

Package ‘Roleswitch’

April 5, 2014

Type Package

Title Infer miRNA-mRNA interactions using paired expression data from a single sample

Version 1.0.0

Date 2013-09-04

Author Yue Li

Maintainer Yue Li <yueli@cs.toronto.edu>

Description Infer Probabilities of MiRNA-mRNA Interaction Signature (ProMiSe) using paired expression data from a single sample. Roleswitch operates in two phases by inferring the probability of mRNA (miRNA) being the targets (“targets”) of miRNA (mRNA), taking into account the expression of all of the mRNAs (miRNAs) due to their potential competition for the same miRNA (mRNA). Due to dynamic miRNA repression in the cell, Roleswitch assumes that the total transcribed mRNA levels are higher than the observed (equilibrium) mRNA levels and iteratively updates the total transcription of each mRNA targets based on the above inference.

Depends R (>= 2.10), pracma, reshape, plotrix, microRNA, biomaRt, Biostrings, Biobase, DBI

Suggests ggplot2

License GPL-2

URL <http://www.cs.utoronto.ca/~yueli/roleswitch.html>

biocViews microRNA

Lazyload yes

R topics documented:

Roleswitch-package	2
diagnosticPlot	3
getSeedMatrix	4
roleswitch	6
tcga_gbm_testdata	8

Index	10
--------------	-----------

Roleswitch-package *Infer miRNA-mRNA interactions using paired expression data from a single sample*

Description

Infer Probabilities of MiRNA-mRNA Interaction Signature (ProMISe) using paired expression data from a single sample. Roleswitch operates in two phases by inferring the probability of mRNA (miRNA) being the targets ("targets") of miRNA (mRNA), taking into account the expression of all of the mRNAs (miRNAs) due to their potential competition for the same miRNA (mRNA). Due to mRNA transcription and miRNA repression events simultaneously happening in the cell, Roleswitch assumes that the total transcribed mRNA levels are higher than the observed (equilibrium) mRNA levels and iteratively updates the total transcription of each mRNA targets based on the above inference.

Details

Package: Roleswitch
Type: Package
Version: 0.99.3
Date: 2013-09-04
License: GPL-2

The main function `roleswitch` takes as inputs the mRNA and miRNA expression as numerical vector and seedmatch (integer) matrix provided by the user or generated from `getSeedMatrix`. The function then outputs a list with `p.xz` containing the ProMISe as the final results and other results for diagnostic purpose.

Author(s)

Yue Li Maintainer: Yue Li <yueli@cs.toronto.edu>

References

Li, Y., ..., Zhang, Z., Inference of personalized miRNA-mRNA interactions toward redefining cancer signatures (in preparation).

Arvey, A., Larsson, E., Sander, C., Leslie, C. S., & Marks, D. S. (2010). Target mRNA abundance dilutes microRNA and siRNA activity. *Molecular systems biology*, 6, 1-7. doi:10.1038/msb.2010.24

See Also

[roleswitch](#)

Examples

```
library(Roleswitch)
ls("package:Roleswitch")
```

diagnosticPlot	<i>Create diagnostic plot for understanding the Roleswitch model outputs.</i>
----------------	---

Description

Create diagnostic plot for understanding the `roleswitch` outputs. Create a 2 by 4 panels of plots. From left to right, the top panel displays the observed N mRNA and M miRNA expression (`x.o` and `z.o`), the N by M seed-match matrix (`c`), and the inferred total mRNA expression; the bottom panel displays the inferred probability of the M miRNAs targeting the N mRNA (miRNA-mRNA; `p.x`), the probability of the N mRNA "targeting" the M miRNAs (mRNA-miRNA; `p.z`), the dot product of the above two matrices (Joint) and the convergence rate (`delta.p.all`).

Usage

```
diagnosticPlot(pred)
```

Arguments

`pred` Results obtained from `roleswitch`.

Author(s)

Yue Li

References

Li, Y., ..., Zhang, Z., Inference of personalized miRNA-mRNA interactions toward redefining cancer signatures (in preparation).

See Also

[roleswitch](#)

Examples

```
x.o <- matrix(abs(rnorm(10, mean=3)),
  dimnames=list(c(1:10), "mRNA")) # mRNA expression

z.o <- matrix(abs(rnorm(4, mean=3)),
  dimnames=list(c(1:4), "miRNA")) # miRNA expression

c <- matrix(rpois(40, lambda=3), nrow=nrow(x.o),
  dimnames=list(c(1:10), c(1:4))) # seed match matrix
```

```
rs.pred <- roleswitch(x.o, z.o, c)

diagnosticPlot(rs.pred)
```

getSeedMatrix	<i>Get seed-match matrix between defined mRNA and miRNA in an organism.</i>
---------------	---

Description

Given N mRNA and M miRNA IDs or simply the species common names (e.g., human), obtain the N by M seed match matrix as a the number of target sites each mRNA has for each miRNA.

Usage

```
getSeedMatrix(mRNA, miRNA, species = "human",
  id_type = "ensembl_transcript_id", mRNA_id_type = id_type,
  miRNA_id_type = id_type, longest3utr = TRUE,
  biomart = "ensembl", dataset = "hsapiens_gene_ensembl",
  returnGeneInfo = FALSE, convert2genesymbol = TRUE, ...)
```

Arguments

mRNA	A character vector of N mRNA ids.
miRNA	A character vector of M miRNA ids
species	Common names for a species. Currently only human and mouse are supported as precompiled target site information.
id_type	A string specifying the id type used for both the mRNAs and miRNA.
mRNA_id_type	A string specifying the id type used for the mRNAs.
miRNA_id_type	A string specifying the id type used for the miRNAs.
longest3utr	For genes having multiple transcripts, whetehr to use only the transcript with the longest 3'UTR (default: TRUE).
biomart	Database for biomart, which is ued to obtain transcript information using getBM (default: ensembl).
dataset	Dataset used to query the biomart database using getBM (default: hsapiens_gene_ensembl).
returnGeneInfo	Binary indicator to return gene information besides seed matrix (default: FALSE); if TRUE, then a list containing seed matrix and gene info is return; otherwise just the seed matrix.
convert2genesymbol	Whether to convert id such as ensembl_gene_id to gene symbols as row names of the seed matrix.
...	Paramters passed to getBM .

Details

Retrieve and process target site information to generate a N by M matrix representing the number of target sites of mRNA *i* for miRNA *k*. If species is specified, then the suggested data package RoleswitchData will be loaded and the pre-compiled seed matrix is used. Currently, only human and mouse are supported with this option. Otherwise, download the sequences based on the specified mRNA and miRNA IDs and obtain the seed matches using [seedRegions](#).

Value

seed match matrix	numeric matrix containing the number of target sites for each miRNA and mRNA pairs
gene info	a data.frame containing miRNA id type, ensembl gene id, gene symbol, start and end of 3'UTR (only returned when returnGeneInfo is TRUE)

Note

This is just a convenience function. Users are encouraged to construct the most up-to-date seed match matrix on their own from other source without using this function.

Author(s)

Yue Li

References

miRBase: tools for microRNA genomics. (2008). miRBase: tools for microRNA genomics., 36(Database issue), D154-8. doi:10.1093/nar/gkm952

R. Gentleman and S. Falcon (2013). microRNA: Data and functions for dealing with microRNAs. R package version 1.18.0.

See Also

[roleswitch](#), [getBM](#), [seedRegions](#)

Examples

```
seedMatrix.human <- getSeedMatrix()  
head(seedMatrix.human)
```

roleswitch	<i>Infer miRNA-mRNA interactions using paired expression data from a single sample.</i>
------------	---

Description

Infer Probabilities of MiRNA-mRNA Interaction Signature (ProMISe) using paired expression data from a single sample. Roleswitch operates in two phases by inferring the probabilities of mRNA (miRNA) being the targets ("targets") of miRNA (mRNA), taking into account the expression of all of the mRNAs (miRNAs) due to their potential competition for the same miRNA (mRNA). Due to mRNA transcription and miRNA repression events simultaneously happening in the cell, Roleswitch assumes that the total transcribed mRNA levels are higher than the observed (equilibrium) mRNA levels and iteratively updates the total transcription of each mRNA targets based on the above inference.

Usage

```
roleswitch(x.o, z.o, c, maxiter = 200, tol = 1e-05,
          eta.z = 0.001, expected.total = 1.3, verbose = TRUE,
          annotation.db, probe2genesymbol=TRUE, ...)
```

Arguments

x.o	A numeric vector as the observed expression of N mRNAs. NOTE: rownames is required for x.o to map each mRNA to the rownames of the seed match matrix.
z.o	A numeric vector as the observed expression of M miRNAs. NOTE: rownames is required for z.o to map each miRNA to the colnames of the seed match matrix.
c	A numeric N x M matrix of integers representing the seed match matrix between N mRNA and M miRNA. NOTE: dimnames is required for the seed match matrix to map each mRNA and each miRNA to its rownames and colnames, respectively.
maxiter	The maximum number of iterations before terminating roleswitch inference (default: 200).
tol	The threshold on the largest absolute difference between the current and the previous probabilities of miRNA-mRNA interactions (default: 1e-5).
eta.z	A scalar decimal value specifying the amount of update applied to the total transcribed mRNA, which is inferred during the iteration (default: 1e-3).
expected.total	The ratio of total transcription over observed transcription (default: 1.3).
verbose	Display progress at each iteration.
annotation.db	Character string specifying the name of the annotation package for microarray platform if the input x.o is an eSet or ExpressionSet object. This is optional only if annotation slot in the eSet x.o is defined.

probe2genesymbol
 Whether to convert probe id to gene symbol (Default: TRUE). This only applies when `x.o` is an `eSet/ExpressionSet`. Probe ID are usually the `featureNames`, an attribute under the class `eSet`. If TRUE, conversion from probe id to gene symbol is performed automatically.

... Arguments passed to `getSeedMatrix`.

Details

The model assumes total expression of mRNA is unobserved and higher than the observed corresponding expression due to RNA degradation induced by miRNA-mRNA interaction. The general algorithm is outlined as follows:

- (1) Infer mRNA i targeted by miRNA k taking into account the hidden total expression of $1 \dots N$ mRNA and miRNA k
- (2) Estimate total transcription level of mRNA i
- (3) Infer miRNA k "targeted" by mRNA i taking into account $1 \dots M$ miRNA and mRNA i expression
- (4) Repeat 1-3 until convergence

User provide roleswitch an $N \times M$ seed-match matrix containing the number of target sites for each mRNA i and miRNA k . Otherwise, `getSeedMatrix` will be used to retrieve seed-match matrix.

Value

An object list (defined class: ProMISE) containing the following items:

`x.t`: inferred total $N \times 1$ mRNA expression vector

`p.x`: inferred $N \times M$ miRNA-mRNA probability matrix (mRNA plays as targets of miRNA)

`p.z`: inferred $N \times M$ mRNA-miRNA probability matrix (miRNA plays as targets of mRNA)

`p.xz`: Dot product of `p.x` and `p.z` as the final Roleswitch prediction or the ProMISE.

`x.o`, `z.t`, `c`: The same as the inputs.

`delta.p.all`: Difference between `p.x` at two adjacent iterations to monitor the progress of convergence.

Note

Warning may be issued if the input seed matrix contain miRNA or mRNA that have zero seed or seed match for any mRNA or miRNA, respectively. Nonetheless, the outputs ProMISE will conform the original input mRNA and miRNA vector in matching their corresponding IDs.

Author(s)

Yue Li

References

Li, Y., ..., Zhang, Z., Inference of probabilistic miRNA-mRNA interaction signatures and its application in novel cancer biomarker discovery (in preparation).

See Also[getSeedMatrix](#)**Examples**

```
x.o <- matrix(abs(rnorm(10, mean=3)),
  dimnames=list(c(1:10),"mRNA")) # mRNA expression

z.o <- matrix(abs(rnorm(4, mean=3)),
  dimnames=list(c(1:4),"miRNA")) # miRNA expression

c <- matrix(rpois(40, lambda=3),nrow=nrow(x.o),
  dimnames=list(c(1:10),c(1:4))) # seed match matrix

rs.pred <- roleswitch(x.o, z.o, c)

rs.pred$p.xz
```

tcga_gbm_testdata	<i>Test data of miRNA and mRNA expression vector.</i>
-------------------	---

Description

Test data of miRNA and mRNA expression vector from the same human individual that is used in the vignette.

Usage

```
data(tcga_gbm_testdata)
```

Format

A data frame with 11884 mRNA expression and 373 miRNA expression on the following 2 variables.

x a numeric vector for expression of 11884 mRNAs

y a numeric vector for expression of 373 miRNA

Details

The miRNA and mRNA expression data for the same individual (barcode ID: TCGA-02-0001-01) were downloaded from TCGA (The Cancer Genome Atlas) GBM (Glioblastoma multiforme). To eliminate negative values in the expression matrices, we linearly transformed all of the data to positive scale with range between 0 and the maximum of the positive values using [rescale](#) R function. The resulting data were further processed by filtering out miRNAs or mRNAs without any seed/seed match based on the Microcosm database. As a result, the test data contain the expression for 11884 distinct mRNA and distinct 373 miRNA.

Source

<https://tcga-data.nci.nih.gov>

References

Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. *Nucleic acids research* 36: D154-8

Examples

```
data(tcga_gbm_testdata)
```

Index

- *Topic **Roleswitch**
 - Roleswitch-package, 2
- *Topic **datasets**
 - tcga_gbm_testdata, 8
- *Topic **expression**
 - roleswitch, 6
- *Topic **miRNA**
 - getSeedMatrix, 4
 - roleswitch, 6
- *Topic **microRNA**
 - roleswitch, 6
- *Topic **plot**
 - diagnosticPlot, 3
- *Topic **roleswitch**
 - roleswitch, 6
- *Topic **seedmatch**
 - getSeedMatrix, 4

diagnosticPlot, 3

eSet, 6, 7

ExpressionSet, 6

getBM, 4, 5

getSeedMatrix, 2, 4, 7, 8

rescale, 8

Roleswitch (Roleswitch-package), 2

roleswitch, 2, 3, 5, 6

Roleswitch-package, 2

seedRegions, 5

tcga_gbm_testdata, 8