariate regression
min.coef the regression coefficient should be larger than min.coef
max.coef the regression coefficient should be smaller than max.coef
output whether the model should be plotted
topdf whether the model should be plotted into a pdf file
xlab xlab of the plot
ylab ylab of the plot
References


**timedelay.lm.batch**

**time delay linear model batch**

**Description**

Batch usage of time delay linear model.

**Usage**

```r
timedelay.lm.batch(bspline.data, expr.data, regulator.list,
                    target.list=rownames(bspline.data), ...)
```

**Arguments**

- `bspline.data`: output of `ts.bspline`, an interpolated expression matrix with genes in the rows, predicted time points in the columns
- `expr.data`: an observed expression matrix with genes in the rows, predicted time points in the columns
- `regulator.list`: a candidate pool of potential regulators
- `target.list`: the list of target genes that will be investigated
- `...`: Additional arguments, passed to `timedelay.lm`

**References**


**timedelay.univariate.lm**

*Internal function for time delay linear model*

**Description**

timedelay.univariate.lm is an internal function for time delay linear model.
Usage

timedelay.univariate.lm(bspline.data, target, regulator,
maxdelay=ncol(bspline.data)*0.25,
single.adj.r.squared=0.8, min.coef=0.25, max.coef=4)

Arguments

bspline.data  an interpolated expression matrix with genes in the rows, predicted time points in the columns


target  target gene


regulator  potential regulators

maxdelay  max time points delay


single.adj.r.squared  the cutoff of adjusted R squared of univariate regression


min.coef  the regression coefficient should be larger than min.coef


max.coef  the regression coefficient should be smaller than max.coef


References


---

```

__ts.bspline__

---

**B spline interpolation**

---

**Description**

ts.bspline is for B spline interpolation.

Usage

ts.bspline(expr, ts.point=NULL, data.predict=100, df=round(ncol(expr)*0.8))

Arguments

expr  an expression matrix with genes in the rows, time points in the columns

ts.point  time points

data.predict  the number of time points to be predicted

df  degrees of freedom, used in B-Spline Basis for Polynomial Splines
**wt.expr.data**

**Time course cell cycle dataset of wild-type cells**

**Description**

wt.expr.data is from a yeast cell cycle datasets which is publicly available at GEO [http://www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo) under accession number GSE8799. The time course gene expression profiles included 15 time points at 16 min resolution in wild type Saccharomyces cerevisiae.

**Usage**

```r
data(wt.expr.data)
```

**Format**

1267 genes expression profiles at 15 time points

**Source**


**References**

Index

GeneReg (GeneReg-package), 2
GeneReg-package, 2

mut.expr.data, 3

plot, 4
plot.GeneReg, 5

tf.list, 5
timedelay.lm, 6, 7
timedelay.lm.batch, 7
timedelay.univariate.lm, 7
ts.bspline, 8

wt.expr.data, 9