

Package ‘planet’

October 18, 2024

Title Placental DNA methylation analysis tools

Version 1.13.0

URL <https://victor.rbind.io/planet>, <http://github.com/wvictor14/planet>

BugReports <http://github.com/wvictor14/planet/issues>

Description This package contains R functions to predict biological variables to from placental DNA methylation data generated from infinium arrays. This includes inferring ethnicity/ancestry, gestational age, and cell composition from placental DNA methylation array (450k/850k) data.

Depends R (>= 4.3)

Imports methods, tibble, magrittr, dplyr

Suggests ggplot2, testthat, tidyr, scales, minfi, EpiDISH, knitr, rmarkdown

License GPL-2

Encoding UTF-8

LazyData false

RoxygenNote 7.2.3

VignetteBuilder knitr

biocViews Software, DifferentialMethylation, Epigenetics, Microarray, MethylationArray, DNAMethylation, CpGIsland

Roxygen list(markdown = TRUE)

git_url <https://git.bioconductor.org/packages/planet>

git_branch devel

git_last_commit 6513ac4

git_last_commit_date 2024-04-30

Repository Bioconductor 3.20

Date/Publication 2024-10-17

Author Victor Yuan [aut, cre],
Wendy P. Robinson [ctb]

Maintainer Victor Yuan <victor.2wy@gmail.com>

Contents

planet-package	2
ageCpGs	3
ethnicityCpGs	3
planet-deprecated	3
plBetas	4
plCellCpGsFirst	4
plCellCpGsThird	5
plColors	5
plPhenoData	6
predictAge	6
predictEthnicity	7
%>%	8
Index	9

planet-package

planet: Placental DNA methylation analysis tools

Description

This package contains R functions to predict biological variables to from placental DNA methylation data generated from infinium arrays. This includes inferring ethnicity/ancestry, gestational age, and cell composition from placental DNA methylation array (450k/850k) data.

Author(s)

Maintainer: Victor Yuan <victor.2wy@gmail.com>

Other contributors:

- Wendy P. Robinson [contributor]

See Also

Useful links:

- <https://victor.rbind.io/planet>
- <http://github.com/wvictor14/planet>
- Report bugs at <http://github.com/wvictor14/planet/issues>

ageCpGs	<i>Placental gestational age CpGs</i>
---------	---------------------------------------

Description

Coefficients from the three placental gestational age clocks from [Lee Y et al. 2019](#).

Reference: Lee Y, Choufani S, Weksberg R, et al. Placental epigenetic clocks: estimating gestational age using placental DNA methylation levels. *Aging (Albany NY)*. 2019;11(12):4238–4253. doi:10.18632/aging.102049. PMID: 31235674

Usage

```
data(ageCpGs)
```

Format

A [tibble](#) with coefficients for the RPC, CPC, and refined RPC.

ethnicityCpGs	<i>CpGs to predict ethnicity</i>
---------------	----------------------------------

Description

1860 CpGs used to predict ethnicity.

See [Yuan et al. 2019](#) for details.

Usage

```
data(ethnicityCpGs)
```

Format

A character vector of length 1860

Source

<https://pubmed.ncbi.nlm.nih.gov/31399127/>

planet-deprecated	<i>Deprecated functions in planet</i>
-------------------	--

Description

These functions still work but will be removed (defunct) in the next version.

Details

- `pl_infer_ethnicity`: This function has been renamed `predictEthnicity`
- `pl_infer_age`: This function has been renamed `predictAge`

plBetas

Example placental DNA methylation data

Description

6 DNA methylation profiles from preeclampsia and healthy control placentas. This data was downloaded from:

- [GSE75196](#)

"Genome wide DNA methylation profiling of normal and preeclampsia placental samples. Illumina Infinium HumanMethylation450 BeadChip (450K array) was used to obtain DNA methylation profiles in placental samples. Samples included 16 samples from healthy uncomplicated pregnancies and 8 samples from pregnancies affected by preeclampsia." - from [Yeung et al.](#)

The DNA methylation data for 24 placental samples were downloaded from [GSE75196](#). After normalizing using `minfi::preprocessNoob` and `watermelon::BMIQ`, the data were filtered to 6/24 samples and 10,000 random CpGs + those CpGs used in the gestational age clock and ethnicity classifier.

Reference: Yeung KR, Chiu CL, Pidsley R, Makris A et al. DNA methylation profiles in preeclampsia and healthy control placentas. *Am J Physiol Heart Circ Physiol* 2016 May 15;310(10):H1295-303. [PMID:26968548](#)

Usage

```
data(plBetas)
```

Format

A matrix

Source

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE75196>

plCellCpGsFirst

First trimester placental cell type coefficients

Description

First trimester coefficients for placental cellular deconvolution from [YuanVetal.2020](#).

Reference: to be edited PMID: to be edited

Usage

```
data(plCellCpGsFirst)
```

Format

A [matrix](#) with coefficients for Trophoblasts, Stromal, Endothelial, Hofbauer cells, nRBCs, and Syncytiotrophoblasts.

plCellCpGsThird	<i>Third trimester placental cell type coefficients</i>
-----------------	---

Description

Third trimester coefficients for placental cellular deconvolution from [YuanVetal.2020](#).

Reference: to be edited PMID: to be edited

Usage

```
data(plCellCpGsThird)
```

Format

A [matrix](#) with coefficients for Trophoblasts, Stromal, Endothelial, Hofbauer cells, nRBCs, and Syncytiotrophoblasts.

plColors	<i>A color palette for placental cell types</i>
----------	---

Description

A nice color palette for placental cell types.

Used in [YuanVetal.2020](#).

Contains colors for:

- Syncytiotrophoblast
- Trophoblast
- Stromal
- Hofbauer
- Endothelial
- nRBCs

Usage

```
data(plColors)
```

Format

An object of class character of length 6.

plPhenoData	<i>Sample information accompanying pl_betas</i>
-------------	---

Description

Sex, disease, and gestational age information associated with pl_betas.

Downloaded from the GEO accession:

- [GSE75196](#)

Reference: Yeung KR, Chiu CL, Pidsley R, Makris A et al. DNA methylation profiles in preeclampsia and healthy control placentas. *Am J Physiol Heart Circ Physiol* 2016 May 15;310(10):H1295-303. PMID: [26968548](#)

Usage

```
data(plPhenoData)
```

Format

A [tibble](#)

Source

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE75196>

predictAge	<i>Predicts gestational age using placental DNA methylation microarray data</i>
------------	---

Description

predictAge Multiplies the coefficients from one of three epigenetic gestational age clocks, by the corresponding CpGs in a supplied betas data.frame.

Usage

```
predictAge(betas, type = "RPC")
```

Arguments

betas	An n by m dataframe of methylation values on the beta scale (0, 1), where the CpGs are arranged in rows, and samples in columns. Should contain all CpGs used in each clock
type	One of the following: "RPC" (Robust), "CPC", (Control) or "RRPC" (Refined Robust).

Details

Predicts gestational age using one of 3 placental gestational age clocks: RPC, CPC, or refined RPC. Requires placental DNA methylation measured on the Infinium 27K/450k/EPIC methylation array. Ensure as many predictive CpGs are present in your data, otherwise accuracy may be impacted.

It's recommended that you have all predictive CpGs, otherwise accuracy may vary.

Value

A vector of length m , containing inferred gestational age.

Examples

```
# Load placenta DNAm data
library(dplyr)
data(plBetas)
data(plPhenoData)

plPhenoData %>%
  mutate(inferred_ga = predictAge(plBetas, type = "RPC"))
```

predictEthnicity

Predicts ethnicity using placental DNA methylation microarray data

Description

Uses 1860 CpGs to predict self-reported ethnicity on placental microarray data.

Usage

```
predictEthnicity(betas, threshold = 0.75, force = FALSE)
```

Arguments

betas	$n \times m$ dataframe of methylation values on the beta scale (0, 1), where the variables are arranged in rows, and samples in columns. Should contain all 1860 predictors and be normalized with NOOB and BMIQ.
threshold	A probability threshold ranging from (0, 1) to call samples 'ambiguous'. Defaults to 0.75.
force	run even if missing predictors. Default is FALSE.

Details

Predicts self-reported ethnicity from 3 classes: Africans, Asians, and Caucasians, using placental DNA methylation data measured on the Infinium 450k/EPIC methylation array. Will return membership probabilities that often reflect genetic ancestry composition.

The input data should contain all 1860 predictors (cpgs) of the final GLMNET model.

It's recommended to use the same normalization methods used on the training data: NOOB and BMIQ.

Value

a [tibble](#)

Examples

```
## To predict ethnicity on 450k/850k samples

# Load placenta DNAm data
data(plBetas)
predictEthnicity(plBetas)
```

%>%

Pipe operator

Description

See [magrittr::%>%](#) for details.

Usage

```
lhs %>% rhs
```

Value

lhs

Examples

```
c(1, 2, 3) %>% sum()
```

Index

* datasets

- ageCpGs, 3
- ethnicityCpGs, 3
- plBetas, 4
- plCellCpGsFirst, 4
- plCellCpGsThird, 5
- plColors, 5
- plPhenoData, 6

* internal

- %>%, 8
- planet-package, 2
- %>%, 8, 8

ageCpGs, 3

ethnicityCpGs, 3

matrix, 4, 5

pl_infer_age, 3

pl_infer_age (predictAge), 6

pl_infer_ethnicity, 3

pl_infer_ethnicity (predictEthnicity), 7

planet (planet-package), 2

planet-deprecated, 3

planet-package, 2

plBetas, 4

plCellCpGsFirst, 4

plCellCpGsThird, 5

plColors, 5

plPhenoData, 6

predictAge, 3, 6

predictEthnicity, 3, 7

tibble, 3, 6, 8