

Package ‘ionflows’

February 20, 2015

Type Package

Title Calculate the Number of Required Flows for Semiconductor Sequencing

Version 1.1

Date 2014-10-21

Author Michael Bockmayr and Jan Budczies

Depends Biostrings

Maintainer Michael Bockmayr <michael.bockmayr@charite.de>

Description Two methods for calculation of the number of required flows for semiconductor sequencing: 1. Using a simulation, the number of flows can be calculated for a concrete list of amplicons. 2. An exact combinatorial model is evaluated to calculate the number of flows for a random ensemble of sequences.

License GPL-3

Encoding UTF-8

NeedsCompilation no

Repository CRAN

Date/Publication 2014-11-10 09:55:34

R topics documented:

ionflows-package	2
chp2	2
flowsPanel	3
flowsRandom	4
readBed	5

Index	6
--------------	----------

ionflows-package	<i>Calculate the Number of Required Flows for Semiconductor Sequencing</i>
------------------	--

Description

This package implements two methods for calculation of the number of required flows for semiconductor sequencing: 1. Using a simulation, the number of flows can be calculated for a concrete list of amplicons. 2. An exact combinatorial model is evaluated to calculate the number of flows for a random ensemble of sequences.

Details

Package:	ionflows
Type:	Package
Version:	1.0
Date:	2014-10-21
License:	GPL-3

Author(s)

Michael Bockmayr <michael.bockmayr@charite.de>,
Jan Budczies <jan.budczies@charite.de>,
Institute of Pathology, Charité University Hospital,
Charitéplatz 1, 10117 Berlin, Germany

References

Budczies, J et al. *Semiconductor sequencing: How many flows do you need?* Submitted.

chp2	<i>An example BED file.</i>
------	-----------------------------

Description

This is an example BED File corresponding to the Ion AmpliSeq (TM) Cancer Hotspot Panel v2. It contains 207 regions that are frequently mutated in cancer.

Usage

```
data(chp2)
```

Format

A data frame with 207 observations on 6 variables.

References

<http://tools.lifetechnologies.com/content/sfs/brochures/Ion-AmpliSeq-Cancer-Hotspot-Panel-Flyer.pdf>

Examples

```
data(chp2, package="ionflows")
```

flowsPanel	<i>Simulation algorithm to calculate the number of required flows for a sequence panel in BED format.</i>
------------	---

Description

This function calculates the number of required flows for a concrete amplicon panel.

Usage

```
flowsPanel(bed.table, genom, order.flows = c("T", "A", "C", "G"),  
           ret.seq = FALSE)
```

Arguments

bed.table	A data frame corresponding to a BED file, which can be read into R with the function readBed .
genom	Genome to be used to extract the sequences. It should be in the BSgenome class. The genome corresponding to the UCSC hg19 build can be obtained from the package BSgenome.Hsapiens.UCSC.hg19 .
order.flows	Character vector indicating the order of the nucleotides used to synthesize the complementary strand during semiconductor sequencing.
ret.seq	logical. If 'TRUE', a matrix containing the sequences extracted from the input BED file and the corresponding forward and backward flow sequence constructed during semiconductor sequencing is returned.

Author(s)

Michael Bockmayr and Jan Budczies

See Also

[readBed](#)

Examples

```

### Load the BED file using the readBed function or use the example dataset:

data(chp2,package="ionflows")

### Load the BSgenome from which the genomic sequences of the amplicons in
### the BED file are to be extracted from, for example the BSgenome.Hsapiens.UCSC.hg19
### package for the UCSC hg19 build, which should be priorly installed .

## Not run: library(BSgenome.Hsapiens.UCSC.hg19)

### Run flowsPanel

## Not run: flowsPanel(chp2,BSgenome.Hsapiens.UCSC.hg19)

### A BED file can also be directly loaded into R:

## Not run: dset <- readBed("path/to/bed")
## Not run: flowsPanel(dset,genom)

```

flowsRandom	<i>A function to compute the percentage of covered random sequences for a given number of flows.</i>
-------------	--

Description

Evaluation of an exact expression to obtain the percentage of all random nucleotide sequences of a given length (n) that can be sequenced with semiconductor sequencing given a certain number of flows (k).

Usage

```
flowsRandom(k = 300, n = 150)
```

Arguments

k	Integer. Number of flows available during semiconductor sequencing.
n	Integer. Length of the random sequences to be sequenced.

Value

P	Percentage of sequences that are completely covered
Q	Percentage of sequences that are not completely covered
δ	$1 - (P + Q)$, should be a small number when calculation is accurate
N	Number of summands for calculation of P
M	Number of summands for calculation of Q
summands	$N + M =$ total number of summands

Author(s)

Michael Bockmayr and Jan Budczies

Examples

```
### To compute the percentage of covered sequences of length 150 with 300 flows run:  
flowsRandom(300,150)
```

readBed

This is a function to read BED files.

Description

This function can be used to read BED files in the standard format into R.

Usage

```
readBed.bed.file = NULL)
```

Arguments

`bed.file` Character string with the path of the BED file which the data are to be read from. If no path is provided, the function [file.choose](#) is called to interactively load the file.

Author(s)

Michael Bockmayr and Jan Budczies

Examples

```
## Not run: dset <- readBed()
```

Index

*Topic **datasets**

chp2, [2](#)

*Topic **package**

ionflows-package, [2](#)

chp2, [2](#)

file.choose, [5](#)

flowsPanel, [3](#)

flowsRandom, [4](#)

ionflows (ionflows-package), [2](#)

ionflows-package, [2](#)

readBed, [3](#), [5](#)