Package ‘RNAstructureModuleMiner’

May 23, 2017

Type Package

Title RNA Secondary Structure Comparison and Module Mining

Version 0.1.0

Author Zheng HeWei

Depends RRNA

Maintainer Zheng HeWei<`zhenghw@hrbmu.edu.cn`>

Description Functions in this program is designed for RNA secondary structure plotting, comparison and module mining. Given a RNA secondary structure, you can obtain stem regions, hairpin loops, internal loops, bulge loops and multibranch loops of this RNA structure using this program. They are the basic modules of RNA secondary structure. For each module you get, you can use this program to label the RNA structure with a specific color. You can also use this program to compare two RNA secondary structures to get a score that represents similarity.

License GPL-2

LazyData TRUE

NeedsCompilation no

Repository CRAN

Date/Publication 2017-05-23 06:32:13 UTC

R topics documented:

`bulge_loop` *Get bulge loops in a RNA secondary structure*

Description

Given a RNA secondary structure,it compute bulge loops in the RNA secondary structure

Usage

```
bulge_loop(ctFile)
```

Arguments

`ctFile`	A RNA secondary structure file without the first line of free energy information

Value

Return a list containing base positions in bulge loops

Author(s)

Zheng Hewei

Examples

```
###Create a RNA secondary structure
col1 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19)
col2 <- c("G","C","C","A","C","C","C","U","G","C","A","G","G","G","U","C","G","G","C")
col3 <- c(0,1,2,3,4,5,6,7,8,0,10,11,12,13,14,15,16,17,18)
col4 <- c(2,3,4,5,6,7,8,9,0,11,12,13,14,15,16,17,18,19,0)
col5 <- c(19,18,17,15,14,13,12,11,10,9,8,7,6,5,4,0,3,2,1)
col6 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19)
data <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol =6)
###Compute bulge loops in the RNA secondary structure
bulge_loop(data)
```

`compare` *compare*

Description

return similarity score of two RNA secondary structure

Usage

```
compare(ctFile1,ctFile2)
```

Arguments

`ctFile1`	A RNA secondary structure file without the first line of free energy information
`ctFile2`	A RNA secondary structure file without the first line of free energy information

Value

Returns a numerical value which represent the similarity of the two RNA secondary structures.The larger the value, the more similar the two RNA structures are.The maximum value is 10, representing the two RNA secondary structures exactly the same,and 0 is the minmum value.

Examples

```
###
col1 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20)
col2 <- c("G","A","A","G","A","T","C","A","T","G","C","A","G","T","C","A","T","C","A","G")
col3 <- c(0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19)
col4 <- c(2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,0)
col5 <- c(18,17,0,15,14,13,0,0,0,0,0,0,6,5,4,0,2,1,0,0)
col6 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20)

col1_2 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13)
col2_2 <- c("G","A","C","U","G","G","G","G","C","G","G","U","C")
col3_2 <- c(0,1,2,3,4,5,6,7,8,9,10,11,12)
col4_2 <- c(2,3,4,5,6,7,8,9,10,11,12,13,0)
col5_2 <- c(13,12,11,10,9,0,0,0,5,4,3,2,1)
col6_2 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13)
data1 <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol = 6)
data2 <- matrix(c(col1_2,col2_2,col3_2,col4_2,col5_2,col6_2),byrow=FALSE,ncol = 6)
compare(data1,data2)
```

`ct2dot` *ct2dot*

Description

Given a RNA secondary structure,it compute the RNA secondary structure in bracket dot form

Usage

```
ct2dot(ctFile)
```

Arguments

`ctFile`	A RNA secondary structure file without the first line of free energy information

Value

return a list including the RNA seqence and the RNA secondary structure in bracket dot form

Examples

```
###
col1 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13)
col2 <- c("G","A","C","U","G","G","G","G","C","G","G","U","C")
col3 <- c(0,1,2,3,4,5,6,7,8,9,10,11,12)
col4 <- c(2,3,4,5,6,7,8,9,10,11,12,13,0)
col5 <- c(13,12,11,10,9,0,0,0,5,4,3,2,1)
col6 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13)
data <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol = 6)
data <- as.data.frame(data)
ct2dot(data)
```

`hairpin_loop` *hairpin_loop*

Description

Given a RNA secondary structure,it compute hairpin loops in the RNA secondary structure

Usage

```
hairpin_loop(ctFile)
```

Arguments

`ctFile`	A RNA secondary structure file without the first line of free energy information

Value

Return a list containing base positions in hairpin loops,and the length of the number of hairpin loops

Examples

```
###
col1 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13)
col2 <- c("G","A","C","U","G","G","G","G","C","G","G","U","C")
col3 <- c(0,1,2,3,4,5,6,7,8,9,10,11,12)
col4 <- c(2,3,4,5,6,7,8,9,10,11,12,13,0)
col5 <- c(13,12,11,10,9,0,0,0,5,4,3,2,1)
col6 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13)
data <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol = 6)
hairpin_loop(data)
```

`internal_loop` *Get Internal Loop*

Description

Given a RNA secondary structure,it compute internal loops in the RNA secondary structure

Usage

```
internal_loop(ctFile)
```

Arguments

`ctFile`	A RNA secondary structure file without the first line of free energy information

Value

Return a list containing base positions in internal loops,and the length of the number of internal loops

Examples

```
###
col1 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20)
col2 <- c("G","A","A","G","A","T","C","A","T","G","C","A","G","T","C","A","T","C","A","G")
col3 <- c(0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19)
col4 <- c(2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,0)
col5 <- c(18,17,0,15,14,13,0,0,0,0,0,0,6,5,4,0,2,1,0,0)
col6 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20)
data <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol =6)
internal_loop(data)
```

`multi_branch_loop` *multi_branch_loop*

Description

Given a RNA secondary structure,it compute multi-branch loops in the RNA secondary structure

Usage

```
multi_branch_loop(ctFile)
```

Arguments

`ctFile`	A RNA secondary structure file without the first line of free energy information

Value

Return a list containing base positions in multi-branch loops,and the length of the number of multi-branch loops

Examples

```
###
col1 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20
,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40
,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60
,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80
,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100
,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115
,116,117,118,119)
col2 <- c(”C”,”C”,”U”,”A”,”G”,”U”,”G”,”A”,”C”,”A”,”A”,”U”,”A”
,”G”,”C”,”G”,”A”,”G”,”A”,”G”,”A”,”G”,”G”,”A”,”A”,”A”,”C”,”A”
,”C”,”C”,”C”,”G”,”U”,”C”,”U”,”C”,”C”,”A”,”U”,”C”,”C”,”C”,”G”
,”A”,”A”,”C”,”A”,”C”,”G”,”A”,”A”,”G”,”G”,”U”,”U”,”A”,”A”,”G”
,”C”,”U”,”C”,”U”,”C”,”C”,”C”,”A”,”G”,”C”,”G”,”C”,”C”,”G”,”A”
,”U”,”G”,”G”,”U”,”A”,”G”,”U”,”U”,”G”,”G”,”G”,”G”,”C”,”C”,”A”
,”G”,”C”,”G”,”C”,”C”,”C”,”C”,”U”,”G”,”C”,”A”,”A”,”G”,”A”,”G”
,”U”,”A”,”G”,”G”,”U”,”U”,”G”,”U”,”C”,”G”,”C”,”U”,”A”,”G”,”G”
,”C”)
col3 <- c(0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19
,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39
,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59
,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79
,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99
,100,101,102,103,104,105,106,107,108,109,110,111,112,113,114
,115,116,117,118)
col4 <- c(2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21
,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41
,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61
,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81
```

```
,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100
,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115
,116,117,118,119,120)
col5 <- c(118,117,116,115,114,113,112,111,0,0,0,0,0,67,66,64
,63,62,0,61,60,59,0,0,0,0,55,54,53,50,49,48,47,0,0,0,0,0,0,0
,0,0,0,0,0,0,33,32,31,30,0,0,29,28,27,0,0,0,22,21,20,18,17,16
,0,15,14,0,0,107,106,105,104,103,102,101,0,0,98,97,96,95,94
,93,92,91,0,0,0,0,86,85,84,83,82,81,80,79,0,0,76,75,74,73,72
,71,70,0,0,0,8,7,6,5,4,3,2,1,0)
col6 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20
,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40
,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60
,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80
,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100
,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115
,116,117,118,119)

data <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol =6)
multi_branch_loop(data)
```

`plot_bulge` *plot_bulge*

Description

Given a RNA secondary structure,it compute bulge loops in the RNA secondary structure and plots the RNA secondary structure

Usage

```
plot_bulge(ctFile)
```

Arguments

`ctFile`	A RNA secondary structure file without the first line of free energy information

Value

Return a list containing base positions in bulge loops

Examples

```
###
col1 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19)
col2 <- c("G","C","C","A","C","C","C","U","G","C","A","G","G","G","U","C","G","G","C")
col3 <- c(0,1,2,3,4,5,6,7,8,0,10,11,12,13,14,15,16,17,18)
col4 <- c(2,3,4,5,6,7,8,9,0,11,12,13,14,15,16,17,18,19,0)
col5 <- c(19,18,17,15,14,13,12,11,10,9,8,7,6,5,4,0,3,2,1)
col6 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19)
data <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol =6)
plot_bulge(data)
```

---

plot_hairpin *plot_hairpin*

Description

Given a RNA secondary structure,it compute hairpin loops in the RNA secondary structure and plots the RNA secondary structure

Usage

```
plot_hairpin(ctFile)
```

Arguments

`ctFile`	A RNA secondary structure file without the first line of free energy information

Value

Return a list containing base positions in hairpin loops

Examples

```
###
col1 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13)
col2 <- c("G","A","C","U","G","G","G","G","C","G","G","U","C")
col3 <- c(0,1,2,3,4,5,6,7,8,9,10,11,12)
col4 <- c(2,3,4,5,6,7,8,9,10,11,12,13,0)
col5 <- c(13,12,11,10,9,0,0,0,5,4,3,2,1)
col6 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13)
data_1 <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol = 6)
plot_hairpin(data_1)
```

---

plot_internal_loop *plot_internal_loop*

Description

Given a RNA secondary structure,it compute internal loops in the RNA secondary structure and plots the RNA secondary structure

Usage

```
plot_internal_loop(ctFile)
```

Arguments

`ctFile`	A RNA secondary structure file without the first line of free energy information

Value

Return a list containing base positions in internal loops

Examples

```
###
col1 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20)
col2 <- c("G","A","A","G","A","T","C","A","T","G","C","A","G","T","C","A","T","C","A","G")
col3 <- c(0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19)
col4 <- c(2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,0)
col5 <- c(18,17,0,15,14,13,0,0,0,0,0,0,6,5,4,0,2,1,0,0)
col6 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20)
data <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol =6)
data <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol =6)
plot_internal_loop(data)
```

`plot_multi_branch_loop`

plot_multi_branch_loop

Description

Given a RNA secondary structure,it compute multi-branch loops in the RNA secondary structure and plots the RNA secondary structure

Usage

```
plot_multi_branch_loop(ctFile)
```

Arguments

`ctFile`	A RNA secondary structure file without the first line of free energy information

Value

Return a list containing base positions in multi-branch loops

Examples

```
###
col1 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20
,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40
,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60
,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80
,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100
,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115
,116,117,118,119)

col2 <- c("C","C","U","A","G","U","G","A","C","A","A","U","A"
```

```
,"G","C","G","A","G","A","G","A","G","G","A","A","A","C","A"
,"C","C","C","G","U","C","U","C","C","A","U","C","C","C","G"
,"A","A","C","A","C","G","A","A","G","G","U","U","A","A","G"
,"C","U","C","U","C","C","C","A","G","C","G","C","C","G","A"
,"U","G","G","U","A","G","U","U","G","G","G","G","C","C","A"
,"G","C","G","C","C","C","C","U","G","C","A","A","G","A","G"
,"U","A","G","G","U","U","G","U","C","G","C","U","A","G","G"
,"C")
col3 <- c(0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19
,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39
,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59
,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79
,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99
,100,101,102,103,104,105,106,107,108,109,110,111,112,113,114
,115,116,117,118)

col4 <- c(2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21
,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41
,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61
,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81
,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100
,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115
,116,117,118,119,120)

col5 <- c(118,117,116,115,114,113,112,111,0,0,0,0,0,67,66,64
,63,62,0,61,60,59,0,0,0,0,55,54,53,50,49,48,47,0,0,0,0,0,0,0
,0,0,0,0,0,0,33,32,31,30,0,0,29,28,27,0,0,0,22,21,20,18,17,16
,0,15,14,0,0,107,106,105,104,103,102,101,0,0,98,97,96,95,94
,93,92,91,0,0,0,0,86,85,84,83,82,81,80,79,0,0,76,75,74,73,72
,71,70,0,0,0,8,7,6,5,4,3,2,1,0)

col6 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20
,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40
,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60
,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80
,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100
,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115
,116,117,118,119)

data <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol =6)
plot_multi_branch_loop(data)
```

`plot_stem` *plot_stem*

Description

Given a RNA secondary structure,it compute stems in the RNA secondary structure and plots the RNA secondary structure

Usage

```
plot_stem(ctFile)
```

Arguments

`ctFile`	A RNA secondary structure file without the first line of free energy information

Value

Return a list containing base positions in stems

Examples

```
###
col1 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13)
col2 <- c("G","A","C","U","G","G","G","G","C","G","G","U","C")
col3 <- c(0,1,2,3,4,5,6,7,8,9,10,11,12)
col4 <- c(2,3,4,5,6,7,8,9,10,11,12,13,0)
col5 <- c(13,12,11,10,9,0,0,0,5,4,3,2,1)
col6 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13)
data <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol =6)
plot_stem(data)
```

`stem` *stem*

Description

Given a RNA secondary structure,it compute stem in the RNA secondary structure

Usage

```
stem(ctFile)
```

Arguments

`ctFile`	A RNA secondary structure file without the first line of free energy information

Value

Return a list containing base positions in stems

Examples

```
###
col1 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13)
col2 <- c("G","A","C","U","G","G","G","G","C","G","G","U","C")
col3 <- c(0,1,2,3,4,5,6,7,8,9,10,11,12)
col4 <- c(2,3,4,5,6,7,8,9,10,11,12,13,0)
col5 <- c(13,12,11,10,9,0,0,0,5,4,3,2,1)
col6 <- c(1,2,3,4,5,6,7,8,9,10,11,12,13)
data <- matrix(c(col1,col2,col3,col4,col5,col6),byrow=FALSE,ncol = 6)
stem(data)
```

Index