

Package ‘NeedlemanWunsch’

July 13, 2021

Type Package

Title Perform Needleman-Wunsch algorithm

Version 0.1.0

Description This package implements a pairwise global sequence alignment method.
The objective is to find the optimal global sequence alignment between two nucleotide sequences using the Needleman-Wunsch algorithm. It also provides functions to calculate separately the score matrix and the direction matrix with alignment.

biocViews Software, Annotation, Alignment

License GPL-2

Encoding UTF-8

LazyData false

RoxygenNote 7.1.1

Suggests BiocStyle,
knitr,
rmarkdown,
testthat (>= 3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

R topics documented:

creatingMatrix	1
globalAlignmentNeedlemanWunsch	2
isItAMatch	3
performTraceback	4

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

creatingMatrix	<i>Creation of the scores matrix</i>
----------------	--------------------------------------

Description

creatingMatrix creates a matrix containing the scores of the alignment for the NW algorithm

Usage

```
creatingMatrix(matchCost, mismatchCost, gapPenalty, sequenceA, sequenceB)
```

Arguments

matchCost	The cost of a match
mismatchCost	The cost of a mismatch
gapPenalty	The gap penalty(cost of gap/indel)
sequenceA	The first nucleotide sequence (positioned on the rows)
sequenceB	The second nucleotide sequence (positioned on the columns)

Value

A matrix containing the scores of the alignment

Examples

```
sequenceA <- "GTT"
sequenceB <- "GCATT"
matchCost <- 7
mismatchCost <- -3
gapPenalty <- -4
scoreMatrix <- creatingMatrix(matchCost,mismatchCost,gapPenalty,sequenceA,sequenceB)
```

```
globalAlignmentNeedlemanWunsch
Needleman Wunsch
```

Description

Core function of the package, performs the Needleman-Wunsch algorithm

Usage

```
globalAlignmentNeedlemanWunsch(
  matchCost,
  mismatchCost,
  gapPenalty,
  sequenceA,
  sequenceB
)
```

Arguments

matchCost	The cost of a match
mismatchCost	The cost of a mismatch
gapPenalty	The gap penalty(cost of gap/indel)
sequenceA	The first nucleotide sequence (positioned on the rows)
sequenceB	The second nucleotide sequence (positioned on the columns)

Value

a list containing the score matrix, the direction matrix, the first alignment, the second alignment

References

https://en.wikipedia.org/wiki/Needleman-Wunsch_algorithm

Examples

```
sequenceA <- "GTT"
sequenceB <- "GCATT"
matchCost <- 7
mismatchCost <- -3
gapPenalty <- -4
result <- globalAlignmentNeedlemanWunsch(matchCost,mismatchCost,gapPenalty,sequenceA,sequenceB)
```

isItAMatch

Checks if the current characters are matching

Description

isItAMatch is a utility function that checks whether the i-th and j-th characters of the two sequences are matching

Usage

```
isItAMatch(sequenceA, sequenceB, i, j)
```

Arguments

sequenceA	The first nucleotide sequence (positioned on the rows)
sequenceB	The second nucleotide sequence (positioned on the columns)
i	The index for sequenceA (rows)
j	The index for sequenceB (columns)

Value

TRUE if the two characters are matching, FALSE otherwise

Examples

```
isItAMatch("GTT", "GCATT", 1, 2)
```

performTraceback	<i>Traceback function</i>
------------------	---------------------------

Description

performTraceback is a function that creates a matrix with the directions of the alignment while computing the alignment itself

Usage

```
performTraceback(matrix, sequenceA, sequenceB)
```

Arguments

matrix	The score matrix computed with the two sequences
sequenceA	The first nucleotide sequence (positioned on the rows)
sequenceB	The second nucleotide sequence (positioned on the columns)

Value

a list containing the directionMatrix, the first alignment, the second alignment

Examples

```
isItAMatch("GTT", "GCATT", 1, 2)
```

Index

`creatingMatrix`, [1](#)

`globalAlignmentNeedlemanWunsch`, [2](#)

`isItAMatch`, [3](#)

`performTraceback`, [4](#)