Package 'FMIndex'

June 26, 2023

Title A package for computing FM index

Version 0.99.0

Description The package computes the FM index of a DNA string contained in a fasta file.
The user can specify the path where he wants to save the FM index data structures.

Imports Biostrings, utils

Suggests knitr, rmarkdown, BiocStyle

License GPL-2

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.3

VignetteBuilder knitr

biocViews Normalization, Preprocessing

Depends R (>= 2.10)

NeedsCompilation no

Author Sara Resta [aut, cre] (YOUR-ORCID-ID)

Maintainer Sara Resta <sara.resta@mail.polimi.it>

R topics documented:

FMIndex-package	*FMIndex - Computer the FM index for a DNA sequence contained in a fasta file The* FMIndex *package implements all the functions necessary to compute the data structures of the FM index*

Description

Package:	FMIndex
Type:	Package
Date:	2023-06-20
License:	GPL (>=2)

Author(s)

Sara Resta [aut, cre]
Politecnico di Milano
Maintainer: Sara Resta
E-mail: sara.resta@mail.polimi.it

References

http://some.url/here/maybe/
Resta et al.

FMindex	*Writes the FM index*

Description

This function writes the FM index of the DNA sequence contained in the .fasta file given in input

Usage

```
FMindex(fastafile, usrpath, tallywidth = 1, includeTerminationCharacter = TRUE)
```

Arguments

`fastafile`	The .fasta file containing one single DNA sequence, files containing multiple sequences are not allowed
`usrpath`	The directory path where the user wants to save the files: it must end with '/' (you can leave it empty if you want to save the files in your working directory).
`tallywidth`	How you want to compress the tally table. Default value=1
`includeTerminationCharacter`	
	default value: true

Value

A list containint all the data structures necessary for the FM index

Examples

```
FMindex(system.file("extdata/seq5.fasta",package="FMIndex"),"")
FMindex(system.file("extdata/seq5.fasta",package="FMIndex"),"",tallywidth=1)
```

---

getBWTcharacter *Compute the i-th character of the BWT*

Description

This function computes the i-th character of the BWT of a given sequence which corresponds to the i-th character in the L column.

Usage

```
getBWTcharacter(index, sequence, suffixarray)
```

Arguments

`index`	The position which you want to find the character
`sequence`	The string you want to know the i-th L column character
`suffixarray`	An integer vector containing the suffix array of the input string

Value

A character containing the i-th character of the input sequence

---

getFcolumn *Compute the F column*

Description

This function gets a DNAString and gives the F column of the FM index

Usage

```
getFcolumn(seq, includeTerminationCharacter = TRUE)
```

Arguments

seq   the DNAString to be analysed

includeTerminationCharacter

                True if you want to include the ".", false if you want to exclude it. Default
                value:TRUE

Value

An integer vector with all the frequences in the dna sequence given in input

Examples

```
getFcolumn(Biostrings::DNAString("ACCT"))
```

---

getLcolumn *Compute the L column of a sequence*

Description

This function gets a character and gives the F column of the FM index

Usage

```
getLcolumn(sequence)
```

Arguments

sequence   the string to be analysed

Value

A DNAString containing the L column of the FM index

Examples

```
getLcolumn(Biostrings::DNAString("ACGT"))
getLcolumn(Biostrings::DNAString("AACCGT"))
```

---

getSuffix   *Get the suffix*

Description

This function gets the suffix of the string starting from a position given in input

Usage

```
getSuffix(index, sequence)
```

Arguments

index	An integer representing the starting position of the string
sequence	The sequence you want to find the suffix

Value

The substring starting at the given position

---

getTallyTable   *Create the Tally Table*

Description

Creates the Tally Table of a given BWT of a string

Usage

```
getTallyTable(l.col, rowwidth = 1)
```

Arguments

l.col	A DNAString containing the L column of the FM index
rowwidth	Step for saving rows of the tally table. The default value is 1 which corresponds to save all the rows of the tally table

Value

The tally table of the input sequence

Examples

```
getTallyTable(Biostrings::DNAString("T.AACCG"))
getTallyTable(Biostrings::DNAString("T.AACCG"),2)
```

---

getTallyTableLine   *Create the i-th row of the Tally Table*

Description

This function creates the i-th row of the Tally Table

Usage

```
getTallyTableLine(index, l.col, alphabet)
```

Arguments

index	An integer representing the index of the i-th row you want to build
l.col	A DNAString object containing the L column of the FM index
alphabet	A character vector containing the alphabet to consider in counting the elements

Value

the i-th row of the Tally Table

Index