Package ‘deltaccd’

October 13, 2022

Type Package

Title Quantify Rhythmic Gene Co-Expression Relative to a Reference

Version 1.0.2

Description Infer progression of circadian rhythms in transcriptome data in which samples are not labeled with time of day and coverage of the circadian cycle may be incomplete. See Shilts et al. (2018) <doi:10.7717/peerj.4327>.

URL https://deltaccd.hugheylab.org,
    https://github.com/hugheylab/deltaccd

License GPL-2

Encoding UTF-8

LazyData true

LazyDataCompression gzip

RoxygenNote 7.1.2

Depends R (>= 3.4)

Imports doRNG (>= 1.6.6), foreach (>= 1.4.3), ggplot2 (>= 2.2.1),
    rlang (>= 0.4.11), scales (>= 0.5.0), statmod (>= 1.4.30),
    data.table (>= 1.14.0)

Suggests doParallel (>= 1.0.11), knitr, rmarkdown, testthat (>= 3.0.0), vdiffr

Config/testthat/edition 3

NeedsCompilation no

Author Jake Hughey [aut, cre],
    Elliot Outland [aut]

Maintainer Jake Hughey <jakejhughey@gmail.com>

Repository CRAN

Date/Publication 2022-02-11 19:30:06 UTC
calcCCD

Calculate clock correlation distance (CCD).

Description

Quantify the similarity of gene co-expression between a reference and a test dataset. Statistical significance is calculated using permutation of the genes.

Usage

```
calcCCD(
  refCor, 
  emat, 
  groupVec = NULL, 
  refEmat = NULL, 
  nPerm = 1000, 
  geneNames = NULL, 
  dopar = FALSE, 
  scale = FALSE 
)
```

Arguments

- **refCor** Correlation matrix to be used as the reference, such as comes from `getRefCor()`. Should contain Spearman correlation values.
- **emat** Matrix of expression values, where each row corresponds to a gene and each column corresponds to a sample. The rownames and colnames of refCor should be present in the rownames of emat. For the p-value calculation, it is important that emat include all measured genes, not just those in refCor.
- **groupVec** Optional vector indicating the group to which group each sample belongs. If not provided, the function assumes all samples belong to the same group.
- **refEmat** Optional expression matrix for calculating co-expression for the reference, with the same organization as emat. Only used if refCor is not provided.
- **nPerm** Number of permutations for assessing statistical significance.
- **geneNames** Optional vector indicating a subset of genes in refCor, emat, and/or refEmat to use for calculating the CCD.
calcDeltaCCD

dopar Logical indicating whether to process features in parallel. Make sure to register a parallel backend first.
scale Logical indicating whether to scale CCD by the number of gene pairs.

Value

A data.table with columns for group name, CCD, and p-value.

See Also

getRefCor(), calcDeltaCCD(), plotHeatmap()

Examples

set.seed(35813)

refCor = getRefCor()
ccdResult = calcCCD(refCor, GSE19188$emat, GSE19188$groupVec, nPerm = 100)

calcDeltaCCD

Calculate delta clock correlation distance.

Description

Calculate the difference between the clock correlation distances (CCDs), relative to a reference, for two groups of samples. Statistical significance is calculated using permutation of the samples that belong to either of those two groups.

Usage

calcDeltaCCD(
    refCor,
    emat,
    groupVec,
    groupNormal,
    refEmat = NULL,
    nPerm = 1000,
    geneNames = NULL,
    dopar = FALSE,
    scale = FALSE
)
Arguments

- **refCor**: Correlation matrix to be used as the reference, such as comes from `getRefCor()`. Should contain Spearman correlation values.
- **emat**: Matrix of expression values, where each row corresponds to a gene and each column corresponds to a sample. The rownames and colnames of `refCor` should be present in the rownames of `emat`. For the p-value calculation, it is important that `emat` include all measured genes, not just those in `refCor`.
- **groupVec**: Vector indicating the group to which group each sample belongs. It’s ok for `groupVec` to have more than two groups.
- **groupNormal**: Value indicating the group in `groupVec` that corresponds to normal or healthy. Other groups will be compared to this group.
- **refEmat**: Optional expression matrix for calculating co-expression for the reference, with the same organization as `emat`. Only used if `refCor` is not provided.
- **nPerm**: Number of permutations for assessing statistical significance.
- **geneNames**: Optional vector indicating a subset of genes in `refCor`, `emat`, and/or `refEmat` to use for calculating the CCD.
- **dopar**: Logical indicating whether to process features in parallel. Make sure to register a parallel backend first.
- **scale**: Logical indicating whether to use scaled CCDs to calculate difference.

Value

A `data.table` with columns for group 1, group 2, deltaCCD, and p-value. In each row, the deltaCCD is the CCD of group 2 minus the CCD of group 1, so group 1 corresponds to `groupNormal`.

See Also

- `getRefCor()`, `calcCCD()`, `plotHeatmap()`

Examples

```r
set.seed(35813)
refCor = getRefCor()
deltaCcdResult = calcDeltaCCD(
  refCor, GSE19188$emat, GSE19188$groupVec, 'healthy', nPerm = 100)
```
getRefCor

Retrieve the reference correlation matrix for circadian gene co-expression.

Description

The pan-tissue reference matrix is based on a fixed-effects meta-analysis of eight circadian transcriptome datasets from mice, as described in Shilts et al. 2018 (doi: 10.7717/peerj.4327). The human blood reference matrix is based on an analysis of three microarray datasets (manuscript in preparation).

Usage

getRefCor(
  species = c("human", "mouse"),
  tissue = c("pan", "blood"),
  useEntrezGeneId = TRUE
)

Arguments

species Currently either 'human' or 'mouse'. Only affects the row and column names of the correlation matrix, not the actual values.
tissue One of either 'pan' or 'blood'.
useEntrezGeneId If FALSE, row and column names of correlation matrix will correspond to gene symbols (e.g., PER2).

Value

A matrix of Spearman correlation values.

See Also

GSE19188, plotRefHeatmap(), calcCCD(), calcDeltaCCD()

GSE19188

Gene expression data for GSE19188.

Description

Data of gene expression measured by microarray for samples from human non-small cell lung cancer.
Usage

GSE19188

Format

A list with two objects:

emat  Matrix of normalized expression values, where each row corresponds to a gene (rownames are Entrez Gene IDs) and each column corresponds to a sample. To save space, genes have been downsampled.

groupVec  Vector of condition (tumor or healthy) for each sample.

Source


See Also

getRefCor(), calcCCD(), calcDeltaCCD()

plotHeatmap  Visualize gene co-expression.

Description

Make heatmaps of the co-expression (Spearman correlation) between pairs of selected genes in a dataset.

Usage

plotHeatmap(geneNames, emat, groupVec = NULL)

Arguments

geneNames  Vector indicating the subset of genes in the rownames of emat for which to calculate the correlations in expression.

emat  Matrix of expression values, where each row corresponds to a gene and each column corresponds to a sample. The elements of geneNames should be present in the rownames of emat.

groupVec  Optional vector indicating the group to which group each sample belongs. If not provided, the function assumes all samples belong to the same group.

Value

A ggplot object, which can be saved using ggplot2::ggsave(). Heatmap colors will be directly comparable to any heatmaps created by this function or by plotRefHeatmap().
plotRefHeatmap

See Also
calcCCD(), calcDeltaCCD(), plotRefHeatmap()

Examples

refCor = getRefCor()
pRef = plotRefHeatmap(refCor)
pTest = plotHeatmap(rownames(refCor), GSE19188$emat, GSE19188$groupVec)

plotRefHeatmap  Visualize the reference pattern of gene co-expression.

Description
Make a heatmap of the reference correlation matrix for gene co-expression.

Usage
plotRefHeatmap(refCor)

Arguments
refCor Correlation matrix, such as comes from getRefCor().

Value
A ggplot object, which can be saved using ggplot2::ggsave(). Heatmap colors will be directly comparable to any heatmaps created by this function or by plotHeatmap().

See Also
getRefCor(), plotHeatmap()

Examples

refCor = getRefCor()
pRef = plotRefHeatmap(refCor)
pTest = plotHeatmap(rownames(refCor), GSE19188$emat, GSE19188$groupVec)
Index

* datasets
  GSE19188, 5

  calcCCD, 2
  calcCCD(), 4–7
  calcDeltaCCD, 3
  calcDeltaCCD(), 3, 5–7

  getRefCor, 5
  getRefCor(), 2–4, 6, 7
  ggplot2::ggsave(), 6, 7
  GSE19188, 5, 5

  plotHeatmap, 6
  plotHeatmap(), 3, 4, 7
  plotRefHeatmap, 7
  plotRefHeatmap(), 5–7