

Package ‘CDROM’

October 12, 2022

Type Package

Title Phylogenetically Classifies Retention Mechanisms of Duplicate Genes from Gene Expression Data

Version 1.1

Date 2016-03-18

Author Brent Perry <brp5173@psu.edu>, Raquel Assis <rassis@psu.edu>

Maintainer Raquel Assis <rassis@psu.edu>

Depends R (>= 3.2.0)

Imports graphics, grDevices, stats, utils

Description Classification is based on the recently developed phylogenetic approach by Assis and Bachtrog (2013). The method classifies the evolutionary mechanisms retaining pairs of duplicate genes (conservation, neofunctionalization, subfunctionalization, or specialization) by comparing gene expression profiles of duplicate genes in one species to those of their single-copy ancestral genes in a sister species.

License GPL-2

RoxygenNote 5.0.1

NeedsCompilation no

Repository CRAN

Date/Publication 2016-04-27 18:56:44

R topics documented:

CDROM	2
Index	4

Description

Classification is based on the recently developed phylogenetic approach by Assis and Bachtrog (2013). The method classifies the evolutionary mechanisms retaining pairs of duplicate genes (conservation, neofunctionalization, subfunctionalization, or specialization) by comparing gene expression profiles of duplicate genes in one species to those of their single-copy ancestral genes in a sister species.

Usage

```
CDROM(dupFile, singleFile, exprFile1, exprFile2, out = "out", Ediv,
      PC = FALSE, useAbsExpr = FALSE, head1 = TRUE, head2 = TRUE,
      head3 = TRUE, head4 = TRUE, legend = "topleft")
```

Arguments

dupFile	a tab-separated file containing duplicate gene pairs and their orthologs (three gene IDs per line).
singleFile	a tab-separated file containing orthologous single-copy genes (two gene IDs per line).
exprFile1	a tab-separated file containing gene expression levels for all genes in species 1 (one gene ID and its expression levels per line).
exprFile2	a tab-separated file containing gene expression levels for all genes in species 2 (one gene ID and its expression levels per line).
out	the prefix to be used in the names of the three output files. Defaults to 'out'.
Ediv	the divergence cutoff to be used in classifications. Defaults to semi-interquartile range of the median (SIQR).
PC	a logical value indicating whether parent and child copies are separated in dupFile. Defaults to FALSE.
useAbsExpr	a logical value indicating whether absolute or relative expression values are used for Euclidean distance calculations. Defaults to FALSE.
head1	a logical value indicating whether exprFile1 contains the names of its variables as the first line. Defaults to TRUE.
head2	a logical value indicating whether exprFile2 contains the names of its variables as the first line. Defaults to TRUE.
head3	a logical value indicating whether dupFile contains the names of its variables as the first line. Defaults to TRUE.
head4	a logical value indicating whether singleFile contains the names of its variables as the first line. Defaults to TRUE.
legend	a keyword indicating the position of the legend in the output plot. Options are 'topright' and 'topleft'. Defaults to 'topleft'.

Value

A table with the classifications of all duplicate gene pairs.

A table with counts of classifications with five Ediv values.

A plot showing distributions of all Euclidean distances and the position of Ediv.

Author(s)

Brent Perry <brp5173@psu.edu>, Raquel Assis <rassis@psu.edu>

References

[1] Assis R and Bachtrog D. Neofunctionalization of young duplicate genes in *Drosophila*. Proc. Natl. Acad. Sci. USA. 110: 17409-17414 (2013).

Examples

```
CDROM(dupFile=system.file("extdata","human_chicken_dups.txt",package="CDROM"),
singleFile=system.file("extdata","human_chicken_singles.txt",package="CDROM"),
exprFile1=system.file("extdata","human_expr.txt",package="CDROM"),
exprFile2=system.file("extdata","chicken_expr.txt",package="CDROM"))
```

Index

* **duplication,**
CDROM, [2](#)

* **gene**
CDROM, [2](#)

* **neofunctionalization,**
CDROM, [2](#)

* **subfunctionalization**
CDROM, [2](#)

CDROM, [2](#)