Package 'MethPed'

November 12, 2026

Type Package

Version 1.40.0

Date 2016-01-01

Title A DNA methylation classifier tool for the identification of pediatric brain tumor subtypes

Depends R (>= 3.0.0), Biobase

Suggests BiocStyle, knitr, markdown, impute

Imports randomForest, grDevices, graphics, stats

VignetteBuilder knitr

biocViews ImmunoOncology, DNAMethylation, Classification, Epigenetics

Encoding UTF-8

Description Classification of pediatric tumors into biologically defined subtypes is challenging and multifaceted approaches are needed. For this aim, we developed a diagnostic classifier based on DNA methylation profiles. We offer MethPed as an easy-to-use toolbox that allows researchers and clinical diagnosticians to test single samples as well as large cohorts for subclass prediction of pediatric brain tumors. The current version of MethPed can classify the following tumor diagnoses/subgroups: Diffuse Intrinsic Pontine Glioma (DIPG), Ependymoma, Embryonal tumors with multilayered rosettes (ETMR), Glioblastoma (GBM), Medulloblastoma (MB) - Group 3 (MB_Gr3), Group 4 (MB_Gr3), Group WNT (MB_WNT), Group SHH (MB_SHH) and Pilocytic Astrocytoma (PiloAstro).

License GPL-2

RoxygenNote 5.0.1

NeedsCompilation no

Author Mohammad Tanvir Ahamed [aut, trl], Anna Danielsson [aut], Szilárd Nemes [aut, trl], Helena Carén [aut, cre, cph]

Maintainer Helena Carén <helenacarenlab@gmail.com>

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

git_branch RELEASE_3_23

git_last_commit 0d2a1ca

git_last_commit_date 2026-04-28

Repository Bioconductor 3.23

Date/Publication 2026-11-11

Contents

checkNA 2
MethPed 3
MethPed_900probes 5
MethPed_sample 6
plot.methped 7
probeMis 8
summary.methped 9

Index **10**

---

checkNA *Check missing valus in dataset*

Description

Check missing valus in dataset

Usage

```
checkNA(data)
```

Arguments

data	Object in class data.frame or matrix or ExpressionSet.

Examples

```
#################### Loading and view sample data
data(MethPed_sample)
head(MethPed_sample,10)

################## Check missing value
checkNA(MethPed_sample)
```

`MethPed` *MethPed: a DNA methylation classifier tool for the identification of pediatric brain tumor subtypes.*

Description

The current version of MethPed can classify the following tumor diagnoses/subgroups: Diffuse Intrinsic Pontine Glioma (DIPG), Ependymoma, Embryonal tumors with multilayered rosettes (ETMR), Glioblastoma (GBM), Medulloblastoma (MB) - Group 3 (MB_Gr3), Group 4 (MB_Gr3), Group WNT (MB_WNT), Group SHH (MB_SHH) and Pilocytic Astrocytoma (PiloAstro).

Usage

```
MethPed(data, TargetID="TargetID", prob = TRUE,...)
```

Arguments

`data`	Data for classification. Data can be in class "ExpressionSet", "matrix" or "data.frame". See MethPed vignette for details.
`TargetID`	Name of the "Probe" column in data
`prob`	If 'TRUE' (Default value), the return value will be conditional probability from which tumor group the sample belongs to. 'FALSE' will only return value for the tumor group which has the maximum conditional probability for a sample and other non-maximum values will be zero.
`...`	More parameter to add

Details

Classification of pediatric tumors into biologically defined subtypes is challenging, and multifaceted approaches are needed. For this aim, we developed a diagnostic classifier based on DNA methylation profiles. We offer MethPed as an easy-to-use toolbox that allows researchers and clinical diagnosticians to test single samples as well as large cohorts for subclass prediction of pediatric brain tumors.

The MethPed classifier uses the Random Forest (RF) algorithm to classify unknown pediatric brain tumor samples into sub-types. The classification proceeds with the selection of the beta values needed for the classification.

The computational process proceeded in two stages. The first stage commences with a reduction of the probe pool or building the training probe dataset for classification. We have named this dataset as "predictors". The second stage is to apply the RF algorithm to classify the probe data of interest based on the training probe dataset (predictors).

For the construction of the training probe pool (predictors), methylation data generated by the Illumina Infinium HumanMethylation 450 BeadChip arrays were downloaded from the Gene Expression Omnibus (GEO). Four hundred seventy-two cases were available, representing several brain tumor diagnoses (DIPG, glioblastoma, ETMR, medulloblastoma, ependymoma, pilocytic astrocytoma) and their further subgroups.

The data sets were merged and probes that did not appear in all data sets were filtered away. In addition, about 190,000 CpGs were removed due to SNPs, repeats and multiple mapping sites. The final data set contained 206,823 unique probes and nine tumor classes including the medulloblastoma subgroups. K–neighbor imputation was used for missing probe data.

After that, a large number of regression analyses were performed to select the 100 probes per tumor class that had the highest predictive power (AUC values). Based on the identified 900 methylation sites, the nine pediatric brain tumor types could be accurately classified using the multiclass classification algorithm MethPed.

Value

The output of the algorithm is partitioned in 6 parts :

"$target_id" Probes name in main data set

"$probes" Number of probes

"$sample" Number of samples

"$probes_missing" Names of the probes that were included in the original classifier but are missing from the data at hand.

$oob_err.rate If a large number of probes are missing this could potentially lead to a drop in the predictive accuracy of the model. Note that, the out-of-bag error is 1.7 percent of the original MethPed classifier.

$predictions The last and most interesting output contains the predictions. We chose not to assign tumors to one or other group but give probabilities of belonging in one or other subtype. See vignette for details.

Author(s)

Anna Danielsson, Mohammad Tanvir Ahamed, Szilárd Nemes and Helena Carén, University of Gothenburg, Sweden.

References

[1] Anna Danielsson, Szilárd Nemes, Magnus Tisell, Birgitta Lannering, Claes Nordborg, Magnus Sabel, and Helena Carén. "MethPed: A DNA Methylation Classifier Tool for the Identification of Pediatric Brain Tumor Subtypes". Clinical Epigenetics 2015, 7:62, 2015

[2] Breiman, Leo (2001). "Random Forests". Machine Learning 45 (1): 5-32. doi:10.1023/A:1010933404324

[3] Troyanskaya, O., M. Cantor, G. Sherlock, P. Brown, T. Hastie, R. Tibshirani, D. Botstein, and R. B. Altman. "Missing Value Estimation Methods for DNA Microarrays." Bioinformatics 17.6 (2001): 520-25.

See Also

See http://www.clinicalepigeneticsjournal.com/content/7/1/62 for more details.

Examples

```
#################### Loading and view sample data
data(MethPed_sample)
head(MethPed_sample,10)

#################### Check dimention of sample data
dim(MethPed_sample)  # Check number pof probes and samples in data

#################### Checking missing value in the data
missingIndex <- checkNA(MethPed_sample)

#################### Apply MethPed to sample data (Probability for all tumor group)
myClassification<-MethPed(MethPed_sample)
myClassification<-MethPed(MethPed_sample,prob=TRUE)

#################### Apply MethPed to sample data (Only maximum probability expected)
myClassification_max<-MethPed(MethPed_sample,prob=FALSE)

#################### Summary of results
summary(myClassification)
summary(myClassification)

#################### Barplot of conditional prediction probability on different samples
par(mai = c(1, 1, 1, 2), xpd=TRUE)
mat<-t(myClassification$predictions)
mycols <- c("green",rainbow(nrow(mat),start=0,end=1)[nrow(mat):1],"red")
barplot(mat,  col = mycols,
        beside=FALSE,axisnames=TRUE,
        ylim=c(0,1),xlab= "Sample",ylab="Probability")
legend( ncol(mat)+0.5,1,
       legend = rownames(mat),fill = mycols,xpd=TRUE, cex = 0.6)

## Generic function to plot
plot(myClassification) # myClassification should be an object in "methPed" class
```

`MethPed_900probes` *List of 900 probes in the predictor data*

Description

List of 900 probes in the predictor (Training data for Random Forest model)

Usage

```
data(MethPed_900probes)
```

Format

Data frame of 900 probes.

Details

For the construction of the training probe pool (predictors), methylation data generated by the Illumina Infinium HumanMethylation 450 BeadChip arrays were downloaded from the Gene Expression Omnibus (GEO). Four hundred seventy-two cases were available, representing several brain tumor diagnoses (DIPG, glioblastoma, ETMR, medulloblastoma, ependymoma, pilocytic astrocytoma) and their further subgroups.

The data sets were merged and probes that did not appear in all data sets were filtered away. In addition, about 190,000 CpGs were removed due to SNPs, repeats and multiple mapping sites. The final data set contained 206,823 unique probes and nine tumor classes including the medulloblastoma subgroups. K–neighbor imputation was used for missing probe data.

After that, a large number of regression analyses were performed to select the 100 probes per tumor class that had the highest predictive power (AUC values). Based on the identified 900 methylation sites, the nine pediatric brain tumor types could be accurately classified using the multiclass classification algorithm MethPed.

Value

450k methylation array probe name

References

[1] Anna Danielsson, Szilárd Nemes, Magnus Tisell, Birgitta Lannering, Claes Nordborg, Magnus Sabel, and Helena Carén. "MethPed: A DNA Methylation Classifier Tool for the Identification of Pediatric Brain Tumor Subtypes". Clinical Epigenetics 2015, 7:62, 2015

See Also

See http://www.clinicalepigeneticsjournal.com/content/7/1/62 for more details.

Examples

```
#################### Loading and view sample data
data(MethPed_900probes)
head(MethPed_900probes)
```

`MethPed_sample` *Sample dataset to MethPed package.*

Description

Methylation beta-values generated with the Infinium HumanMethylation450 BeadChips (Illumina).

Usage

```
data(MethPed_sample)
```

Format

A data frame with 468821 probes and 2 tumor samples

Value

DNA Methylation beta-values

References

[1] Anna Danielsson, Szilárd Nemes, Magnus Tisell, Birgitta Lannering, Claes Nordborg, Magnus Sabel, and Helena Carén. "MethPed: A DNA Methylation Classifier Tool for the Identification of Pediatric Brain Tumor Subtypes". Clinical Epigenetics 2015, 7:62, 2015

See Also

See http://www.clinicalepigeneticsjournal.com/content/7/1/62 for more details.

Examples

```
#################### Loading and view sample data
data(MethPed_sample)
head(MethPed_sample)

#################### Check dimention of sample data
dim(MethPed_sample)  # Check number of probes and samples in data

#################### Checking missing value in the data
missingIndex <- checkNA(MethPed_sample)
```

`plot.methped` *Plot conditional probability of samples that belongs to different tumor subtypes.*

Description

Plot conditional probability of samples that belongs to different tumor subtypes.

Usage

```
## S3 method for class 'methped'
plot(x, ...)
```

Arguments

`x`	Object in "methped" class. Output of function MethPed.
`...`	More arguments from function barplot.

Value

Object in "methped" class. Output of function MethPed.

Examples

```
#################### Loading sample data
data(MethPed_sample)

#################### Applying MethPed to sample data
res<-MethPed(MethPed_sample)

#################### Plot conditional probability
plot(res)
```

---

probeMis *Missing probe names from training data compared with input samples.*

Description

Missing probe names from training data compared with input samples.

Usage

```
probeMis(x)
```

Arguments

x	Object in MethPed class. Output of function MethPed.

Value

Object in "methped" class. Output of function MethPed.

Examples

```
#################### Loading sample data
data(MethPed_sample)

#################### Applying MethPed to sample data
res<-MethPed(MethPed_sample)

#################### The names of the probes that were included in the original
                  # classifier but are missing from the data at hand
probeMis(res)
```

summary.methped *Summary of conditional probability or binary classification of samples that belong to different tumor subtypes.*

Description

Summary of conditional probability or binary classification of samples that belong to different tumor subtypes.

Usage

```
## S3 method for class 'methped'
summary(object, ...)
```

Arguments

object	Object in methped class. Output of function MethPed.
...	Additional arguments affecting the summary produced

Value

Object in "methped" class. Output of function MethPed.

Examples

```
#################### Loading sample data
data(MethPed_sample)

#################### Applying MethPed to sample data
res<-MethPed(MethPed_sample)

#################### Summary function of MethPed output
summary (res)
```

Index

∗ **data**
MethPed_900probes, 5
MethPed_sample, 6

barplot, 7

checkNA, 2

MethPed, 3, 7–9
MethPed_900probes, 5
MethPed_sample, 6

plot.methped, 7
probeMis, 8

summary.methped, 9