# Package 'proActiv'

July 17, 2025

Title Estimate Promoter Activity from RNA-Seq data

**Version** 1.19.2

Description Most human genes have multiple promoters that control the expression of different isoforms. The use of these alternative promoters enables the regulation of isoform expression pre-transcriptionally.

Alternative promoters have been found to be important in a wide number of cell types and diseases. proActiv is an R package that enables the analysis of promoters from RNA-seq data. proActiv uses aligned reads as input, and generates counts and normalized promoter activity estimates for each annotated promoter. In particular, proActiv accepts junction files from TopHat2 or STAR or BAM files as inputs. These estimates can then be used to identify which promoter is active, which promoter is inactive, and which promoters change their activity across conditions. proActiv also allows visualization of promoter activity across conditions.

License MIT + file LICENSE

**Encoding** UTF-8

LazyData true

**Depends** R (>= 4.0.0)

RoxygenNote 7.3.1

Imports AnnotationDbi, BiocParallel, data.table, dplyr, DESeq2, IRanges, GenomicRanges, GenomicFeatures, GenomicAlignments, GenomeInfoDb, GenomeInfoDbData, ggplot2, gplots, graphics, methods, rlang, scales, S4Vectors, SummarizedExperiment, stats, tibble, txdbmaker

Suggests testthat, rmarkdown, knitr, Rtsne, gridExtra

URL https://github.com/GoekeLab/proActiv

**biocViews** RNASeq, GeneExpression, Transcription, AlternativeSplicing, GeneRegulation, DifferentialSplicing, FunctionalGenomics, Epigenetics, Transcriptomics, Preprocessing

VignetteBuilder knitr

git\_url https://git.bioconductor.org/packages/proActiv

git\_branch devel

git\_last\_commit 1ed6022

git\_last\_commit\_date 2025-07-14

2 boxplotPromoters

# **Contents**

Index		17
	reexports	16
	promoterAnnotation.gencode.v34.subset	
	PromoterAnnotation-class	14
	proActiv	12
	preparePromoterAnnotation	12
	plotPCA	11
	plotHeatmap	9
	normalizePromoterReadCounts	9
	integrateProactiv	8
	getRelativePromoterActivity	7
	getGeneExpression	7
	getAlternativePromoters	6
	getAbsolutePromoterActivity	5
	calculatePromoterReadCounts	4
	calculateJunctionReadCounts	3
	boxplotPromoters	

 ${\tt boxplotPromoters}$ 

Visualizes promoter activity and gene expression with boxplots

# Description

Visualizes promoter activity and gene expression with boxplots

# Usage

```
boxplotPromoters(
  result,
  geneId,
  geneName = NULL,
  filterInternal = TRUE,
  col = NULL
)
```

calculateJunctionReadCounts

### **Arguments**

resul <sup>.</sup>	t	A SummarizedExperiment object return by proActiv, with assays giving promoter counts and activity with gene expression stored as metadata. rowData contains promoter metadata and absolute promoter activity summarized across conditions. Condition must be provided.
geneI	d	A character vector. A single gene id. This identifier must correspond to the identifier in the promoter annotation.
geneNa	ame	A character vector. Common gene name to be displayed on plot. Optional. Defaults to NULL.
filte	rInternal	A boolean. Determines if internal promoters should be removed from the plot. Defaults to TRUE.
col		A character vector of colours to be used for plotting.

### Value

A list of length 3. Each entry is a plot corresponding to absolute promoter activity, relative promoter activity and gene expression.

#### **Examples**

calculateJunctionReadCounts

Calculate the total number of junction reads overlapping with the introns of each promoter for the input junction file

# **Description**

Calculate the total number of junction reads overlapping with the introns of each promoter for the input junction file

# Usage

```
calculateJunctionReadCounts(
  promoterCoordinates,
  intronRanges,
  file = "",
  fileType = "",
  genome = ""
)
```

#### **Arguments**

promoterCoordinates

A GRanges object containing promoter coordinates and reduced exon coordi-

nates by gene

intronRanges A Granges object containing the annotated unique intron ranges. These ranges

will be used for counting the reads

file character path for the input junction bed or bam file

fileType character type of the junction bed file. Either 'tophat', 'star' or 'bam'

genome character genome version

#### Value

The total number of junction reads overlapping with each promoter for the input annotated intron ranges

 ${\tt calculatePromoterReadCounts}$ 

Calculate the promoter read counts using junction read counts approach for all the input junction files

# **Description**

Calculate the promoter read counts using junction read counts approach for all the input junction files

### Usage

```
calculatePromoterReadCounts(
  promoterAnnotation,
  files = NULL,
  fileLabels = NULL,
  fileType = NULL,
  genome = NULL,
  numberOfCores = 1
)
```

# **Arguments**

promoterAnnotation

A PromoterAnnotation object containing the reduced exon ranges, annotated

intron ranges, promoter coordinates and the promoter id mapping

files A character vector. The list of junction or BAM files for which the junction read

counts will be calculated

fileLabels A character vector. The labels of junction or BAM files for which the junction

read counts will be calculated. These labels will be used as column names for

the output data.frame object

fileType A character. Type of the junction bed or bam file, either 'tophat', 'star' or 'bam

genome A character. Genome version used. Must be specified if input is a BAM file.

Defaults to NULL

numberOfCores

A numeric value. The number of cores to be used for counting junction reads. Defaults to 1 (no parallelization). This parameter will be used as an argument to BiocParallel::bplapply

# Value

A data frame object. The number of junction reads per promoter (rows) for each sample (cols)

```
getAbsolutePromoterActivity
```

Prepare the absolute promoter activity table including the promoter and gene ids

# **Description**

Prepare the absolute promoter activity table including the promoter and gene ids

# Usage

```
getAbsolutePromoterActivity(
   junctionReadCounts,
   promoterAnnotation,
   log2 = TRUE,
   pseudocount = 1
)
```

#### **Arguments**

junctionReadCounts

Matrix of junction read counts (rows: promoters, cols: samples)

promoterAnnotation

A PromoterAnnotation object containing the intron ranges, promoter coordi-

nates and the promoter id mapping

log2 Logical indicating whether log2 read counts should be used (default: TRUE) or

not

pseudocount Number to be used for log2 as pseudocount if log2 is TRUE

#### Value

data.frame of absolute promoter activity with promoter and gene ids

```
getAlternativePromoters
```

Identifies alternative promoters.

### **Description**

Identifies alternative promoters.

# Usage

```
getAlternativePromoters(
  result,
  referenceCondition,
 minAbs = 0.25,
 minRel = 0.05,
 maxPval = 0.05,
 promoterFC = 2,
 geneFC = 1.5
)
```

# **Arguments**

result

A SummarizedExperiment object with assays giving promoter counts, activity and gene expression (output from proActiv). rowData contains promoter metadata and absolute promoter activity summarized across conditions. Condition must be provided.

referenceCondition

A character vector. The reference condition to be compared. Samples corresponding to all other conditions will be compared to this samples in this current

condition.

A numeric value. Minimum value for promoter to be active in absolute terms. minAbs

Defaults to 0.25.

minRel A numeric value. Minimum value for promoter to be active in relative terms.

Defaults to 0.05.

maxPval A numeric value. Adjusted p-value threshold for detecting alternative promot-

ers. Defaults to 0.05.

A numeric value. Minimum fold change for a promoter in the current condition promoterFC

compared to all other conditions. Promoters must have at least this magnitude

of fold change for alternative usage.

geneFC A numeric value. Maximum fold change for gene expression. To identify al-

ternative promoter usage independent of changes in gene expression, limit the

gene expression fold change.

# Value

A list of length 2. Each entry is a dataframe summarizing up-regulated and down-regulated promoters and their corresponding genes, if any.

getGeneExpression 7

### **Examples**

getGeneExpression

Prepare the gene expression table including the gene ids

# Description

Prepare the gene expression table including the gene ids

# Usage

```
getGeneExpression(absolutePromoterActivity)
```

# **Arguments**

absolutePromoterActivity

data.frame of absolute promoter activity with promoter and gene ids

### Value

data.frame of gene expression with gene ids#'

```
getRelativePromoterActivity
```

Prepare the relative promoter activity table including the promoter and gene ids

# **Description**

Prepare the relative promoter activity table including the promoter and gene ids

# Usage

```
getRelativePromoterActivity(absolutePromoterActivity, geneExpression)
```

#### **Arguments**

```
absolutePromoterActivity
```

data.frame of absolute promoter activity with promoter and gene ids geneExpression data.frame of gene expression with gene ids

8 integrateProactiv

#### Value

data.frame of relative promoter activity with promoter and gene ids

integrateProactiv In

Integrate multiple proActiv runs

# **Description**

Integrate multiple proActiv runs

# Usage

```
integrateProactiv(res1, res2, ..., promoterAnnotation, renormalize = TRUE)
```

# Arguments

res1	A summarizedExperiment object returned by proActiv			
res2	A summarizedExperiment object returned by proActiv			
	$Additional\ summarized Experiment\ objects\ returned\ by\ proActiv$			
promoterAnnotation				
	Promoter annotation object used to create proActiv runs			
renormalize	Whether to renormalize counts after merging. Defaults to TRUE			

#### Value

A SummarizedExperiment object with assays giving promoter counts and activity with gene expression. rowData contains promoter metadata and absolute promoter activity summarized across conditions (if condition is provided)

normalize Promoter Read Counts

Normalize promoter read counts using DESeq2

# **Description**

Normalize promoter read counts using DESeq2

# Usage

normalizePromoterReadCounts(promoterReadCounts)

# **Arguments**

promoterReadCounts

A data.frame object. The number of junction reads per promoter (rows) for each sample (cols)

#### Value

A data.frame object. The normalized number of junction reads per promoter (rows) for each sample (cols) using DESeq2 counts function. Requires 'DESeq2' package to be installed

plotHeatmap

Visualizes heatmap of features for samples

# Description

Visualizes heatmap of features for samples

# Usage

```
plotHeatmap(
  result,
  by = "absolutePromoterActivity",
  features = NULL,
  cex.legend = 0.75,
  cex.row = NULL,
  cex.col = NULL,
  row.margin = 5,
  col.margin = 12,
  col = NULL,
  breaks = NULL,
  palette = "bluered"
```

10 plotHeatmap

#### **Arguments**

result A SummarizedExperiment object return by proActiv, with assays giving promoter counts and activity with gene expression stored as metadata. rowData contains promoter metadata and absolute promoter activity summarized across

conditions. Condition must be provided.

by A character vector. The assay to visualize the heatmap for. One of promoter-

Counts, normalizedPromoterCounts, absolutePromoterActivity and geneExpression (unambiguous substrings can be supplied). Defaults to absolutePromoter-

Activity.

features Features to visualize. Either a list of promoterIds or geneIds. The features must

correspond to the assay is used, i.e., if promoter assays are used, features must be promoterIds, while if gene expression assay is used, features must be geneIds.

Defaults to NULL (visualizes all features of the assay).

cex.legend A numeric value. Legend size.

cex.row A numeric value. Row label size.

cex.col A numeric value. Column label size.

row.margin A numeric value. Row margins.

col.margin A numeric value. Column margins.

col A vector of colours. Length should correspond to number of experimental con-

ditions. Defaults to NULL.

breaks A numeric vector. Breaks for heatmap plotting.

palette A character vector. One of bluered, redblue, redgreen, greenred. Defaults to

bluered.

# Value

Displays heatmap.

plotPCA 11

plotPCA

Performs principal component analysis

# **Description**

Performs principal component analysis

# Usage

```
plotPCA(
  result,
  by = "absolutePromoterActivity",
  main = NULL,
  col = NULL,
  alpha = 0.75,
  cex.size = 2
)
```

# **Arguments**

result A SummarizedExperiment object return by proActiv, with assays giving pro-

moter counts and activity with gene expression stored as metadata. rowData contains promoter metadata and absolute promoter activity summarized across

conditions. Condition must be provided.

by A character vector. The assay to perform principal component analysis by. One

of promoterCounts, normalizedPromoterCounts, absolutePromoterActivity and geneExpression (unambiguous substrings can be supplied). Defaults to abso-

lutePromoterActivity.

main A character vector. Plot title (optional). Defaults to NULL.

col A vector of colours. If NULL, uses standard ggplot colours. Defaults to NULL.

alpha A numeric value in between 0 and 1. Determines point transparency.

cex.size A numeric value. Determines point size.

# Value

PCA plot.

12 proActiv

```
preparePromoterAnnotation
```

Prepares promoter annotation from a gtf or txdb

# Description

Prepares promoter annotation from a gtf or txdb

# Usage

```
preparePromoterAnnotation(txdb, file, species)
```

# **Arguments**

txdb	A txdb object. The txdb of the annotation version for which promoters will be identified. Either 'txdb' or 'file' argument must be specified, but not both.
file	A character object. The file path to a gtf/gff or txdb of the annotation version for which promoters will be identified. Either 'txdb' or 'file' argument must be specified, but not both.
species	A character object. The genus and species of the organism to be used in keep-StandardChromosomes(), Supported species can be seen with names(genomeStyles()),

# Value

A PromoterAnnotation object. The annotated intron ranges, promoter coordinates and the promoter id mapping are attributes of the promoter annotation data.

# Examples

proActiv

Estimates promoter counts and activity in a single command

# Description

Estimates promoter counts and activity in a single command

proActiv 13

#### Usage

```
proActiv(
   files,
   promoterAnnotation,
   fileLabels = NULL,
   condition = NULL,
   genome = NULL,
   ncores = 1
)
```

### Arguments

files A character vector. The list of input files for which the junction read counts will

be calculated

promoterAnnotation

A PromoterAnnotation object containing the intron ranges, promoter coordi-

nates and the promoter id mapping

fileLabels A character vector. The labels of input files for which the junction read counts

will be calculated. These labels will be used as column names for each output data.frame object. If not provided, filenames will be used as labels. Defaults to

NULL

condition A character vector. The condition to which each sample belong to. Must corre-

spond to the order of the files. If supplied, results are summarized by condition.

Defaults to NULL

genome A character. Genome version. Must be specified if input file type is a BAM file.

Defaults to NULL

ncores A numeric value. The number of cores to be used for counting junction reads.

Defaults to 1 (no parallelization). This parameter will be used as an argument to

BiocParallel::bplapply

### Value

A SummarizedExperiment object with assays giving promoter counts and activity with gene expression. rowData contains promoter metadata and absolute promoter activity summarized across conditions (if condition is provided)

14 PromoterAnnotation-class

PromoterAnnotation-class

S4 class for promoter annotation data for a specific annotation version

#### **Description**

S4 class for promoter annotation data for a specific annotation version

# Usage

```
PromoterAnnotation(
  intronRanges = GenomicRanges::GRanges(),
  promoterIdMapping = data.frame(),
  promoterCoordinates = GenomicRanges::GRanges()
)
intronRanges(x)
## S4 method for signature 'PromoterAnnotation'
intronRanges(x)
promoterIdMapping(x)
## S4 method for signature 'PromoterAnnotation'
promoterIdMapping(x)
promoterCoordinates(x)
## S4 method for signature 'PromoterAnnotation'
promoterCoordinates(x)
intronRanges(x) <- value</pre>
## S4 replacement method for signature 'PromoterAnnotation'
intronRanges(x) <- value</pre>
promoterIdMapping(x) <- value</pre>
## S4 replacement method for signature 'PromoterAnnotation'
promoterIdMapping(x) <- value</pre>
promoterCoordinates(x) <- value</pre>
## S4 replacement method for signature 'PromoterAnnotation'
promoterCoordinates(x) <- value</pre>
```

# Arguments

intronRanges A GRanges object containing annotated intron ranges

PromoterAnnotation-class 15

promoterIdMapping

A data.frame containing mapping between transcript, TSS, promoter and gene ids

promoterCoordinates

A GRanges object containing promoter coordinates

x A PromoterAnnotation object

value intronRanges, promoterIdMapping or promoterCoordinates to be assigned

#### Value

A promoter annotation object with three slots: intronRanges, promoterIdMapping and promoter Coordinates

#### **Functions**

- intronRanges: Getter for intronRanges
- intronRanges, PromoterAnnotation-method: Getter for intronRanges
- promoterIdMapping: Getter for promoterIdMapping
- promoterIdMapping, PromoterAnnotation-method: Getter for promoterIdMapping
- promoterCoordinates: Getter for promoterCoordinates
- $\bullet \ \, \text{promoterCoordinates}, \\ \text{PromoterAnnotation-method: Getter for promoterCoordinates}$
- intronRanges<-: Setter for intronRanges
- intronRanges<-, PromoterAnnotation-method: Setter for intronRanges
- promoterIdMapping<-: Setter for promoterIdMapping
- promoterIdMapping<-, PromoterAnnotation-method: Setter for promoterIdMapping
- promoterCoordinates<-: Setter for promoterCoordinates
- promoterCoordinates<-, PromoterAnnotation-method: Setter for promoterCoordinates

### **Slots**

intronRanges A GRanges object. The intron ranges annotated with the promoter information.

promoterIdMapping A data.frame object. The id mapping between transcript ids, names, TSS ids, promoter ids and gene ids.

promoterCoordinates A GRanges object. Promoter coordinates (TSS) with gene id and internal promoter state

16 reexports

```
promoterAnnotation.gencode.v34.subset
```

Promoter annotation for Gencode.v34 (subset)

# **Description**

Promoter annotation for Gencode.v34 (chr1:10,000,000 - 30,000,000)

# Usage

promoterAnnotation.gencode.v34.subset

#### **Format**

A PromoterAnnotation (S4 Class) object containing all promoter annotation objects for Gencode.v34 chr1:10,000,000-30,000,000. The object has 3 slots:

**intronRanges** A GRanges object of 4,523 ranges corresponding to introns, annotated with the associated transcript.

**promoterIdMapping** The id mapping between transcript names, promoter ids and gene ids for Gencode v34.

**promoterCoordinates** A GRanges object of 1,380 ranges showing the tss coordinate for each promoter of Gencode v34 chr1:10,000,000-30,000,000, annotated with the associated gene id, coordinate of the 3' end of the first reduced exon, and intron id.

reexports

Objects exported from other packages

# **Description**

These objects are imported from other packages. Follow the links below to see their documentation.

AnnotationDbi loadDb

dplyr %>%, arrange, as\_tibble, filter, group\_by, mutate, n

SummarizedExperiment assays, colData, rowData

tibble rownames\_to\_column

# **Index**

```
* datasets
                                                intronRanges<-,PromoterAnnotation-method</pre>
    promoterAnnotation.gencode.v34.subset,
                                                         (PromoterAnnotation-class), 14
                                                loadDb, 16
* internal
                                                loadDb (reexports), 16
    reexports, 16
'intronRanges<-',PromoterAnnotation-method
                                                mutate, 16
        (PromoterAnnotation-class), 14
'promoterCoordinates<-',PromoterAnnotation-method (reexports), 16
        (PromoterAnnotation-class), 14
'promoterIdMapping<-', PromoterAnnotation-method n (reexports), 16
        (PromoterAnnotation-class), 14
                                                normalizePromoterReadCounts, 9
%>% (reexports), 16
%>%, 16
                                                plotHeatmap, 9
                                                plotPCA, 11
arrange, 16
                                                preparePromoterAnnotation, 12
arrange (reexports), 16
                                                proActiv, 12
as_tibble, 16
                                                PromoterAnnotation
as_tibble (reexports), 16
                                                         (PromoterAnnotation-class), 14
assays, 16
                                                PromoterAnnotation-class, 14
assays (reexports), 16
                                                promoterAnnotation.gencode.v34.subset,
                                                         16
boxplotPromoters, 2
                                                promoterCoordinates
                                                         (PromoterAnnotation-class), 14
calculateJunctionReadCounts, 3
                                                {\tt promoterCoordinates, PromoterAnnotation-method}
calculatePromoterReadCounts, 4
                                                         (PromoterAnnotation-class), 14
colData, 16
                                                promoterCoordinates<-
colData (reexports), 16
                                                         (PromoterAnnotation-class), 14
filter, 16
                                                promoterCoordinates<-,PromoterAnnotation-method</pre>
filter (reexports), 16
                                                         (PromoterAnnotation-class), 14
                                                promoterIdMapping
getAbsolutePromoterActivity, 5
                                                         (PromoterAnnotation-class), 14
getAlternativePromoters, 6
                                                promoterIdMapping,PromoterAnnotation-method
getGeneExpression, 7
                                                         (PromoterAnnotation-class), 14
getRelativePromoterActivity, 7
                                                promoterIdMapping<-</pre>
group_by, 16
                                                         (PromoterAnnotation-class), 14
group_by (reexports), 16
                                                promoterIdMapping<-,PromoterAnnotation-method</pre>
                                                         (PromoterAnnotation-class), 14
integrateProactiv, 8
intronRanges
                                                reexports, 16
        (PromoterAnnotation-class), 14
                                                rowData, 16
intronRanges, PromoterAnnotation-method
                                                rowData (reexports), 16
        (PromoterAnnotation-class), 14
                                                rownames_to_column, 16
intronRanges<-
                                                rownames_to_column (reexports), 16
        (PromoterAnnotation-class), 14
```