

Package ‘FusionMetaCaller’

April 21, 2015

Type Package

Title FusionMetaCaller

Version 1.0

Date 2015-04-20

Author Silvia Liu and George C. Tseng

Maintainer Silvia Liu <silvia.shuchang.liu@gmail.com>

Description A meta-caller algorithm by combining top performing methods to re-prioritize candidate fusion transcripts with high confidence that can be followed by experimental validation.

License GPL-2

R topics documented:

FusionMetaCaller-package	1
FusionMetaCaller	2
GenerateReadCount	3

FusionMetaCaller-package

A package to combine to combine the fusion results detected by multiple tools

Description

A meta-caller algorithm by combining top performing methods to re-prioritize candidate fusion transcripts with high confidence that can be followed by experimental validation.

Details

Package:	FusionMetaCaller
Type:	Package
Version:	1.0
Date:	2015-04-20
License:	University of Pittsburgh

FusionMetaCaller

Author(s)

Silvia Liu Maintainer: Silvia Liu <silvia.shuchang.liu@gmail.com>

References

Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. (2015) Silvia Liu[^], Wei-Hsiang Tsai[^], Ying Ding[^], Rui Chen, Zhou Fang, SungHwan Kim, Tianzhou Ma, Ting-Yu Chang, Jianhua Luo, Hsei-Wei Wang*, I-Fang Chung* and George C. Tseng*.

FusionMetaCaller *A meta-caller to combine and reprioritize the fusion transcripts detected by multiple tools*

Description

Top performing methods likely had complementary advantages to accurately detect different types of fusion events. First of all, we selected fusion events detected by at least a certain number of tools. We next ranked the detected fusion events from each method by the number of supporting reads. Rank sums of the selected fusion events were calculated and the fusion events were reprioritized accordingly.

Usage

```
FusionMetaCaller(countMatrix, vote, plot, trueFusion)
```

Arguments

<code>countMatrix</code>	An integer matrix for supporting read counts with rows representing the fusion events and columns representing the fusion detection tools.
<code>vote</code>	Minimum number of supporting tools to keep the fusion events in the combined result. The default value is 2.
<code>plot</code>	A boolean variable to indicate whether or not to plot the precision-recall curve. The plot will be saved as 'precision_recall_plot.pdf'. The default value is FALSE.
<code>trueFusion</code>	A vector to list all the true fusion events. This argument is only needed when plot is TRUE.

Value

<code>sortMatrix</code>	The combined results to reprioritize the fusion events.
-------------------------	---

Author(s)

Silvia Liu <silvia.shuchang.liu@gmail.com>

References

Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. (2015) Silvia Liu[^], Wei-Hsiang Tsai[^], Ying Ding[^], Rui Chen, Zhou Fang, SungHwan Kim, Tianzhou Ma, Ting-Yu Chang, Jianhua Luo, Hsei-Wei Wang*, I-Fang Chung* and George C. Tseng*.

Examples

```
## generate synthetic read count matrix
syn <- GenerateReadCount(fusionNum=100,toolNum=5,trueFusionNum=70,missingRate=c(0.4,0.6,0.8),
countMatrix <- syn$countMatrix
trueFusion <- syn$trueFusion

## the meta-caller to reprioritize the fusion events
metaRes <- FusionMetaCaller(countMatrix, vote=2)

## plot the precision-recall curve
metaRes <- FusionMetaCaller(countMatrix, vote=2, plot=TRUE,trueFusion)
```

GenerateReadCount *Simulation of supporting read counts detected by multiple tools*

Description

Generate synthetic read count matrix with rows representing fusion events and columns representing fusion detection tools.

Usage

```
GenerateReadCount(fusionNum, toolNum, trueFusionNum, missingRate, seed)
```

Arguments

fusionNum	An integer indicating the total number of fusion events
toolNum	An integer indicating the total nubmer of fusion detection tools
trueFusionNum	An integer for the total number of true fusion events. This value should be less than or equal to fusionNum.
missingRate	A vector representing the missing detection rate of each tool. The true fusion events have 2/3 missing rate. This value should be in [0,1]. The length of the vector is the same as toolNum.
seed	A seed to generate random number.

Value

countMatrix	A read count matrix with integer element (i,j) representing the number of supporting reads for fusion event i detected by tool j.
trueFusion	A vector showing the name of true fusion events, the length of which is indicated by trueFusionNum.

Author(s)

Silvia Liu <silvia.shuchang.liu@gmail.com>

References

Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. (2015) Silvia Liu[^], Wei-Hsiang Tsai[^], Ying Ding[^], Rui Chen, Zhou Fang, SungHwan Kim, Tianzhou Ma, Ting-Yu Chang, Jianhua Luo, Hsei-Wei Wang*, I-Fang Chung* and George C. Tseng*.

Examples

```
fusionNum <- 100
toolNum <- 5
trueFusionNum <- 70 # high supporting reads
missingRate <- c(0.4,0.6,0.8,0.8,0.9) # missing fusion rate
seed <- 12345

## simulate read counts
syn <- GenerateReadCount(fusionNum,toolNum,trueFusionNum,missingRate,seed)

## check output
syn$countMatrix
syn>trueFusion
```