

Package ‘ELMER.data’

October 31, 2024

Title Data for the ELMER package

Version 2.31.0

Author Tiago Chedraoui Silva,
Lijing Yao
Simon Coetzee,
Benjamin P. Berman

Maintainer Tiago Chedraoui Silva <tiagochst@gmail.com>

Description Supporting data for the ELMER package.

It includes:

- elmer.data.example.promoter: mae.promoter
- elmer.data.example: data
- EPIC.hg38.manifest
- EPIC.hg19.manifest
- hm450.hg38.manifest
- hm450.hg19.manifest
- hocomoco.table
- human.TF
- LUSC_meth_refined: Meth
- LUSC_RNA_refined: GeneExp
- Probes.motif.hg19.450K
- Probes.motif.hg19.EPIC
- Probes.motif.hg38.450K
- Probes.motif.hg38.EPIC
- TF.family
- TF.subfamily
- Human_genes__GRCh37_p13
- Human_genes__GRCh38_p12
- Human_genes__GRCh37_p13__tss
- Human_genes__GRCh38_p12__tss

License GPL-3

LazyData false

Depends R (>= 3.5.0)

Imports GenomicRanges

Suggests BiocStyle, knitr, dplyr, devtools, DT, rmarkdown

biocViews AssayDomainData, TechnologyData, OrganismData

VignetteBuilder knitr

RoxygenNote 6.1.1

git_url <https://git.bioconductor.org/packages/ELMER.data>

git_branch devel

git_last_commit c160a58

git_last_commit_date 2024-10-29

Repository Bioconductor 3.21

Date/Publication 2024-10-31

Contents

data	2
ELMER.data	3
EPIC.hg19.manifest	4
EPIC.hg38.manifest	4
GeneExp	5
hm450.hg19.manifest	5
hm450.hg38.manifest	6
hocomoco.table	6
human.TF	7
Human_genes__GRCh37_p13	7
Human_genes__GRCh37_p13__tss	8
Human_genes__GRCh38_p12	8
Human_genes__GRCh38_p12__tss	9
mae.promoter	9
Meth	10
Probes.motif.hg19.450K	10
Probes.motif.hg19.EPIC	11
Probes.motif.hg38.450K	11
Probes.motif.hg38.EPIC	12
TF.family	13
TF.subfamily	13
Index	14

data

A MultiAssayExperiment containing DNA methylation data: 101 probes from platform 450K Gene Expression data: 1026 genes for 234 samples from TCGA-LUSC. This data is used in the examples of ELMER package

Description

A MultiAssayExperiment containing DNA methylation data: 101 probes from platform 450K Gene Expression data: 1026 genes for 234 samples from TCGA-LUSC. This data is used in the examples of ELMER package

Usage

```
data
```

Format

A MultiAssayExperiment for 234 Samples (8 normal samples, 226 Primary solid tumor)

Examples

```
## Not run:
data("elmer.data.example")

## End(Not run)
```

ELMER.data

Data for ELMER package

Description

ELMER is package using DNA methylation to identify enhancers, and correlates enhancer state with expression of nearby genes to identify one or more transcriptional targets. Transcription factor (TF) binding site analysis of enhancers is coupled with expression analysis of all TFs to infer upstream regulators. ELMER.data provide the necessary data for ELMER analysis:

- Probes.motif: motif occurrences within $-/+250$ bp of probe sites on HM450K/EPIC array aligned against hg19/hg38.
- DNA methylation platform manifest: from <http://zwdzwd.github.io/InfiniumAnnotation>
- TF.family TFs family from TFClass
- TF.subfamily TFs subfamily from TFClass

For more information how to create these objects please read the vignette of this package with the following command: `browseVignettes("ELMER.data")`

See Also

[EPIC.hg19.manifest](#), [EPIC.hg38.manifest](#), [hm450.hg19.manifest](#), [hm450.hg38.manifest](#), [Probes.motif.hg19.450K](#), [Probes.motif.hg38.450K](#), [Probes.motif.hg38.EPIC](#), [Probes.motif.hg19.EPIC](#), [Human_genes__GRCh37_p13__tss](#), [Human_genes__GRCh37_p13](#), [Human_genes__GRCh38_p12](#), [Human_genes__GRCh38_p12__tss](#), [TF.subfamily](#), [TF.family](#), and [hocomoco.table](#)

Examples

```
# Please see the datasets
```

EPIC.hg19.manifest *A GRanges containing hg19 annotation with suggested overall masking for EPIC platform*

Description

A GRanges containing hg19 annotation with suggested overall masking for EPIC platform

Usage

EPIC.hg19.manifest

Format

A GRanges with 866895 elements

Examples

```
## Not run:  
data("EPIC.hg19.manifest")  
  
## End(Not run)
```

EPIC.hg38.manifest *A GRanges containing hg38 annotation with suggested overall masking for EPIC platform*

Description

A GRanges containing hg38 annotation with suggested overall masking for EPIC platform

Usage

EPIC.hg38.manifest

Format

A GRanges with 866895 elements

Examples

```
## Not run:  
data("EPIC.hg38.manifest")  
  
## End(Not run)
```

GeneExp	<i>A matrix containing gene expression data from TCGA Gene Expression data: 3842 genes This data is used in the examples of ELMER package</i>
---------	-----------------------------------------------------------------------------------------------------------------------------------------------

Description

A matrix containing gene expression data from TCGA Gene Expression data: 3842 genes This data is used in the examples of ELMER package

Usage

GeneExp

Format

A gene expression matrix for 234 Samples and 3842 genes

hm450.hg19.manifest	<i>A GRanges containing hg19 annotation with suggested overall masking for hm450 platform</i>
---------------------	-----------------------------------------------------------------------------------------------

Description

A GRanges containing hg19 annotation with suggested overall masking for hm450 platform

Usage

hm450.hg19.manifest

Format

A GRanges with 485577 elements

Examples

```
## Not run:  
data("hm450.hg19.manifest")  
  
## End(Not run)
```

hm450.hg38.manifest *A GRanges containing hg38 annotation with suggested overall masking for hm450 platform*

Description

A GRanges containing hg38 annotation with suggested overall masking for hm450 platform

Usage

```
hm450.hg38.manifest
```

Format

A GRanges with 485577 elements

Examples

```
## Not run:  
data("hm450.hg38.manifest")  
  
## End(Not run)
```

hocomoco.table *Table parsed from hocomoco v11*

Description

Table parsed from hocomoco v11

Usage

```
hocomoco.table
```

Format

A dataframe with 771 rows (motifs) and 20 columns

human.TF	<i>Table parsed from Lambert, Samuel A., et al. "The human transcription factors." Cell 172.4 (2018): 650-665.</i>
----------	--------------------------------------------------------------------------------------------------------------------

Description

Table parsed from Lambert, Samuel A., et al. "The human transcription factors." Cell 172.4 (2018): 650-665.

Usage

human.TF

Format

A dataframe with 1639 rows (motifs) and 27 columns

Human_genes__GRCh37_p13	<i>A matrix containing ENSEMBL hg19 gene metadata accessed using biomart This data is used if ensembl cannot be reached</i>
-------------------------	-----------------------------------------------------------------------------------------------------------------------------

Description

A matrix containing ENSEMBL hg19 gene metadata accessed using biomart This data is used if ensembl cannot be reached

Usage

Human_genes__GRCh37_p13

Format

A matrix with metadata for 60482 genes

Human_genes__GRCh37_p13__tss

A matrix containing ENSEMBL hg19 transcripts metadata accessed using biomaRT This data is used if ensembl cannot be reached

Description

A matrix containing ENSEMBL hg19 transcripts metadata accessed using biomaRT This data is used if ensembl cannot be reached

Usage

Human_genes__GRCh37_p13__tss

Format

A matrix with metadata for 196317 transcripts

Human_genes__GRCh38_p12

A matrix containing ENSEMBL hg38 gene metadata accessed using biomaRT This data is used if ensembl cannot be reached

Description

A matrix containing ENSEMBL hg38 gene metadata accessed using biomaRT This data is used if ensembl cannot be reached

Usage

Human_genes__GRCh38_p12

Format

A matrix with metadata for 58639 genes

 Human_genes__GRCh38_p12__tss

A matrix containing ENSEMBL hg38 transcripts metadata accessed using biomaRt This data is used if ensembl cannot be reached

Description

A matrix containing ENSEMBL hg38 transcripts metadata accessed using biomaRt This data is used if ensembl cannot be reached

Usage

Human_genes__GRCh38_p12__tss

Format

A matrix with metadata for 208423 transcripts

 mae.promoter

A MultiAssayExperiment containing DNA methylation data: 16 promoter probes from platform 450K Gene Expression data: 3808 genes for 234 samples from TCGA-LUSC. This data is used in the examples of ELMER package

Description

A MultiAssayExperiment containing DNA methylation data: 16 promoter probes from platform 450K Gene Expression data: 3808 genes for 234 samples from TCGA-LUSC. This data is used in the examples of ELMER package

Usage

mae.promoter

Format

A MultiAssayExperiment for 234 Samples (8 normal samples, 226 Primary solid tumor)

Examples

```
## Not run:
data("elmer.data.example.promoter")

## End(Not run)
```

Meth *A matrix containing DNA methylation beta-values from TCGA DNA methylation data: 1728 probes This data is used in the examples of ELMER package*

Description

A matrix containing DNA methylation beta-values from TCGA DNA methylation data: 1728 probes This data is used in the examples of ELMER package

Usage

Meth

Format

A MultiAssayExperiment for 268 Samples and 1728 probes

Probes.motif.hg19.450K

A matrix with 1 if the probe (row) has a motif (column)

Description

It was generated using HOMER with a p-value $< 1e-4$ to scan a +/- 250bp region around each probe using HOmo sapiens COmprehensive MOdel COllection [<http://hocomoco.autosome.ru/>](HOCOMOCO) v10 position weight matrices (PWMs). HOCOMOCO offers 640 PMWs each has a quality rating from A to D where A represents motifs with the highest confidence, and D motifs only weakly describe the pattern with a limited applications for quantitative analyses. By default only quality A and B will be used for the Motif enrichment analysis, but the minimum quality score can be selected by the user. (Additional information [<http://hocomoco.autosome.ru/help>](Source) [<http://nar.oxfordjournals.org/content/44/D1/D1> information]). The DNA methylation information was retrieved from: <http://zwdzwd.github.io/InfiniumAnnotation> For more information check the vignette.

Usage

Probes.motif.hg19.450K

Format

A matrix with 466007 rows and 640 columns

Examples

```
## Not run:
data("Probes.motif.hg19.450K")

## End(Not run)
```

Probes.motif.hg19.EPIC

A matrix with 1 if the probe (row) has a motif (column)

Description

It was generated using HOMER with a p-value < 1e-4 to scan a +/- 250bp region around each probe using HOmo sapiens COmprehensive MOdel COllection [<http://hocomoco.autosome.ru/>](HOCOMOCO) v10 position weight matrices (PWMs). HOCOMOCO offers 640 PMWs each has a quality rating from A to D where A represents motifs with the highest confidence, and D motifs only weakly describe the pattern with a limited applications for quantitative analyses. By default only quality A and B will be used for the Motif enrichment analysis, but the minimum quality score can be selected by the user. (Additional information [<http://hocomoco.autosome.ru/help>](Source) [<http://nar.oxfordjournals.org/content/44/D1/D> information]). The DNA methylation information was retrieved from: <http://zwdzwd.github.io/InfiniumAnnotation> For more information check the vignette.

Usage

```
Probes.motif.hg19.EPIC
```

Format

A matrix with 838881 rows and 640 columns

Examples

```
## Not run:  
data("Probes.motif.hg19.EPIC")  
  
## End(Not run)
```

Probes.motif.hg38.450K

A matrix with 1 if the probe (row) has a motif (column)

Description

It was generated using HOMER with a p-value < 1e-4 to scan a +/- 250bp region around each probe using HOmo sapiens COmprehensive MOdel COllection [<http://hocomoco.autosome.ru/>](HOCOMOCO) v10 position weight matrices (PWMs). HOCOMOCO offers 640 PMWs each has a quality rating from A to D where A represents motifs with the highest confidence, and D motifs only weakly describe the pattern with a limited applications for quantitative analyses. By default only quality A and B will be used for the Motif enrichment analysis, but the minimum quality score can be selected by the user. (Additional information [<http://hocomoco.autosome.ru/help>](Source) [<http://nar.oxfordjournals.org/content/44/D1/D> information]). The DNA methylation information was retrieved from: <http://zwdzwd.github.io/InfiniumAnnotation> For more information check the vignette.

Usage

```
Probes.motif.hg38.450K
```

Format

A matrix with 466007 rows and 640 columns

Examples

```
## Not run:  
data("Probes.motif.hg38.450K")  
  
## End(Not run)
```

```
Probes.motif.hg38.EPIC
```

A matrix with 1 if the probe (row) has a motif (column)

Description

It was generated using HOMER with a p-value < 1e-4 to scan a +/- 250bp region around each probe using HOmo sapiens COmprehensive MOdel COllection [<http://hocomoco.autosome.ru/>](HOCOMOCO) v10 position weight matrices (PWMs). HOCOMOCO offers 640 PMWs each has a quality rating from A to D where A represents motifs with the highest confidence, and D motifs only weakly describe the pattern with a limited applications for quantitative analyses. By default only quality A and B will be used for the Motif enrichment analysis, but the minimum quality score can be selected by the user. (Additional information [<http://hocomoco.autosome.ru/help>](Source) [<http://nar.oxfordjournals.org/content/44/D1/D> information]). The DNA methylation information was retrieved from: <http://zwdzwd.github.io/InfiniumAnnotation> For more information check the vignette.

Usage

```
Probes.motif.hg38.EPIC
```

Format

A matrix with 838881 rows and 640 columns

Examples

```
## Not run:  
data("Probes.motif.hg38.EPIC")  
  
## End(Not run)
```

TF.family	<i>A list of 641 motifs with TF families (with similar bindings) from TFClass Created with the following function from ELMER pacakge</i> <i>TF.family <- createMotifRelevantTfs()</i>
-----------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Description

A list of 641 motifs with TF families (with similar bindings) from TFClass Created with the following function from ELMER pacakge TF.family <- createMotifRelevantTfs()

Usage

TF.family

Format

A list of 641 motifs with TF families (with similar bindings)

TF.subfamily	<i>A list of 641 motifs with TF subfamilies (with similar bindings) from TFClass Created with the following function from ELMER pacakge</i> <i>TF.family <- createMotifRelevantTfs("subfamily")</i>
--------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Description

A list of 641 motifs with TF subfamilies (with similar bindings) from TFClass Created with the following function from ELMER pacakge TF.family <- createMotifRelevantTfs("subfamily")

Usage

TF.subfamily

Format

A list of 641 motifs with TF subfamilies (with similar bindings)

Index

* internal

- data, [2](#)
- EPIC.hg19.manifest, [4](#)
- EPIC.hg38.manifest, [4](#)
- GeneExp, [5](#)
- hm450.hg19.manifest, [5](#)
- hm450.hg38.manifest, [6](#)
- hocomoco.table, [6](#)
- human.TF, [7](#)
- Human_genes__GRCh37_p13, [7](#)
- Human_genes__GRCh37_p13__tss, [8](#)
- Human_genes__GRCh38_p12, [8](#)
- Human_genes__GRCh38_p12__tss, [9](#)
- mae.promoter, [9](#)
- Meth, [10](#)
- Probes.motif.hg19.450K, [10](#)
- Probes.motif.hg19.EPIC, [11](#)
- Probes.motif.hg38.450K, [11](#)
- Probes.motif.hg38.EPIC, [12](#)
- TF.family, [13](#)
- TF.subfamily, [13](#)

* utilities

- ELMER.data, [3](#)

data, [2](#)

- ELMER.data, [3](#)
- ELMER.data-package (ELMER.data), [3](#)
- EPIC.hg19.manifest, [3, 4](#)
- EPIC.hg38.manifest, [3, 4](#)

GeneExp, [5](#)

- hm450.hg19.manifest, [3, 5](#)
- hm450.hg38.manifest, [3, 6](#)
- hocomoco.table, [3, 6](#)
- human.TF, [7](#)
- Human_genes__GRCh37_p13, [3, 7](#)
- Human_genes__GRCh37_p13__tss, [3, 8](#)
- Human_genes__GRCh38_p12, [3, 8](#)

Human_genes__GRCh38_p12__tss, [3, 9](#)

mae.promoter, [9](#)

Meth, [10](#)

Probes.motif.hg19.450K, [3, 10](#)

Probes.motif.hg19.EPIC, [3, 11](#)

Probes.motif.hg38.450K, [3, 11](#)

Probes.motif.hg38.EPIC, [3, 12](#)

TF.family, [3, 13](#)

TF.subfamily, [3, 13](#)