# Package 'FMIndex'

June 25, 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.

---

computeFMIndex                *Writes the FM index*

---

## Description

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

## Usage

```
computeFMIndex(
  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 path where the user wants to save the files. |
| tallywidth | How you want to compress the tally table. Default value=1 |
| includeTerminationCharacter | |
| | default value: true |

## Examples

```
computeFMIndex(system.file("extdata/seq5.fasta",package="FMIndex"),"test01.txt")
computeFMIndex(system.file("extdata/seq5.fasta",package="FMIndex"),"test02.txt",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

## Examples

```
getBWTcharacter(2,"ACGT",c(5,1,2,3,4))
getBWTcharacter(3,"ACG",c(4,1,2,3))
```

---

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 sting containing the L column of the FM index

### Examples

```
getLcolumn("ACGT")
getLcolumn("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

### Examples

```
getSuffix(2,"ACGT")
getSuffix(5,"ACCCAGT")
```

# Index