Package ‘methylclock’

January 12, 2024

Type Package

Title Methylclock - DNA methylation-based clocks

Version 1.8.0

Description This package allows to estimate chronological and gestational DNA methylation (DNAm) age as well as biological age using different methylation clocks.

Chronological DNAm age (in years) : Horvath's clock, Hannum's clock, BNN, Horvath's skin+blood clock, PedBE clock and Wu's clock.

Gestational DNAm age : Knight's clock, Bohlin's clock, Mayne's clock and Lee's clocks.

Biological DNAm clocks : Levine's clock and Telomere Length's clock.

biocViews DNAMethylation, BiologicalQuestion, Preprocessing, StatisticalMethod, Normalization

License MIT + file LICENSE

Depends R (>= 4.1.0), methylclockData, devtools, quadprog

Imports Rcpp (>= 1.0.6), ExperimentHub, dplyr, impute,
       PerformanceAnalytics, Biobase, ggpmisc, tidyverse, ggplot2,
       ggrepur, minfi, tibble, RPMM, stats, graphics, tidyr, gridExtra,
       preprocessCore, dynamicTreeCut, planet

Suggests BiocStyle, knitr, GEOquery, rmarkdown

LinkingTo Rcpp

Encoding UTF-8

RoxygenNote 7.1.2

URL https://github.com/isglobal-brge/methylclock

BugReports https://github.com/isglobal-brge/methylclock/issues

VignetteBuilder knitr

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

git_branch RELEASE_3_18

git_last_commit bcf2548

git_last_commit_date 2023-10-24

Repository Bioconductor 3.18
checkClocks

Check whether input data contains the required CpGs for the implemented clocks.

Usage

checkClocks(x, ...)

Arguments

x

data.frame or tibble (Individual in columns, CpGs in rows, CpG names in first column - i.e. Horvath’s format), ExpressionSet or GenomicRatioSet. A matrix is also possible having the CpG names in the rownames.

... other parameters
checkClocksGA

Details
To be supplied

Value
a list with the different clocks when there are more than 80 the required CpGs

Examples
TestDataset <- get_TestDataset()
checkClocksGA(TestDataset)

Description
Check whether input data contains the required CpGs for the implemented clocks for Gestational Age.

Usage
checkClocksGA(x, ...)

Arguments
x          data.frame or tibble (Individual in columns, CpGs in rows, CpG names in first column - i.e. Horvath’s format), ExpressionSet or GenomicRatioSet. A matrix is also possible having the CpG names in the rownames.
...
other parameters

Details
To be supplied

Value
a list with the different GA clocks when there are more than 80

Examples
TestDataset <- get_TestDataset()
checkClocksGA(TestDataset)
commonClockCpgs | Get common CpGs

**Description**

Show the required CpGs contained on input data for the implemented clocks

**Usage**

commonClockCpgs(object, clock)

**Arguments**

- **object**
  - resulting object from checkClocks functions
- **clock**
  - string with the implemented clock, possible values are: "Knight", "Bohlin", "Mayne" and "Lee", "Horvath", "Hannum", "Levine", "skinHorvath", "PedBE", "Wu" and "TL"

**Value**

The common CpGs between input data and defined GA clock

**Examples**

```r
TestDataset <- get_TestDataset()
cpgs.missing.GA <- checkClocksGA(TestDataset)
cpgs.missing <- checkClocks(TestDataset)
commonClockCpgs(cpgs.missing.GA, "Bohlin")
commonClockCpgs(cpgs.missing, "Hannum")
```

**DNAmAge**

*DNAm age estimation using different DNA methylation clocks.*

**Description**

DNAm age estimation using different DNA methylation clocks.

**Usage**

DNAmAge(
  x,
  clocks = "all",
  toBetas = FALSE,
  fastImp = FALSE,
  normalize = FALSE,
  age,
)
DNAmAge

cell.count = TRUE,
cell.count.reference = "blood gse35069 complete",
min.perc = 0.8,
...
)

Arguments

x data.frame (Individual in columns, CpGs in rows, CpG names in first column - i.e. Horvath’s format), matrix (individuals in columns and Cpgs in rows having CpG names in the rownames), ExpressionSet or GenomicRatioSet.
clocks the methods used for estimating DNAmAge. Currently "Horvath", "Hannum", "Levine", "BNN", "skinHorvath", "PedBE", "Wu", "TL", "BLUP", "EN" and "all" are available. Default is "all" and all clocks are estimated.
toBetas Should data be transformed to beta values? Default is FALSE. If TRUE, it implies data are M values.
fastImp Is fast imputation performed if necessary? (see details). Default is FALSE
normalize Is Horvath’s normalization performed? By default is FALSE
age individual’s chronological age.
cell.count Are cell counts estimated? Default is TRUE.
cell.count.reference Used when 'cell.count' is TRUE. Default is "blood gse35069 complete". See 'mefil::mefil.list.cell.count.references()' for possible values.
min.perc Indicates the minimum coincidence percentage required between CpGs in or dataframe x and Cpgs in clock coefficients to perform the calculation. If min.perc is too low, the estimated gestational DNAm age can be poor
...
Other arguments to be passed through impute package

Details

Imputation is performed when having missing data. Fast imputation is performed by ... what about imputing only when CpGs for the clock are missing?

Value

The estimated chronological and biological mDNA age

Examples

MethylationData <- get_MethylationDataExample()
age.example55 <- DNAmAge(MethylationData)
DNAmGA

Gestational DNAm age estimation using different DNA methylation clocks.

Description

Gestational DNAm age estimation using different DNA methylation clocks.

Usage

```r
DNAmGA(
  x,
  toBetas = FALSE,
  fastImp = FALSE,
  normalize = FALSE,
  age,
  cell.count = TRUE,
  cell.count.reference = "andrews and bakulski cord blood",
  min.perc = 0.8,
  ...
)
```

Arguments

- `x`: data.frame (Individual in columns, CpGs in rows, CpG names in first column - i.e. Horvath’s format), matrix (individuals in columns and CpGs in rows having CpG names in the rownames), ExpressionSet or GenomicRatioSet.
- `toBetas`: Should data be transformed to beta values? Default is FALSE. If TRUE, it implies data are M values.
- `fastImp`: Is fast imputation performed if necessary? (see details). Default is FALSE.
- `normalize`: Is Horvath’s normalization performed? By default is FALSE.
- `age`: individual’s chronological age. Required to compute gestational age difference output.
- `cell.count`: Are cell counts estimated? Default is TRUE.
- `cell.count.reference`: Used when 'cell.count' is TRUE. Default is "blood gse35069 complete". See 'meffil::meffil.list.cell.count.references()' for possible values.
- `min.perc`: Indicates the minimum coincidence percentage required between CpGs in or dataframe x and CpGs in clock coefficients to perform the calculation. If min.perc is too low, the estimated gestational DNAm age can be poor.
- `...`: Other arguments to be passed through impute package

Details

Imputation is performed when having missing data. Fast imputation is performed by ... what about imputing only when CpGs for the clock are missing?
value

the estimated gestational DNAm age

examples

TestDataset <- get_TestDataset()
TestDataset[1:5, ]
ga.test <- DNAmGA(TestDataset)

description

get cell type reference

usage

getCellTypeReference(name)

arguments

name, string with predefined datasets andrews and bakulski cord blood, blood gse35069, blood gse35069 chen, blood gse35069 complete, "combined cord blood", "cord blood gse68456", "gervin and lyle cord blood", "guintivano dlpfc" or "saliva gse48472"

details

ORIGINAL AUTHOR: Matthew Suderman at github : https://github.com/perishky/meffil The original meffilListCellTypeReferences and getCellTypeReference function from meffil v1.0.0

value

name and reference.globals

examples

name <- "andrews and bakulski cord blood"
getCellTypeReference(name)
load_DNAm_Clocks_data

Description

Loads DNAm clock data from methylclockData

Usage

load_DNAm_Clocks_data()

Value

void

Examples

load_DNAm_Clocks_data()
mefilEstimateCellCountsFromBetas

Estimate cell counts for a beta matrix from a reference

Description

Estimate cell type ratios from methylation profiles of purified cell populations (Infinium Human-Methylation450 BeadChip).

Usage

mefilEstimateCellCountsFromBetas(beta, cellTypeReference, verbose = FALSE)

Arguments

beta
Matrix of Illumina 450K methylation levels (rows = CpG sites, columns = subjects).

cellTypeReference
Character string name of the cell type reference to use for estimating cell counts. See meffilListCellTypeReferences() for a list of available references. New references can be created using

verbose
If TRUE, then status messages are printed during execution (Default: FALSE).

Details

ORIGINAL AUTHOR: Matthew Suderman The original meffil.list.cellTypeReferences and get.cellTypeReference function from meffil v1.0.0 downloaded from github : https://github.com/perishky/meffil

Value

A matrix of cell count estimates.

Results should be nearly identical to minfi::estimateCellCounts()

betas

Examples

cell.count.reference <- "andrews and bakulski cord blood"
TestDataset <- get_TestDataset()
cpgs <- t(as.matrix(TestDataset[, -1]))
colnames(cpgs) <- TestDataset$CpGName
mefilEstimateCellCountsFromBetas(t(cpgs), cell.count.reference)
mefillListCellTypeReferences

*List of available cell type references*

**Description**

List of available cell type references

**Usage**

mefillListCellTypeReferences()

**Details**

ORIGINAL AUTHOR: Matthew Suderman The original mefillListCellTypeReferences and get-CellTypeReference function from mefill v1.0.0 at github: https://github.com/perishky/meffil

**Value**

a list with reference globals

**Examples**

mefillListCellTypeReferences()

---

methylclock

**Description**

Package to estimate DNA methylation age (DNAmAge) using different methylation clocks.

**Author(s)**

Juan R Gonzalez <juan.r.gonzalez@isglobal.org>
plotCorClocks

Description

Plot correlation among DNAm clock

Usage

plotCorClocks(x, ...)

Arguments

x a tibble or data.frame with the different DNAm clocks

... other arguments to be passs through function 'chart.Correlation' from 'PerformanceAnalytics' package

Details

To be supplied

Value

Plot with Correlation Clocks

Examples

library(Biobase)
library(GEOquery)

dd <- GEOquery::getGEO("GSE109446")
gse109446 <- dd[[1]]
controls <- Biobase::pData(gse109446)$"diagnosis:ch1" == "control"
gse <- gse109446[, controls]
age <- as.numeric(Biobase::pData(gse)$"age:ch1")
age.gse <- DNAmAge(gse, age = age)
plotCorClocks(age.gse)
plotDNAmAge

Plot DNAm age estimation vs chronological age.

Description

Plot DNAm age estimation vs chronological age.

Usage

plotDNAmAge(x, y, tit = "Horvath's method", clock = "chronological", ...)

Arguments

x   DNAm age estimation
y   Chronological age
tit  Plot title. Default is "Horvath's method".
clock  Type of clock 'chronological' or 'GA', default 'chronological'
... Other plot parameters for ggplot

Value

Plot with estimated DNAmAge

Examples

library(tidyverse)

path <- system.file("extdata", package = "methylclock")
covariates <- read_csv(file.path(
  path, 
  "SampleAnnotationExample55.csv"
))
age <- covariates$Age
MethylationData <- get_MethylationDataExample()

age.example55 <- DNAmAge(MethylationData)
plotDNAmAge(age.example55$Horvath, age)

Usage

data(progress_data)

Format

A data frame with 148 obs. and 151 variables

Details

A dataset containing data from the PROGRESS (Programming Research in Obesity, Growth, Environment and Social Stressors) cohort

Examples

data(progress_data)


Usage

data(progress_vars)

Format

A data frame with 150 obs. and 3 variables

Details

A dataset containing data from the PROGRESS (Programming Research in Obesity, Growth, Environment and Social Stressors) cohort
Examples

data(progress_vars)
Index

* datasets
  - progress_data, 13
  - progress_vars, 13
checkClocks, 2
checkClocksGA, 3
commonClockCpgs, 4
DNAmAge, 4
DNAmGA, 6
ggetCellTypeReference, 7
load_DNAm_Clocks_data, 8
load_DNAmGA_Clocks_data, 8
meffilEstimateCellCountsFromBetas, 9
meffilListCellTypeReferences, 9, 10
methylclock, 10
minfi::estimateCellCounts(), 9
plotCorClocks, 11
plotDNAmAge, 12
progress_data, 13
progress_vars, 13