Package 'MUTSCOT'

January 15, 2019

Title MutScot

Version 0.0.1

Date 2019-01-14

Author Bin Zhu, Ho-Hsiang Wu

Description The MutScot package.

Maintainer Bill Wheeler <wheelerb@imsweb.com>

Imports stats, utils, MASS, dndscv

License GPL-2

NeedsCompilation yes

Depends R (>= 2.10)

R topics documented:

MutScot-package   *MutScot*

Description

The MutScot package for identifying driver genes

Details

Identifying cancer driver genes is essential for understanding mechanisms of carcinogenesis and designing therapeutic strategies. Consequently, a set of driver genes has been identified for each cancer types, assumed to be identical across subtypes. This assumption may not hold, and the sets of driver genes are possibly distinct across cancer subtypes. The MutScot package includes a statistical framework that identifies driver genes and utilizes patient information to investigate subtype heterogeneity of driver genes. The function `MutScot` identifies driver genes, and the function `MutScotPatientHeter` is used for conducting a subtype heterogeneity test.

1

Author(s)

Ho-Hsiang Wu and Bin Zhu <bin.zhu@nih.gov>

---

data *Data for the* MutScot *example*

Description

Data for the MutScot example.

Details

The object data is a data frame with columns for the sample id, chromosome, position (hg19), reference allele, and alternate allele.

Examples

```
data(data, package="MUTSCOT")
data[1:5, ]
```

---

getBenchmarkGenes *Benchmark Genes*

Description

Get a vector of gene names given by Bailey et al (2018)

Usage

```
getBenchmarkGenes()
```

Details

This is the list of driver gene names given by Bailey et al (2018).

Value

A character vector of gene names.

Examples

```
getBenchmarkGenes()
```

MutScot   *MutScot*

Description

MutScot

Usage

```
MutScot(data, MutSigCV.out=NULL, options=NULL)
```

Arguments

data	Data frame contining five columns for sample id, chromosome, position (hg19), reference allele, and alternate allele. The columns must be in the order given above. No default
MutSigCV.out	Information provided by MutSigCV algorithm (see details). The default is NULL.
options	List of options (see details). The default is NULL.

Details

See data(MutSigCV) for an example format of MutSigCV.out. Currently, the only columns of this object that are used are: gene, X, x, and q.

Options

Option	Description	Default
maxMutPerPat	maximum number of mutations per patient	300
multitesting	method used for multiple tests (see p.adjust)	"fdr"
alpha.q	significance level	0.1
benchmarkGenes	vector of genes for summary	genes returned from getBenchmarkGenes

Value

A list of class "MutScot" with the following names and descriptions:

- driverGenes Vector of identified driver genes.
- significantGenes Data frame of significant genes among the methods determined by options$alpha.q.
- MutScotGenes Data frame of p-values and adjusted p-values for the MutScot method.
- countMat Matrix of TP, FP, F1 counts by method.
- dndsout Output from the dndscv algorithm.
- patientIds Vector of patient ids.
- genes Vector of gene names.
- mutationRates List containing relative SNV mutation rates of patients (PatRat), weight parameter of patients w.r.t SNV mutations (wt), relative Indel mutation rates of patients (PatRatI), weight parameter of patients w.r.t Indel mutations (wtI), annotated Indel mutations (Indels_record) , annotated observed/expected Indel mutations (geneindels), IVW estimated gene-specific background mutation rates (Lambda_g).

- obsExp List containing gene-specific observed and expected missense mutation counts (`miss`), gene-specific observed and expected nonsense mutation counts (`nsen`), gene-specific observed and expected splicing site mutation counts (`splc`), gene-specific observed and expected Indel mutation counts (`indel`).

Author(s)

Ho-Hsiang Wu and Bin Zhu <bin.zhu@nih.gov>

See Also

`MutScotPatientHeter`

Examples

```
data(data, package="MUTSCOT")
data(MutSigCV.out, package="MUTSCOT")
# Not run due to the time it takes
#MutScot(data, MutSigCV.out=MutSigCV.out)
```

`MutScotPatientHeter`

Test for heterogeneity

Description

Test genes for subtype heterogeneity

Usage

```
MutScotPatientHeter(obj, patientInfo, testGenes, options=NULL)
```

Arguments

`obj`	Object returned from `MutScot`. No default.
`patientInfo`	Data frame containing sample ids and numeric columns of adjusted continuous covariates. The first column must be the ids and there must be at least one covariate (see details). No default.
`testGenes`	Character vector of gene names to test. No default.
`options`	List of options (see details). The default is NULL.

Details

The first column of patient ids in the `patientInfo` data frame are matched with `obj$patientIds`.

Options

Option	Description	Default
MCT	number of Monte-Carlo simulations	1e4

Value

A list with the following names and descriptions:

- pvalues Data frame of gene-specific p-values for missense, nonsense, splicing, indel, and the test of subtype heterogeneity. A row of NAs will appear for any gene not found in obj$genes.
- glm_miss List of the glm fitting of missense mutation for each gene.
- glm_nsen List of the glm fitting of nonsense mutation for each gene.
- glm_splc List of the glm fitting of splicing site mutation for each gene.
- glm_indel List of the glm fitting of indel mutation for each gene.

Author(s)

Ho-Hsiang Wu and Bin Zhu <bin.zhu@nih.gov>

See Also

MutScot

Examples

```
data(data, package="MUTSCOT")
data(MutSigCV.out, package="MUTSCOT")
data(PatInfo, package="MUTSCOT")

# Not run due to the time it takes
#retM    <- MutScot(data, MutSigCV.out=MutSigCV.out)
#x       <- retM$significantGenes
#tmp     <- (x[, "Benchmark"]==TRUE) & (x[, "MutScot"]==TRUE)
#TestSet <- as.vector(x[tmp, "Gene"])
#retH    <- MutScotPatientHeter(retM, PatInfo, TestSet)
```

MutSigCV.out   *Example of the* MutSigCV.out *data object*

Description

Example of the MutSigCV.out data object

Details

The object is a data frame named MutSigCV.out. Currently, the only columns of this data that are used are: gene, X, x, and q.

Examples

```
# Look at the first five rows
data(MutSigCV.out, package="MUTSCOT")
MutSigCV.out[1:5, ]
```

PatInfo   *Data for the* MutScotPatientHeter *example*

Description

Data for the MutScotPatientHeter example.

Details

The object data is a data frame with columns for the sample id and ER status.

Examples

```
data(PatInfo, package="MUTSCOT")
PatInfo[1:5, ]
```