

Package ‘RNASeqR’

March 20, 2023

Type Package

Title RNASeqR: an R package for automated two-group RNA-Seq analysis workflow

Version 1.16.0

Date 2018-8-7

Author Kuan-Hao Chao

Maintainer Kuan-Hao Chao <ntueeb05howard@gmail.com>

biocViews Genetics, Infrastructure, DataImport, Sequencing, RNASeq, GeneExpression, GeneSetEnrichment, Alignment, QualityControl, DifferentialExpression, FunctionalPrediction, ExperimentalDesign, GO, KEGG, Visualization, Normalization, Pathways, Clustering, ImmunoOncology

Description This R package is designed for case-control RNA-Seq analysis (two-group). There are six steps: ``RNASeqRParam S4 Object Creation'', ``Environment Setup'', ``Quality Assessment'', ``Reads Alignment & Quantification'', ``Gene-level Differential Analyses'' and ``Functional Analyses''. Each step corresponds to a function in this package. After running functions in order, a basic RNASeq analysis would be done easily.

License Artistic-2.0

Encoding UTF-8

RoxygenNote 6.1.1

Depends R(>= 3.5.0), ggplot2, pathview, edgeR, methods

Imports Rsamtools, tools, reticulate, ballgown, gridExtra, rafalib, FactoMineR, factoextra, corrplot, PerformanceAnalytics, reshape2, DESeq2, systemPipeR, systemPipeRdata, clusterProfiler, org.Hs.eg.db, org.Sc.sgd.db, stringr, pheatmap, grDevices, graphics, stats, utils, DOSE, Biostrings, parallel

Suggests knitr, rmarkdown, png, grid, RNASeqRData

VignetteBuilder knitr

SystemRequirements RNASeqR only support Linux and macOS. Window is not supported. Python2 is highly recommended. If your machine is Python3, make sure '2to3' command is available.

BugReports <https://github.com/HowardChao/RNASeqR/issues>

URL <https://github.com/HowardChao/RNASeqR>

NeedsCompilation no

PackageStatus Deprecated

git_url <https://git.bioconductor.org/packages/RNASeqR>

git_branch RELEASE_3_16

git_last_commit 0f6d741

git_last_commit_date 2022-11-01

Date/Publication 2023-03-20

R topics documented:

| | |
|--|-----------|
| All_Steps_Interface | 3 |
| All_Steps_Interface_CMD | 4 |
| CheckToolAll | 6 |
| RNASeqDifferentialAnalysis | 7 |
| RNASeqDifferentialAnalysis_CMD | 8 |
| RNASeqEnvironmentSet | 10 |
| RNASeqEnvironmentSet_CMD | 12 |
| RNASeqGoKegg | 13 |
| RNASeqGoKegg_CMD | 15 |
| RNASeqQualityAssessment | 16 |
| RNASeqQualityAssessment_CMD | 17 |
| RNASeqR | 18 |
| RNASeqReadProcess | 18 |
| RNASeqReadProcess_CMD | 27 |
| RNASeqRParam-class | 37 |
| RNASeqRParam-constructor | 38 |
| RNASeqRParam_Bam-class | 39 |
| RNASeqRParam_Bam-constructor | 40 |
| RNASeqRParam_Sam-class | 41 |
| RNASeqRParam_Sam-constructor | 42 |
| Update_Fastq_gz | 43 |
| yeast | 44 |
| Index | 45 |

All_Steps_Interface *RNASeqEnvironmentSet*

Description

Set up the environment for the following RNA-Seq workflow in R shell
This function do 4 things :

1. Create file directories.
2. Install necessary tools.
3. Export 'RNASeq_bin/' to the R environment.
4. Check command of tools.

First it will create 'gene_data/', 'RNASeq_bin/', 'RNASeq_results/', 'Rscript/', 'Rscript_out/' directories.

Afterwards, 'Hisat2', 'Stringtie', 'Gffcompare' will be installed under 'RNASeq_bin/Download/' and be unpacked under 'RNASeq_bin/Unpacked/'.

'RNASeq_bin/' will be added to the R environment and validity of tools will be checked.

Any ERROR occurs will be reported and the program will be terminated.

If you want to set up the environment for the following RNA-Seq workflow in background, please see `RNASeqEnvironmentSet_CMD()` function.

Usage

```
All_Steps_Interface(RNASeqRParam, which.trigger = "OUTSIDE",  
  INSIDE.path.prefix = NA, RNASeqQualityAssessment.RUN = TRUE,  
  RNASeqReadProcess.RUN = TRUE, RNASeqDifferentialAnalysis.RUN = TRUE,  
  RNASeqGoKegg.RUN = TRUE, OrgDb.species, go.level = 3, input.TYPE.ID,  
  KEGG.organism, check.s4.print = TRUE)
```

Arguments

`RNASeqRParam` S4 object instance of experiment-related parameters

`which.trigger` Default value is OUTSIDE. User should not change this value.

`INSIDE.path.prefix`
 Default value is NA. User should not change this value.

`RNASeqQualityAssessment.RUN`
 Default value is TRUE. Set FALSE to skip "Quality Assessment" step.

`RNASeqReadProcess.RUN`
 Default value is TRUE. Set FALSE to skip "RNASeq Read Process" step.

RNASeqDifferentialAnalysis.RUN
 Default value is TRUE. Set FALSE to skip "RNASeq Differential Analysis" step.
 RNASeqGoKegg.RUN
 Default value is TRUE. Set FALSE to skip "RNASeq Go & Kegg" step.
 OrgDb.species the genome wide annotation packages of species on Bioconductor. Currently,
 there are 19 supported genome wide annotation packages of species.
 go.level the depth of acyclic graph in GO analysis
 input.TYPE.ID The gene name type in OrgDb.species annotation package.
 KEGG.organism the species that are supported for KEGG analysis. Currently, there are more than
 5000 supported species genome. Check the valid species terms on <https://www.genome.jp/kegg/catalog/>
 check.s4.print Default TRUE. If TRUE, the result of checking RNASeqRParam will be reported in
 'Rscript_out/Environment_Set.Rout'. If FALSE, the result of checking RNASeqRParam
 will not be in 'Rscript_out/Environment_Set.Rout'.

Value

None

Author(s)

Kuan-Hao Chao

Examples

```

data(yeast)
## Not run:
RNASeqEnvironmentSet(RNASeqRParam = yeast)
## End(Not run)

```

All_Steps_Interface_CMD

All_Steps_Interface_CMD

Description

A function to run all the steps with in one function. This function execute in the background:

1. Create file directories.
2. Install necessary tools.
3. Export 'RNASeq_bin/' to the R environment.
4. Check command of tools.

First it will create 'gene_data/', 'RNASeq_bin/', 'RNASeq_results/', 'Rscript/', 'Rscript_out/' directories.

Afterwards, 'Hisat2', 'Stringtie', 'Gffcompare' will be installed under 'RNASeq_bin/Download/' and be unpacked under 'RNASeq_bin/Unpacked/'.

'RNASeq_bin/' will be added to the R environment and validity of tools will be checked.

Any ERROR occurs will be reported and the program will be terminated.

If you want to set up the environment for the following RNA-Seq workflow in R shell, please see RNASeqEnvironmentSet() function.

Usage

```
All_Steps_Interface_CMD(RNASeqRParam, RNASeqQualityAssessment.RUN = TRUE,
  RNASeqReadProcess.RUN = TRUE, RNASeqDifferentialAnalysis.RUN = TRUE,
  RNASeqGoKegg.RUN = TRUE, OrgDb.species, go.level = 3, input.TYPE.ID,
  KEGG.organism, run = TRUE, check.s4.print = TRUE)
```

Arguments

| | |
|--------------------------------|--|
| RNASeqRParam | S4 object instance of experiment-related parameters |
| RNASeqQualityAssessment.RUN | Default value is TRUE. Set FALSE to skip "Quality Assessment" step. |
| RNASeqReadProcess.RUN | Default value is TRUE. Set FALSE to skip "RNASeq Read Process" step. |
| RNASeqDifferentialAnalysis.RUN | Default value is TRUE. Set FALSE to skip "RNASeq Differential Analysis" step. |
| RNASeqGoKegg.RUN | Default value is TRUE. Set FALSE to skip "RNASeq Go & Kegg" step. |
| OrgDb.species | the genome wide annotation packages of species on Bioconductor. Currently, there are 19 supported genome wide annotation packages of species. |
| go.level | the depth of acyclic graph in GO analysis |
| input.TYPE.ID | The gene name type in OrgDb.species annotation packahge. |
| KEGG.organism | the species that are supported for KEGG analysis. Currently, there are more than 5000 supported species genome. Check the valid species terms on https://www.genome.jp/kegg/catalog/or |
| run | Default value is TRUE. If TRUE, 'Rscript/Environment_Set.R' will be created and executed. The output log will be stored in 'Rscript_out/Environment_Set.Rout'. If False, 'Rscript/Environment_Set.R' will be created without executed. |
| check.s4.print | Default TRUE. If TRUE, the result of checking RNASeqRParam will be reported in 'Rscript_out/Environment_Set.Rout'. If FALSE, the result of checking RNASeqRParam will not be in 'Rscript_out/Environment_Set.Rout'. |

Value

None

Author(s)

Kuan-Hao Chao

Examples

```
data(yeast)
## Not run:
RNASeqEnvironmentSet_CMD(yeast)
## End(Not run)
```

CheckToolAll

CheckToolAll

Description

Check whether 'Hisat2', 'Stringtie' and 'Gffcompare' are installed on the workstation

Usage

```
CheckToolAll(path.prefix, print = TRUE)
```

Arguments

| | |
|-------------|---|
| path.prefix | path prefix of 'gene_data/', 'RNASeq_bin/', 'RNASeq_results/', 'Rscript/' and 'Rscript_out/' directories. |
| print | If TRUE, detailed information will be printed. If FALSE, detailed information will not be printed. |

Value

None

Examples

```
data(yeast)
## Not run:
CheckToolAll(yeast@path.prefix,
              print=TRUE)
## End(Not run)
```

RNASeqDifferentialAnalysis
RNASeqDifferentialAnalysis

Description

This function will run differential analysis on ballgown, DESeq2 and edgeR in background.
This function do following things :

1. ballgown analysis
Raw reads are normalized into FPKM values
The main statistic test in ballgown is paramatic F-test comparing nested linear models
2. DESeq2 analysis
Median of rations normalization(MRN) is used in DESeq2 for raw reads count normalization.
Sequencing depth and RNA composition is taken into consideration is this normalization method.
The main statistic test in DESeq2 is negative binomial distribution.
3. edgeR analysis
Raw reads are normalized by TMM and library size. (run `calcNormFactors()` to get a DGE-List, and then run `cpm()` on that DGEList)
The main statistic test in edgeR is trimmed mean of M-values(TMM).

If you want to run differential analysis on ballgown, DESeq2, edgeR for the following RNA-Seq workflow in background, please see `RNASeqDifferentialAnalysis()` function.

Usage

```
RNASeqDifferentialAnalysis(RNASeqRParam, which.trigger = "OUTSIDE",  
  INSIDE.path.prefix = NA, Pre_DE.visualization = TRUE,  
  Post_DE.visualization = TRUE, ballgown.run = TRUE,  
  ballgown.pval = 0.05, ballgown.log2FC = 1, DESeq2.run = TRUE,  
  DESeq2.pval = 0.1, DESeq2.log2FC = 1, edgeR.run = TRUE,  
  edgeR.pval = 0.05, edgeR.log2FC = 1, check.s4.print = TRUE)
```

Arguments

| | |
|-----------------------------------|--|
| <code>RNASeqRParam</code> | S4 object instance of experiment-related parameters |
| <code>which.trigger</code> | Default value is OUTSIDE. User should not change this value. |
| <code>INSIDE.path.prefix</code> | Default value is NA. User should not change this value. |
| <code>Pre_DE.visualization</code> | Default TRUE. Whether to visualize pre-DE analysis results. |

| | |
|-----------------------|---|
| Post_DE.visualization | Default TRUE. Whether to visualize post-DE analysis results. |
| ballgown.run | Default TRUE. Logical value whether to run ballgown differential analysis. |
| ballgown.pval | Default 0.05. Set the threshold of ballgown p-value to filter out differential expressed gene. |
| ballgown.log2FC | Default 1. Set the threshold of ballgown log2 fold change to filter out differential expressed gene. |
| DESeq2.run | Default TRUE. Logical value whether to run DESeq2 differential analysis. |
| DESeq2.pval | Default 0.05. Set the threshold of DESeq2 p-value to filter out differential expressed gene. |
| DESeq2.log2FC | Default 1. Set the threshold of DESeq2 log2 fold change to filter out differential expressed gene. |
| edgeR.run | Default TRUE. Logical value whether to run edgeR differential analysis. |
| edgeR.pval | Default 0.05. Set the threshold of edgeR p-value to filter out differential expressed gene. |
| edgeR.log2FC | Default 1. Set the threshold of edgeR log2 fold change to filter out differential expressed gene. |
| check.s4.print | Default TRUE. If TRUE, the result of checking RNASeqRParam will be reported in 'Rscript_out/Environment_Set.Rout'. If FALSE, the result of checking RNASeqRParam will not be in 'Rscript_out/Environment_Set.Rout'. |

Value

None

Author(s)

Kuan-Hao Chao

Examples

```
data(yeast)
## Not run:
RNASeqDifferentialAnalysis(RNASeqRParam = yeast)
## End(Not run)
```

 RNASeqDifferentialAnalysis_CMD

RNASeqDifferentialAnalysis_CMD

Description

This function will run differential analysis on ballgown, DESeq2 and edgeR in R shell.

This function do following things :

1. ballgown analysis
Raw reads are normalized into FPKM values
The main statistic test in ballgown is paramatic F-test comparing nested linear models
2. DESeq2 analysis
Median of rations normalization(MRN) is used in DESeq2 for raw reads count normalization. Sequencing depth and RNA composition is taken into consideration is this normalization method.
The main statistic test in DESeq2 is negative binomial distribution.
3. edgeR analysis
Raw reads are normalized by TMM and library size. (run `calcNormFactors()` to get a DGEList, and then run `cpm()` on that DGEList)
The main statistic test in edgeR is trimmed mean of M-values(TMM).

If you want to run differential analysis on ballgown, DESeq2, edgeR for the following RNA-Seq workflow in R shell, please see `RNASeqDifferentialAnalysis()` function.

Usage

```
RNASeqDifferentialAnalysis_CMD(RNASeqRParam, which.trigger = "OUTSIDE",
  INSIDE.path.prefix = NA, Pre_DE.visualization = TRUE,
  Post_DE.visualization = TRUE, ballgown.run = TRUE,
  ballgown.pval = 0.05, ballgown.log2FC = 1, DESeq2.run = TRUE,
  DESeq2.pval = 0.1, DESeq2.log2FC = 1, edgeR.run = TRUE,
  edgeR.pval = 0.05, edgeR.log2FC = 1, run = TRUE,
  check.s4.print = TRUE)
```

Arguments

| | |
|------------------------------------|--|
| <code>RNASeqRParam</code> | S4 object instance of experiment-related parameters |
| <code>which.trigger</code> | Default value is OUTSIDE. User should not change this value. |
| <code>INSIDE.path.prefix</code> | Default value is NA. User should not change this value. |
| <code>Pre_DE.visualization</code> | Default TRUE. Whether to visualize pre-DE analysis results. |
| <code>Post_DE.visualization</code> | Default TRUE. Whether to visualize post-DE analysis results. |
| <code>ballgown.run</code> | Default TRUE. Logical value whether to run ballgown differential analysis. |
| <code>ballgown.pval</code> | Default 0.05. Set the threshold of ballgown p-value to filter out differential expressed gene. |

| | |
|-----------------|--|
| ballgown.log2FC | Default 1. Set the threshold of ballgown log2 fold change to filter out differential expressed gene. |
| DESeq2.run | Default TRUE. Logical value whether to run DESeq2 differential analysis. |
| DESeq2.pval | Default 0.05. Set the threshold of DESeq2 p-value to filter out differential expressed gene. |
| DESeq2.log2FC | Default 1. Set the threshold of DESeq2 log2 fold change to filter out differential expressed gene. |
| edgeR.run | Default TRUE. Logical value whether to run edgeR differential analysis. |
| edgeR.pval | Default 0.05. Set the threshold of edgeR p-value to filter out differential expressed gene. |
| edgeR.log2FC | Default 1. Set the threshold of edgeR log2 fold change to filter out differential expressed gene. |
| run | Default value is TRUE. If TRUE, 'Rscript/Environment_Set.R' will be created and executed. The output log will be stored in 'Rscript_out/Environment_Set.Rout'. If False, 'Rscript/Environment_Set.R' will be created without executed. |
| check.s4.print | Default TRUE. If TRUE, the result of checking RNASeqRParam will be reported in 'Rscript_out/Environment_Set.Rout'. If FALSE, the result of checking RNASeqRParam will not be in 'Rscript_out/Environment_Set.Rout'. |

Value

None

Author(s)

Kuan-Hao Chao

Examples

```
data(yeast)
## Not run:
RNASeqDifferentialAnalysis_CMD(RNASeqRParam = yeast)
## End(Not run)
```

RNASeqEnvironmentSet *RNASeqEnvironmentSet*

Description

Set up the environment for the following RNA-Seq workflow in R shell
This function do 4 things :

1. Create file directories.

2. Install necessary tools.
3. Export 'RNASeq_bin/' to the R environment.
4. Check command of tools.

First it will create 'gene_data/', 'RNASeq_bin/', 'RNASeq_results/', 'Rscript/', 'Rscript_out/' directories.

Afterwards, 'Hisat2', 'Stringtie', 'Gffcompare' will be installed under 'RNASeq_bin/Download/' and be unpacked under 'RNASeq_bin/Unpacked/'.

'RNASeq_bin/' will be added to the R environment and validity of tools will be checked.

Any ERROR occurs will be reported and the program will be terminated.

If you want to set up the environment for the following RNA-Seq workflow in background, please see RNASeqEnvironmentSet_CMD() function.

Usage

```
RNASeqEnvironmentSet(RNASeqRParam, which.trigger = "OUTSIDE",
  INSIDE.path.prefix = NA, install.hisat2 = TRUE,
  install.STAR = TRUE, install.stringtie = TRUE,
  install.gffcompare = TRUE, check.s4.print = TRUE)
```

Arguments

| | |
|--------------------|---|
| RNASeqRParam | S4 object instance of experiment-related parameters |
| which.trigger | Default value is OUTSIDE. User should not change this value. |
| INSIDE.path.prefix | Default value is NA. User should not change this value. |
| install.hisat2 | Whether to install 'HISAT2' in this function step. Default value is TRUE. Set FALSE to skip 'HISAT2' installation. |
| install.STAR | Whether to install 'STAR' in this function step. Default value is TRUE. Set FALSE to skip 'STAR' installation. |
| install.stringtie | Whether to install 'StringTie' in this function step. Default value is TRUE. Set FALSE to skip 'StringTie' installation. |
| install.gffcompare | Whether to install 'Gffcompare' in this function step. Default value is TRUE. Set FALSE to skip 'Gffcompare' installation. |
| check.s4.print | Default TRUE. If TRUE, the result of checking RNASeqRParam will be reported in 'Rscript_out/Environment_Set.Rout'. If FALSE, the result of checking RNASeqRParam will not be in 'Rscript_out/Environment_Set.Rout'. |

Value

None

Author(s)

Kuan-Hao Chao

Examples

```
data(yeast)
## Not run:
RNASeqEnvironmentSet(RNASeqRParam = yeast)
## End(Not run)
```

RNASeqEnvironmentSet_CMD

RNASeqEnvironmentSet_CMD

Description

Set up the environment for the following RNA-Seq workflow in background.
This function do 4 things :

1. Create file directories.
2. Install necessary tools.
3. Export 'RNASeq_bin/' to the R environment.
4. Check command of tools.

First it will create 'gene_data/', 'RNASeq_bin/', 'RNASeq_results/', 'Rscript/', 'Rscript_out/' directories.

Afterwards, 'Hisat2', 'Stringtie', 'Gffcompare' will be installed under 'RNASeq_bin/Download/' and be unpacked under 'RNASeq_bin/Unpacked/'.

'RNASeq_bin/' will be added to the R environment and validity of tools will be checked.

Any ERROR occurs will be reported and the program will be terminated.

If you want to set up the environment for the following RNA-Seq workflow in R shell, please see RNASeqEnvironmentSet() function.

Usage

```
RNASeqEnvironmentSet_CMD(RNASeqRParam, install.hisat2 = TRUE,
  install.STAR = TRUE, install.stringtie = TRUE,
  install.gffcompare = TRUE, run = TRUE, check.s4.print = TRUE)
```

Arguments

| | |
|--------------------|--|
| RNASeqRParam | S4 object instance of experiment-related parameters |
| install.hisat2 | Whether to install 'HISAT2' in this function step. Default value is TRUE. Set FALSE to skip 'HISAT2' installation. |
| install.STAR | Whether to install 'STAR' in this function step. Default value is TRUE. Set FALSE to skip 'STAR' installation. |
| install.stringtie | Whether to install 'StringTie' in this function step. Default value is TRUE. Set FALSE to skip 'StringTie' installation. |
| install.gffcompare | Whether to install 'Gffcompare' in this function step. Default value is TRUE. Set FALSE to skip 'Gffcompare' installation. |
| run | Default value is TRUE. If TRUE, 'Rscript/Environment_Set.R' will be created and executed. The output log will be stored in 'Rscript_out/Environment_Set.Rout'. If FALSE, 'Rscript/Environment_Set.R' will be created without executed. |
| check.s4.print | Default TRUE. If TRUE, the result of checking RNASeqRParam will be reported in 'Rscript_out/Environment_Set.Rout'. If FALSE, the result of checking RNASeqRParam will not be in 'Rscript_out/Environment_Set.Rout'. |

Value

None

Author(s)

Kuan-Hao Chao

Examples

```
data(yeast)
## Not run:
RNASeqEnvironmentSet_CMD(yeast)
## End(Not run)
```

 RNASeqGoKegg

RNASeqGoKegg

Description

Run Gene Ontology(GO) and Kyoto Encyclopedia of Genes and Genomes(KEGG) analysis in R shell.

This function do Gene Ontology(GO) and Kyoto Encyclopedia of Genes and Genomes(KEGG) analysis :

1. Gene Ontology(GO) :

- (a) Do GO function classification analysis.
- (b) Do GO function enrichment analysis.
- (c) Visualization : bar plot, dot plot etc.

2. Kyoto Encyclopedia of Genes and Genomes(KEGG) :

- (a) Do KEGG pathway enrichment analysis
- (b) Pathway visulization with pathview package. KEGG webpage pathway url will also be created

If you want to do GO functional analysis and KEGG pathway analysis for the following RNA-Seq workflow in background, please see RNASeqGoKegg_CMD() function.

Usage

```
RNASeqGoKegg(RNASeqRParam, which.trigger = "OUTSIDE",
             INSIDE.path.prefix = NA, OrgDb.species, go.level = 3, input.TYPE.ID,
             KEGG.organism, check.s4.print = TRUE)
```

Arguments

| | |
|--------------------|--|
| RNASeqRParam | S4 object instance of experiment-related parameters |
| which.trigger | Default value is OUTSIDE. User should not change this value. |
| INSIDE.path.prefix | Default value is NA. User should not change this value. |
| OrgDb.species | the genome wide annotation packages of species on Bioconductor. Currently, there are 19 supported genome wide annotation packages of species. |
| go.level | the depth of acyclic graph in GO analysis |
| input.TYPE.ID | The gene name type in OrgDb.species annotation packahge. |
| KEGG.organism | the species that are supported for KEGG analysis. Currently, there are more than 5000 supported species genome. Check the valid species terms on https://www.genome.jp/kegg/catalog/or |
| check.s4.print | Default TRUE. If TRUE, the result of checking RNASeqRParam will be reported in 'Rscript_out/Environment_Set.Rout'. If FALSE, the result of checking RNASeqRParam will not be in 'Rscript_out/Environment_Set.Rout' |

Value

None

Author(s)

Kuan-Hao Chao

Examples

```
data(yeast)
## Not run:
RNASeqGoKegg(RNASeqRParam = yeast,
             OrgDb.species = "org.Sc.sgd.db",
             go.level = 3,
             input.TYPE.ID = "GENENAME",
             KEGG.organism = "sce")
## End(Not run)
```

`RNASeqGoKegg_CMD``RNASeqGoKegg_CMD`

Description

Run Gene Ontology(GO) and Kyoto Encyclopedia of Genes and Genomes(KEGG) analysis in background.

This function do Gene Ontology(GO) and Kyoto Encyclopedia of Genes and Genomes(KEGG) analysis :

1. Gene Ontology(GO) :
 - (a) Do GO function classification analysis.
 - (b) Do GO function enrichment analysis.
 - (c) Visualization : bar plot, dot plot etc.
2. Kyoto Encyclopedia of Genes and Genomes(KEGG) :
 - (a) Do KEGG pathway enrichment analysis
 - (b) Pathway visualization with pathview package. KEGG webpage pathway url will also be created

If you want to do GO functional analysis and KEGG pathway analysis for the following RNA-Seq workflow in R shell, please see `RNASeqGoKegg()` function.

Usage

```
RNASeqGoKegg_CMD(RNASeqRParam, OrgDb.species, go.level = 3,
                 input.TYPE.ID, KEGG.organism, run = TRUE, check.s4.print = TRUE)
```

Arguments

| | |
|----------------|--|
| RNASeqRParam | S4 object instance of experiment-related parameters |
| OrgDb.species | the genome wide annotation packages of species on Bioconductor. Currently, there are 19 supported genome wide annotation packages of species. |
| go.level | the depth of acyclic graph in GO analysis |
| input.TYPE.ID | The gene name type in OrgDb.species annotation package. |
| KEGG.organism | the species that are supported for KEGG analysis. Currently, there are more than 5000 supported species genome. Check the valid species terms on https://www.genome.jp/kegg/catalog/or |
| run | Default value is TRUE. If TRUE, 'Rscript/Environment_Set.R' will be created and executed. The output log will be stored in 'Rscript_out/Environment_Set.Rout'. If FALSE, 'Rscript/Environment_Set.R' will be created without executed. |
| check.s4.print | Default TRUE. If TRUE, the result of checking RNASeqRParam will be reported in 'Rscript_out/Environment_Set.Rout'. If FALSE, the result of checking RNASeqRParam will not be in 'Rscript_out/Environment_Set.Rout' |

Value

None

Author(s)

Kuan-Hao Chao

Examples

```

data(yeast)
## Not run:
RNASeqGoKegg_CMD(RNASeqRParam = yeast,
                  OrgDb.species = "org.Sc.sgd.db",
                  go.level = 3,
                  input.TYPE.ID = "GENENAME",
                  KEGG.organism = "sce")

## End(Not run)

```

 RNASeqQualityAssessment

RNASeqQualityAssessment

Description

Assess the quality of '.fastq.gz' files for RNA-Seq workflow in R shell. This step is optional in the whole RNA-Seq workflow.

This function reports the quality assessment result in packages systemPipeR. For systemPipeR, 'RNASeq_results/QA_results/Rqc/systemPipeR/fastqReport.pdf' will be created.

If you want to assess the quality of '.fastq.gz' files for the following RNA-Seq workflow in background, please see RNASeqQualityAssessment_CMD() function.

Usage

```
RNASeqQualityAssessment(RNASeqRParam, which.trigger = "OUTSIDE",
  INSIDE.path.prefix = NA, check.s4.print = TRUE)
```

Arguments

RNASeqRParam S4 object instance of experiment-related parameters

which.trigger Default value is OUTSIDE. User should not change this value.

INSIDE.path.prefix Default value is NA. User should not change this value.

check.s4.print Default TRUE. If TRUE, the result of checking RNASeqRParam will be reported in 'Rscript_out/Environment_Set.Rout'. If FALSE, the result of checking RNASeqRParam will not be in 'Rscript_out/Environment_Set.Rout'

Value

None

Author(s)

Kuan-Hao Chao

Examples

```
data(yeast)
## Not run:
RNASeqQualityAssessment(RNASeqRParam = yeast)
## End(Not run)
```

RNASeqQualityAssessment_CMD

RNASeqQualityAssessment_CMD

Description

Assess the quality of '.fastq.gz' files for RNA-Seq workflow in background. This step is optional in the whole RNA-Seq workflow.

This function reports the quality assessment result in packages systemPipeR. For systemPipeR, 'RNASeq_results/QA_results/Rqc/systemPipeR/fastqReport.pdf' will be created.

If you want to assess the quality of '.fastq.gz' files for the following RNA-Seq workflow in R shell, please see RNASeqQualityAssessment() function.

Usage

```
RNASeqQualityAssessment_CMD(RNASeqRParam, run = TRUE,
  check.s4.print = TRUE)
```

Arguments

| | |
|----------------|--|
| RNASeqRParam | S4 object instance of experiment-related parameters |
| run | Default value is TRUE. If TRUE, 'Rscript/Environment_Set.R' will be created and executed. The output log will be stored in 'Rscript_out/Environment_Set.Rout'. If FALSE, 'Rscript/Environment_Set.R' will be created without executed. |
| check.s4.print | Default TRUE. If TRUE, the result of checking RNASeqRParam will be reported in 'Rscript_out/Environment_Set.Rout'. If FALSE, the result of checking RNASeqRParam will not be in 'Rscript_out/Environment_Set.Rout' |

Value

None

Author(s)

Kuan-Hao Chao

Examples

```
data(yeast)
## Not run:
RNASeqQualityAssessment_CMD(RNASeqRParam = yeast)
## End(Not run)
```

RNASeqR

RNASeqR-package

Description

RNASeqR-package

RNASeqReadProcess

RNASeqReadProcess

Description

Process raw reads for RNA-Seq workflow in R shell

This function do 5 things :

1. 'Hisat2' : aligns raw reads to reference genome. If indices.optional in RNASeqRParam is FALSE, Hisat2 indices will be created.
2. 'Rsamtools': converts '.sam' files to '.bam' files.

3. 'Stringtie': assembles alignments into transcript.
4. 'Gffcompare': examines how transcripts compare with the reference annotation.
5. 'Stringtie': creates input files for ballgown, edgeR and DESeq2.
6. raw reads count: create raw reads count for DESeq2 and edgeR

Before running this function, `RNASeqEnvironmentSet_CMD()` or `RNASeqEnvironmentSet()` must be executed successfully. If you want to process raw reads for the following RNA-Seq workflow in background, please see `RNASeqReadProcess_CMD()` function.

Usage

```
RNASeqReadProcess(RNASeqRParam, which.trigger = "OUTSIDE",
  INSIDE.path.prefix = NA, SAMtools.or.Rsamtools = "Rsamtools",
  Hisat2.Index.run = TRUE, Hisat2.Index.num.parallel.threads = "1",
  Hisat2.Index.large.index = FALSE,
  Hisat2.Index.local.ftab.chars = "6",
  Hisat2.Index.local.off.rate = "3", Hisat2.Index.ftab.chars = "10",
  Hisat2.Index.off.rate = "4", Hisat2.Alignment.run = TRUE,
  Hisat2.Alignment.num.parallel.threads = "1",
  Hisat2.Alignment.skip = "0", Hisat2.Alignment.trim5 = "0",
  Hisat2.Alignment.trim3 = "0",
  Hisat2.Alignment.n.ceil.1.function.type = "L",
  Hisat2.Alignment.n.ceil.2.constant.term = "0",
  Hisat2.Alignment.n.ceil.3.coefficient = "0.15",
  Hisat2.Alignment.mp.MX = "6", Hisat2.Alignment.mp.MN = "2",
  Hisat2.Alignment.sp.MX = "2", Hisat2.Alignment.sp.MN = "1",
  Hisat2.Alignment.np = "1", Hisat2.Alignment.rdg.1 = "5",
  Hisat2.Alignment.rdg.2 = "3", Hisat2.Alignment.rfg.1 = "5",
  Hisat2.Alignment.rfg.2 = "3",
  Hisat2.Alignment.score.min.1.function.type = "L",
  Hisat2.Alignment.score.min.2.constant.term = "0",
  Hisat2.Alignment.score.min.3.coefficient = "-0.2",
  Hisat2.Alignment.pen.cansplice = "0",
  Hisat2.Alignment.penc.noncansplice = "12",
  Hisat2.Alignment.pen.canintronlen.1.function.type = "G",
  Hisat2.Alignment.pen.canintronlen.2.constant.term = "-8",
  Hisat2.Alignment.pen.canintronlen.3.coefficient = "1",
  Hisat2.Alignment.pen.noncanintronlen.1.function.type = "G",
  Hisat2.Alignment.pen.noncanintronlen.2.constant.term = "-8",
  Hisat2.Alignment.pen.noncanintronlen.3.coefficient = "1",
  Hisat2.Alignment.min.intronlen = "20",
  Hisat2.Alignment.max.intronlen = "500000",
  Hisat2.Alignment.rna.strandness = "None", Hisat2.Alignment.k = "5",
  Hisat2.Alignment.max.seeds = "5", Hisat2.Alignment.secondary = FALSE,
```

```
Hisat2.Alignment.minins = "0", Hisat2.Alignment.maxins = "500",
Hisat2.Alignment.seed = "0", STAR.Index.num.parallel.threads = "1",
STAR.Index.sjdbOverhang.Read.length = "100",
STAR.Index.genomeSAindexNbases = "14",
STAR.Index.genomeChrBinNbits = "18",
STAR.Index.genomeSAsparseD = "1", STAR.Alignment.run = FALSE,
STAR.Alignment.num.parallel.threads = "1",
STAR.Alignment.genomeLoad = "NoSharedMemory",
STAR.Alignment.readMapNumber = "-1",
STAR.Alignment.clip3pNbases = "0", STAR.Alignment.clip5pNbases = "0",
STAR.Alignment.clip3pAdapterSeq = "-",
STAR.Alignment.clip3pAdapterMMp = "0.1",
STAR.Alignment.clip3pAfterAdapterNbases = "0",
STAR.Alignment.limitGenomeGenerateRAM = "31000000000",
STAR.Alignment.limitIObufferSize = "150000000",
STAR.Alignment.limitOutSAMoneReadBytes = "100000",
STAR.Alignment.limitOutSJoneRead = "1000",
STAR.Alignment.limitOutSJcollapsed = "1000000",
STAR.Alignment.limitBAMsortRAM = "0",
STAR.Alignment.outReadsUnmapped = "None",
STAR.Alignment.outQsconversionAdd = "0",
STAR.Alignment.outSAMprimaryFlag = "OneBestScore",
STAR.Alignment.outSAMmapqUnique = "255",
STAR.Alignment.scoreGap = "0", STAR.Alignment.scoreGapNoncan = "-8",
STAR.Alignment.scoreGapGCAG = "-4",
STAR.Alignment.scoreGapATAC = "-8",
STAR.Alignment.scoreGenomicLengthLog2scale = "-0.25",
STAR.Alignment.scoreDelOpen = "-2",
STAR.Alignment.scoreDelBase = "-2",
STAR.Alignment.scoreInsOpen = "-2",
STAR.Alignment.scoreInsBase = "-2",
STAR.Alignment.scoreStitchSJshift = "1",
STAR.Alignment.seedSearchStartLmax = "50",
STAR.Alignment.seedSearchStartLmaxOverLread = "1.0",
STAR.Alignment.seedSearchLmax = "0",
STAR.Alignment.seedMultimapNmax = "10000",
STAR.Alignment.seedPerReadNmax = "1000",
STAR.Alignment.seedPerWindowNmax = "50",
STAR.Alignment.seedNoneLociPerWindow = "10",
STAR.Alignment.alignIntronMin = "21",
STAR.Alignment.alignIntronMax = "0",
STAR.Alignment.alignMatesGapMax = "0",
STAR.Alignment.alignSJoverhangMin = "5",
STAR.Alignment.alignSJDBoverhangMin = "3",
STAR.Alignment.alignSplicedMateMapLmin = "0",
STAR.Alignment.alignSplicedMateMapLminOverLmate = "0.66",
STAR.Alignment.alignWindowsPerReadNmax = "10000",
STAR.Alignment.alignTranscriptsPerWindowNmax = "100",
```

```

STAR.Alignment.alignTranscriptsPerReadNmax = "10000",
STAR.Alignment.alignEndsType = "Local",
STAR.Alignment.winAnchorMultimapNmax = "50",
STAR.Alignment.winBinNbits = "16",
STAR.Alignment.winAnchorDistNbins = "9",
STAR.Alignment.winFlankNbins = "4", Rsamtools.Bam.run = TRUE,
Samtools.Bam.num.parallel.threads = "1", Rsamtools.nCores = "1",
StringTie.Assembly.run = TRUE,
StringTie.Assembly.num.parallel.threads = "1",
StringTie.Assembly.f = "0.1", StringTie.Assembly.m = "200",
StringTie.Assembly.c = "2.5", StringTie.Assembly.g = "50",
StringTie.Assembly.M = "0.95", StringTie.Merge.Trans.run = TRUE,
StringTie.Merge.num.parallel.threads = "1",
Gffcompare.Ref.Sample.run = TRUE, StringTie.Ballgown.run = TRUE,
StringTie.2.Ballgown.num.parallel.threads = "1",
PreDECountTable.run = TRUE, check.s4.print = TRUE)

```

Arguments

RNASeqRParam S4 object instance of experiment-related parameters

which.trigger Default value is OUTSIDE. User should not change this value.

INSIDE.path.prefix
Default value is NA. User should not change this value.

SAMtools.or.Rsamtools
Default value is Rsamtools. User can set to SAMtools to use command-line-based 'samtools' instead.

Hisat2.Index.run
Whether to run 'HISAT2 index' step in this function step. Default value is TRUE. Set FALSE to skip 'HISAT2 index' step.

Hisat2.Index.num.parallel.threads
Specify the number of processing threads (CPUs) to use for Hisat2 index step. The default is "1"

Hisat2.Index.large.index
Hisat2 index terminal '-large-index' option. Default value is FALSE

Hisat2.Index.local.ftab.chars
Hisat2 index terminal '-t/-ftabchars' option. Default value is "6"

Hisat2.Index.local.off.rate
Hisat2 index terminal '-localoffrate' option. Default value is "3"

Hisat2.Index.ftab.chars
Hisat2 index terminal '-localftabchars' option. Default value is "10"

Hisat2.Index.off.rate
Hisat2 index terminal '-offrate' option. Default value is "4"

Hisat2.Alignment.run
Whether to run 'HISAT2 alignment' step in this function step. Default value is TRUE. Set FALSE to skip 'HISAT2 alignment' step.

Hisat2.Alignment.num.parallel.threads
Specify the number of processing threads (CPUs) to use for Hisat2 alignment step. The default is "1"

Hisat2.Alignment.skip
Hisat2 alignment terminal '-s/--skip' option. Default value is "0"

Hisat2.Alignment.trim5
Hisat2 alignment terminal '-5/--trim5' option. Default value is "0"

Hisat2.Alignment.trim3
Hisat2 alignment terminal '-3/--trim3' option. Default value is "0"

Hisat2.Alignment.n.ceil.1.function.type
Hisat2 alignment terminal '-n-ceil' option. Default value is "L"

Hisat2.Alignment.n.ceil.2.constant.term
Hisat2 alignment terminal '-n-ceil' option. Default value is "0"

Hisat2.Alignment.n.ceil.3.coefficient
Hisat2 alignment terminal '-n-ceil' option. Default value is "0.15"

Hisat2.Alignment.mp.MX
Hisat2 alignment terminal '-mp MX' option. Default value is "6"

Hisat2.Alignment.mp.MN
Hisat2 alignment terminal '-mp MN' option. Default value is "2"

Hisat2.Alignment.sp.MX
Hisat2 alignment terminal '-sp MX' option. Default value is "2"

Hisat2.Alignment.sp.MN
Hisat2 alignment terminal '-sp MN' option. Default value is "1"

Hisat2.Alignment.np
Hisat2 alignment terminal '-np' option. Default value is "1"

Hisat2.Alignment.rdg.1
Hisat2 alignment terminal '-rdg' first option. Default value is "5"

Hisat2.Alignment.rdg.2
Hisat2 alignment terminal '-rdg' first option. Default value is "3"

Hisat2.Alignment.rfg.1
Hisat2 alignment terminal '-rfg' first option. Default value is "5"

Hisat2.Alignment.rfg.2
Hisat2 alignment terminal '-rfg' first option. Default value is "3"

Hisat2.Alignment.score.min.1.function.type
Hisat2 alignment terminal '-rdg' first option. Default value is "L"

Hisat2.Alignment.score.min.2.constant.term
Hisat2 alignment terminal '-rdg' first option. Default value is "0"

Hisat2.Alignment.score.min.3.coefficient
Hisat2 alignment terminal '-rdg' first option. Default value is "-0.2"

Hisat2.Alignment.pen.cansplice
Hisat2 alignment terminal '-pen-cansplice' first option. Default value is "-0"

Hisat2.Alignment.penc.noncansplice
Hisat2 alignment terminal '-pen-noncansplice' option. Default value is "12"

Hisat2.Alignment.pen.canintronlen.1.function.type
Hisat2 alignment terminal '-pen-canintronlen' first option. Default value is "G"

Hisat2.Alignment.pen.canintronlen.2.constant.term
 Hisat2 alignment terminal '-pen-canintronlen' second option. Default value is "-8"

Hisat2.Alignment.pen.canintronlen.3.coefficient
 Hisat2 alignment terminal '-pen-canintronlen' third option. Default value is "1"

Hisat2.Alignment.pen.noncanintronlen.1.function.type
 Hisat2 alignment terminal '-pen-noncanintronlen' first option. Default value is "G"

Hisat2.Alignment.pen.noncanintronlen.2.constant.term
 Hisat2 alignment terminal '-pen-noncanintronlen' second option. Default value is "-8"

Hisat2.Alignment.pen.noncanintronlen.3.coefficient
 Hisat2 alignment terminal '-pen-noncanintronlen' third option. Default value is "1"

Hisat2.Alignment.min.intronlen
 Hisat2 alignment terminal '-min-intronlen' option. Default value is "20"

Hisat2.Alignment.max.intronlen
 Hisat2 alignment terminal '-max-intronlen' option. Default value is "20"

Hisat2.Alignment.rna.strandness
 Hisat2 alignment terminal '-rna-strandness' option. Default value is "None"

Hisat2.Alignment.k
 Hisat2 alignment terminal '-k' option. Default value is "5"

Hisat2.Alignment.max.seeds
 Hisat2 alignment terminal '-max-seeds' option. Default value is "5"

Hisat2.Alignment.secondary
 Hisat2 alignment terminal '-secondary' option. Default value is "FALSE"

Hisat2.Alignment.minins
 Hisat2 alignment terminal '-I/-minins' option. Default value is "0"

Hisat2.Alignment.maxins
 Hisat2 alignment terminal '-X/-maxins' option. Default value is "500"

Hisat2.Alignment.seed
 Hisat2 alignment terminal '-X/-maxins' option. Default value is "0"

STAR.Index.num.parallel.threads
 Specify the number of processing threads (CPUs) to use for STAR index step.
 The default is "1"

STAR.Index.sjdbOverhang.Read.length
 STAR index terminal '-sjdbOverhang' option. Default value is "100"

STAR.Index.genomeSAindexNbases
 STAR index terminal '-genomeSAindexNbases' option. Default value is "14"

STAR.Index.genomeChrBinNbits
 STAR index terminal '-genomeChrBinNbits' option. Default value is "18"

STAR.Index.genomeSAsparseD
 STAR index terminal '-genomeSAsparseD' option. Default value is "1"

STAR.Alignment.run
 Whether to run 'STAR index' step in this function step. Default value is FALSE.
 Set TRUE to run STAR alignment step. (need to set Hisat2.Index.run to FALSE)

STAR.Alignment.num.parallel.threads
Specify the number of processing threads (CPUs) to use for STAR alignment step. The default is "1"

STAR.Alignment.genomeLoad
STAR alignment terminal '-genomeLoad' option. Default value is "NoSharedMemory"

STAR.Alignment.readMapNumber
STAR alignment terminal '-readMapNumber' option. Default value is "-1"

STAR.Alignment.clip3pNbases
STAR alignment terminal '-clip3pNbases' option. Default value is "0"

STAR.Alignment.clip5pNbases
STAR alignment terminal '-clip5pNbases' option. Default value is "0"

STAR.Alignment.clip3pAdapterSeq
STAR alignment terminal '-clip3pAdapterSeq' option. Default value is "-"

STAR.Alignment.clip3pAdapterMMp
STAR alignment terminal '-clip3pAdapterMMp' option. Default value is "0.1"

STAR.Alignment.clip3pAfterAdapterNbases
STAR alignment terminal '-clip3pAfterAdapterNbases' option. Default value is "0"

STAR.Alignment.limitGenomeGenerateRAM
STAR alignment terminal '-limitGenomeGenerateRAM' option. Default value is "31000000000"

STAR.Alignment.limitIObufferSize
STAR alignment terminal '-limitIObufferSize' option. Default value is "150000000"

STAR.Alignment.limitOutSAMoneReadBytes
STAR alignment terminal '-limitOutSAMoneReadBytes' option. Default value is "100000"

STAR.Alignment.limitOutSJoneRead
STAR alignment terminal '-limitOutSJoneRead' option. Default value is "1000"

STAR.Alignment.limitOutSJcollapsed
STAR alignment terminal '-limitOutSJcollapsed' option. Default value is "1000000"

STAR.Alignment.limitBAMsortRAM
STAR alignment terminal '-limitBAMsortRAM' option. Default value is "0"

STAR.Alignment.outReadsUnmapped
STAR alignment terminal '-outReadsUnmapped' option. Default value is "None"

STAR.Alignment.outQSconversionAdd
STAR alignment terminal '-outQSconversionAdd' option. Default value is "0"

STAR.Alignment.outSAMprimaryFlag
STAR alignment terminal '-outSAMprimaryFlag' option. Default value is "OneBestScore"

STAR.Alignment.outSAMmapqUnique
STAR alignment terminal '-outSAMmapqUnique' option. Default value is "255"

STAR.Alignment.scoreGap
STAR alignment terminal '-scoreGap' option. Default value is "0"

STAR.Alignment.scoreGapNoncan
STAR alignment terminal '-scoreGapNoncan' option. Default value is "-8"

STAR.Alignment.scoreGapGCAG
STAR alignment terminal '-scoreGapGCAG' option. Default value is "-4"

STAR.Alignment.scoreGapATAC
STAR alignment terminal '-scoreGapATAC' option. Default value is "-8"

STAR.Alignment.scoreGenomicLengthLog2scale
STAR alignment terminal '-scoreGenomicLengthLog2scale' option. Default value is "-0.25"

STAR.Alignment.scoreDelOpen
STAR alignment terminal '-scoreDelOpen' option. Default value is "-2"

STAR.Alignment.scoreDelBase
STAR alignment terminal '-scoreDelBase' option. Default value is "-2"

STAR.Alignment.scoreInsOpen
STAR alignment terminal '-scoreInsOpen' option. Default value is "-2"

STAR.Alignment.scoreInsBase
STAR alignment terminal '-scoreInsBase' option. Default value is "-2"

STAR.Alignment.scoreStitchSJshift
STAR alignment terminal '-scoreStitchSJshift' option. Default value is "1"

STAR.Alignment.seedSearchStartLmax
STAR alignment terminal '-scoreStitchSJshift' option. Default value is "50"

STAR.Alignment.seedSearchStartLmaxOverLread
STAR alignment terminal '-seedSearchStartLmaxOverLread' option. Default value is "1.0"

STAR.Alignment.seedSearchLmax
STAR alignment terminal '-seedSearchLmax' option. Default value is "0"

STAR.Alignment.seedMultimapNmax
STAR alignment terminal '-seedMultimapNmax' option. Default value is "10000"

STAR.Alignment.seedPerReadNmax
STAR alignment terminal '-seedPerReadNmax' option. Default value is "1000"

STAR.Alignment.seedPerWindowNmax
STAR alignment terminal '-seedPerWindowNmax' option. Default value is "50"

STAR.Alignment.seedNoneLociperWindow
STAR alignment terminal '-seedNoneLociperWindow' option. Default value is "10"

STAR.Alignment.alignIntronMin
STAR alignment terminal '-alignIntronMin' option. Default value is "21"

STAR.Alignment.alignIntronMax
STAR alignment terminal '-alignIntronMax' option. Default value is "0"

STAR.Alignment.alignMatesGapMax
STAR alignment terminal '-alignMatesGapMax' option. Default value is "0"

STAR.Alignment.alignSJoverhangMin
STAR alignment terminal '-alignSJoverhangMin' option. Default value is "5"

STAR.Alignment.alignSJDBoverhangMin
STAR alignment terminal '-alignSJDBoverhangMin' option. Default value is "3"

STAR.Alignment.alignSplicedMateMapLmin
STAR alignment terminal '-alignSplicedMateMapLmin' option. Default value is "0"

STAR.Alignment.alignSplicedMateMapLminOverLmate
 STAR alignment terminal '-alignSplicedMateMapLminOverLmate' option. Default value is "0.66"

STAR.Alignment.alignWindowsPerReadNmax
 STAR alignment terminal '-alignWindowsPerReadNmax' option. Default value is "10000"

STAR.Alignment.alignTranscriptsPerWindowNmax
 STAR alignment terminal '-alignTranscriptsPerWindowNmax' option. Default value is "100"

STAR.Alignment.alignTranscriptsPerReadNmax
 STAR alignment terminal '-alignTranscriptsPerReadNmax' option. Default value is "10000"

STAR.Alignment.alignEndsType
 STAR alignment terminal '-alignEndsType' option. Default value is "Local"

STAR.Alignment.winAnchorMultimapNmax
 STAR alignment terminal '-winAnchorMultimapNmax' option. Default value is "50"

STAR.Alignment.winBinNbits
 STAR alignment terminal '-winBinNbits' option. Default value is "16"

STAR.Alignment.winAnchorDistNbins
 STAR alignment terminal '-winAnchorDistNbins' option. Default value is "9"

STAR.Alignment.winFlankNbins
 STAR alignment terminal '-winFlankNbins' option. Default value is "4"

Rsamtools.Bam.run
 Whether to run 'Rsamtools SAM to BAM' step in this function step. Default value is TRUE. Set FALSE to skip 'Rsamtools SAM to BAM' step.

Samtools.Bam.num.parallel.threads
 Specify the number of processing threads (CPUs) to use for Samtools sam to bam step. The default is "1"

Rsamtools.nCores
 The number of cores to use when running 'Rsamtools' step. Default value is 1

StringTie.Assemble.run
 Whether to run 'StringTie assembly' step in this function step. Default value is TRUE. Set FALSE to skip 'StringTie assembly' step.

Stringtie.Assembly.num.parallel.threads
 Specify the number of processing threads (CPUs) to use for Stringtie assembly. The default is "1"

Stringtie.Assembly.f
 Stringtie assembly terminal '-f' option. Default value is "0.1"

Stringtie.Assembly.m
 Stringtie assembly terminal '-m' option. Default value is "200"

Stringtie.Assembly.c
 Stringtie assembly terminal '-c' option. Default value is "2.5"

Stringtie.Assembly.g
 Stringtie assembly terminal '-g' option. Default value is "50"

Stringtie.Assembly.M
 Stringtie assembly terminal '-M' option. Default value is "0.95"

`StringTie.Merge.Trans.run`
 Whether to run 'StringTie GTF merging' step in this function step. Default value is TRUE. Set FALSE to skip 'StringTie GTF merging' step.

`Stringtie.Merge.num.parallel.threads`
 Specify the number of processing threads (CPUs) to use for Stringtie merge step. The default is "1"

`Gffcompare.Ref.Sample.run`
 Whether to run 'Gffcompare comparison' step in this function step. Default value is TRUE. Set FALSE to skip 'Gffcompare comparison' step.

`StringTie.Ballgown.run`
 Whether to run 'StringTie ballgown creation' step in this function step. Default value is TRUE. Set FALSE to skip 'StringTie ballgown creation' step.

`Stringtie.2.Ballgown.num.parallel.threads`
 Specify the number of processing threads (CPUs) to use for Stringtie to ballgown step. The default is "1"

`PreDECountTable.run`
 Whether to run 'gene raw reads count creation' step in this function step. Default value is TRUE. Set FALSE to skip 'gene raw reads count creation' step.

`check.s4.print` Default TRUE. If TRUE, the result of checking RNASeqRParam will be reported in 'Rscript_out/Environment_Set.Rout'. If FALSE, the result of checking RNASeqRParam will not be in 'Rscript_out/Environment_Set.Rout'.

Value

None

Author(s)

Kuan-Hao Chao

Examples

```
data(yeast)
## Not run:
## Before run this function, make sure \code{RNASeqEnvironmentSet_CMD()}
##(or\code{RNASeqEnvironmentSet()}) is executed successfully.
RNASeqReadProcess(RNASeqRParam      = yeast,
                  num.parallel.threads = 10)
## End(Not run)
```

Description

Process raw reads for RNA-Seq workflow in background.

This function do 5 things :

1. 'Hisat2' : aligns raw reads to reference genome. If indices.optional in RNASeqRParam is FALSE, Hisat2 indices will be created.
2. 'Rsamtools': converts '.sam' files to '.bam' files.
3. 'Stringtie': assembles alignments into transcript.
4. 'Gffcompare': examines how transcripts compare with the reference annotation.
5. 'Stringtie': creates input files for ballgown, edgeR and DESeq2.
6. raw reads count: create raw reads count for DESeq2 and edgeR

Before running this function, RNASeqEnvironmentSet_CMD() or RNASeqEnvironmentSet() must be executed successfully.

If you want to process raw reads for the following RNA-Seq workflow in R shell, please see RNASeqReadProcess() function.

Usage

```
RNASeqReadProcess_CMD(RNASeqRParam, SAMtools.or.Rsamtools = "Rsamtools",
  Hisat2.Index.run = TRUE, Hisat2.Index.num.parallel.threads = "1",
  Hisat2.Index.large.index = FALSE,
  Hisat2.Index.local.ftab.chars = "6",
  Hisat2.Index.local.off.rate = "3", Hisat2.Index.ftab.chars = "10",
  Hisat2.Index.off.rate = "4", Hisat2.Alignment.run = TRUE,
  Hisat2.Alignment.num.parallel.threads = "1",
  Hisat2.Alignment.skip = "0", Hisat2.Alignment.trim5 = "0",
  Hisat2.Alignment.trim3 = "0",
  Hisat2.Alignment.n.ceil.1.function.type = "L",
  Hisat2.Alignment.n.ceil.2.constant.term = "0",
  Hisat2.Alignment.n.ceil.3.coefficient = "0.15",
  Hisat2.Alignment.mp.MX = "6", Hisat2.Alignment.mp.MN = "2",
  Hisat2.Alignment.sp.MX = "2", Hisat2.Alignment.sp.MN = "1",
  Hisat2.Alignment.np = "1", Hisat2.Alignment.rdg.1 = "5",
  Hisat2.Alignment.rdg.2 = "3", Hisat2.Alignment.rfg.1 = "5",
  Hisat2.Alignment.rfg.2 = "3",
  Hisat2.Alignment.score.min.1.function.type = "L",
  Hisat2.Alignment.score.min.2.constant.term = "0",
  Hisat2.Alignment.score.min.3.coefficient = "-0.2",
```

```
Hisat2.Alignment.pen.cansplice = "0",
Hisat2.Alignment.penc.noncansplice = "12",
Hisat2.Alignment.pen.canintronlen.1.function.type = "G",
Hisat2.Alignment.pen.canintronlen.2.constant.term = "-8",
Hisat2.Alignment.pen.canintronlen.3.coefficient = "1",
Hisat2.Alignment.pen.noncanintronlen.1.function.type = "G",
Hisat2.Alignment.pen.noncanintronlen.2.constant.term = "-8",
Hisat2.Alignment.pen.noncanintronlen.3.coefficient = "1",
Hisat2.Alignment.min.intronlen = "20",
Hisat2.Alignment.max.intronlen = "500000",
Hisat2.Alignment.rna.strandness = "None", Hisat2.Alignment.k = "5",
Hisat2.Alignment.max.seeds = "5", Hisat2.Alignment.secondary = FALSE,
Hisat2.Alignment.minins = "0", Hisat2.Alignment.maxins = "500",
Hisat2.Alignment.seed = "0", STAR.Index.num.parallel.threads = "1",
STAR.Index.sjdbOverhang.Read.length = "100",
STAR.Index.genomeSAindexNbases = "14",
STAR.Index.genomeChrBinNbits = "18",
STAR.Index.genomeSAsparseD = "1", STAR.Alignment.run = FALSE,
STAR.Alignment.num.parallel.threads = "1",
STAR.Alignment.genomeLoad = "NoSharedMemory",
STAR.Alignment.readMapNumber = "-1",
STAR.Alignment.clip3pNbases = "0", STAR.Alignment.clip5pNbases = "0",
STAR.Alignment.clip3pAdapterSeq = "-",
STAR.Alignment.clip3pAdapterMMp = "0.1",
STAR.Alignment.clip3pAfterAdapterNbases = "0",
STAR.Alignment.limitGenomeGenerateRAM = "31000000000",
STAR.Alignment.limitIObufferSize = "150000000",
STAR.Alignment.limitOutSAMoneReadBytes = "100000",
STAR.Alignment.limitOutSJoneRead = "1000",
STAR.Alignment.limitOutSJcollapsed = "1000000",
STAR.Alignment.limitBAMsortRAM = "0",
STAR.Alignment.outReadsUnmapped = "None",
STAR.Alignment.outQSconversionAdd = "0",
STAR.Alignment.outSAMprimaryFlag = "OneBestScore",
STAR.Alignment.outSAMmapqUnique = "255",
STAR.Alignment.scoreGap = "0", STAR.Alignment.scoreGapNoncan = "-8",
STAR.Alignment.scoreGapGCAG = "-4",
STAR.Alignment.scoreGapATAC = "-8",
STAR.Alignment.scoreGenomicLengthLog2scale = "-0.25",
STAR.Alignment.scoreDelOpen = "-2",
STAR.Alignment.scoreDelBase = "-2",
STAR.Alignment.scoreInsOpen = "-2",
STAR.Alignment.scoreInsBase = "-2",
STAR.Alignment.scoreStitchSJshift = "1",
STAR.Alignment.seedSearchStartLmax = "50",
STAR.Alignment.seedSearchStartLmaxOverLread = "1.0",
STAR.Alignment.seedSearchLmax = "0",
STAR.Alignment.seedMultimapNmax = "10000",
```

```

STAR.Alignment.seedPerReadNmax = "1000",
STAR.Alignment.seedPerWindowNmax = "50",
STAR.Alignment.seedNoneLociperWindow = "10",
STAR.Alignment.alignIntronMin = "21",
STAR.Alignment.alignIntronMax = "0",
STAR.Alignment.alignMatesGapMax = "0",
STAR.Alignment.alignSJoverhangMin = "5",
STAR.Alignment.alignSJDBoverhangMin = "3",
STAR.Alignment.alignSplicedMateMapLmin = "0",
STAR.Alignment.alignSplicedMateMapLminOverLmate = "0.66",
STAR.Alignment.alignWindowsPerReadNmax = "10000",
STAR.Alignment.alignTranscriptsPerWindowNmax = "100",
STAR.Alignment.alignTranscriptsPerReadNmax = "10000",
STAR.Alignment.alignEndsType = "Local",
STAR.Alignment.winAnchorMultimapNmax = "50",
STAR.Alignment.winBinNbits = "16",
STAR.Alignment.winAnchorDistNbins = "9",
STAR.Alignment.winFlankNbins = "4", Rsamtools.Bam.run = TRUE,
Samtools.Bam.num.parallel.threads = "1", Rsamtools.nCores = "1",
StringTie.Assemble.run = TRUE,
Stringtie.Assembly.num.parallel.threads = "1",
Stringtie.Assembly.f = "0.1", Stringtie.Assembly.m = "200",
Stringtie.Assembly.c = "2.5", Stringtie.Assembly.g = "50",
Stringtie.Assembly.M = "0.95", StringTie.Merge.Trans.run = TRUE,
Stringtie.Merge.num.parallel.threads = "1",
Gffcompare.Ref.Sample.run = TRUE, StringTie.Ballgown.run = TRUE,
Stringtie.2.Ballgown.num.parallel.threads = "1",
PreDECountTable.run = TRUE, run = TRUE, check.s4.print = TRUE)

```

Arguments

- RNASeqRParam** S4 object instance of experiment-related parameters
- SAMtools.or.Rsamtools**
Default value is Rsamtools. User can set to SAMtools to use command-line-based 'samtools' instead.
- Hisat2.Index.run**
Whether to run 'HISAT2 index' step in this function step. Default value is TRUE. Set FALSE to skip 'HISAT2 index' step.
- Hisat2.Index.num.parallel.threads**
Specify the number of processing threads (CPUs) to use for Hisat2 index step. The default is "1"
- Hisat2.Index.large.index**
Hisat2 index terminal '-large-index' option. Default value is FALSE
- Hisat2.Index.local.ftab.chars**
Hisat2 index terminal '-t/-ftabchars' option. Default value is "6"
- Hisat2.Index.local.off.rate**
Hisat2 index terminal '-localoffrate' option. Default value is "3"

Hisat2.Index.ftab.chars
Hisat2 index terminal '-localftabchars' option. Default value is "10"

Hisat2.Index.off.rate
Hisat2 index terminal '-offrate' option. Default value is "4"

Hisat2.Alignment.run
Whether to run 'HISAT2 alignment' step in this function step. Default value is TRUE. Set FALSE to skip 'HISAT2 alignment' step.

Hisat2.Alignment.num.parallel.threads
Specify the number of processing threads (CPUs) to use for Hisat2 alignment step. The default is "1"

Hisat2.Alignment.skip
Hisat2 alignment terminal '-s/-skip' option. Default value is "0"

Hisat2.Alignment.trim5
Hisat2 alignment terminal '-5/-trim5' option. Default value is "0"

Hisat2.Alignment.trim3
Hisat2 alignment terminal '-3/-trim3' option. Default value is "0"

Hisat2.Alignment.n.ceil.1.function.type
Hisat2 alignment terminal '-n-ceil' option. Default value is "L"

Hisat2.Alignment.n.ceil.2.constant.term
Hisat2 alignment terminal '-n-ceil' option. Default value is "0"

Hisat2.Alignment.n.ceil.3.coefficient
Hisat2 alignment terminal '-n-ceil' option. Default value is "0.15"

Hisat2.Alignment.mp.MX
Hisat2 alignment terminal '-mp MX' option. Default value is "6"

Hisat2.Alignment.mp.MN
Hisat2 alignment terminal '-mp MN' option. Default value is "2"

Hisat2.Alignment.sp.MX
Hisat2 alignment terminal '-sp MX' option. Default value is "2"

Hisat2.Alignment.sp.MN
Hisat2 alignment terminal '-sp MN' option. Default value is "1"

Hisat2.Alignment.np
Hisat2 alignment terminal '-np' option. Default value is "1"

Hisat2.Alignment.rdg.1
Hisat2 alignment terminal '-rdg' first option. Default value is "5"

Hisat2.Alignment.rdg.2
Hisat2 alignment terminal '-rdg' first option. Default value is "3"

Hisat2.Alignment.rfg.1
Hisat2 alignment terminal '-rfg' first option. Default value is "5"

Hisat2.Alignment.rfg.2
Hisat2 alignment terminal '-rfg' first option. Default value is "3"

Hisat2.Alignment.score.min.1.function.type
Hisat2 alignment terminal '-rdg' first option. Default value is "L"

Hisat2.Alignment.score.min.2.constant.term
Hisat2 alignment terminal '-rdg' first option. Default value is "0"

Hisat2.Alignment.score.min.3.coefficient
Hisat2 alignment terminal '-rdg' first option. Default value is "-0.2"

Hisat2.Alignment.pen.cansplice
Hisat2 alignment terminal '-pen-cansplice' first option. Default value is "-0"

Hisat2.Alignment.penc.noncansplice
Hisat2 alignment terminal '-pen-noncansplice' option. Default value is "12"

Hisat2.Alignment.pen.canintronlen.1.function.type
Hisat2 alignment terminal '-pen-canintronlen' first option. Default value is "G"

Hisat2.Alignment.pen.canintronlen.2.constant.term
Hisat2 alignment terminal '-pen-canintronlen' second option. Default value is "-8"

Hisat2.Alignment.pen.canintronlen.3.coefficient
Hisat2 alignment terminal '-pen-canintronlen' third option. Default value is "1"

Hisat2.Alignment.pen.noncanintronlen.1.function.type
Hisat2 alignment terminal '-pen-noncanintronlen' first option. Default value is "G"

Hisat2.Alignment.pen.noncanintronlen.2.constant.term
Hisat2 alignment terminal '-pen-noncanintronlen' second option. Default value is "-8"

Hisat2.Alignment.pen.noncanintronlen.3.coefficient
Hisat2 alignment terminal '-pen-noncanintronlen' third option. Default value is "1"

Hisat2.Alignment.min.intronlen
Hisat2 alignment terminal '-min-intronlen' option. Default value is "20"

Hisat2.Alignment.max.intronlen
Hisat2 alignment terminal '-max-intronlen' option. Default value is "20"

Hisat2.Alignment.rna.strandness
Hisat2 alignment terminal '-rna-strandness' option. Default value is "None"

Hisat2.Alignment.k
Hisat2 alignment terminal '-k' option. Default value is "5"

Hisat2.Alignment.max.seeds
Hisat2 alignment terminal '-max-seeds' option. Default value is "5"

Hisat2.Alignment.secondary
Hisat2 alignment terminal '-secondary' option. Default value is "FALSE"

Hisat2.Alignment.minins
Hisat2 alignment terminal '-I/-minins' option. Default value is "0"

Hisat2.Alignment.maxins
Hisat2 alignment terminal '-X/-maxins' option. Default value is "500"

Hisat2.Alignment.seed
Hisat2 alignment terminal '-X/-maxins' option. Default value is "0"

STAR.Index.num.parallel.threads
Specify the number of processing threads (CPUs) to use for STAR index step.
The default is "1"

STAR.Index.sjdbOverhang.Read.length
STAR index terminal '-sjdbOverhang' option. Default value is "100"

STAR.Index.genomeSAindexNbases
 STAR index terminal '-genomeSAindexNbases' option. Default value is "14"

STAR.Index.genomeChrBinNbits
 STAR index terminal '-genomeChrBinNbits' option. Default value is "18"

STAR.Index.genomeSAsparseD
 STAR index terminal '-genomeSAsparseD' option. Default value is "1"

STAR.Alignment.run
 Whether to run 'STAR index' step in this function step. Default value is FALSE.
 Set TRUE to run STAR alignment step. (need to set Hisat2.Index.run to FALSE)

STAR.Alignment.num.parallel.threads
 Specify the number of processing threads (CPUs) to use for STAR alignment
 step. The default is "1"

STAR.Alignment.genomeLoad
 STAR alignment terminal '-genomeLoad' option. Default value is "NoSharedMemory"

STAR.Alignment.readMapNumber
 STAR alignment terminal '-readMapNumber' option. Default value is "-1"

STAR.Alignment.clip3pNbases
 STAR alignment terminal '-clip3pNbases' option. Default value is "0"

STAR.Alignment.clip5pNbases
 STAR alignment terminal '-clip5pNbases' option. Default value is "0"

STAR.Alignment.clip3pAdapterSeq
 STAR alignment terminal '-clip3pAdapterSeq' option. Default value is "-"

STAR.Alignment.clip3pAdapterMMp
 STAR alignment terminal '-clip3pAdapterMMp' option. Default value is "0.1"

STAR.Alignment.clip3pAfterAdapterNbases
 STAR alignment terminal '-clip3pAfterAdapterNbases' option. Default value
 is "0"

STAR.Alignment.limitGenomeGenerateRAM
 STAR alignment terminal '-limitGenomeGenerateRAM' option. Default value
 is "31000000000"

STAR.Alignment.limitIObufferSize
 STAR alignment terminal '-limitIObufferSize' option. Default value is "150000000"

STAR.Alignment.limitOutSAMoneReadBytes
 STAR alignment terminal '-limitOutSAMoneReadBytes' option. Default value
 is "100000"

STAR.Alignment.limitOutSJoneRead
 STAR alignment terminal '-limitOutSJoneRead' option. Default value is "1000"

STAR.Alignment.limitOutSJcollapsed
 STAR alignment terminal '-limitOutSJcollapsed' option. Default value is "1000000"

STAR.Alignment.limitBAMsortRAM
 STAR alignment terminal '-limitBAMsortRAM' option. Default value is "0"

STAR.Alignment.outReadsUnmapped
 STAR alignment terminal '-outReadsUnmapped' option. Default value is "None"

STAR.Alignment.outQScnversionAdd
 STAR alignment terminal '-outQScnversionAdd' option. Default value is "0"

STAR.Alignment.outSAMprimaryFlag
STAR alignment terminal '-outSAMprimaryFlag' option. Default value is "OneBestScore"

STAR.Alignment.outSAMmapqUnique
STAR alignment terminal '-outSAMmapqUnique' option. Default value is "255"

STAR.Alignment.scoreGap
STAR alignment terminal '-scoreGap' option. Default value is "0"

STAR.Alignment.scoreGapNoncan
STAR alignment terminal '-scoreGapNoncan' option. Default value is "-8"

STAR.Alignment.scoreGapGCAG
STAR alignment terminal '-scoreGapGCAG' option. Default value is "-4"

STAR.Alignment.scoreGapATAC
STAR alignment terminal '-scoreGapATAC' option. Default value is "-8"

STAR.Alignment.scoreGenomicLengthLog2scale
STAR alignment terminal '-scoreGenomicLengthLog2scale' option. Default value is "-0.25"

STAR.Alignment.scoreDelOpen
STAR alignment terminal '-scoreDelOpen' option. Default value is "-2"

STAR.Alignment.scoreDelBase
STAR alignment terminal '-scoreDelBase' option. Default value is "-2"

STAR.Alignment.scoreInsOpen
STAR alignment terminal '-scoreInsOpen' option. Default value is "-2"

STAR.Alignment.scoreInsBase
STAR alignment terminal '-scoreInsBase' option. Default value is "-2"

STAR.Alignment.scoreStitchSJshift
STAR alignment terminal '-scoreStitchSJshift' option. Default value is "1"

STAR.Alignment.seedSearchStartLmax
STAR alignment terminal '-scoreStitchSJshift' option. Default value is "50"

STAR.Alignment.seedSearchStartLmaxOverLread
STAR alignment terminal '-seedSearchStartLmaxOverLread' option. Default value is "1.0"

STAR.Alignment.seedSearchLmax
STAR alignment terminal '-seedSearchLmax' option. Default value is "0"

STAR.Alignment.seedMultimapNmax
STAR alignment terminal '-seedMultimapNmax' option. Default value is "10000"

STAR.Alignment.seedPerReadNmax
STAR alignment terminal '-seedPerReadNmax' option. Default value is "1000"

STAR.Alignment.seedPerWindowNmax
STAR alignment terminal '-seedPerWindowNmax' option. Default value is "50"

STAR.Alignment.seedNoneLociPerWindow
STAR alignment terminal '-seedNoneLociPerWindow' option. Default value is "10"

STAR.Alignment.alignIntronMin
STAR alignment terminal '-alignIntronMin' option. Default value is "21"

STAR.Alignment.alignIntronMax
STAR alignment terminal '-alignIntronMax' option. Default value is "0"

STAR.Alignment.alignMatesGapMax
STAR alignment terminal '-alignMatesGapMax' option. Default value is "0"

STAR.Alignment.alignSJoverhangMin
STAR alignment terminal '-alignSJoverhangMin' option. Default value is "5"

STAR.Alignment.alignSJDBoverhangMin
STAR alignment terminal '-alignSJDBoverhangMin' option. Default value is "3"

STAR.Alignment.alignSplicedMateMapLmin
STAR alignment terminal '-alignSplicedMateMapLmin' option. Default value is "0"

STAR.Alignment.alignSplicedMateMapLminOverLmate
STAR alignment terminal '-alignSplicedMateMapLminOverLmate' option. Default value is "0.66"

STAR.Alignment.alignWindowsPerReadNmax
STAR alignment terminal '-alignWindowsPerReadNmax' option. Default value is "10000"

STAR.Alignment.alignTranscriptsPerWindowNmax
STAR alignment terminal '-alignTranscriptsPerWindowNmax' option. Default value is "100"

STAR.Alignment.alignTranscriptsPerReadNmax
STAR alignment terminal '-alignTranscriptsPerReadNmax' option. Default value is "10000"

STAR.Alignment.alignEndsType
STAR alignment terminal '-alignEndsType' option. Default value is "Local"

STAR.Alignment.winAnchorMultimapNmax
STAR alignment terminal '-winAnchorMultimapNmax' option. Default value is "50"

STAR.Alignment.winBinNbits
STAR alignment terminal '-winBinNbits' option. Default value is "16"

STAR.Alignment.winAnchorDistNbins
STAR alignment terminal '-winAnchorDistNbins' option. Default value is "9"

STAR.Alignment.winFlankNbins
STAR alignment terminal '-winFlankNbins' option. Default value is "4"

Rsamtools.Bam.run
Whether to run 'Rsamtools SAM to BAM' step in this function step. Default value is TRUE. Set FALSE to skip 'Rsamtools SAM to BAM' step.

Samtools.Bam.num.parallel.threads
Specify the number of processing threads (CPUs) to use for Samtools sam to bam step. The default is "1"

Rsamtools.nCores
The number of cores to use when running 'Rsamtools' step. Default value is 1

StringTie.Assemble.run
Whether to run 'StringTie assembly' step in this function step. Default value is TRUE. Set FALSE to skip 'StringTie assembly' step.

Stringtie.Assembly.num.parallel.threads
Specify the number of processing threads (CPUs) to use for Stringtie assembly. The default is "1"

| | |
|---|--|
| Stringtie.Assembly.f | Stringtie assembly terminal '-f' option. Default value is "0.1" |
| Stringtie.Assembly.m | Stringtie assembly terminal '-m' option. Default value is "200" |
| Stringtie.Assembly.c | Stringtie assembly terminal '-c' option. Default value is "2.5" |
| Stringtie.Assembly.g | Stringtie assembly terminal '-g' option. Default value is "50" |
| Stringtie.Assembly.M | Stringtie assembly terminal '-M' option. Default value is "0.95" |
| StringTie.Merge.Trans.run | Whether to run 'StringTie GTF merging' step in this function step. Default value is TRUE. Set FALSE to skip 'StringTie GTF merging' step. |
| Stringtie.Merge.num.parallel.threads | Specify the number of processing threads (CPUs) to use for Stringtie merge step. The default is "1" |
| Gffcompare.Ref.Sample.run | Whether to run 'Gffcompare comparison' step in this function step. Default value is TRUE. Set FALSE to skip 'Gffcompare comparison' step. |
| StringTie.Ballgown.run | Whether to run 'StringTie ballgown creation' step in this function step. Default value is TRUE. Set FALSE to skip 'StringTie ballgown creation' step. |
| Stringtie.2.Ballgown.num.parallel.threads | Specify the number of processing threads (CPUs) to use for Stringtie to ballgown step. The default is "1" |
| PreDECountTable.run | Whether to run 'gene raw reads count creation' step in this function step. Default value is TRUE. Set FALSE to skip 'gene raw reads count creation' step. |
| run | Default value is TRUE. If TRUE, 'Rscript/Environment_Set.R' will be created and executed. The output log will be stored in 'Rscript_out/Environment_Set.Rout'. If False, 'Rscript/Environment_Set.R' will be created without executed. |
| check.s4.print | Default TRUE. If TRUE, the result of checking RNASeqRParam will be reported in 'Rscript_out/Environment_Set.Rout'. If FALSE, the result of checking RNASeqRParam will not be in 'Rscript_out/Environment_Set.Rout'. |

Value

None

Author(s)

Kuan-Hao Chao

Examples

```

data(yeast)
## Not run:
## Before run this function, make sure \code{RNASeqEnvironmentSet_CMD()}
## (or\code{RNASeqEnvironmentSet()}) is executed successfully.
RNASeqReadProcess_CMD(RNASeqRParam = yeast,
                      num.parallel.threads = 10)
## End(Not run)

```

RNASeqRParam-class *RNASeqR*

Description

An S4 class for checking and storing RNA-Seq workflow parameters of this package.

Slots

`os.type` 'linux' or 'osx'. The operating system type.

`python.variable` A list storing python environment. (`check.answer`, `python.version`)

`python.2to3` Logical value whether 2to3 command is available on the workstation.

`path.prefix` Path prefix of 'gene_data/', 'RNASeq_bin/', 'RNASeq_results/', 'Rscript/' and 'Rscript_out/' directories.

`input.path.prefix` Path prefix of 'input_files/' directory,

`genome.name` Variable of genome name defined in this RNA-Seq workflow (ex. `genome.name.fa`, `genome.name.gtf`).

`sample.pattern` Regular expression of paired-end fastq.gz files under 'input_files/raw_fastq.gz'. Expression not includes `_[1,2].fastq.gz`.

`independent.variable` Independent variable for the biological. experiment design of two-group RNA-Seq workflow.

`case.group` Group name of the case group.

`control.group` Group name of the control group.

`indices.optional` Logical value whether 'indices/' is exit in 'input_files/'.

`fastq.gz.type` Specify the fastq.gz file type. 'PE' represents paired-end and 'SE' represents single-end.

Author(s)

Kuan-Hao Chao

Examples

```

data(yeast)
"@"(yeast, os.type)
"@"(yeast, python.variable)
"@"(yeast, python.2to3)
"@"(yeast, path.prefix)
"@"(yeast, input.path.prefix)
"@"(yeast, genome.name)
"@"(yeast, sample.pattern)
"@"(yeast, independent.variable)
"@"(yeast, case.group)
"@"(yeast, control.group)
"@"(yeast, indices.optional)
"@"(yeast, fastq.gz.type)

```

RNASeqRParam-constructor

RNASeqRParam

Description

Constructor function for RNASeqRParam objects

Usage

```

RNASeqRParam(path.prefix = NA, input.path.prefix = NA,
             genome.name = NA, sample.pattern = NA, independent.variable = NA,
             case.group = NA, control.group = NA, fastq.gz.type = NA)

```

Arguments

| | |
|-----------------------------------|--|
| <code>path.prefix</code> | Path prefix of 'gene_data/', 'RNASeq_bin/', 'RNASeq_results/', 'Rscript/' and 'Rscript_out/' directories. |
| <code>input.path.prefix</code> | Path prefix of 'input_files/' directory. |
| <code>genome.name</code> | variable of genome name defined in this RNA-Seq workflow (ex. <code>genome.name.fa</code> , <code>genome.name.gtf</code>). |
| <code>sample.pattern</code> | Regular expression of paired-end fastq.gz files under 'input_files/raw_fastq.gz'. Expression not includes <code>_[1,2].fastq.gz</code> . |
| <code>independent.variable</code> | Independent variable for the biological experiment design of two-group RNA-Seq workflow. |
| <code>case.group</code> | Group name of the case group. |
| <code>control.group</code> | Group name of the control group. |
| <code>fastq.gz.type</code> | Specify the fastq.gz file type. 'PE' represents paired-end and 'SE' represents single-end. |

Value

an object of class RNASeqRParam

Author(s)

kuan-hao Chao

Kuan-Hao Chao

Examples

```
input_files.path <- system.file("extdata/", package = "RNASeqRData")
rnaseq_result.path <- tempdir(check = TRUE)
exp <- RNASeqRParam(path.prefix = rnaseq_result.path,
                    input.path.prefix = input_files.path,
                    genome.name = "Saccharomyces_cerevisiae_XV_Ensembl",
                    sample.pattern = "SRR[0-9]*_XV",
                    independent.variable = "state",
                    case.group = "60mins_ID20_amphotericin_B",
                    control.group = "60mins_ID20_control",
                    fastq.gz.type = "PE")
```

RNASeqRParam_Bam-class

RNASeqRParam_Bam

Description

An S4 class for checking and storing RNA-Seq workflow parameters starting with BAM files.

Slots

`os.type` 'linux' or 'osx'. The operating system type.

`python.variable` A list storing python environment. (`check.answer`, `python.version`)

`python.2to3` Logical value whether 2to3 command is available on the workstation.

`path.prefix` Path prefix of 'gene_data/', 'RNASeq_bin/', 'RNASeq_results/', 'Rscript/' and 'Rscript_out/' directories.

`input.path.prefix` Path prefix of 'input_files/' directory,

`genome.name` Variable of genome name defined in this RNA-Seq workflow (ex. `genome.name.fa`, `genome.name.gtf`).

`sample.pattern` Regular expression of paired-end fastq.gz files under 'input_files/raw_bam'. Expression not includes `_[1,2].fastq.gz`.

`independent.variable` Independent variable for the biological. experiment design of two-group RNA-Seq workflow.

`case.group` Group name of the case group.

`control.group` Group name of the control group.

Author(s)

Kuan-Hao Chao

Examples

```

data(yeast)
"@"(yeast, os.type)
"@"(yeast, python.variable)
"@"(yeast, python.2to3)
"@"(yeast, path.prefix)
"@"(yeast, input.path.prefix)
"@"(yeast, genome.name)
"@"(yeast, sample.pattern)
"@"(yeast, independent.variable)
"@"(yeast, case.group)
"@"(yeast, control.group)

```

 RNASeqRParam_Bam-constructor

RNASeqR_Bam

Description

Constructor function for RNASeqRParam_Bam objects

Usage

```

RNASeqRParam_Bam(path.prefix = NA, input.path.prefix = NA,
  genome.name = NA, sample.pattern = NA, independent.variable = NA,
  case.group = NA, control.group = NA)

```

Arguments

| | |
|-----------------------------------|---|
| <code>path.prefix</code> | Path prefix of 'gene_data/', 'RNASeq_bin/', 'RNASeq_results/', 'Rscript/' and 'Rscript_out/' directories. |
| <code>input.path.prefix</code> | Path prefix of 'input_files/' directory. |
| <code>genome.name</code> | variable of genome name defined in this RNA-Seq workflow (ex. <code>genome.name.fa</code> , <code>genome.name.gtf</code>). |
| <code>sample.pattern</code> | Regular expression of paired-end fastq.gz files under 'input_files/raw_bam'. Expression not includes <code>_[1,2].fastq.gz</code> . |
| <code>independent.variable</code> | Independent variable for the biological experiment design of two-group RNA-Seq workflow. |
| <code>case.group</code> | Group name of the case group. |
| <code>control.group</code> | Group name of the control group. |

Value

an object of class RNASeqRParam_Bam

Author(s)

Kuan-Hao Chao

Kuan-Hao Chao

Examples

```
input_files.path <- system.file("extdata/", package = "RNASeqRData")
rnaseq_result.path <- tempdir(check = TRUE)
## Not run:
exp <- RNASeqRParam_Bam(path.prefix          = rnaseq_result.path,
                        input.path.prefix    = input_files.path,
                        genome.name         = "Saccharomyces_cerevisiae_XV_Ensembl",
                        sample.pattern      = "SRR[0-9]*_XV",
                        independent.variable = "state",
                        case.group          = "60mins_ID20_amphotericin_B",
                        control.group       = "60mins_ID20_control")

## End(Not run)
```

RNASeqRParam_Sam-class

RNASeqRParam_Sam

Description

An S4 class for checking and storing RNA-Seq workflow parameters starting with SAM files.

Slots

`os.type` 'linux' or 'osx'. The operating system type.

`python.variable` A list storing python environment. (`check.answer`, `python.version`)

`python.2to3` Logical value whether 2to3 command is available on the workstation.

`path.prefix` Path prefix of 'gene_data/', 'RNASeq_bin/', 'RNASeq_results/', 'Rscript/' and 'Rscript_out/' directories.

`input.path.prefix` Path prefix of 'input_files/' directory,

`genome.name` Variable of genome name defined in this RNA-Seq workflow (ex. `genome.name.fa`, `genome.name.gtf`).

`sample.pattern` Regular expression of paired-end fastq.gz files under 'input_files/raw_sam'. Expression not includes `_[1,2].fastq.gz`.

`independent.variable` Independent variable for the biological. experiment design of two-group RNA-Seq workflow.

`case.group` Group name of the case group.

`control.group` Group name of the control group.

Author(s)

Kuan-Hao Chao

Examples

```

data(yeast)
"@"(yeast, os.type)
"@"(yeast, python.variable)
"@"(yeast, python.2to3)
"@"(yeast, path.prefix)
"@"(yeast, input.path.prefix)
"@"(yeast, genome.name)
"@"(yeast, sample.pattern)
"@"(yeast, independent.variable)
"@"(yeast, case.group)
"@"(yeast, control.group)

```

 RNASeqRParam_Sam-constructor

RNASeqR_Sam

Description

Constructor function for RNASeqRParam_Sam objects

Usage

```

RNASeqRParam_Sam(path.prefix = NA, input.path.prefix = NA,
  genome.name = NA, sample.pattern = NA, independent.variable = NA,
  case.group = NA, control.group = NA)

```

Arguments

| | |
|-----------------------------------|---|
| <code>path.prefix</code> | Path prefix of 'gene_data/', 'RNASeq_bin/', 'RNASeq_results/', 'Rscript/' and 'Rscript_out/' directories. |
| <code>input.path.prefix</code> | Path prefix of 'input_files/' directory. |
| <code>genome.name</code> | variable of genome name defined in this RNA-Seq workflow (ex. <code>genome.name.fa</code> , <code>genome.name.gtf</code>). |
| <code>sample.pattern</code> | Regular expression of paired-end fastq.gz files under 'input_files/raw_sam'. Expression not includes <code>_[1,2].fastq.gz</code> . |
| <code>independent.variable</code> | Independent variable for the biological experiment design of two-group RNA-Seq workflow. |
| <code>case.group</code> | Group name of the case group. |
| <code>control.group</code> | Group name of the control group. |

Value

an object of class RNASeqRParam_Sam

Author(s)

Kuan-Hao Chao

Kuan-Hao Chao

Examples

```
input_files.path <- system.file("extdata/", package = "RNASeqRData")
rnaseq_result.path <- tempdir(check = TRUE)
## Not run:
exp <- RNASeqRParam_Sam(path.prefix          = rnaseq_result.path,
                        input.path.prefix    = input_files.path,
                        genome.name          = "Saccharomyces_cerevisiae_XV_Ensembl",
                        sample.pattern       = "SRR[0-9]*_XV",
                        independent.variable = "state",
                        case.group           = "60mins_ID20_amphotericin_B",
                        control.group       = "60mins_ID20_control")

## End(Not run)
```

Update_Fastq_gz

Update_Fastq_gz

Description

This function let users update their trimmed fastq.gz files automatically.

Usage

```
Update_Fastq_gz(RNASeqRParam, prepared_fastq_gz, target_samples = "ALL")
```

Arguments

RNASeqRParam S4 object instance of experiment-related parameters

prepared_fastq_gz

absolute path to the prepared 'raw_fastq.gz' directory.

target_samples list of samples that are going to update. Default value is ALL

Value

None

Author(s)

Kuan-Hao Chao

Examples

```
data(yeast)
## Not run:
RNASeqDifferentialAnalysis(RNASeqRParam = yeast)
## End(Not run)
```

yeast

Toy RNASeqRParam object

Description

Small RNASeqRParam S4 object created with checked valid parameters for demonstration purposes

Author(s)

Kuan-Hao Chao

Examples

```
data(yeast)
yeast
# RNASeqRParam S4 object for example demonstration.
```

Index

All_Steps_Interface, [3](#)
All_Steps_Interface_CMD, [4](#)

CheckToolAll, [6](#)

RNASeqDifferentialAnalysis, [7](#)
RNASeqDifferentialAnalysis_CMD, [8](#)
RNASeqEnvironmentSet, [10](#)
RNASeqEnvironmentSet_CMD, [12](#)
RNASeqGoKegg, [13](#)
RNASeqGoKegg_CMD, [15](#)
RNASeqQualityAssessment, [16](#)
RNASeqQualityAssessment_CMD, [17](#)
RNASeqR, [18](#)
RNASeqReadProcess, [18](#)
RNASeqReadProcess_CMD, [27](#)
RNASeqRParam
 (RNASeqRParam-constructor), [38](#)
RNASeqRParam-class, [37](#)
RNASeqRParam-constructor, [38](#)
RNASeqRParam_Bam
 (RNASeqRParam_Bam-constructor),
 [40](#)
RNASeqRParam_Bam-class, [39](#)
RNASeqRParam_Bam-constructor, [40](#)
RNASeqRParam_Sam
 (RNASeqRParam_Sam-constructor),
 [42](#)
RNASeqRParam_Sam-class, [41](#)
RNASeqRParam_Sam-constructor, [42](#)

Update_Fastq_gz, [43](#)

yeast, [44](#)