

Package ‘AlphaBeta’

October 16, 2020

Type Package

Title Computational inference of epimutation rates and spectra from high-throughput DNA methylation data in plants

Version 1.2.3

Description

AlphaBeta is a computational method for estimating epimutation rates and spectra from high-throughput DNA methylation data in plants.

The method has been specifically designed to:

1. analyze 'germline' epimutations in the context of multi-generational mutation accumulation lines (MA-lines).
2. analyze 'somatic' epimutations in the context of plant development and aging.

License GPL-3

Depends R (>= 3.6.0)

Imports dplyr (>= 0.7), data.table (>= 1.10), stringr (>= 1.3), utils (>= 3.6.0), gtools (>= 3.8.0), optimx (>= 2018-7.10), expm (>= 0.999-4), stats (>= 3.6), BiocParallel (>= 1.18), igraph (>= 1.2.4), graphics (>= 3.6), ggplot2 (>= 3.2), grDevices (>= 3.6), plotly (>= 4.9)

Encoding UTF-8

LazyData false

LazyLoad yes

VignetteBuilder knitr

RoxygenNote 7.1.0

Suggests knitr, rmarkdown

biocViews Epigenetics, FunctionalGenomics, Genetics, MathematicalBiology

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

git_branch RELEASE_3_11

git_last_commit 666df14

git_last_commit_date 2020-06-06

Date/Publication 2020-10-16

Author Yadollah Shahryary Dizaji [cre, aut],
Frank Johannes [aut],
Rashmi Hazarika [aut]

Maintainer Yadollah Shahryary Dizaji <shahryary@gmail.com>

R topics documented:

ABneutral	2
ABneutralSOMA	3
ABnull	4
ABplot	5
ABselectMM	6
ABselectMMSOMA	7
ABselectUU	8
ABselectUUSOMA	9
BOOTmodel	10
buildPedigree	10
dMatrix	11
FtestRSS	12
plotPedigree	12
rc.meth.lvl	13
Index	15

 ABneutral

Run Model with no selection (ABneutral)

Description

This model assumes that heritable gains and losses in cytosine methylation are selectively neutral.

Usage

```
ABneutral(pedigree.data, p0uu, eqp, eqp.weight, Nstarts, out.dir, out.name)
```

Arguments

pedigree.data	pedigree data.
p0uu	initial proportion of unmethylated cytosines.
eqp	equilibrium proportion of unmethylated cytosines.
eqp.weight	weight assigned to equilibrium function.
Nstarts	iterations for non linear LSQ optimization.
out.dir	output directory.
out.name	output file name.

Value

ABneutral RData file.

Examples

```
#Get some toy data
inFile <- readRDS(system.file("extdata/dm/", "output.rds", package="AlphaBeta"))
pedigree <- inFile$Pdata
p0uu_in <- inFile$tmp0
eqp.weight <- 1
Nstarts <- 2
out.name <- "CG_global_estimates_ABneutral"
out <- ABneutral(pedigree.data = pedigree,
                 p0uu=p0uu_in,
                 eqp=p0uu_in,
                 eqp.weight=eqp.weight,
                 Nstarts=Nstarts,
                 out.dir=getwd(),
                 out.name=out.name)

summary(out)
```

ABneutralSOMA

Model with no selection (outneutral)

Description

This model assumes that somatically heritable gains and losses in cytosine methylation are selectively neutral.

Usage

```
ABneutralSOMA(pedigree.data, p0uu, eqp, eqp.weight, Nstarts, out.dir, out.name)
```

Arguments

pedigree.data	pedigree data.
p0uu	initial proportion of unmethylated cytosines.
eqp	equilibrium proportion of unmethylated cytosines.
eqp.weight	weight assigned to equilibrium function.
Nstarts	iterations for non linear LSQ optimization.
out.dir	output directory.
out.name	output file name.

Value

ABneutralSoma RData file.

Examples

```
#Get some toy data
inFile <- readRDS(system.file("extdata/soma/", "outputSoma.rds", package="AlphaBeta"))
pedigree <- inFile$Pdata
p0uu_in <- inFile$tmp0
eqp.weight <- 0.001
Nstarts <- 2
out.name <- "ABneutralSOMA_CG_estimates"
out <- ABneutralSOMA(pedigree.data = pedigree,
                    p0uu=p0uu_in,
                    eqp=p0uu_in,
                    eqp.weight=eqp.weight,
                    Nstarts=Nstarts,
                    out.dir=getwd(),
                    out.name=out.name)

summary(out)
```

ABnull

Run model that considers no accumulation of epimutations (ABnull)

Description

Run model that considers no accumulation of epimutations (ABnull)

Usage

```
ABnull(pedigree.data, out.dir, out.name)
```

Arguments

pedigree.data	Generation table name, you can find sample file in
out.dir	outputdirectory
out.name	name of file

Value

ABnull RData file.

Examples

```
#Get some toy data
inFile <- readRDS(system.file("extdata/dm/", "output.rds", package="AlphaBeta"))
pedigree <- inFile$Pdata
out.name <- "CG_global_estimates_ABnull"
out <- ABnull(pedigree.data = pedigree,
             out.dir=getwd(),
             out.name=out.name)

summary(out)
```

 ABplot

Plotting estimates

Description

Plotting Estimating epimutation

Usage

```
ABplot(
  pedigree.names,
  output.dir,
  out.name,
  alpha = 0.5,
  geom.point.size = 2,
  geom.line.size = 0.9,
  plot.height = 8,
  plot.width = 11,
  plot.type = "both",
  lsq.line = "theory",
  interact = FALSE
)
```

Arguments

pedigree.names	Models output AB*.Rdata
output.dir	output directory
out.name	filename
alpha	ggplot parameters
geom.point.size	ggplot parameters
geom.line.size	ggplot parameters
plot.height	ggplot parameters
plot.width	ggplot parameters
plot.type	type of plot (data.only, fit.only, both)
lsq.line	Least Square Regression line (theory or pred)
interact	to see intarctive plot. (useing plotly)

Value

plot

Examples

```
# Get some toy data
file <- system.file("extdata/dm/", "Col_CG_global_estimates_ABneutral.Rdata", package="AlphaBeta")
ABplot(pedigree.names=file, output.dir=getwd(), out.name="ABneutral")
```

ABselectMM	<i>Run model with selection against spontaneous gain of methylation (ABselectMM)</i>
------------	--

Description

This model assumes that heritable losses of cytosine methylation are under negative selection.

Usage

```
ABselectMM(pedigree.data, p0uu, eqp, eqp.weight, Nstarts, out.dir, out.name)
```

Arguments

pedigree.data	pedigree data.
p0uu	initial proportion of unmethylated cytosines.
eqp	equilibrium proportion of unmethylated cytosines.
eqp.weight	nweight assigned to equilibrium function.
Nstarts	iterations for non linear LSQ optimization.
out.dir	output directory.
out.name	output file name.

Value

ABselectMM RData file.

Examples

```
#Get some toy data
inFile <- readRDS(system.file("extdata/dm/", "output.rds", package="AlphaBeta"))
pedigree <- inFile$Pdata
p0uu_in <- inFile$tmpp0
eqp.weight <- 1
Nstarts <- 2
out.name <- "CG_global_estimates_ABselectMM"
out <- ABselectMM(pedigree.data = pedigree,
                  p0uu=p0uu_in,
                  eqp=p0uu_in,
                  eqp.weight=eqp.weight,
                  Nstarts=Nstarts,
                  out.dir=getwd(),
                  out.name=out.name)

summary(out)
```

ABselectMMSOMA	<i>Model with selection against spontaneous gain of methylation (outselectMM)</i>
----------------	---

Description

This model assumes that somatically heritable gains of cytosine methylation are under negative selection.

Usage

```
ABselectMMSOMA(  
  pedigree.data,  
  p0uu,  
  eqp,  
  eqp.weight,  
  Nstarts,  
  out.dir,  
  out.name  
)
```

Arguments

pedigree.data	pedigree data.
p0uu	initial proportion of unmethylated cytosines.
eqp	equilibrium proportion of unmethylated cytosines.
eqp.weight	weight assigned to equilibrium function.
Nstarts	iterations for non linear LSQ optimization.
out.dir	output directory.
out.name	output file name.

Value

ABneutralSoma RData file.

Examples

```
#Get some toy data  
inFile <- readRDS(system.file("extdata/soma/", "outputSoma.rds", package="AlphaBeta"))  
pedigree <- inFile$Pdata  
p0uu_in <- inFile$tmpp0  
eqp.weight <- 0.001  
Nstarts <- 2  
out.name <- "ABselectMMSOMA_CG_estimates"  
out <- ABselectMMSOMA(pedigree.data = pedigree,  
  p0uu=p0uu_in,  
  eqp=p0uu_in,  
  eqp.weight=eqp.weight,  
  Nstarts=Nstarts,  
  out.dir=getwd(),  
  out.name=out.name)
```

```
summary(out)
```

ABselectUU	<i>Run model with selection against spontaneous loss of methylation (ABselectUU)</i>
------------	--

Description

This model assumes that heritable gains of cytosine methylation are under negative selection.

Usage

```
ABselectUU(pedigree.data, p0uu, eqp, eqp.weight, Nstarts, out.dir, out.name)
```

Arguments

pedigree.data	pedigree data.
p0uu	initial proportion of unmethylated cytosines.
eqp	equilibrium proportion of unmethylated cytosines.
eqp.weight	weight assigned to equilibrium function.
Nstarts	iterations for non linear LSQ optimization.
out.dir	output directory.
out.name	output file name.

Value

ABselectMM RData file.

Examples

```
#Get some toy data
inFile <- readRDS(system.file("extdata/dm/", "output.rds", package="AlphaBeta"))
pedigree <- inFile$Pdata
p0uu_in <- inFile$tmp0
eqp.weight <- 1
Nstarts <- 2
out.name <- "CG_global_estimates_ABselectUU"
out3 <- ABselectUU(pedigree.data = pedigree,
                   p0uu=p0uu_in,
                   eqp=p0uu_in,
                   eqp.weight=eqp.weight,
                   Nstarts=Nstarts,
                   out.dir=getwd(),
                   out.name=out.name)

summary(out3)
```

ABselectUUSOMA	<i>Model with selection against spontaneous loss of methylation (outselectUU)</i>
----------------	---

Description

This model assumes that somatically heritable gains of cytosine methylation are under negative selection.

Usage

```
ABselectUUSOMA(
  pedigree.data,
  p0uu,
  eqp,
  eqp.weight,
  Nstarts,
  out.dir,
  out.name
)
```

Arguments

pedigree.data	pedigree data.
p0uu	initial proportion of unmethylated cytosines.
eqp	equilibrium proportion of unmethylated cytosines.
eqp.weight	weight assigned to equilibrium function.
Nstarts	iterations for non linear LSQ optimization.
out.dir	output directory.
out.name	output file name.

Value

ABneutralSoma RData file.

Examples

```
#Get some toy data
inFile <- readRDS(system.file("extdata/soma/", "outputSoma.rds", package="AlphaBeta"))
pedigree <- inFile$Pdata
p0uu_in <- inFile$tmpp0
eqp.weight <- 0.001
Nstarts <- 2
out.name <- "ABselectUUSOMA_CG_estimates"
out <- ABselectUUSOMA(pedigree.data = pedigree,
  p0uu=p0uu_in,
  eqp=p0uu_in,
  eqp.weight=eqp.weight,
  Nstarts=Nstarts,
  out.dir=getwd(),
  out.name=out.name)
```

```
summary(out)
```

```
BOOTmodel
```

```
Bootstrap analysis with the best model
```

Description

Bootstrap analysis with the best model

Usage

```
BOOTmodel(pedigree.data, Nboot, out.dir, out.name)
```

Arguments

```
pedigree.data  pedigree data.
Nboot          number of boot.
out.dir        output directory.
out.name       output file name.
```

Value

bootstrap result.

Examples

```
## Get some toy data
inFile <- system.file("extdata/models/", "ABneutral_CG_global_estimates.Rdata", package="AlphaBeta")
Nboot <- 4
out.name <- "Boot_CG_global_estimates_ABneutral"
Bout <- BOOTmodel(pedigree.data=inFile,
                  Nboot=Nboot,
                  out.dir=getwd(),
                  out.name=out.name)
summary(Bout)
```

```
buildPedigree
```

```
Building Pedigree
```

Description

calculate divergence times of the pedigree

Usage

```
buildPedigree(nodelist, edgelist, cytosine = "CG", posteriorMaxFilter = 0.99)
```

Arguments

nodelist	input file containing information on generation times and pedigree lineages "ext-data" called "nodelist.fn"
edgelist	input file containing edges
cytosine	Type of cytosine (CHH/CHG/CG)
posteriorMaxFilter	Filter value, based on posteriorMax

Value

generating divergence matrices file.

Examples

```
# Get some toy data
file <- system.file("extdata/dm/", "nodelist.fn", package="AlphaBeta")
df<-read.csv(file)
df$filename <- gsub("^", paste0(dirname(dirname(file)), "/"), df$filename )
write.csv(df, file = paste0(dirname(file), "/", "tmp_nodelist.fn"), row.names=FALSE, quote=FALSE)
file <- system.file("extdata/dm/", "tmp_nodelist.fn", package="AlphaBeta")
file2 <- system.file("extdata/dm/", "edgelist.fn", package="AlphaBeta")
buildPedigree(nodelist = file, edgelist=file2, cytosine="CG", posteriorMaxFilter=0.99)
```

dMatrix

Constructing D-Matrices

Description

Estimating epimutation rates from high-throughput DNA methylation data

Usage

```
dMatrix(nodelist, cytosine, posteriorMaxFilter)
```

Arguments

nodelist	list of samples, you can find sample file in "extdata" called "nodelist.fn"
cytosine	Type of cytosine (CHH/CHG/CG)
posteriorMaxFilter	Filter value, based on posteriorMax ex: >= 0.95 or 0.99

Value

generating divergence matrices file.

Examples

```
# Get some toy data
file <- system.file("extdata/dm/", "nodelist.fn", package="AlphaBeta")
df<-read.csv(file)
df$filename<-sub("^",paste0(dirname(file),"/"),df$filename )
write.csv(df, file = paste0(dirname(file),"tmp_nodelist.fn"),row.names=FALSE,quote=FALSE)
file <- system.file("extdata/dm/", "tmp_nodelist.fn", package="AlphaBeta")
dMatrix(file, "CG", 0.99)
```

FtestRSS

Comparison of different models and selection of best model

Description

Comparison of different models and selection of best model

Usage

```
FtestRSS(pedigree.select, pedigree.null)
```

Arguments

```
pedigree.select
                pedigree model.
pedigree.null  ABnull pedigree.
```

Value

result of Ftest.

Examples

```
## Get some toy data
file1 <- system.file("extdata/models/", "ABneutral_CG_global_estimates.Rdata", package="AlphaBeta")
file2 <- system.file("extdata/models/", "ABnull_CG_global_estimates.Rdata", package="AlphaBeta")
out <- FtestRSS(pedigree.select=file1,
                pedigree.null=file2)
```

plotPedigree

Plot Pedigree

Description

Plotting Pedigree tree

Usage

```
plotPedigree(
  nodelist,
  edgelist,
  sampling.design,
  out.pdf = NULL,
  output.dir = NULL,
  plot.width = 11,
  plot.height = 8,
  vertex.label = NULL,
  vertex.size = 12,
  aspect.ratio = 2.5
)
```

Arguments

nodelist	input file containing information on generation times and pedigree lineages "ext-data" called "nodelist.fn"
edgelist	input file containing edges "edgelist.fn"
sampling.design	"progenitor.intermediate"; "sibling"; "progenitor.endpoint"; "tree"
out.pdf	output file name
output.dir	output directory
plot.width	plotting width
plot.height	plotting height
vertex.label	label vertex
vertex.size	size of vertex
aspect.ratio	aspect.ration

Value

plot pedigree matrices file.

Examples

```
# Get some toy data
file <- system.file("extdata/dm/", "nodelist.fn", package="AlphaBeta")
file2 <- system.file("extdata/dm/", "edgelist.fn", package="AlphaBeta")
plotPedigree(nodelist = file, edgelist=file2, sampling.design="sibling", vertex.label=TRUE,
  out.pdf="Plot", output.dir=getwd() )
```

Description

Estimating epimutation rates from high-throughput DNA methylation data

Usage

```
rc.meth.lvl(nodelist, cytosine, posteriorMaxFilter)
```

Arguments

nodelist	List of samples, you can find sample file in "extdata" called "nodelist.fn"
cytosine	Type of cytosine (CHH/CHG/CG)
posteriorMaxFilter	Filter value, based on posteriorMax

Value

rc meth lvl.

Examples

```
## Get some toy data  
file <- system.file("extdata/dm/", "tmp_nodelist.fn", package="AlphaBeta")  
rc.meth.lvl(file, "CG", 0.99)
```

Index

ABneutral, [2](#)
ABneutralSOMA, [3](#)
ABnull, [4](#)
ABplot, [5](#)
ABselectMM, [6](#)
ABselectMMSOMA, [7](#)
ABselectUU, [8](#)
ABselectUUSOMA, [9](#)

BOOTmodel, [10](#)
buildPedigree, [10](#)

dMatrix, [11](#)

FtestRSS, [12](#)

plotPedigree, [12](#)

rc.meth.lvl, [13](#)