

Package ‘strawr’

March 29, 2023

Title Fast Implementation of Reading/Dump for .hic Files

Version 0.0.91

Description API for fast data extraction for .hic files that provides programmatic access to the matrices. It doesn't store the pointer data for all the matrices, only the one queried, and currently we are only supporting matrices (not vectors).

Depends R (>= 3.4.0)

License MIT + file LICENSE

URL <https://github.com/aidenlab/straw/tree/master/R>

SystemRequirements libcurl: libcurl-devel (rpm) or
libcurl4-openssl-dev (deb)

Encoding UTF-8

Imports Rcpp

LinkingTo Rcpp

NeedsCompilation yes

RoxygenNote 7.2.0

Author Neva Cherniavsky Durand [aut, cre],
Muhammad Saad Shamim [aut],
Aiden Lab [cph]

Maintainer Neva Cherniavsky Durand <neva@broadinstitute.org>

Repository CRAN

Date/Publication 2023-03-29 20:00:02 UTC

R topics documented:

readHicBpResolutions	2
readHicChroms	2
readHicNormTypes	3
straw	3
strawr	4

Index	5
--------------	----------

readHicBpResolutions *Function for reading basepair resolutions from .hic file*

Description

Function for reading basepair resolutions from .hic file

Usage

```
readHicBpResolutions(fname)
```

Arguments

fname path to .hic file

Value

Vector of basepair resolutions

Examples

```
readHicBpResolutions(system.file("extdata", "test.hic", package = "strawr"))
```

readHicChroms *Function for reading chromosomes from .hic file*

Description

Function for reading chromosomes from .hic file

Usage

```
readHicChroms(fname)
```

Arguments

fname path to .hic file

Value

Data frame of chromosome names and lengths

Examples

```
readHicChroms(system.file("extdata", "test.hic", package = "strawr"))
```

readHicNormTypes	<i>Function for reading available normalizations from .hic file</i>
------------------	---

Description

Function for reading available normalizations from .hic file

Usage

```
readHicNormTypes(fname)
```

Arguments

fname	path to .hic file
-------	-------------------

Value

Vector of available normalizations

Examples

```
readHicNormTypes(system.file("extdata", "test.hic", package = "straw"))
```

straw	<i>Straw Quick Dump</i>
-------	-------------------------

Description

fast C++ implementation of dump. Not as fully featured as the Java version. Reads the .hic file, finds the appropriate matrix and slice of data, and outputs as data.frame in sparse upper triangular format. Currently only supporting matrices.

Usage

```
straw(norm, fname, chr1loc, chr2loc, unit, binsize, matrix = "observed")
```

Arguments

norm	Normalization to apply. Must be one of NONE/VC/VC_SQRT/KR. VC is vanilla coverage, VC_SQRT is square root of vanilla coverage, and KR is Knight-Ruiz or Balanced normalization.
fname	path to .hic file
chr1loc	first chromosome location
chr2loc	second chromosome location
unit	BP (BasePair) or FRAG (FRAGment)

binsize	The bin size. By default, for BP, this is one of <2500000, 1000000, 500000, 250000, 100000, 50000, 25000, 10000, 5000> and for FRAG this is one of <500, 200, 100, 50, 20, 5, 2, 1>.
matrix	Type of matrix to output. Must be one of observed/oe/expected. observed is observed counts, oe is observed/expected counts, expected is expected counts.

Details

Usage: straw <NONE/VC/VC_SQRT/KR> <hicFile(s)> <chr1>[:x1:x2] <chr2>[:y1:y2] <BP/FRAG> <binsize> [observed/oe/expected]

Value

Data.frame of a sparse matrix of data from hic file. x,y,counts

Examples

```
straw("NONE", system.file("extdata", "test.hic", package = "strawr"), "1", "1", "BP", 2500000)
```

strawr	<i>strawr</i>
--------	---------------

Description

API for fast data extraction for .hic files that provides programmatic access to the matrices. It doesn't store the pointer data for all the matrices, only the one queried, and currently we are only supporting matrices (not vectors).

Index

readHicBpResolutions, [2](#)
readHicChroms, [2](#)
readHicNormTypes, [3](#)

straw, [3](#)
strawr, [4](#)