## Package 'regionReport'

October 12, 2016

Type Package

Title Generate HTML or PDF reports for a set of genomic regions or DESeq2/edgeR results

Version 1.6.5

Date 2016-05-12

**Description** Generate HTML or PDF reports to explore a set of regions such as the results from annotation-agnostic expression analysis of RNA-seq data at base-pair resolution performed by derfinder. You can also create reports for DESeq2 or edgeR results.

License Artistic-2.0

LazyData true

URL https://github.com/leekgroup/regionReport

BugReports https://github.com/leekgroup/regionReport/issues

#### VignetteBuilder knitr

**biocViews** DifferentialExpression, Sequencing, RNASeq, Software, Visualization, Transcription, Coverage, ReportWriting, DifferentialMethylation, DifferentialPeakCalling

#### **Depends** R(>= 3.2)

- **Imports** derfinder (>= 1.1.0), DEFormats, DESeq2, GenomeInfoDb, GenomicRanges, knitcitations (>= 1.0.1), knitr (>= 1.6), knitrBootstrap (>= 0.9.0), RefManageR, rmarkdown (>= 0.9.5), S4Vectors, SummarizedExperiment
- Suggests BiocStyle, biovizBase, bumphunter (>= 1.7.6), Cairo, derfinderPlot (>= 1.3.2), devtools (>= 1.6), DT, DESeq, edgeR, ggbio (>= 1.13.13), ggplot2, grid, gridExtra, IRanges, mgcv, pasilla, pheatmap, RColorBrewer, TxDb.Hsapiens.UCSC.hg19.knownGene, whisker

RoxygenNote 5.0.1

#### NeedsCompilation no

Author Leonardo Collado-Torres [aut, cre], Andrew E. Jaffe [aut], Jeffrey T. Leek [aut, ths]

Maintainer Leonardo Collado-Torres <lcollado@jhu.edu>

## **R** topics documented:

| 1:  |
|---|
| with_wd $\ldots$ $\ldots$ $\ldots$ $\ldots$ $1^4$ |
| renderReport                                      |
| regionReport-defunct                              |
| oad_install                                       |
| edgeReport  |
| DESeq2Report                                      |
| lerfinderReport                                   |
| egionReport-package                               |

## Index

regionReport-package Generate HTML or PDF reports for a set of regions or DESeq2/edgeR results.

#### Description

Generate an HTML reports to explore a set of regions such as the results from annotation-agnostic expression analysis of RNA-seq data at base-pair resolution performed by derfinder. The package can also be used to generate reports for DESeq2 or edgeR results. The HTML and PDF reports are generated using rmarkdown (http://rmarkdown.rstudio.com/).

#### Author(s)

Leonardo Collado-Torres <lcollado@jhsph.edu>

| derfinderReport | Generate a HTML/PDF report exploring the basic results from derfinder |
|-----------------|---|
|-----------------|---|

#### Description

This function generates a HTML report exploring the basic results from single base-level approach derfinder analysis results (www.bioconductor.org/packages/derfinder). The HTML report itself is generated using rmarkdown (http://rmarkdown.rstudio.com/). It works best after using mergeResults.

## derfinderReport

#### Usage

```
derfinderReport(prefix, outdir = "basicExploration",
  output = "basicExploration", project = prefix, browse = interactive(),
  nBestRegions = 100, makeBestClusters = TRUE, nBestClusters = 2,
  fullCov = NULL, hg19 = TRUE, p.ideos = NULL, txdb = NULL,
  device = "png", significantVar = "qvalue", customCode = NULL,
  template = NULL, theme = NULL, digits = 2, ...)
```

#### Arguments

| prefix          | The main data directory path where mergeResults was run. It should be the same as mergeResults(prefix).   |  |
|-----------------|---|--|
| outdir          | The name of output directory relative to prefix.  |  |
| output          | The name of output HTML file (without the html extension).  |  |
| project         | The title of the project.   |  |
| browse          | If TRUE the HTML report is opened in your browser once it's completed.  |  |
| nBestRegions    | The number of region plots to make, ordered by area.  |  |
| makeBestCluster | S   |  |
|                 | If TRUE, plotCluster is used on the nBestClusters regions by area. Note that these plots take some time to make.  |  |
| nBestClusters   | The number of region cluster plots to make by taking the nBestClusters re-<br>gions ranked by area of the cluster.  |  |
| fullCov         | A list where each element is the result from loadCoverage used with cutoff=NULL.<br>Can be generated using fullCoverage.  |  |
| hg19            | If TRUE then the reference is assumed to be hg19 and chromosome lengths as well as the default transcription database (TxDb.Hsapiens.UCSC.hg19.knownGene) will be used.   |  |
| p.ideos         | A list where each element is the result of plotIdeogram. If it's NULL and hg19=TRUE then they are created for the hg19 human reference.   |  |
| txdb            | Specify the transcription database to use for making the plots for the top regions by area. If NULL and hg19=TRUE then TxDb.Hsapiens.UCSC.hg19.knownGene is used.   |  |
| device          | The graphical device used when knitting. See more at http://yihui.name/knitr/options (dev argument).  |  |
| significantVar  | A character variable specifying whether to use the p-values, the FDR adjusted p-values or the FWER adjusted p-values to determine significance. Has to be either 'pvalue', 'qvalue' or 'fwer'.  |  |
| customCode      | An absolute path to a child R Markdown file with code to be evaluated before<br>the reproducibility section. Its useful for users who want to customize the report<br>by adding conclusions derived from the data and/or further quality checks and<br>plots. |  |
| template        | Template file to use for the report. If not provided, will use the default file found in basicExploration/basicExploration.Rmd within the package source.   |  |

derfinderReport

| theme  | A ggplot2 theme to use for the plots made with ggplot2.  |
|--------|--|
| digits | The number of digits to round to in the interactive table of the top nBestRegions.<br>Note that p-values and adjusted p-values won't be rounded. |
|        | Arguments passed to other methods and/or advanced arguments.   |

#### Details

Set output\_format to 'knitrBootstrap::bootstrap\_document' or 'pdf\_document' if you want a HTML report styled by knitrBootstrap or a PDF report respectively. If using knitrBootstrap, we recommend the version available only via GitHub at https://github.com/jimhester/knitrBootstrap which has nicer features than the current version available via CRAN.

If you modify the YAML front matter of template, you can use other values for output\_format.

The HTML report styled with knitrBootstrap can be smaller in size than the default HTML report.

#### Value

An HTML report with a basic exploration of the derfinder results.

#### Author(s)

Leonardo Collado-Torres

#### See Also

mergeResults, analyzeChr, fullCoverage

#### Examples

```
## Load derfinder
library('derfinder')
```

```
## If you prefer, you can generate the output from derfinder
initialPath <- getwd()
setwd(file.path(initialPath, 'derfinderReport-example'))
```

```
## Collapse the coverage information
collapsedFull <- collapseFullCoverage(list(genomeData$coverage),
    verbose=TRUE)
```

```
## Calculate library size adjustments
sampleDepths <- sampleDepth(collapsedFull, probs=c(0.5), nonzero=TRUE,</pre>
```

```
verbose=TRUE)
## Build the models
group <- genomeInfo$pop</pre>
adjustvars <- data.frame(genomeInfo$gender)</pre>
models <- makeModels(sampleDepths, testvars=group, adjustvars=adjustvars)</pre>
## Analyze chromosome 21
analyzeChr(chr='21', coverageInfo=genomeData, models=models,
    cutoffFstat=1, cutoffType='manual', seeds=20140330, groupInfo=group,
   mc.cores=1, writeOutput=TRUE, returnOutput=FALSE)
## Change the directory back to the original one
setwd(initialPath)
## End(Not run)
## Merge the results from the different chromosomes. In this case, there's
## only one: chr21
mergeResults(chrs = '21', prefix = 'derfinderReport-example',
    genomicState = genomicState$fullGenome)
## Load the options used for calculating the statistics
load(file.path('derfinderReport-example', 'chr21', 'optionsStats.Rdata'))
## Generate the HTML report
report <- derfinderReport(prefix='derfinderReport-example', browse=FALSE,
   nBestRegions=1, makeBestClusters=FALSE,
    fullCov=list('21'=genomeDataRaw$coverage), optionsStats=optionsStats)
if(interactive()) {
    ## Browse the report
   browseURL(report)
}
## Not run:
## Note that you can run the example using:
example('derfinderReport', 'regionReport', ask=FALSE)
## End(Not run)
```

DESeq2Report

Generate a HTML/PDF report exploring DESeq2 results

#### Description

This function generates a HTML report with exploratory data analysis plots for DESeq2 results created with DESeq. Other output formats are possible such as PDF but lose the interactivity. Users can easily append to the report by providing a R Markdown file to customCode, or can customize the entire template by providing an R Markdown file to template.

```
DESeq2Report(dds, project = "", intgroup, colors = NULL, res = NULL,
  nBest = 500, nBestFeatures = 20, customCode = NULL,
  outdir = "DESeq2Exploration", output = "DESeq2Exploration",
  browse = interactive(), device = "png", template = NULL,
  searchURL = "http://www.ncbi.nlm.nih.gov/gene/?term=", theme = NULL,
  digits = 2, ...)
```

#### Arguments

| dds           | A DESeqDataSet object with the results from running DESeq.  |
|---------------|---|
| project       | The title of the project.   |
| intgroup      | interesting groups: a character vector of names in colData(x) to use for group-<br>ing. This parameter is passed to functions such as plotPCA.  |
| colors        | vector of colors used in heatmap. If NULL, then a a default set of colors will be used. This argument is passed to pheatmap.  |
| res           | A DESeqResults object. If NULL, then results will be used on dds with default parameters.   |
| nBest         | The number of features to include in the interactive table. Features are ordered by their adjusted p-values.  |
| nBestFeatures | The number of best features to make plots of their counts. We recommend a small number, say 20.   |
| customCode    | An absolute path to a child R Markdown file with code to be evaluated before<br>the reproducibility section. Its useful for users who want to customize the report<br>by adding conclusions derived from the data and/or further quality checks and<br>plots. |
| outdir        | The name of output directory.   |
| output        | The name of output HTML file (without the html extension).  |
| browse        | If TRUE the HTML report is opened in your browser once it's completed.  |
| device        | The graphical device used when knitting. See more at http://yihui.name/knitr/options (dev argument).  |
| template      | Template file to use for the report. If not provided, will use the default file found in DESeq2Exploration/DESeq2Exploration.Rmd within the package source.   |
| searchURL     | A url used for searching the name of the features in the web. By default http://www.ncbi.nlm.nih.gov/<br>is used which is the recommended option when features are genes. It's only used<br>when the output is a HTML file.                                   |
| theme         | A ggplot2 theme to use for the plots made with ggplot2.   |
| digits        | The number of digits to round to in the interactive table of the top nBestFeatures.<br>Note that p-values and adjusted p-values won't be rounded.   |
|               | Arguments passed to other methods and/or advanced arguments.  |

#### DESeq2Report

#### Details

Set output\_format to 'knitrBootstrap::bootstrap\_document' or 'pdf\_document' if you want a HTML report styled by knitrBootstrap or a PDF report respectively. If using knitrBootstrap, we recommend the version available only via GitHub at https://github.com/jimhester/knitrBootstrap which has nicer features than the current version available via CRAN.

If you modify the YAML front matter of template, you can use other values for output\_format.

#### Value

An HTML report with a basic exploration for the given set of DESeq2 results.

#### Author(s)

Leonardo Collado-Torres

#### Examples

```
## Load example data from the pasilla package
library('pasilla')
library('DESeq')
library('DESeq2')
## Create DESeqDataSet object from the pasilla package
data('pasillaGenes')
countData <- counts(pasillaGenes)</pre>
colData <- pData(pasillaGenes)[, c('condition', 'type')]</pre>
dds <- DESeqDataSetFromMatrix(countData = countData,</pre>
   colData = colData,
   design = ~ condition)
dds <- DESeq(dds)
## The output will be saved in the 'DESeq2Report-example' directory
dir.create('DESeq2Report-example', showWarnings = FALSE, recursive = TRUE)
## Generate the HTML report
report <- DESeq2Report(dds, 'DESeq2-example', c('condition', 'type'),</pre>
    outdir = 'DESeq2Report-example')
if(interactive()) {
    ## Browse the report
    browseURL(report)
}
## Not run:
## Note that you can run the example using:
example('DESeq2Report', 'regionReport', ask=FALSE)
## End(Not run)
```

```
edgeReport
```

#### Description

This function generates a HTML report with exploratory data analysis plots for edgeR results created. Other output formats are possible such as PDF reports but they lose the interactivity. Users can easily append to the report by providing a R Markdown file to customCode, or can customize the entire template by providing an R Markdown file to template.

#### Usage

```
edgeReport(dge, object, project = "", intgroup, colors = NULL,
pAdjustMethod = "BH", alpha = 0.1, independentFiltering = FALSE, filter,
theta, filterFun, nBest = 500, nBestFeatures = 20, customCode = NULL,
outdir = "edgeRexploration", output = "edgeRexploration",
browse = interactive(), device = "png", template = NULL,
searchURL = "http://www.ncbi.nlm.nih.gov/gene/?term=", theme = NULL,
digits = 2, ...)
```

#### Arguments

| dge             | A DGEList object.   |
|-----------------|---|
| object          | A DGEExact or DGELRT object that contains p-values stored in object\$table\$PValue.   |
| project         | The title of the project.   |
| intgroup        | interesting groups: a character vector of names in colData(x) to use for group-<br>ing. This parameter is passed to functions such as plotPCA.  |
| colors          | vector of colors used in heatmap. If NULL, then a a default set of colors will be used. This argument is passed to pheatmap.  |
| pAdjustMethod   | the method to use for adjusting p-values, see p.adjust. This argument will be passed to results.  |
| alpha           | the significance cutoff used for optimizing the independent filtering (by default 0.1). If the adjusted p-value cutoff (FDR) will be a value other than 0.1, alpha should be set to that value. This argument will be passed to results.                                      |
| independentFilt | tering  |
|                 | logical, whether independent filtering should be applied automatically. By de-<br>fault it's set to FALSE in contrast with the default used in results to match edgeR's<br>behavior.  |
| filter          | the vector of filter statistics over which the independent filtering will be opti-<br>mized. By default the logCPM will be used if independentFiltering is set<br>to TRUE. It can also be a length 1 character vector specifying one of the column<br>names of object\$table. |
| theta           | the quantiles at which to assess the number of rejections from independent fil-<br>tering. This argument is passed results.   |

#### edgeReport

| filterFun     | an optional custom function as described in results.  |
|---------------|---|
| nBest         | The number of features to include in the interactive table. Features are ordered by their adjusted p-values.  |
| nBestFeatures | The number of best features to make plots of their counts. We recommend a small number, say 20.   |
| customCode    | An absolute path to a child R Markdown file with code to be evaluated before<br>the reproducibility section. Its useful for users who want to customize the report<br>by adding conclusions derived from the data and/or further quality checks and<br>plots. |
| outdir        | The name of output directory.   |
| output        | The name of output HTML file (without the html extension).  |
| browse        | If TRUE the HTML report is opened in your browser once it's completed.  |
| device        | The graphical device used when knitting. See more at http://yihui.name/knitr/options (dev argument).  |
| template      | Template file to use for the report. If not provided, will use the default file found in DESeq2Exploration/DESeq2Exploration.Rmd within the package source.   |
| searchURL     | A url used for searching the name of the features in the web. By default http://www.ncbi.nlm.nih.gov/<br>is used which is the recommended option when features are genes. It's only used<br>when the output is a HTML file.                                   |
| theme         | A ggplot2 theme to use for the plots made with ggplot2.   |
| digits        | The number of digits to round to in the interactive table of the top nBestFeatures.<br>Note that p-values and adjusted p-values won't be rounded.   |
|               | Arguments passed to other methods and/or advanced arguments.  |
|               |   |

## Details

Set output\_format to 'knitrBootstrap::bootstrap\_document' or 'pdf\_document' if you want a HTML report styled by knitrBootstrap or a PDF report respectively. If using knitrBootstrap, we recommend the version available only via GitHub at https://github.com/jimhester/knitrBootstrap which has nicer features than the current version available via CRAN.

If you modify the YAML front matter of template, you can use other values for output\_format.

This report is similar to the one created by DESeq2Report with two additional plots exclusive for edgeR results. We designed the reports to be very similar intentionally and use the Bioconductor package DEFormats to achieve this goal.

#### Value

An HTML report with a basic exploration for the given set of edgeR results.

#### Author(s)

Leonardo Collado-Torres

#### Examples

```
## Create example data using DEFormats
library('DEFormats')
set.seed(20160407)
counts <- simulateRnaSeqData()</pre>
group <- rep(c("A", "B"), each = 3)</pre>
## Create DGEList object
library('edgeR')
dge <- DGEList(counts, group = group)</pre>
## Perform DE analysis with edgeR
design <- model.matrix( ~ group)</pre>
dge <- estimateDisp(dge, design)</pre>
fit <- glmFit(dge, design)</pre>
lrt <- glmLRT(fit, coef = 2)</pre>
## The output will be saved in the 'edgeReport-example' directory
dir.create('edgeReport-example', showWarnings = FALSE, recursive = TRUE)
## Generate the HTML report
report <- edgeReport(dge, lrt, project = 'edgeR-example', intgroup = 'group',</pre>
    outdir = 'edgeReport-example')
if(interactive()) {
    ## Browse the report
    browseURL(report)
}
## Not run:
## Note that you can run the example using:
example('edgeReport', 'regionReport', ask=FALSE)
## End(Not run)
```

load\_install Attempt to load the namespace of a package and install it if it's missing

#### Description

This function uses requireNamespace to try to load a package. But if it's misssing it will then install it via Bioconductor.

#### Usage

load\_install(pkg, quietly = TRUE)

10

#### Arguments

| pkg     | A single character vector with the name of the package.      |
|---------|--|
| quietly | Whether to run requireNamespace and biocLite quietly or not. |

regionReport-defunct Defunct functions in package 'regionReport'

#### Description

These functions are defunct and no longer available.

#### Usage

derfinder\_report()

render\_report()

#### Details

The following functions are defunct; use the replacements indicated below:

- derfinder\_report: derfinderReport
- render\_report: renderReport

renderReport

Generate a HTML/PDF report exploring a set of genomic regions

#### Description

This function generates a HTML report with quality checks, genome location exploration, and an interactive table with the results. Other output formats are possible such as PDF but lose the interactivity. Users can easily append to the report by providing a R Markdown file to customCode, or can customize the entire template by providing an R Markdown file to template.

#### Usage

```
renderReport(regions, project = "", pvalueVars = c(`P-values` = "pval"),
densityVars = NULL, significantVar = mcols(regions)$pval <= 0.05,
annotation = NULL, nBestRegions = 500, customCode = NULL,
outdir = "regionExploration", output = "regionExploration",
browse = interactive(), txdb = NULL, device = "png",
densityTemplates = list(Pvalue = templatePvalueDensity, Common =
templateDensity, Manhattan = templateManhattan), template = NULL,
theme = NULL, digits = 2, ...)
```

## renderReport

templatePvalueDensity

templateDensity

templateManhattan

templatePvalueHistogram

templateHistogram

## Arguments

| regions          | The set of genomic regions of interest as a GRanges object. All sequence lengths must be provided.  |  |
|------------------|---|--|
| project          | The title of the project.   |  |
| pvalueVars       | The names of the variables with values between 0 and 1 to plot density values by chromosome and a table for commonly used cutoffs. Most commonly used to explore p-value distributions. If a named character vector is provided, the names are used in the plot titles. |  |
| densityVars      | The names of variables to use for making density plots by chromosome. Com-<br>monly used to explore scores and other variables given by region. If a named<br>character vector is provided, the names are used in the plot titles.                                      |  |
| significantVar   | A logical variable differentiating statistically significant regions from the rest.<br>When provided, both types of regions are compared against each other to see<br>differences in width, location, etc.  |  |
| annotation       | The output from matchGenes used on regions. Note that this can take time for a large set of regions so it's better to pre-compute this information and save it.   |  |
| nBestRegions     | The number of regions to include in the interactive table.  |  |
| customCode       | An absolute path to a child R Markdown file with code to be evaluated before<br>the reproducibility section. Its useful for users who want to customize the report<br>by adding conclusions derived from the data and/or further quality checks and<br>plots.           |  |
| outdir           | The name of output directory.   |  |
| output           | The name of output HTML file (without the html extension).  |  |
| browse           | If TRUE the HTML report is opened in your browser once it's completed.  |  |
| txdb             | Specify the transcription database to use for identifying the closest genes via matchGenes. If NULL it will use TxDb.Hsapiens.UCSC.hg19.knownGene by default.   |  |
| device           | The graphical device used when knitting. See more at http://yihui.name/knitr/options (dev argument).  |  |
| densityTemplates |   |  |
|                  | A list of length 3 with templates for the p-value density plots (variables from pvalueVars), the continuous variables density plots (variables from densityVars), and Manhattan plots for the p-value variables (pvalueVars). These templates                           |  |

12

|          | are processed by whisker.render. Check the default templates for more informa-<br>tion. The densityTemplates argument is available for those users interested<br>in customizing these plots. For example, to show histograms instead of den-<br>sity plots use templatePvalueHistogram and templateHistogram instead of<br>templatePvalueDensity and templateDensity respectively. |
|----------|--|
| template | Template file to use for the report. If not provided, will use the default file found in regionExploration/regionExploration.Rmd within the package source.  |
| theme    | A ggplot2 theme to use for the plots made with ggplot2.  |
| digits   | The number of digits to round to in the interactive table of the top nBestRegions.<br>Note that p-values and adjusted p-values won't be rounded.   |
|          | Arguments passed to other methods and/or advanced arguments.   |

#### Format

An object of class character of length 1.

#### Details

Set output\_format to 'knitrBootstrap::bootstrap\_document' or 'pdf\_document' if you want a HTML report styled by knitrBootstrap or a PDF report respectively. If using knitrBootstrap, we recommend the version available only via GitHub at https://github.com/jimhester/knitrBootstrap which has nicer features than the current version available via CRAN.

If you modify the YAML front matter of template, you can use other values for output\_format.

#### Value

An HTML report with a basic exploration for the given set of genomic regions.

#### Author(s)

Leonardo Collado-Torres

#### Examples

```
## Load derfinder for an example set of regions
library('derfinder')
regions <- genomeRegions$regions</pre>
```

```
## Assign chr length
library('GenomicRanges')
seqlengths(regions) <- c('chr21' = 48129895)</pre>
```

```
## The output will be saved in the 'renderReport-example' directory
dir.create('renderReport-example', showWarnings = FALSE, recursive = TRUE)
```

```
## Generate the HTML report
report <- renderReport(regions, 'Example run', pvalueVars = c(
    'Q-values' = 'qvalues', 'P-values' = 'pvalues'), densityVars = c(</pre>
```

```
'Area' = 'area', 'Mean coverage' = 'meanCoverage'),
    significantVar = regions$qvalues <= 0.05, nBestRegions = 20,</pre>
    outdir = 'renderReport-example')
if(interactive()) {
    ## Browse the report
   browseURL(report)
}
## Not run:
## Note that you can run the example using:
example('renderReport', 'regionReport', ask=FALSE)
## End(Not run)
## Check the default templates. For users interested in customizing these
## plots.
## For p-value variables:
cat(templatePvalueDensity)
## For continous variables:
cat(templateDensity)
## For Manhattan plots
cat(templateManhattan)
```

| with_wd | Temporarily evaluate an ex | xpression in a directory |
|---------|----------------------------|--------------------------|
|---------|----------------------------|--------------------------|

#### Description

Temporarily evaluate an expression in a directory, then set the directory back to the original.

#### Usage

```
with_wd(dir, expr)
```

#### Arguments

| dir  | a directory to perform an expression within |
|------|---|
| expr | expression to evaluate                      |

#### Details

See here: http://plantarum.ca/code/setwd-part2/

#### Author(s)

Tyler Smith, contributed to regionReport by David Robinson https://github.com/dgrtwo

14

# Index

\*Topic **datasets** renderReport, 11 \*Topic package regionReport-package, 2 analyzeChr, 4 derfinder\_report (regionReport-defunct), 11 derfinderReport, 2, 11 DESeq, 5, 6 DESeq2Report, 5, 9 DESeqDataSet, 6 DESeqResults, 6 DGEExact, 8 DGEList, 8 DGELRT, 8 edgeReport, 8fullCoverage, 3, 4 load\_install, 10 loadCoverage, 3 matchGenes, 12 mergeResults, 2-4 p.adjust, 8 pheatmap, 6, 8plotCluster, 3 plotIdeogram, 3 plotPCA, 6, 8 regionReport-defunct, 11 regionReport-package, 2 render\_report (regionReport-defunct), 11 renderReport, 11, 11 results, 6, 8, 9

templateDensity(renderReport), 11

whisker.render, *13* with\_wd, 14