Package ‘consensusOV’

March 25, 2024

Type Package

Title Gene expression-based subtype classification for high-grade serous ovarian cancer

Version 1.24.0

Date 2019-12-04

Author Gregory M Chen, Lavanya Kannan, Ludwig Geistlinger, Victor Kofia, Levi Waldron, Benjamin Haibe-Kains

Maintainer Benjamin Haibe-Kains <benjamin.haibe.kains@utoronto.ca>

Description This package implements four major subtype classifiers for high-grade serous (HGS) ovarian cancer as described by Helland et al. (PLoS One, 2011), Bentink et al. (PLoS One, 2012), Verhaak et al. (J Clin Invest, 2013), and Konecny et al. (J Natl Cancer Inst, 2014). In addition, the package implements a consensus classifier, which consolidates and improves on the robustness of the proposed subtype classifiers, thereby providing reliable stratification of patients with HGS ovarian tumors of clearly defined subtype.

License Artistic-2.0

Depends R (>= 3.6)

Imports Biobase, GSVA, gdata, genefu, limma, matrixStats, randomForest, stats, utils, methods

URL http://www.pmgenomics.ca/bhklab/software/consensusOV

Suggests BiocStyle, ggplot2, knitr, rmarkdown

VignetteBuilder knitr

Encoding UTF-8

RoxygenNote 6.1.1

LazyData true

biocViews Classification, Clustering, DifferentialExpression, GeneExpression, Microarray, Transcriptomics

BugReports https://github.com/bhklab/consensusOV/issues
dataset.merging

Merging all individual esets and merging them into a big eset

description

Merging all individual esets and merging them into a big eset

Usage

dataset.merging(esets, method = c("union", "intersect"),
standardization = c("quantile", "robust.scaling", "scaling", "none"),
nthread = 1)

Arguments

esets The list containing all GSE file that need to be merged.
method either "unique" or "intersect" is use to for selecting geneid
standardization choose between "quantile", "robust.scaling", "scaling" or "none"
nthread number of threads (1 by default)

Value

The merging eset
get.bentink.subtypes

Get ovarian cancer subtypes as defined by Bentink et al., 2012

Description

Get ovarian cancer subtypes as defined by Bentink et al., 2012

Usage

get.bentink.subtypes(expression.matrix, entrez.ids)

Arguments

expression.matrix
A matrix of gene expression values with rows as genes, columns as samples.
entrez.ids
A vector of Entrez Gene IDs, corresponding to the rows of expression.matrix

Value

A list with first value Bentink.subtypes containing a factor of subtype names; and second value
angio containing the output of genefu::ovcAngiogenic

References

Bentink et al. Angiogenic mRNA and microRNA gene expression signature predicts a novel subtype

Examples

library(Biobase)
library(genefu)
data(GSE14764.eset)
expression.matrix <- exprs(GSE14764.eset)
entrez.ids <- as.character(fData(GSE14764.eset)$EntrezGene.ID)
get.bentink.subtypes(expression.matrix, entrez.ids)

get.consensus.subtypes

Get consensusOV ovarian cancer subtypes

Description

Get consensusOV ovarian cancer subtypes
get.consensus.subtypes

Usage

get.consensus.subtypes(expression.matrix, entrez.ids,
  concordant.tumors.only = TRUE, remove.using.cutoff = FALSE,
  percentage.dataset.removed = 0.75,
  .training.dataset = consensus.training.dataset.full,
  .dataset.names.to.keep = names(esets.rescaled.classified.filteredgenes))

margin(rf.probs)

Arguments

expression.matrix
  A matrix of gene expression values with rows as genes, columns as samples.
entrez.ids
  A vector of Entrez Gene IDs, corresponding to the rows of expression.matrix
concordant.tumors.only
  Logical. Should the classifier trained only on tumors that are concordantly classified by Helland, Konecny, and Verhaak? Defaults to TRUE.
remove.using.cutoff
  Specify whether to classify NA for samples that do not meet a margin cutoff
percentage.dataset.removed
  If remove.using.cutoff is TRUE, then classify this percentage of samples to NA based on margin values
.training.dataset
  ExpressionSet containing the training data. Defaults to the pooled dataset across selected MetaGxOvarian datasets.
.dataset.names.to.keep
  Names of MetaGxOvarian datasets to use for training
rf.probs
  random forest probabilities for each subtype as returned by get.consensus.subtypes

Value

get.consensus.subtypes returns a list with first value consensusOV.subtypes containing a factor of subtype labels; and second value rf.probs containing a matrix of subtype probabilities.

margin returns a numeric vector containing the classification margin scores, i.e. the difference between the top two subtype scores for each tumor.

Examples

library(Biobase)
data(GSE14764.eset)
expression.matrix <- exprs(GSE14764.eset)
entrez.ids <- as.character(fData(GSE14764.eset)$EntrezGene.ID)
sts <- get.consensus.subtypes(expression.matrix, entrez.ids)
margins <- margin(sts$rf.probs)
get.hao.subtypes

Get ovarian cancer subtypes as defined by Hao et al., 2017

Description
Get ovarian cancer subtypes as defined by Hao et al., 2017

Usage
get.hao.subtypes(expression.matrix, entrez.ids)

Arguments
expression.matrix
A matrix of gene expression values with genes as rows, samples as columns.
entrez.ids A vector of Entrez Gene IDs, corresponding to the rows of expression.matrix.

Details
Hao et al., 2017 derived a gene signature to predict the tissue of origin of ovarian tumors as either fallopian tube (FT) or ovarian surface epithelium (OSE).
The authors found that expression patterns of tissue-specific genes, prognostic genes, and molecular markers support a dualistic tissue origin of ovarian cancer, from either FT or OSE.
The subtype classifier considers 112 signature genes including 37 genes upregulated in FT and 75 genes upregulated in OSE. A score is computed that is designed to range from 0 to 1 for FT tumors, while OSE tumors have a score ranging from -1 to 0.

Value
A list with first value tissue containing a factor of subtype names (tissue of origin); and second value score containing the tissue-of-origin score.

Author(s)
Ludwig Geistlinger

References

Examples
library(Biobase)
data(GSE14764.eset)
expression.matrix <- exprs(GSE14764.eset)
entrez.ids <- as.character(fData(GSE14764.eset)$EntrezGene.ID)
get.hao.subtypes(expression.matrix, entrez.ids)
get.helland.subtypes  Get ovarian cancer subtypes as defined by Helland et al., 2011

Description
Get ovarian cancer subtypes as defined by Helland et al., 2011

Usage
get.helland.subtypes(expression.matrix, entrez.ids)

Arguments
expression.matrix
A matrix of gene expression values with rows as genes, columns as samples.
entrez.ids A vector of Entrez Gene IDs, corresponding to the rows of expression.matrix

Value
A list with first value Helland.subtypes containing a factor of subtype names; and second value subtype.scores containing a matrix of subtype scores

References

Examples
library(Biobase)
data(GSE14764.eset)
expression.matrix <- exprs(GSE14764.eset)
entrez.ids <- as.character(fData(GSE14764.eset)$EntrezGene.ID)
get.helland.subtypes(expression.matrix, entrez.ids)

get.konecny.subtypes  Get ovarian cancer subtypes as defined by Konecny et al., 2014

Description
Get ovarian cancer subtypes as defined by Konecny et al., 2014

Usage
get.konecny.subtypes(expression.matrix, entrez.ids)
get.subtypes

Arguments

expression.matrix

A matrix of gene expression values with rows as genes, columns as samples.

entrez.ids

A vector of Entrez Gene IDs, corresponding to the rows of expression.matrix

Value

A list with first value Konecny.subtypes containing a factor of subtype names; and second value spearman.cc.vals containing the Spearman correlation values per subtype

References


Examples

library(Biobase)
data(GSE14764.eset)
expression.matrix <- exprs(GSE14764.eset)
entrez.ids <- as.character(fData(GSE14764.eset)$EntrezGene.ID)
get.konecny.subtypes(expression.matrix, entrez.ids)

Description

Get ovarian cancer subtypes

Usage

get.subtypes(expression.dataset, entrez.ids = NULL,
method = c("consensusOV", "Helland", "Verhaak", "Konecny", "Bentink"),
...

Arguments

expression.dataset

Either a matrix of gene expression values with rows as genes, columns as samples; or a BioBase::ExpressionSet object from MetaGxOvarian. If expression.dataset is a matrix, then entrez.ids must have length equal to the number of rows of expression.dataset.

entrez.ids

A vector of Entrez Gene IDs, corresponding to the rows of expression.dataset

method

The subtyping method to use

...

Optional parameters to be passed to the low level function
get.verhaak.subtypes

Value

A list with first value Konecny.subtypes containing a factor of subtype names; and second value spearman.cc.vals containing the Spearman correlation values per subtype.

Examples

```r
library(Biobase)
data(GSE14764.eset)
expression.matrix <- exprs(GSE14764.eset)
entrez.ids <- as.character(fData(GSE14764.eset)$EntrezGene.ID)
get.subtypes(expression.matrix, entrez.ids, method="Konecny")
```

get.verhaak.subtypes  Get ovarian cancer subtypes as defined by Verhaak et al., 2013

Description

Get ovarian cancer subtypes as defined by Verhaak et al., 2013.

Usage

```r
get.verhaak.subtypes(expression.matrix, entrez.ids)
```

Arguments

expression.matrix

A matrix of gene expression values with rows as genes, columns as samples.

entrez.ids

A vector of Entrez Gene IDs, corresponding to the rows of expression.matrix.

Value

A list with first value Verhaak.subtypes containing a factor of subtype names; and second value gsve containing the GSVA subtype scores.

References


Examples

```r
library(Biobase)
data(GSE14764.eset)
expression.matrix <- exprs(GSE14764.eset)
entrez.ids <- as.character(fData(GSE14764.eset)$EntrezGene.ID)
get.konecny.subtypes(expression.matrix, entrez.ids)
```
GSE14764.eset

Sample ExpressionSet from MetaGxOvarian

Description

A Biobase::ExpressionSet from package MetaGxOvarian for the dataset GSE14764

Usage

GSE14764.eset

Format

A Biobase::ExpressionSet object

Source

http://biorxiv.org/content/biorxiv/early/2016/05/12/052910.full.pdf
Index

* datasets
  GSE14764.eset, 9

dataset.merging, 2

get.bentink.subtypes, 3
get.consensus.subtypes, 3, 4
get.hao.subtypes, 5
get.helland.subtypes, 6
get.konecny.subtypes, 6
get.subtypes, 7
get.verhaak.subtypes, 8
GSE14764.eset, 9

margin(get.consensus.subtypes), 3