Package ‘QDNAseq.hg19’

Type Package

Title QDNAseq bin annotation for hg19

Version 1.4.0

Date 2015-09-30

Author Daoud Sie [aut, cre]

Maintainer Daoud Sie <d.sie@vumc.nl>

Description This package provides QDNAseq bin annotations for the human genome build hg19.

Depends R (>= 3.2.1), QDNAseq

biocViews ExperimentData, OrganismData, Homo_sapiens_Data

License GPL

URL https://github.com/tgac-vumc/QDNAseq.hg19

BugReports https://github.com/tgac-vumc/QDNAseq.hg19/issues

NeedsCompilation no

R topics documented:

QDNAseq.hg19-package ........................................ 2
hg19.1000kbp.SR50 ........................................ 2
hg19.100kbp.SR50 ........................................ 4
hg19.10kbp.SR50 ........................................ 5
hg19.15kbp.SR50 ........................................ 6
hg19.1kbp.SR50 ........................................ 7
hg19.30kbp.SR50 ........................................ 8
hg19.500kbp.SR50 ........................................ 9
hg19.50kbp.SR50 ......................................... 10
hg19.5kbp.SR50 ......................................... 11

Index 13
Description

This package provides QDNAseq bin annotations for the mouse genome build hg19 for bin sizes 1, 5, 10, 15, 30, 50, 100, 500 and 1000 kbp (kilobasepair).

The datasets are named as follows:

hg19.1kbp.SR50
hg19.5kbp.SR50
hg19.10kbp.SR50
hg19.15kbp.SR50
hg19.30kbp.SR50
hg19.50kbp.SR50
hg19.100kbp.SR50
hg19.500kbp.SR50
hg19.1000kbp.SR50

License

This package is licensed under GPL.

Author(s)

Daoud Sie

Examples

```r
data("hg19.30kbp.SR50")
assign("bins", get("hg19.30kbp.SR50"))
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")

# or

bins <- getBinAnnotations(binSize=30, genome="hg19")
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")
```

Description

Bin annotations are calculated for non-overlapping 1000kbp bins generated as described in Scheinin et al. (see references). The annotated data frame contains:

- chromosome: Chromosome name,
- start: Base pair start position,
- end: Base pair end position,
• bases: Percentage of non-N nucleotides (of full bin size),
• gc: Percentage of C and G nucleotides (of non-N nucleotides),
• mappability: Average mappability of 50mers with a maximum of 2 mismatches as described in by Derrien et al. (see references),
• blacklist: Percent overlap with ENCODE blacklisted regions (see references),
• residual: Median loess residual calculated from 1000 Genomes (see references),
• use: Whether the bin should be used in subsequent analysis steps,

Value

Returns an AnnotatedDataFrame object.

Author(s)

Daoud Sie

References


Examples

data("hg19.1000kbp.SR50")
assign("bins", get("hg19.1000kbp.SR50"))
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")

# or

bins <- getBinAnnotations(binSize=1000, genome="hg19")
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")
Bin annotations are calculated for non-overlapping 100kbp bins generated as described in Scheinin et al. (see references). The annotated data frame contains:

- chromosome: Chromosome name,
- start: Base pair start position,
- end: Base pair end position,
- bases: Percentage of non-N nucleotides (of full bin size),
- gc: Percentage of C and G nucleotides (of non-N nucleotides),
- mappability: Average mappability of 50mers with a maximum of 2 mismatches as described in by Derrien et al. (see references),
- blacklist: Percent overlap with ENCODE blacklisted regions (see references),
- residual: Median loess residual calculated from 1000 Genomes (see references),
- use: Whether the bin should be used in subsequent analysis steps,

Value

Returns an `AnnotatedDataFrame` object.

Author(s)

Daoud Sie

References


Examples

data("hg19.100kbp.SR50")
assign("bins", get("hg19.100kbp.SR50"))
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")

# or

bins <- getBinAnnotations(binSize=100, genome="hg19")
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")

---

**Description**

Bin annotations are calculated for non-overlapping 10kbp bins generated as described in Scheinin et al. (see references). The annotated data frame contains:

- chromosome: Chromosome name,
- start: Base pair start position,
- end: Base pair end position,
- bases: Percentage of non-N nucleotides (of full bin size),
- gc: Percentage of C and G nucleotides (of non-N nucleotides),
- mappability: Average mappability of 50mers with a maximum of 2 mismatches as described in by Derrien et al. (see references),
- blacklist: Percent overlap with ENCODE blacklisted regions (see references),
- residual: Median loess residual calculated from 1000 Genomes (see references),
- use: Whether the bin should be used in subsequent analysis steps,

**Value**

Returns an AnnotatedDataFrame object.

**Author(s)**

Daoud Sie

**References**


Examples

data("hg19.10kbp.SR50")
assign("bins", get("hg19.10kbp.SR50"))
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")

# or

bins <- getBinAnnotations(binSize=10, genome="hg19")
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")

### Description

Bin annotations are calculated for non overlapping 15kbp bins generated as described in Scheinin et al. (see references). The annotated data frame contains:

- chromosome: Chromosome name,
- start: Base pair start position,
- end: Base pair end position,
- bases: Percentage of non-N nucleotides (of full bin size),
- gc: Percentage of C and G nucleotides (of non-N nucleotides),
- mappability: Average mappability of 50mers with a maximum of 2 mismatches as described in by Derrien et al. (see references),
- blacklist: Percent overlap with ENCODE blacklisted regions (see references),
- residual: Median loess residual calculated from 1000 Genomes (see references),
- use: Whether the bin should be used in subsequent analysis steps,

### Value

Returns an AnnotatedDataFrame object.

### Author(s)

Daoud Sie

### References


Examples

data("hg19.15kbp.SR50")
assign("bins", get("hg19.15kbp.SR50"))
## Not run: readCounts <- binReadCounts(bins=bins, path="/bam")

# or

bins <- getBinAnnotations(binSize=15, genome="hg19")
## Not run: readCounts <- binReadCounts(bins=bins, path="/bam")

---

Description

Bin annotations are calculated for non-overlapping 1kbp bins generated as described in Scheinin et al. (see references). The annotated data frame contains:

- chromosome: Chromosome name,
- start: Base pair start position,
- end: Base pair end position,
- bases: Percentage of non-N nucleotides (of full bin size),
- gc: Percentage of C and G nucleotides (of non-N nucleotides),
- mappability: Average mappability of 50mers with a maximum of 2 mismatches as described in by Derrien et al. (see references),
- blacklist: Percent overlap with ENCODE blacklisted regions (see references),
- residual: Median loess residual calculated from 1000 Genomes (see references),
- use: Whether the bin should be used in subsequent analysis steps,

Value

Returns an AnnotatedDataFrame object.

Author(s)

Daoud Sie

References


Examples

```r
data("hg19.1kbp.SR50")
assign("bins", get("hg19.1kbp.SR50"))
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")

# or

bins <- getBinAnnotations(binSize=1, genome="hg19")
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")
```

---

**Description**

Bin annotations are calculated for non-overlapping 30kbp bins generated as described in Scheinin et al. (see references). The annotated data frame contains:

- chromosome: Chromosome name,
- start: Base pair start position,
- end: Base pair end position,
- bases: Percentage of non-N nucleotides (of full bin size),
- gc: Percentage of C and G nucleotides (of non-N nucleotides),
- mappability: Average mappability of 50mers with a maximum of 2 mismatches as described in by Derrien et al. (see references),
- blacklist: Percent overlap with ENCODE blacklisted regions (see references),
- residual: Median loess residual calculated from 1000 Genomes (see references),
- use: Whether the bin should be used in subsequent analysis steps,

**Value**

Returns an `AnnotatedDataFrame` object.

**Author(s)**

Daoud Sie

**References**


Examples

data("hg19.30kbp.SR50")
assign("bins", get("hg19.30kbp.SR50"))
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")

# or

bins <- getBinAnnotations(binSize=30, genome="hg19")
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")

---

**hg19.500kbp.SR50**  
**Hg19 500kbp bin annotations**

**Description**

Bin annotations are calculated for non-overlapping 500kbp bins generated as described in Scheinin et al. (see references). The annotated data frame contains:

- chromosome: Chromosome name,
- start: Base pair start position,
- end: Base pair end position,
- bases: Percentage of non-N nucleotides (of full bin size),
- gc: Percentage of C and G nucleotides (of non-N nucleotides),
- mappability: Average mappability of 50mers with a maximum of 2 mismatches as described in by Derrien et al. (see references),
- blacklist: Percent overlap with ENCODE blacklisted regions (see references),
- residual: Median loess residual calculated from 1000 Genomes (see references),
- use: Whether the bin should be used in subsequent analysis steps,

**Value**

Returns an AnnotatedDataFrame object.

**Author(s)**

Daoud Sie

**References**

DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly.  


```r
data("hg19.500kbp.SR50")
assign("bins", get("hg19.500kbp.SR50"))
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")

bins <- getBinAnnotations(binSize=500, genome="hg19")
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")
```

### hg19.50kbp.SR50

**Hg19 50kbp bin annotations**

#### Description

Bin annotations are calculated for non overlapping 50kbp bins generated as described in Scheinin et al. (see references). The annotated data frame contains:

- chromosome: Chromosome name,
- start: Base pair start position,
- end: Base pair end position,
- bases: Percentage of non-N nucleotides (of full bin size),
- gc: Percentage of C and G nucleotides (of non-N nucleotides),
- mappability: Average mappability of 50mers with a maximum of 2 mismatches as described in by Derrien et al. (see references),
- blacklist: Percent overlap with ENCODE blacklisted regions (see references),
- residual: Median loess residual calculated from 1000 Genomes (see references),
- use: Whether the bin should be used in subsequent analysis steps,

#### Value

Returns an `AnnotatedDataFrame` object.

#### Author(s)

Daoud Sie

#### References


Examples

data("hg19.5kbp.SR50")
assign("bins", get("hg19.5kbp.SR50"))
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")

# or

bins <- getBinAnnotations(binSize=50, genome="hg19")
## Not run: readCounts <- binReadCounts(bins=bins, path="./bam")

Description

Bin annotations are calculated for non overlapping 5kbp bins generated as described in Scheinin et al. (see references). The annotated data frame contains:

- chromosome: Chromosome name,
- start: Base pair start position,
- end: Base pair end position,
- bases: Percentage of non-N nucleotides (of full bin size),
- gc: Percentage of C and G nucleotides (of non-N nucleotides),
- mappability: Average mappability of 50mers with a maximum of 2 mismatches as described in by Derrien et al. (see references),
- blacklist: Percent overlap with ENCODE blacklisted regions (see references),
- residual: Median loess residual calculated from 1000 Genomes (see references),
- use: Whether the bin should be used in subsequent analysis steps,

Value

Returns an AnnotatedDataFrame object.

Author(s)

Daoud Sie

References


Examples

```r
data("hg19.5kbp.SR50")
assign("bins", get("hg19.5kbp.SR50"))
## Not run: readCounts <- binReadCounts(bins=bins, path="/bam")

# or

bins <- getBinAnnotations(binSize=5, genome="hg19")
## Not run: readCounts <- binReadCounts(bins=bins, path="/bam")
```
Index

*Topic **datasets**
- hg19.1000kbp.SR50, 2
- hg19.100kbp.SR50, 4
- hg19.10kbp.SR50, 5
- hg19.15kbp.SR50, 6
- hg19.1kbp.SR50, 7
- hg19.30kbp.SR50, 8
- hg19.500kbp.SR50, 9
- hg19.50kbp.SR50, 10
- hg19.5kbp.SR50, 11

*Topic **package**
- QDNaseq.hg19-package, 2

AnnotatedDataFrame, 3–11

- hg19.1000kbp.SR50, 2
- hg19.100kbp.SR50, 4
- hg19.10kbp.SR50, 5
- hg19.15kbp.SR50, 6
- hg19.1kbp.SR50, 7
- hg19.30kbp.SR50, 8
- hg19.500kbp.SR50, 9
- hg19.50kbp.SR50, 10
- hg19.5kbp.SR50, 11

QDNaseq.hg19 (QDNaseq.hg19-package), 2
QDNaseq.hg19-package, 2