

# Package ‘ChIC.data’

October 14, 2018

**Title** ChIC package data

**Version** 1.0.0

**Maintainer** Carmen Maria Livi <carmen.livi@ifom.eu>

**Description** This package contains annotation and metagene profile data for the ChIC package.

**Depends** R (>= 3.5)

**Imports** caret (>= 6.0-78)

**biocViews** ExperimentData, ENCODE

**License** GPL-2

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 6.0.1

**NeedsCompilation** no

**git\_url** <https://git.bioconductor.org/packages/ChIC.data>

**git\_branch** RELEASE\_3\_7

**git\_last\_commit** ec766d5

**git\_last\_commit\_date** 2018-04-30

**Date/Publication** 2018-10-14

**Author** Carmen Maria Livi [aut, cre, dtc],  
Endre Sebestyen [aut]

## R topics documented:

chipSubset . . . . .	2
compendium_db . . . . .	2
compendium_db_tf . . . . .	3
compendium_profiles . . . . .	3
crossvalues_Chip . . . . .	4
hg19_chrom_info . . . . .	5
hg19_refseq_genes_filtered_granges . . . . .	5
inputSubset . . . . .	6
mm9_chrom_info . . . . .	6
mm9_refseq_genes_filtered_granges . . . . .	7
rf_models . . . . .	7
<b>Index</b>	<b>8</b>

---

chipSubset	<i>ChIP-seq bam file stored as spp tag-list for a subset of chromosomes for the chip</i>
------------	--

---

**Description**

Example data for manual and vignette. tag-list created with the read.bam.tags() function from spp package. The original bam file has been downloaded from ENCODE (ID: ENCFF000BFX).

**Usage**

```
data(chipSubset)
```

**Format**

list of 2 elements containing the reads and the read quality of the ChIP.

- tags : list containing the start coordinates of each read aligned (ChIP) (3'end)
- quality: list containing the read quality of each read

---

compendium_db	<i>Histone mark compendium</i>
---------------	--------------------------------

---

**Description**

Histone mark ChIP-seq compendium for the ChIC package. The compendium contains quality control metrics and metadata for 2329 histone mark samples analysed from ENCODE and Roadmap Epigenomics.

**Usage**

```
data(compendium_db)
```

**Format**

Data frame with 366 variables (quality control metrics, metadata) for each analysed sample.

**Source**

XX

---

compendium_db_tf	<i>Transcription factor compendium</i>
------------------	--

---

**Description**

Transcription factor ChIP-seq compendium for the ChIC package. The compendium contains quality control metrics and metadata for 1427 transcription factors analysed from ENCODE.

**Usage**

```
data(compendium_db_tf)
```

**Format**

Data frame with 366 variables (quality control metrics, metadata) for each analysed sample.

**Source**

XX

---

compendium_profiles	<i>Metagene profile data</i>
---------------------	------------------------------

---

**Description**

Compendium of averaged metagene profiles for the ChIC package. Contains averaged metagene profiles for following ChIP-seq datasets from ENCODE and Roadmap Epigenomics.

**Usage**

```
data(compendium_profiles)
```

**Format**

A list of data frames with the coordinates of the metagene profile of the respective chromatin mark.  
x : genomic coordinates mean : the mean of the signal intensity in the compendium sd : the standard deviation of the signal intensity in the compendium q1..q5 : being the respective quantile of the value distribution sderr : standard error

**Details**

- H2A.Z
- H2AFZ
- H2AK5ac
- H2AK9ac
- H2BK120ac
- H2BK12ac
- H2BK15ac

- H2BK20ac
- H2BK5ac
- H3K14ac
- H3K18ac
- H3K23ac
- H3K23me2
- H3K27ac
- H3K27me3
- H3K36me3
- H3K4ac
- H3K4me1
- H3K4me2
- H3K4me3
- H3K56ac
- H3K79me1
- H3K79me2
- H3K9ac
- H3K9me1
- H3K9me3
- H3T11ph
- H4K12ac
- H4K20me1
- H4K5ac
- H4K8ac
- H4K91ac
- POLR2A
- POLR2AphosphoS2
- POLR2AphosphoS5

**Source**

XX

---

`crossvalues_Chip`*CrsosCorrelation values for example CHIP-seq data for Vignette*

---

**Description**

Example data to be used in the vignette for chrom2

**Usage**`data(crossvalues_Chip)`**Format**

list of 20 elements containing EM scores

---

hg19_chrom_info	<i>hg19 chromosome information</i>
-----------------	------------------------------------

---

**Description**

hg19 chromosome information for the ChIC package.

**Usage**

```
data(hg19_chrom_info)
```

**Format**

A named list of int vectors with 2 elements, the start and end position of all hg19 chromosomes.

**Source**

<http://hgdownload.cse.ucsc.edu/goldenPath/hg19/bigZips/hg19.chrom.sizes>

---

hg19_refseq_genes_filtered_granges	<i>Filtered RefSeq gene annotation as GRanges</i>
------------------------------------	---

---

**Description**

RefSeq gene annotation data for the ChIC package in GRanges format, filtered by the gene length and by overlaps.

**Usage**

```
data(hg19_refseq_genes_filtered_granges)
```

**Format**

A GRanges object.

---

inputSubset	<i>ChIP-seq bam file stored as spp tag-list for a subset of chromosomes for the input</i>
-------------	---

---

### Description

Example data for manual and vignette. Tag-list created with the `read.bam.tags()` function from `spp` package for the input data. The original bam file has been downloaded from ENCODE (ID: ENCFF000BDQ).

### Usage

```
data(inputSubset)
```

### Format

list of 2 elements containing the reads and the read quality of the input.

- tags : list containing the start coordinates of each read aligned (input) (3'end)
- quality: list containing the read quality of each read

---

mm9_chrom_info	<i>mm9 chromosome information</i>
----------------	-----------------------------------

---

### Description

mm9 chromosome information for the ChIC package.

### Usage

```
data(mm9_chrom_info)
```

### Format

A named list of int vectors with 2 elements, the start and end position of all mm9 chromosomes.

### Source

<http://hgdownload.cse.ucsc.edu/goldenPath/mm9/bigZips/mm9.chrom.sizes>

---

`mm9_refseq_genes_filtered_granges`*Filtered RefSeq gene annotation as GRanges for mm9*

---

**Description**

RefSeq gene annotation data for the ChIC package in GRanges format, filtered by the gene length and by overlaps.

**Usage**

```
data(mm9_refseq_genes_filtered_granges)
```

**Format**

A GRanges object.

---

`rf_models`*Random forest models for chromatin marks ChIP-seq experiment classification*

---

**Description**

Random forest models based on ENCODE and Roadmap data for ChIP-seq experiment classification using the ChIC package.

**Usage**

```
data(rf_models)
```

**Format**

A list of 7 random forest models for the different chromatin marks and transcription factors:

- `broadEncode` : model for broad binding marks
- `H3K9Encode` : model for H3K9me3
- `H3K27Encode` : model for H3K27me3
- `H3K36Encode` : model for H3K36me3
- `RNAPol2Encode` : model for RNAPol2
- `sharpEncode` : model for sharp binding marks
- `TFmodel` : model for transcription factors

# Index

- \*Topic **annotation**
    - hg19\_refseq\_genes\_filtered\_granges, 5
    - mm9\_refseq\_genes\_filtered\_granges, 7
  - \*Topic **bamfile**
    - chipSubset, 2
    - inputSubset, 6
  - \*Topic **classification**
    - rf\_models, 7
  - \*Topic **compendium**
    - compendium\_db, 2
    - compendium\_db\_tf, 3
    - compendium\_profiles, 3
  - \*Topic **control**
    - compendium\_db, 2
    - compendium\_db\_tf, 3
  - \*Topic **crossvalues\_Chip**
    - crossvalues\_Chip, 4
  - \*Topic **hg19**
    - hg19\_chrom\_info, 5
  - \*Topic **histone**
    - compendium\_db, 2
  - \*Topic **marks**
    - compendium\_db, 2
  - \*Topic **metagene**
    - compendium\_profiles, 3
  - \*Topic **metrics**
    - compendium\_db, 2
    - compendium\_db\_tf, 3
  - \*Topic **mm9**
    - mm9\_chrom\_info, 6
  - \*Topic **profiles**
    - compendium\_profiles, 3
  - \*Topic **quality**
    - compendium\_db, 2
    - compendium\_db\_tf, 3
  - \*Topic **tf**
    - compendium\_db\_tf, 3
- chipSubset, 2  
compendium\_db, 2  
compendium\_db\_tf, 3  
compendium\_profiles, 3  
crossvalues\_Chip, 4  
hg19\_chrom\_info, 5  
hg19\_refseq\_genes\_filtered\_granges, 5  
inputSubset, 6  
mm9\_chrom\_info, 6  
mm9\_refseq\_genes\_filtered\_granges, 7  
rf\_models, 7