Package 'RNAAgeCalc'

November 12, 2026

Type Package

Title A multi-tissue transcriptional age calculator

Version 1.24.0

Description It has been shown that both DNA methylation and RNA transcription are linked to chronological age and age related diseases. Several estimators have been developed to predict human aging from DNA level and RNA level. Most of the human transcriptional age predictor are based on microarray data and limited to only a few tissues. To date, transcriptional studies on aging using RNASeq data from different human tissues is limited. The aim of this package is to provide a tool for across-tissue and tissue-specific transcriptional age calculation based on GTEx RNASeq data.

License GPL-2

URL https://github.com/reese3928/RNAAgeCalc

BugReports https://github.com/reese3928/RNAAgeCalc/issues

Encoding UTF-8

Imports ggplot2, recount, impute, AnnotationDbi, org.Hs.eg.db, stats, SummarizedExperiment, methods

Depends R (>= 3.6)

Suggests knitr, rmarkdown, testthat

RoxygenNote 7.0.2

VignetteBuilder knitr

biocViews RNASeq,GeneExpression

git_url https://git.bioconductor.org/packages/RNAAgeCalc

git_branch RELEASE_3_23

git_last_commit 8b80095

git_last_commit_date 2026-04-28

Repository Bioconductor 3.23

Date/Publication 2026-11-11

Author Xu Ren [aut, cre], Pei Fen Kuan [aut]

Maintainer Xu Ren <xuren2120@gmail.com>

Contents

count2FPKM 2
fpkm 3
makeplot 3
predict_age 4
predict_age_fromse 7
rawcount 9

Index **10**

`count2FPKM` *Converting gene expression data from raw count to FPKM*

Description

This function converts gene expression data from raw count to FPKM by using `getRPKM`

Usage

```
count2FPKM(rawcount, genelength = NULL, idtype = "SYMBOL")
```

Arguments

`rawcount`	a matrix or data frame which contains gene expression counts data.
`genelength`	gene length in bp. The size of 'genelength' should be equal to the number of rows in 'rawcount'. This argument is optional. If not provided, gene length is obtained from the internal database.
`idtype`	a string which indicates the gene id type in rawcount matrix. It should be one of "SYMBOL", "ENSEMBL", "ENTREZID" or "REFSEQ". Default is "SYMBOL".

Value

a data frame contains FPKM.

References

Carlson M (2019). org.Hs.eg.db: Genome wide annotation for Human. R package version 3.8.2.

Collado-Torres, Leonardo, et al. "Reproducible RNA-seq analysis using recount2." Nature biotechnology 35.4 (2017): 319-321.

Examples

```
data(countExample)
head(rawcount)
fpkm = count2FPKM(rawcount)
head(fpkm)
```

fpkm *An example of FPKM data*

Description

This is an example of FPKM data. It is the gene expression FPKM data from The genotype-tissue expression (GTEx) project brain samples. The data was downloaded from recount2(https://jhubiostatistics.shinyapps.io/recou For illustration purpose, only two samples were included. The R script to generate this example data can be found in inst/scripts/createExamples.R.

Usage

```
fpkm
```

Format

A data frame which contains the fpkm data. The dimension of this data frame is 24989 by 2. Each row is a gene and each column is a sample.

References

Collado-Torres, Leonardo, et al. "Reproducible RNA-seq analysis using recount2." Nature biotechnology 35.4 (2017): 319-321.

Lonsdale, John, et al. "The genotype-tissue expression (GTEx) project." Nature genetics 45.6 (2013): 580.

makeplot *Make plot to visualize RNA age*

Description

This function makes plots to visualize the relationship between chronological age and RNA age.

Usage

```
makeplot(
  res,
  main = "RNA age vs chronological age",
  xlab = "chronological age",
  ylab = "RNA Age"
)
```

Arguments

res	a data frame returned by ‘predict_age‘ function. If the chronological age is not provided when using ‘predict_age‘ function, visulization cannot be made.
main	title of the plot
xlab	label of x-axis
ylab	label of y-axis

Value

the plot which shows RNA age vs chronological age

Examples

```
data(fpkmExample)
fpkm_large = cbind(fpkm, fpkm, fpkm, fpkm)
fpkm_large = cbind(fpkm_large, fpkm_large, fpkm_large, fpkm_large)
colnames(fpkm_large) = paste0("sample",1:32)
chronage = data.frame(sampleid = colnames(fpkm_large), age = 1:32)
res = predict_age(exprdata = fpkm_large, exprtype = "FPKM",
chronage = chronage)
makeplot(res)
```

---

predict_age *Calculate RNA age*

Description

This function calculates RNA age based on pre-trained calculators.

Usage

```
predict_age(
  exprdata,
  tissue,
  exprtype = c("FPKM", "counts"),
  idtype = c("SYMBOL", "ENSEMBL", "ENTREZID", "REFSEQ"),
  stype = c("all", "caucasian"),
  signature = NULL,
  genelength = NULL,
  chronage = NULL,
  maxp = NULL
)
```

Arguments

exprdata	a matrix or data frame which contains gene expression data with each row represents a gene and each column represents a sample. Use the argument ‘exprtype‘ to specify raw count or FPKM. The rownames of ‘exprdata‘ should be gene ids and colnames of ‘exprdata‘ should be sample ids.
tissue	a string indicate which tissue the gene expression data is obtained from. Users are expected to provide one of the following tissues. If the tissue argument is not provide or the provided tissue is not in this list, then the age predictor trained on all tissues will be used to calculate RNA age.

- adipose_tissue
- adrenal_gland
- blood
- blood_vessel
- brain
- breast
- colon
- esophagus
- heart
- liver
- lung
- muscle
- nerve
- ovary
- pancreas
- pituitary
- prostate
- salivary_gland
- skin
- small_intestine
- spleen
- stomach
- testis
- thyroid
- uterus
- vagina

exprtype	either "counts" or "FPKM". For RPKM data, please use ‘exprtype‘ = "FPKM".
idtype	a string which indicates the gene id type in ‘exprdata‘. It should be one of "SYMBOL", "ENSEMBL", "ENTREZID" or "REFSEQ". Default is "SYMBOL".
stype	a string which specifies which version of pre-trained calculators to be used. It should be either "all" or "Caucasian". "all" means samples from all races (American Indian/Alaska Native, Asian, Black/African American, and Caucasian) are used to obtain the pre-trained calculator. "Caucasian" means only the Caucasian samples are used to build up the pre-trained calculator.

`signature` a string which indicates the age signature to use when calculating RNA age. This argument is not required.
In the case that this argument is not provided, if ‘tissue‘ argument is also provided and the tissue is in the list above, the tissue specific age signature given by our DESeq2 analysis result on GTEx data will be used. Otherwise, the across tissue signature "GTExAge" will be used.
In the case that this argument is provided, it should be one of the following signatures. A detailed description of the meaning of these signatures is given in the package vignette.

- DESeq2
- Pearson
- Dev
- deMagalhaes
- GenAge
- GTExAge
- Peters
- all

`genelength` a vector which contains gene length in bp. The size of ‘genelength‘ should be equal to the number of rows in ‘exprdata‘. This argument is optional. If using ‘exprtype = "FPKM"‘, ‘genelength‘ argument is ignored. If using ‘exprtype = "counts"‘, the raw count will be converted to FPKM. If ‘genelength‘ is provided, the function will convert raw count to FPKM by the user-supplied gene length. Otherwise, gene length is obtained from the internal database.

`chronage` a data frame which contains the chronological age of each sample. This argument is optional. If provided, it should be a dataframe with 1st column sample id and 2nd column chronological age. The sample order in ‘chronage‘ doesn’t have to be in the same order as in ‘exprdata‘. However, the samples in ‘chronage‘ and ‘exprdata‘ should be the same. If some samples’ chronological age are not available, users are expected to set the chronological age in ‘chronage‘ to NA. If ‘chronage‘ contains more than 2 columns, only the first 2 columns will be considered. If this argument is not provided, the age acceleration residual will not be calculated. See package vignette for the definition of age acceleration residual.

`maxp` the maxp argument used in `impute.knn` function. This is optional.

Value

a data frame contains RNA age.

Examples

```
data(fpkmExample)
res = predict_age(exprdata = fpkm, exprtype = "FPKM")
```

predict_age_fromse *Calculate RNA age using SummarizedExperiment*

Description

This function takes SummarizedExperiment object as input and calculates RNA age.

Usage

```
predict_age_fromse(
  se,
  tissue,
  exprtype = c("FPKM", "counts"),
  idtype = c("SYMBOL", "ENSEMBL", "ENTREZID", "REFSEQ"),
  stype = c("all", "caucasian"),
  signature = NULL,
  maxp = NULL
)
```

Arguments

se	a SummarizedExperiment object. The assays(se) should contain gene expression data. The name of assays(se) should be either "FPKM" or "counts". Use 'exprtype' argument to specify the type of gene expression data provided. Users are able to provide the chronological age of samples using colData(se). This is optional. If provided, the column name for chronological age in colData(se) should be "age". If some samples' chronological age are not available, users are expected to set the chronological age in colData(se) to NA. If chronological age is not provided, the age acceleration residual will not be calculated. See package vignette for the definition of age acceleration residual. In addition, users are able to provide their own gene length using rowData(se). This is also optional. If using 'exprtype = "FPKM"', the provided gene length will be ignored. If provided, the column name for gene length in rowData(se) should be "bp_length". The function will convert raw count to FPKM by the user-supplied gene length. Otherwise, gene length is obtained from the internal database. See below for an example of se object.
tissue	a string indicate which tissue the gene expression data is obtained from. Users are expected to provide one of the following tissues. If the tissue argument is not provide or the provided tissue is not in this list, then the age predictor trained on all tissues will be used to calculate RNA age. • adipose_tissue • adrenal_gland • blood • blood_vessel • brain • breast

- colon
- esophagus
- heart
- liver
- lung
- muscle
- nerve
- ovary
- pancreas
- pituitary
- prostate
- salivary_gland
- skin
- small_intestine
- spleen
- stomach
- testis
- thyroid
- uterus
- vagina

`exprtype` either "counts" or "FPKM". For RPKM data, please use ‘exprtype‘ = "FPKM".

`idtype` a string which indicates the gene id type in ‘exprdata‘. It should be one of "SYMBOL", "ENSEMBL", "ENTREZID" or "REFSEQ". Default is "SYMBOL".

`stype` a string which specifies which version of pre-trained calculators to be used. It should be either "all" or "Caucasian". "all" means samples from all races (American Indian/Alaska Native, Asian, Black/African American, and Caucasian) are used to obtain the pre-trained calculator. "Caucasian" means only the Caucasian samples are used to build up the pre-trained calculator.

`signature` a string which indicates the age signature to use when calculating RNA age. This argument is not required.
In the case that this argument is not provided, if ‘tissue‘ argument is also provided and the tissue is in the list above, the tissue specific age signature given by our DESeq2 analysis result on GTEx data will be used. Otherwise, the across tissue signature "GTExAge" will be used.
In the case that this argument is provided, it should be one of the following signatures. A detailed description of the meaning of these signatures is given in the package vignette.

- DESeq2
- Pearson
- Dev
- deMagalhaes
- GenAge

- GTExAge
- Peters
- all

maxp the maxp argument used in impute.knn function. This is optional.

Value

a data frame contains RNA age.

Examples

```
library(SummarizedExperiment)
data(fpkmExample)
colData = data.frame(age = c(40, 50))
se = SummarizedExperiment(assays=list(FPKM=fpkm),
colData=colData)
res = predict_age_fromse(se = se, exprtype = "FPKM")
```

---

rawcount *An example of RNASeq counts data*

Description

An example of RNASeq counts data. It is the gene expression counts data from The genotype-tissue expression (GTEx) project brain samples. The data was downloaded from recount2(https://jhubiostatistics.shinyapps.io/recou For illustration purpose, only two samples were included. The R script to generate this example data can be found in inst/scripts/createExamples.R.

Usage

```
rawcount
```

Format

A data frame which contains the RNASeq counts data. The dimension of this data frame is 24989 by 2. Each row is a gene and each column is a sample.

References

Collado-Torres, Leonardo, et al. "Reproducible RNA-seq analysis using recount2." Nature biotechnology 35.4 (2017): 319-321.

Lonsdale, John, et al. "The genotype-tissue expression (GTEx) project." Nature genetics 45.6 (2013): 580.

Index

∗ **datasets**
fpkm, 3
rawcount, 9

count2FPKM, 2

fpkm, 3

getRPKM, *2*

impute.knn, *6*, *9*

makeplot, 3

predict_age, 4
predict_age_fromse, 7

rawcount, 9

SummarizedExperiment, *7*