Package ‘gDNAinRNAseqData’

May 14, 2024

**Title**  RNA-seq data with different levels of gDNA contamination

**Version**  1.4.0

**Description**  Provides access to BAM files generated from RNA-seq data produced with different levels of gDNA contamination. It currently allows one to download a subset of the data published by Li et al., BMC Genomics, 23:554, 2022. This subset of data is formed by BAM files with about 100,000 alignments with three different levels of gDNA contamination.

**Depends**  R (>= 4.3)

**Imports**  RCurl, XML, ExperimentHub, BiocGenerics, Rsamtools

**Suggests**  BiocStyle, knitr, rmarkdown

**VignetteBuilder**  knitr

**License**  Artistic-2.0

**NeedsCompilation**  no

**Encoding**  UTF-8

**biocViews**  ExperimentHub, ExperimentData, ExpressionData, SequencingData, RNASeqData, Homo_sapiens_Data

**URL**  https://github.com/functionalgenomics/gDNAinRNAseqData

**BugReports**  https://github.com/functionalgenomics/gDNAinRNAseqData

**RoxygenNote**  7.2.3

**git_url**  https://git.bioconductor.org/packages/gDNAinRNAseqData

**git_branch**  RELEASE_3_19

**git_last_commit**  5d45f06

**git_last_commit_date**  2024-04-30

**Repository**  Bioconductor 3.19

**Date/Publication**  2024-05-14

**Author**  Robert Castelo [aut, cre] (<https://orcid.org/0000-0003-2229-4508>)

**Maintainer**  Robert Castelo <robert.castelo@upf.edu>
**Contents**

RNA-seq data with different levels of gDNA contamination

**Description**

This package provides access to RNA-seq BAM files containing different levels of genomic DNA (gDNA) contamination.

**Usage**

```r
LiYu22subsetBAMfiles(path = tempdir(), offline = FALSE)
LiYu22phenoData(bamfiles)
```

**Arguments**

- **path**  
  (Default='tempdir()') Filesystem path where to store the BAM files.
- **offline**  
  (Default='FALSE') If there is no internet connection, but the data has been previously downloaded, setting `offline=TRUE` allows one to retrieve the data from the ExperimentHub cache.
- **bamfiles**  
  full filesystem paths to where the BAM files were downloaded with `LiYu22subsetBAMfiles()`.

**Details**

Currently, this package allows one to download a subset of the data published in:


The subset of the data accessible through this package corresponds to BAM files containing about 100,000 alignments sampled uniformly at random for the RNA-seq experiments produced from total RNA libraries mixed with different concentrations of gDNA, concretely 0% (no contamination), 1% and 10%; see Fig. 2 from Li et al. (2022).

**Value**

- `LiYu22subsetBAMfiles()` returns a string character vector of filesystem paths to the downloaded BAM files.
- `LiYu22phenoData()` returns a `data.frame` object with the gDNA contamination levels for the BAM files specified in the `bamfiles` parameter, according to the publication by Li et al. (2022).
Functions

- `LiYu22subsetBAMfiles()`: downloads the BAM files from the RNA-seq data through the ExperimentHub, and returns the path in the filesystem where the BAM files are stored.
- `LiYu22phenoData()`: retrieves phenotypic data from the BAM files downloaded with `LiYu22subsetBAMfiles()`.

References


Examples

```r
## for LiYu22subsetBAMfiles()
bamfiles <- LiYu22subsetBAMfiles()
bamfiles

## for LiYu22phenoData()
bamfiles <- LiYu22subsetBAMfiles()
LiYu22phenoData(bamfiles)
```
Index

* package
  gDNAinRNAseqData-package, 2

  gDNAinRNAseqData
   (gDNAinRNAseqData-package), 2
   gDNAinRNAseqData-package, 2

  LiYu22phenoData
   (gDNAinRNAseqData-package), 2
  LiYu22subsetBAMfiles
   (gDNAinRNAseqData-package), 2