

# Package ‘planet’

April 9, 2025

**Title** Placental DNA methylation analysis tools

**Version** 1.15.5

**URL** <http://github.com/wvictor14/planet>, <http://victoryuan.com/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** ExperimentHub, mixOmics, ggplot2, testthat, tidyr, scales, minfi, EpiDISH, knitr, rmarkdown

**License** GPL-2

**Encoding** UTF-8

**LazyData** false

**RoxygenNote** 7.3.2

**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** faa2413

**git\_last\_commit\_date** 2025-02-01

**Repository** Bioconductor 3.21

**Date/Publication** 2025-04-08

**Author** Victor Yuan [aut, cre],  
Wendy P. Robinson [aut, ctb],  
Icíaar Fernández-Boyano [aut, ctb]

**Maintainer** Victor Yuan <[victor.2wy@gmail.com](mailto:victor.2wy@gmail.com)>

## Contents

planet-package . . . . .	2
ageCpGs . . . . .	3
ethnicityCpGs . . . . .	3
plBetas . . . . .	4
plCellCpGsFirst . . . . .	4
plCellCpGsThird . . . . .	5
plColors . . . . .	5
plPhenoData . . . . .	6
predict . . . . .	6
predictAge . . . . .	8
predictEthnicity . . . . .	9
predictPreeclampsia . . . . .	10
%>% . . . . .	11
<b>Index</b>	<b>12</b>

---

planet-package	<i>planet: Placental DNA methylation analysis tools</i>
----------------	---

---

## 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>

Authors:

- Wendy P. Robinson [contributor]
- Icíar Fernández-Boyano <iciarfernandez@outlook.com> [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`*Placental gestational age CpGs*

---

**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`*CpGs to predict ethnicity*

---

**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/>

---

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      *Third trimester placental cell type coefficients*

---

**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      *A color palette for placental cell types*

---

**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

*Sample information accompanying pl\_betas*

---

**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>

---

predict

*predict mixxo from*

---

**Description**

copied from mixOmicsTeam/mixOmics/refs/heads/master/R/predict.R on 2025 Jan 30. Some components omitted that are not used in planet.

**Usage**

```
## S3 method for class 'mixo_pls'  
predict(  
  object,  
  newdata,  
  study.test,  
  dist = c("all", "max.dist", "centroids.dist", "mahalanobis.dist"),  
  multilevel = NULL,  
  ...  
)
```

```
## S3 method for class 'mixo_spls'
predict(
  object,
  newdata,
  study.test,
  dist = c("all", "max.dist", "centroids.dist", "mahalanobis.dist"),
  multilevel = NULL,
  ...
)
```

### Arguments

<code>object</code>	object of class inheriting from " <code>(mint).(block).(s)pls(da)</code> ".
<code>newdata</code>	data matrix in which to look for explanatory variables to be used for prediction. Please note that this method does not perform multilevel decomposition or log ratio transformations, which need to be processed beforehand.
<code>study.test</code>	For MINT objects, grouping factor indicating which samples of <code>newdata</code> are from the same study. Overlap with <code>object\$study</code> are allowed.
<code>dist</code>	distance to be applied for discriminant methods to predict the class of new data, should be a subset of " <code>centroids.dist</code> ", " <code>mahalanobis.dist</code> " or " <code>max.dist</code> " (see Details). Defaults to " <code>all</code> ".
<code>multilevel</code>	Design matrix for multilevel analysis (for repeated measurements). A numeric matrix or data frame. For a one level factor decomposition, the input is a vector indicating the repeated measures on each individual, i.e. the individuals ID. For a two level decomposition with <code>splsda</code> models, the two factors are included in <code>Y</code> . Finally for a two level decomposition with <code>spls</code> models, 2nd AND 3rd columns in design indicate those factors (see example in <code>?splsda</code> and <code>?spls</code> ).
<code>...</code>	not used currently.

### Value

`predict` produces a list with the following components:

<code>predict</code>	predicted response values. The dimensions correspond to the observations, the response variables and the model dimension, respectively. For a supervised model, it corresponds to the predicted dummy variables.
<code>variates</code>	matrix of predicted variates.
<code>B.hat</code>	matrix of regression coefficients (without the intercept).
<code>AveragedPredict</code>	if more than one block, returns the average predicted values over the blocks (using the <code>predict</code> output)
<code>WeightedPredict</code>	if more than one block, returns the weighted average of the predicted values over the blocks (using the <code>predict</code> and <code>weights</code> outputs)
<code>class</code>	predicted class of <code>newdata</code> for each 1, ..., <code>ncomp</code> components.

MajorityVote	if more than one block, returns the majority class over the blocks. NA for a sample means that there is no consensus on the predicted class for this particular sample over the blocks.
WeightedVote	if more than one block, returns the weighted majority class over the blocks. NA for a sample means that there is no consensus on the predicted class for this particular sample over the blocks.
weights	Returns the weights of each block used for the weighted predictions, for each nrepeat and each fold
centroids	matrix of coordinates for centroids.
dist	type of distance requested.
vote	majority vote result for multi block analysis (see details above).

**Author(s)**

Florian Rohart, Sébastien Déjean, Ignacio González, Kim-Anh Lê Cao, Al J Abadi

**References**

Rohart F, Gautier B, Singh A, Lê Cao K-A. mixOmics: an R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol* 13(11): e1005752

Tenenhaus, M. (1998). *La regression PLS: theorie et pratique*. Paris: Editions Technic.

**See Also**

<http://www.mixOmics.org> for more details.

**Examples**

```
# example code
```

---

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	<i>Predicts ethnicity using placental DNA methylation microarray data</i>
------------------	---

---

**Description**

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

**Usage**

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

**Arguments**

betas	n x 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)
```

---

predictPreeclampsia    *predictPreeclampsia*

---

**Description**

Uses 45 CpGs to predict early preeclampsia (PE delivered before or at 34 weeks of gestation) on placental DNA methylation microarray data.

**Usage**

```
predictPreeclampsia(betas, ...)
```

**Arguments**

betas	matrix or array of methylation values on the beta scale (0, 1), where the variables are arranged in rows, and samples in columns.
...	feeds into outersect function

**Details**

Assigns the class labels "early-PE" or "normotensive" to each sample and returns a class probability.

**Value**

produces a list with components detailed in the `mixOmics::predict` R documentation

**It is recommended that users apply beta-mixture quantile normalization (BMIQ) to their data** prior to prediction. This was the normalization method used on the training data.

### Examples

```
# To predict early preeclampsia on 450k/850k samples

# Load data
library(ExperimentHub)
eh <- ExperimentHub()
query(eh, "eoPredData")

# test object
x_test <- eh[['EH8403']]
x_test %>% predictPreeclampsia()
```

---

%>%

*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**
  - %>%, [11](#)
  - planet-package, [2](#)
- \* **multivariate**
  - predict, [6](#)
- \* **regression**
  - predict, [6](#)
- %>%, [11](#), [11](#)
- ageCpGs, [3](#)
- ethnicityCpGs, [3](#)
- matrix, [5](#)
- planet (planet-package), [2](#)
- planet-package, [2](#)
- plBetas, [4](#)
- plCellCpGsFirst, [4](#)
- plCellCpGsThird, [5](#)
- plColors, [5](#)
- plPhenoData, [6](#)
- predict, [6](#)
- predictAge, [8](#)
- predictEthnicity, [9](#)
- predictPreeclampsia, [10](#)
- tibble, [3](#), [6](#), [10](#)