Package ‘FANTOM3and4CAGE’

January 22, 2026

Title CAGE data from FANTOM3 and FANTOM4 projects

Version 1.46.0

Date 02-06-2013

Author Vanja Haberle, Department of Biology, University of Bergen, Norway <`vanja.haberle@bio.uib.no`>

Maintainer Vanja Haberle <`vanja.haberle@bio.uib.no`>

Depends R (>= 2.15.0)

Description CAGE (Cap Analysis Gene Expression) data from FANTOM3 and FANTOM4 projects produced by RIKEN Omics Science Center.

License GPL-3

biocViews ExperimentData, Tissue

git_url https://git.bioconductor.org/packages/FANTOM3and4CAGE

git_branch RELEASE_3_22

git_last_commit ab85357

git_last_commit_date 2025-10-29

Repository Bioconductor 3.22

Date/Publication 2026-01-22

Contents

`FANTOMhumanSamples` *List of FANTOM CAGE samples for human*

Description

A `data.frame` with information on all human CAGE samples from FANTOM3 and FANTOM4 projects available in this package.

Usage

```
data(FANTOMhumanSamples)
```

Details

`FANTOMhumanSamples` is a `data.frame` with 3 columns:

`dataset`: the name of the dataset that can be loaded using data() function
`group`: the name of the group of samples that originate from the same tissue (*e.g.* blood)
`sample`: the name of the specific sample

Source

The annotation of CAGE samples was obtained from the FANTOM web resource (http://fantom.gsc.riken.jp/4/download/

References

Kawaji *et al.* (2010) Update of the FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation, *Nucleic Acids Research* **39**:D856-D860.

Examples

```
data(FANTOMhumanSamples)
head(FANTOMhumanSamples)
```

`FANTOMmouseSamples` *List of FANTOM CAGE samples for mouse*

Description

A `data.frame` with information on all mouse CAGE samples from FANTOM3 and FANTOM4 projects available in this package.

Usage

```
data(FANTOMmouseSamples)
```

Details

`FANTOMmouseSamples` is a `data.frame` with 3 columns:

`dataset`: the name of the dataset that can be loaded using data() function
`group`: the name of the group of samples that originate from the same tissue (*e.g.* brain)
`sample`: the name of the specific sample

Source

The annotation of CAGE samples was obtained from the FANTOM web resource (http://fantom.gsc.riken.jp/4/download/

References

Kawaji *et al.* (2010) Update of the FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation, *Nucleic Acids Research* **39**:D856-D860.

Examples

```
data(FANTOMmouseSamples)
head(FANTOMmouseSamples)
```

`FANTOMtimecourseCAGEhuman`

Human timecourse CAGE data produced in FANTOM4 project

Description

This dataset contains all transcription start sites (TSSs) detected by CAGE and their frequency of usage in the human THP-1 cells induction timecourse from FANTOM4 project.

Usage

```
data(FANTOMtimecourseCAGEhuman)
```

Details

`FANTOMtimecourseCAGEhuman` is a list with only one element (timecourse) named `THP-1_monocytic_induction`. This element is a `data.frame` with genomic positions of TSSs detected by CAGE and frequency of their usage in various timepoints during stimulation of human THP-1 myelomonocytic leukemia cells by phorbol myristate acetate (PMA). First three columns contain the chromosome (`chr`), the 1-based coordinate (`pos`) and the strand (`strand`) of the TSS, and the following columns contain the number of CAGE tags supporting that TSS in each specified timepoint.

Source

The CAGE detected TSSs (mapped to the hg18 human genome assembly) and numbers of CAGE tags for THP-1 monocytic induction were obtained from the FANTOM web resource (http://fantom.gsc.riken.jp/4/downlo
The data were originally published by The FANTOM Consortium (Suzuki *et al.*, Nature Genetics 2009).

References

Kawaji *et al.* (2010) Update of the FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation, *Nucleic Acids Research* **39**:D856-D860.
Suzuki *et al.* (2009) The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line, *Nature Genetics* **41**:553-562.

Examples

```
data(FANTOMtimecourseCAGEhuman)
names(FANTOMtimecourseCAGEhuman)
head(FANTOMtimecourseCAGEhuman[["THP-1_monocytic_induction"]])
```

`FANTOMtimecourseCAGEmouse` *Mouse timecourse CAGE data produced in FANTOM projects*

Description

This dataset contains all transcription start sites (TSSs) detected by CAGE and their frequency of usage in two mouse timecourses: adipogenic induction of DFAT-D1 preadipocytes and liver under constant darkness condition.

Usage

```
data(FANTOMtimecourseCAGEmouse)
```

Details

`FANTOMtimecourseCAGEmouse` is a list with two elements (timecourses) named `adipogenic_induction` and `liver_under_constant_darkness`. Each element is a `data.frame` with genomic positions of TSSs detected by CAGE and frequency of their usage in various timepoints during timecourse. First three columns contain the chromosome (`chr`), the 1-based coordinate (`pos`) and the strand (`strand`) of the TSS, and the following columns contain the number of CAGE tags supporting that TSS in each specified timepoint.

Source

The CAGE detected TSSs (mapped to the mm9 mouse genome assembly) and numbers of CAGE tags for both timecourses were obtained from the FANTOM web resource (http://fantom.gsc.riken.jp/4/download/). The data were originally published by The FANTOM Consortium (Carninci *et al.*, Science 2005).

References

Kawaji *et al.* (2010) Update of the FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation, *Nucleic Acids Research* **39**:D856-D860.
Carninci *et al.* (2005) The Transcriptional Landscape of the Mammalian Genome, *Science* **309**(5740):1559-1563.

Examples

```
data(FANTOMtimecourseCAGEmouse)
names(FANTOMtimecourseCAGEmouse)
head(FANTOMtimecourseCAGEmouse[["adipogenic_induction"]])
```

`FANTOMtissueCAGEhuman` *CAGE data produced in FANTOM projects for various human tissues and cell lines*

Description

This dataset contains all transcription start sites (TSSs) detected by CAGE and their frequency of usage in the various human tissues and cell lines generated in FANTOM3 and FANTOM4 projects.

Usage

```
data(FANTOMtissueCAGEhuman)
```

Details

`FANTOMtissueCAGEhuman` is a list of groups of human samples, which are grouped by the tissue of origin. Each element is named according to the corresponding tissue and it is a `data.frame` with genomic positions of TSSs detected by CAGE and frequency of their usage in various samples originating from that tissue. First three columns in every `data.frame` contain the chromosome (`chr`), the 1-based coordinate (`pos`) and the strand (`strand`) of the TSS, and the following columns contain the number of CAGE tags supporting that TSS in each specific sample. These columns are named according to the corresponding sample.

Source

The CAGE detected TSSs (mapped to the hg18 human genome assembly) and numbers of CAGE tags were obtained from the FANTOM web resource (http://fantom.gsc.riken.jp/4/download/). The data were originally published by The FANTOM Consortium (Carninci *et al.* Science 2005, and Carninci *et al.*, Nature Genetics 2006, and Faulkner *et al.*, Nature Genetics 2009).

References

Kawaji *et al.* (2010) Update of the FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation, *Nucleic Acids Research* **39**:D856-D860.
Carninci *et al.* (2005) The transcriptional landscape of the mammalian genome, *Science* **309**:1559-1563.
Carninci *et al.* (2006) Genome-wide analysis of mammalian promoter architecture and evolution, *Nature Genetics* **6**:626-635.
Faulkner *et al.* (2009) The regulated retrotransposon transcriptome of mammalian cells, *Nature Genetics* **41**:563-571.

Examples

```
data(FANTOMtissueCAGEhuman)
names(FANTOMtissueCAGEhuman)
head(FANTOMtissueCAGEhuman[["liver"]])
```

`FANTOMtissueCAGEmouse` *CAGE data produced in FANTOM projects for various mouse tissues and cell lines*

Description

This dataset contains all transcription start sites (TSSs) detected by CAGE and their frequency of usage in the various mouse tissues and cell lines generated in FANTOM3 and FANTOM4 projects.

Usage

```
data(FANTOMtissueCAGEmouse)
```

Details

`FANTOMtissueCAGEmouse` is a list of groups of mouse samples, which are grouped by the tissue of origin. Each element is named according to the corresponding tissue and it is a `data.frame` with genomic positions of TSSs detected by CAGE and frequency of their usage in various samples originating from that tissue. First three columns in every `data.frame` contain the chromosome (`chr`), the 1-based coordinate (`pos`) and the strand (`strand`) of the TSS, and the following columns contain the number of CAGE tags supporting that TSS in each specific sample. These columns are named according to the corresponding sample.

Source

The CAGE detected TSSs (mapped to the mm9 mouse genome assembly) and numbers of CAGE tags were obtained from the FANTOM web resource (http://fantom.gsc.riken.jp/4/download/). The data were originally published by The FANTOM Consortium (Carninci *et al.*, Science 2005, and Faulkner *et al.*, Nature Genetics 2009).

References

Kawaji *et al.* (2010) Update of the FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation, *Nucleic Acids Research* **39**:D856-D860.
Carninci *et al.* (2005) The transcriptional landscape of the mammalian genome, *Science* **309**:1559-1563.
Faulkner *et al.* (2009) The regulated retrotransposon transcriptome of mammalian cells, *Nature Genetics* **41**:563-571.

Examples

```
data(FANTOMtissueCAGEmouse)
names(FANTOMtissueCAGEmouse)
head(FANTOMtissueCAGEmouse[["lung"]])
```

Index